



# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse III - Paul Sabatier  
Discipline ou spécialité : *Traitement du signal et des images*

---

Présentée et soutenue par  
*Emma Villeneuve*

Le 10 Décembre 2012

**Titre :**  
*Déconvolution de données hyperspectrales  
pour l'instrument MUSE du VLT*

---

### JURY

*Yannick Deville  
Jean-François Giovannelli  
Laurent Mugnier  
Eric Slezak  
Hervé Carfantan  
Eric Anterrieu*

*Professeur, IRAP  
Professeur, IMS  
Maître de Recherche, ONERA  
Astronome, Lagrange  
Maître de Conférence, IRAP  
Ingénieur de Recherche, IRAP*

*Président  
Rapporteur  
Rapporteur  
Examineur  
Directeur de Thèse  
Directeur de Thèse*

---

**Ecole doctorale :** *Mathématiques, Informatique et Télécommunications de Toulouse*

**Unité de recherche :** *IRAP - Institut de Recherche en Astrophysique et Planétologie*

**Directeurs de Thèse :** *Hervé Carfantan, Eric Anterrieu*



# Remerciements

J'aimerais tout d'abord remercier Hervé Carfantan, mon directeur de thèse qui m'a encadrée pendant ces trois années. Merci pour ta passion de la recherche, tes nombreux conseils, tant sur le plan de la thèse que sur celui de l'enseignement. J'ai sincèrement apprécié ton implication et ta grande disponibilité. Je te remercie de m'avoir embarquée dans cette épopée. Je remercie également mon co-directeur de thèse Eric Anterrieu pour son aide précieuse à la rédaction et à la préparation de la soutenance.

Ensuite, je souhaite exprimer ma reconnaissance envers les membres du jury de thèse. Je tiens à remercier Yannick Deville d'avoir accepté la fonction de Président du jury, ainsi que pour les nombreux conseils qu'il a bien voulu me prodiguer, tout au long de ma thèse. Je remercie également mes rapporteurs Jean-François Giovannelli et Laurent Mugnier pour l'attention qu'ils ont bien voulu porter à mon manuscrit, ainsi que pour leurs questions et conseils lors de la soutenance. Enfin, j'adresse mes remerciements à Eric Slezak, responsable du projet DAHLIA, pour sa collaboration dans le cadre de l'ANR ainsi que pour sa participation au jury de thèse.

Je garderai un excellent souvenir de toutes les personnes que j'ai côtoyées durant ces trois années de thèse. Mes pensées se tournent vers Inès, ma colocataire de bureau et mon amie. Merci pour ta bonne humeur, tes conseils et ton soutien. Tu t'es montrée présente lorsque j'en ai eu besoin et je t'en suis reconnaissante. Un grand merci à tous les membres de l'équipe SISU et anciens membres de l'équipe S2I : Laurent, Sylvie, Ahmed, Shahram, Jean-François, Karine, José, Djaouad, Jean-Louis, Khalid, Chahinez et Guilhem. J'ai beaucoup apprécié de partager ces moments avec vous ; gardez cette bonne atmosphère qui fait la force de l'équipe ! Merci Camille pour l'énergie que tu as consacrée à l'adaptation de mon travail en Python pour le projet MUSE. J'aimerais saluer la disponibilité et la réactivité des membres de l'équipe informatique de l'IRAP, en particulier Loïc et Nicolas qui m'ont souvent dépannée dans l'urgence. Merci également à Christine pour son sourire et sa sympathie qui rendent les journées plus faciles.

Je remercie les personnes impliquées dans le projet MUSE avec lesquelles j'ai pu échanger et collaborer, en particulier Denis Serre, Thierry Contini, Roland Bacon, Nicolas Bouché, Jihane Moulta et Benoît Epinat. Ce fut pour moi très enrichissant de pouvoir travailler avec vous, de partager un point de vue différent et d'acquérir des connaissances en astrophysique. Grâce à vous, j'ai pu saisir l'utilité de mon travail en le mettant en application et cela a été une réelle motivation. Je remercie les membres de l'ANR DAHLIA avec lesquels j'ai pris plaisir à travailler, notamment Vincent Mazet et Sébastien Bourguignon que j'ai eu la chance de rencontrer à de nombreuses reprises.

J'aimerais exprimer ma sincère gratitude envers mes parents qui m'ont soutenue durant ces longues années d'étude. C'est la fin d'une histoire, mais c'est surtout le début d'une autre ! Merci d'avoir cru en moi et de m'avoir permis de réaliser ce rêve.

*Last but not least*, merci Alex pour ton indéfectible soutien quotidien, dans les bons et les mauvais moments, des premiers jours jusqu'à la dernière minute. Je te dédie mes étoiles.



# Introduction

Notre compréhension de l'univers s'appuie avant tout sur la comparaison entre les observations et les modèles physiques des astres. En particulier, l'étude des composants chimiques et de la cinématique des étoiles, des galaxies, etc, a été rendue possible par l'invention de la spectroscopie. Avec la construction d'instruments hyperspectraux, tels que le spectro-imageur MUSE (pour *Multi-Unit Spectroscopic Explorer*), c'est-à-dire capables d'acquérir simultanément des images à un grand nombre de longueurs d'onde, le traitement et l'analyse des données deviennent un challenge pour la communauté signal-image. Pour l'analyse et le traitement de données hyperspectrales, une des difficultés réside dans la réponse impulsionnelle 3D du système ou fonction d'étalement (*Point Spread Function* : PSF), à la fois spatiale et spectrale.

L'objectif initial de cette thèse était d'aborder le problème de la déconvolution de données hyperspectrales de type MUSE, ce qui nécessite au préalable la modélisation et l'estimation de la PSF. Notons que la connaissance de cette PSF est également importante pour d'autres types de traitements que la déconvolution, par exemple la fusion de cubes de données.

Pour commencer, nous situerons le contexte dans lequel a été réalisée cette thèse, du point de vue de l'instrument et du traitement de données. Le premier chapitre présentera les domaines de l'imagerie multi-spectrale et de la spectroscopie, ainsi que l'instrument MUSE qui est au centre de nos travaux. Nous introduirons également la notion de fonction d'étalement et nous détaillerons l'influence des différents composants de la chaîne d'acquisition de MUSE sur la qualité des images. Les différents thèmes de recherche du projet DAHLIA (*Dedicated Algorithms for HyperspectraL Imaging in Astronomy*) auquel nous avons participé mettront en avant la nécessité de modéliser et d'estimer la PSF spatiale.

Dans le second chapitre, nous nous intéresserons à la modélisation de la PSF spatiale en tout point du cube de données hyperspectrales. Sans correction par optique adaptative (OA), cette PSF spatiale, essentiellement due aux turbulences atmosphériques, varie uniquement en fonction de la longueur d'onde, alors qu'avec OA, la variation est également spatiale. À partir de simulations numériques de la PSF de l'atmosphère et du télescope, nous étudierons ces variations spectrales et spatiales de la PSF spatiale que nous approximerons par une fonction de Moffat. Puis nous proposerons une modélisation analytique de la variation des paramètres de cette fonction.

Nous étudierons, dans le troisième chapitre, l'estimation de la PSF spatiale lorsque celle-ci n'est pas corrigée par l'optique adaptative. Cette étude sera conduite pour l'ensemble des longueurs d'onde, à partir de l'observation d'une étoile isolée. Pour cela, nous exploiterons le modèle de variation de la PSF développé dans le chapitre 2. Puis nous nous pencherons sur le problème de l'estimation des paramètres de nuisance que sont le spectre de l'étoile et la variance du bruit. Ainsi, nous proposerons des schémas d'estimation pour l'ensemble de ces paramètres que nous évaluerons en fonction de la brillance de l'étoile. Notre objectif est de tirer profit de la nature hyperspectrale des données afin d'optimiser les performances d'estimation, en particulier pour une étoile faiblement brillante.

Enfin, nous nous concentrerons sur le problème particulièrement délicat de la déconvolution hyperspectrale pour l'étude de la cinématique des galaxies. Par effet Doppler, ce problème se traduit en l'estimation de cartes de flux, de position et de largeur d'une raie spectrale d'émission de la galaxie. Nous étudierons l'estimation de ces cartes en considérant les pixels spatiaux indépendamment, c'est-à-dire sans prendre en compte la PSF 3D, puis l'estimation de ces cartes par déconvolution, en prenant en compte la PSF.

Cette thèse s'inscrit dans une approche de traitement de signal avec pour objectif le développement de nouvelles méthodes de traitement de données hyperspectrales. Ces méthodes seront appliquées sur des données simulées de l'instrument MUSE. Dans cette thèse, nous aborderons les divers problèmes dans le cadre de l'estimation et de l'inférence statistique. Une telle approche nécessite de décrire précisément les modèles liant les données aux phénomènes observés, les statistiques des perturbations intervenant sur les données ainsi que les informations *a priori* disponibles sur les paramètres à estimer. Nous utiliserons notamment l'estimateur du maximum de vraisemblance et celui de l'espérance *a posteriori*, dans un cadre bayésien. Le calcul des paramètres estimés nécessitera la mise en œuvre d'outils d'optimisation et d'échantillonnage stochastique dont les simulations de Monte-Carlo par chaînes de Markov.

# Table des matières

<b>Introduction</b>	<b>2</b>
<b>Abréviations - Acronymes</b>	<b>9</b>
<b>I Acquisition de données astronomiques hyperspectrales</b>	<b>11</b>
I.1 L'observation hyperspectrale . . . . .	12
I.1.1 L'imagerie multi-spectrale . . . . .	12
I.1.2 Quelques notions de spectroscopie . . . . .	12
I.1.3 Le spectro-imageur MUSE . . . . .	16
I.2 L'observation astronomique au sol : MUSE . . . . .	18
I.2.1 Effets de l'atmosphère . . . . .	19
I.2.2 Effet du télescope . . . . .	20
I.2.3 Optique Adaptative . . . . .	21
I.2.4 Effet de l'instrument . . . . .	22
I.2.5 Logiciel de réduction des données . . . . .	22
I.2.6 Bilan et approximations pour la PSF globale . . . . .	22
I.3 Le projet DAHLIA . . . . .	25
I.3.1 Simulation de données . . . . .	25
I.3.2 Fusion de données . . . . .	25
I.3.3 Estimation du fond du ciel . . . . .	26
I.3.4 Séparation de sources . . . . .	26
I.3.5 Détection de sources, segmentation et déconvolution . . . . .	27
I.3.6 Visualisation de données hyperspectrales . . . . .	27
I.3.7 Modélisation et estimation de la PSF . . . . .	27
<b>II Modélisation de la FSF</b>	<b>29</b>
Notations . . . . .	30
Introduction . . . . .	31
II.1 Méthodologie . . . . .	31
II.1.1 FSF simulées par PAOLA . . . . .	32
II.1.2 Indices d'erreur . . . . .	32
II.2 Modèles courants pour les PSF spatiales . . . . .	33
II.2.1 Modèles gaussien et Moffat . . . . .	33
II.2.2 Choix de la normalisation . . . . .	34
II.2.3 Fonction de Moffat elliptique . . . . .	35
II.2.4 Paramètres de la fonction de Moffat . . . . .	35
II.3 Modélisation de la FSF en mode WFM sans OA . . . . .	38
II.3.1 Prise en compte de la FSF du télescope . . . . .	38
II.3.2 Modèle elliptique à <i>seeing</i> fixé . . . . .	39
II.3.3 Modèles circulaires à <i>seeing</i> fixé . . . . .	40

II.3.4	Modèles circulaires en fonction du <i>seeing</i> . . . . .	40
II.3.5	Influence du $C_n^2$ . . . . .	42
II.3.6	Influence de l'échelle externe . . . . .	43
II.3.7	Synthèse . . . . .	44
II.4	Modélisation de la FSF en mode WFM avec optique adaptative . . . . .	44
II.4.1	Données et hypothèses . . . . .	44
II.4.2	Modèle circulaire constant dans le champ . . . . .	46
II.4.3	Modèle elliptique variable dans le champ . . . . .	48
Conclusions	. . . . .	55
<b>III</b>	<b>Estimation de la FSF</b> . . . . .	<b>57</b>
Notations	. . . . .	58
Introduction	. . . . .	59
III.1	Modèle de données . . . . .	59
III.1.1	FSF . . . . .	59
III.1.2	FSF simulées . . . . .	60
III.1.3	Simulations de données réalistes . . . . .	60
III.1.4	Magnitude et rapport signal sur bruit . . . . .	62
III.1.5	Méthodologie . . . . .	62
III.2	Estimation par maximum de vraisemblance . . . . .	64
III.2.1	Estimation des hyper-paramètres . . . . .	64
III.2.2	Estimations des paramètres de nuisance . . . . .	65
III.2.3	Résultats d'expérimentations . . . . .	67
III.3	Amélioration de l'estimation des paramètres de nuisance . . . . .	71
III.3.1	Normalisation des données par le spectre . . . . .	72
III.3.2	Moyenne non pondérée . . . . .	72
III.3.3	Moyenne pondérée par la variance du bruit . . . . .	74
III.3.4	Résultats d'expérimentations . . . . .	75
III.3.5	Choix de l'intervalle moyenné . . . . .	76
III.4	Estimation jointe . . . . .	77
III.5	Estimation par espérance <i>a posteriori</i> . . . . .	79
III.5.1	Modèle bayésien . . . . .	80
III.5.2	Méthode d'échantillonnage de Metropolis-Hastings à marche aléatoire . . . . .	81
III.5.3	Estimation conjointe du spectre de l'étoile et de la FSF . . . . .	82
III.5.4	Résultats d'expérimentations . . . . .	83
III.6	Synthèse des méthodes d'estimation du spectre . . . . .	83
III.7	Estimation à partir d'une étoile non isolée . . . . .	84
III.7.1	Données . . . . .	84
III.7.2	FSF de référence . . . . .	84
III.7.3	Masques . . . . .	88
III.7.4	Résultats d'expérimentations . . . . .	89
Conclusions	. . . . .	91
<b>IV</b>	<b>Déconvolution de données hyperspectrales pour la cinématique des galaxies</b> . . . . .	<b>93</b>
Notations	. . . . .	94
Introduction	. . . . .	95
IV.1	Données . . . . .	95
IV.1.1	Modèle d'objet . . . . .	95
IV.1.2	Modèle de PSF . . . . .	97
IV.1.3	Modèle de données . . . . .	97
IV.1.4	Estimation et déconvolution . . . . .	98

IV.2 Méthodes usuelles d'estimation de cartes de cinématique . . . . .	98
IV.2.1 Méthode des moments . . . . .	99
IV.2.2 Estimation par maximum de vraisemblance . . . . .	100
IV.3 Estimation par espérance <i>a posteriori</i> . . . . .	101
IV.3.1 Lois <i>a priori</i> . . . . .	101
IV.3.2 Échantillonnage stochastique de la loi <i>a posteriori</i> . . . . .	102
IV.3.3 Échantillonneur de Gibbs et lois conditionnelles . . . . .	102
IV.3.4 Choix des lois de proposition pour l'échantillonnage de Metropolis-Hastings . .	104
IV.4 Déconvolution de cartes de cinématique . . . . .	105
IV.4.1 Maximum de vraisemblance conjointe . . . . .	105
IV.4.2 Échantillonnage stochastique de la loi <i>a posteriori</i> . . . . .	106
IV.4.3 Lois <i>a posteriori</i> conditionnelles . . . . .	106
IV.4.4 Déconvolution par espérance <i>a posteriori</i> conjointe . . . . .	108
IV.5 Expérimentations sur données 1D1D . . . . .	108
IV.5.1 Description des données . . . . .	108
IV.5.2 Méthodologie . . . . .	109
IV.5.3 Indicateurs de convergence pour PM et JPM . . . . .	109
IV.5.4 Résultats d'estimation . . . . .	112
IV.5.5 Résultats de déconvolution . . . . .	112
IV.6 Expérimentations sur données 3D . . . . .	113
IV.6.1 Description des données . . . . .	113
IV.6.2 Méthodologie . . . . .	114
IV.6.3 Résultats d'estimation . . . . .	114
IV.6.4 Résultats de déconvolution . . . . .	115
Conclusions . . . . .	125
<b>Conclusion générale</b>	<b>127</b>
<b>Bibliographie</b>	<b>132</b>
<b>Publications et communications de l'auteur</b>	<b>139</b>
<b>Annexes</b>	<b>141</b>
<b>A Compléments astrophysiques</b>	<b>143</b>
A.1 Spectres de galaxies . . . . .	143
A.2 Influence du mouvement des galaxies sur leur spectre . . . . .	143
A.2.1 Effet Doppler-Fizeau . . . . .	143
A.2.2 Décalage spectral . . . . .	144
<b>B Mesures de la qualité d'image pour l'astronomie</b>	<b>145</b>
B.1 Notions de résolution . . . . .	145
B.1.1 Résolution spatiale . . . . .	145
B.1.2 Résolution spectrale . . . . .	145
B.2 Les critères de qualité d'une observation . . . . .	146
B.2.1 Rapport de Strehl . . . . .	146
B.2.2 Énergie encerclée . . . . .	146
B.2.3 Paramètre de Fried . . . . .	146
B.2.4 Seeing . . . . .	146
B.2.5 Rapport signal sur bruit . . . . .	146

<b>C</b>	<b>Modélisation de la FSF</b>	<b>149</b>
C.1	Compléments pour la modélisation de la FSF en mode WFM avec optique adaptative	149
C.2	Influence du <i>seeing</i> sur la FSF en mode WFM avec optique adaptative . . . . .	150
<b>D</b>	<b>Compléments méthodologiques</b>	<b>153</b>
D.1	Méthode des moments : extension . . . . .	153
D.2	Intégration du spectre de l'étoile pour l'estimation de la FSF par espérance <i>a posteriori</i>	154

# Abréviations - Acronymes

**OA** optique adaptative : système de correction par miroir déformable des effets de la turbulence atmosphérique dans l'observation astronomique depuis le sol.

**DAHLIA** *Dedicated Algorithms for HyperspectraL Imaging in Astronomy*, projet n°08-BLAN-0253-01 financé par l'ANR de 2009 à 2012.

**DRS** *Data Reduction Software* : logiciel de réduction des données arrivant en fin de chaîne d'acquisition des données MUSE ; il permet notamment de construire le cube de données hyperspectrales à partir des acquisitions sur les CCD.

**FSF** *Field Spread Function* : fonction d'étalement spatial.

**FWHM** *Full Width at Half Maximum* : largeur totale à mi-hauteur.

**GLAO** *Ground Layer Adaptive Optics* : système de correction par optique adaptative utilisé dans le cas du mode grand champ de MUSE (WFM) ; il corrige la couche turbulente de faible altitude.

**LSF** *Line Spread Function* : fonction d'étalement spectral.

**MUSE** *Multi-Unit Spectroscopic Explorer* : spectro-imageur de seconde génération du VLT. Première lumière prévue en 2013.

**NFM** *Narrow Field Mode* : mode petit champ de l'instrument MUSE, permettant d'observer un champ de dimension  $7.5'' \times 7.5''$ .

**PSF** *Point Spread Function* : fonction d'étalement d'un point, réponse impulsionnelle.

**pxl** pixel

**QSIM** logiciel de simulation permettant de simuler des cubes de données d'observation MUSE ; la LSF est le résultat de la convolution d'une fonction de Gauss-Hermite par une porte de largeur  $1.35 \text{ \AA}$  ; dans le cas WFM sans OA, la FSF est une fonction de Moffat circulaire variable en longueur d'onde ; le spectre du ciel est formé d'un continuum et de raies d'émission ; le bruit est poissonnien.

**RSB** rapport signal sur bruit (cf. section B.2.5).

**WFM** *Wide Field Mode* : mode grand champ de l'instrument MUSE, permettant d'observer un champ de dimension  $60'' \times 60''$ .

'' seconde d'arc.



# Chapitre I

## Acquisition de données astronomiques hyperspectrales

### Sommaire

---

<b>I.1</b>	<b>L’observation hyperspectrale</b>	<b>12</b>
I.1.1	L’imagerie multi-spectrale	12
I.1.2	Quelques notions de spectroscopie	12
I.1.3	Le spectro-imageur MUSE	16
<b>I.2</b>	<b>L’observation astronomique au sol : MUSE</b>	<b>18</b>
I.2.1	Effets de l’atmosphère	19
I.2.2	Effet du télescope	20
I.2.3	Optique Adaptative	21
I.2.4	Effet de l’instrument	22
I.2.5	Logiciel de réduction des données	22
I.2.6	Bilan et approximations pour la PSF globale	22
<b>I.3</b>	<b>Le projet DAHLIA</b>	<b>25</b>
I.3.1	Simulation de données	25
I.3.2	Fusion de données	25
I.3.3	Estimation du fond du ciel	26
I.3.4	Séparation de sources	26
I.3.5	Détection de sources, segmentation et déconvolution	27
I.3.6	Visualisation de données hyperspectrales	27
I.3.7	Modélisation et estimation de la PSF	27

---

Ce premier chapitre pose le contexte dans lequel a été réalisée cette thèse, du point de vue de l’instrument mais aussi de celui du traitement des données. La première partie présente l’instrument MUSE qui est au centre de nos travaux et le situe à la croisée des domaines de l’imagerie multi-spectrale et de la spectroscopie. Nous introduisons ensuite la notion de fonction d’étalement (ou PSF pour *Point Spread Function*) liée à la qualité des images, notamment par la résolution spatiale et spectrale. Nous détaillons l’effet des différents composants de la chaîne d’acquisition de MUSE sur la qualité des images, en termes de PSF. Enfin, nous présentons les différents thèmes de recherche du projet DAHLIA, auquel nous avons participé. Certains de ces thèmes d’analyse et de traitements de données nécessitent de connaître la PSF, que nous étudions dans les Chapitres II et III.

## I.1 L'observation hyperspectrale

### I.1.1 L'imagerie multi-spectrale

L'imagerie multi-spectrale est pratiquée au quotidien, notamment par l'usage des appareils photos numériques, grâce à l'utilisation de filtres de couleurs, dont les plus communs sont le rouge, le vert et le bleu (RGB). Généralement organisés selon un schéma matriciel représenté par la Fig. I.1, ces filtres permettent d'observer une scène dans des bandes spectrales spécifiques. L'imagerie multi-spectrale diffère donc de l'observation panchromatique où chaque pixel intègre l'ensemble des longueurs d'onde visibles. Toutefois, ces deux modes d'observation sont couramment combinés, notamment dans le domaine de la télédétection, afin d'exploiter la haute résolution apportée par l'image panchromatique et les informations spectrales apportées par les filtres RGB.

Avec la multiplication des satellites d'observation de la Terre, les filtres spectraux se sont diversifiés. Ainsi, aux filtres standards RGB s'ajoutent le proche infra-rouge (NIR - *Near InfraRed*), l'infrarouge moyen (*Middle IR*) et l'infrarouge lointain ou thermique (*Far IR*). En effet, la motivation est de détecter des informations qui ne sont pas perceptibles dans le visible, telles que la salinité de l'eau ou sa profondeur. On peut citer le satellite de télédétection LANDSAT 7, qui observe la Terre à travers 7 bandes (RGB, NIR, MIR, FIR) et en panchromatique.

Dans le domaine de l'observation astrophysique, la caméra FOCAM du Canada-France-Hawaii Telescope (CFHT) explore le ciel à travers 5 filtres. Depuis le *Hubble Space Telescope*, l'imageur WFC3 (*Wide Field Camera 3*) ausculte les astres, avec une précision spectrale largement accrue puisqu'il est équipé de 80 filtres dans l'ultra-violet, le visible et l'infrarouge.

Ainsi, on voit apparaître des instruments détectant plusieurs centaines de longueurs d'onde différentes. À titre d'exemple, on peut citer l'imageur de type *pushbroom* AISA de Channel Systems, pouvant détecter jusqu'à 244 bandes spectrales, et l'instrument CASI-1500 de ITRES doté de 288 canaux spectraux. Les performances de l'imagerie multi-spectrale se rapprochent alors progressivement de celles obtenues dans le domaine de la spectroscopie.

### I.1.2 Quelques notions de spectroscopie

En 1802, le physicien anglais William Hyde Wollaston observe pour la première fois plusieurs raies noires dans le spectre du Soleil. Puis en 1814, Joseph von Fraunhofer, physicien allemand, invente le spectroscopie, constitué d'un prisme doté des graduations qui correspondent aux longueurs d'onde. Peu après, il réalise la cartographie méthodique des raies du spectre solaire, dont il mesure la longueur d'onde. Il faut attendre 1860 pour que Kirchhoff et Bunsen [KIR 61] conçoivent un spectroscopie moderne, qui leur permet de découvrir deux nouveaux métaux, le césium et le rubidium. Ce spectroscopie, illustré par la Fig. I.2 est composé d'un collimateur (B) constitué d'une fente placée au foyer objet d'une lentille, puis d'un prisme rotatif (F) séparant les rayons lumineux, et enfin d'un télescope (C) permettant d'observer le spectre [WEE 32].

De manière générale, un spectromètre<sup>2</sup> est un instrument capable de décomposer une onde en éléments simples et de mesurer leur intensité. Ces éléments sont exprimés par des énergies, des fréquences, des

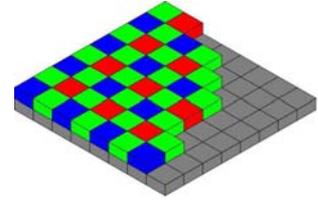


FIG. I.1 – Filtre de Bayer placé devant un capteur CCD, composé d'une matrice de filtres rouges, verts et bleus, permettant de prendre des photographies en couleur.<sup>1</sup>

<sup>1</sup>Figure de Colin M.L. Burnett.

<sup>2</sup>D'après C.R.Kitchin [KIT 95], un spectroscopie est un instrument destiné à afficher le spectre d'une source ; on parle d'un spectrographe lorsqu'une plaque photographique est utilisée pour détecter le spectre ; on parle de spectromètre lorsque l'instrument est associé à tout autre type de détecteur, tel que les CCD. Le terme spectromètre est donc plus général.

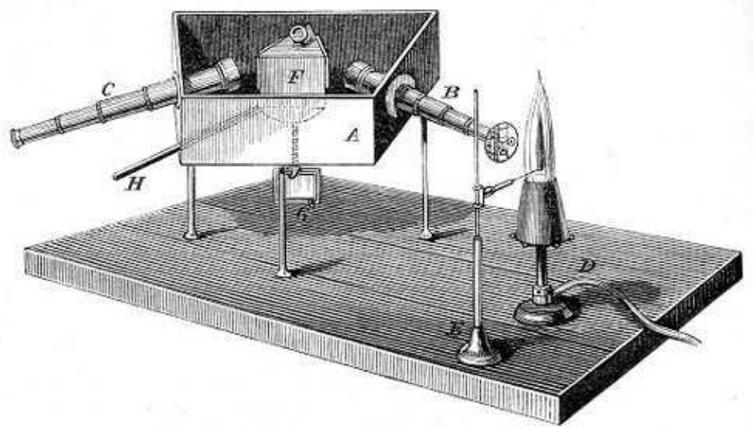


FIG. I.2 – Spectroscopie de Kirchhoff et Bunsen [KIR 61].

électrons-volts ou des longueurs d'onde en fraction de mètre, ce qui est le cas dans ce manuscrit. Chaque bande du spectre électromagnétique (ondes radio, micro-ondes, infrarouge, visible, ultraviolet, rayons X, rayons  $\gamma$ ) a suscité la création d'une technologie de spectroscopie spécifique. Dans le cadre de cette thèse, nous nous intéressons uniquement à la séparation des ondes du visible et du proche infrarouge. Le pouvoir de séparation d'un spectrographe est appelé résolution spectrale (définie dans l'annexe B.1). Il peut être quantifié par le ratio  $R = \frac{\lambda}{\Delta\lambda}$  [LEN 09], sans dimension, et dépend donc de la longueur d'onde. Les champs d'application variés tels que la chimie, la télédétection spatiale et l'astronomie, ont engendré l'apparition de différents modes d'observation, gérant de différentes manières les domaines spatiaux et spectraux. On peut ainsi distinguer [LEN 08] :

**la spectrométrie séquentielle** : le spectre d'une source unique est balayé et détecté par un détecteur simple ; ce type de spectrométrie ne fournit pas d'information spatiale. Les spectroscopes de Fraunhofer et Kirchhoff-Bunsen en font partie.

**la spectrométrie multi-canaux** : un élément dispersif sépare les différents éléments du spectre observé, qui sont détectés par un récepteur doté de plusieurs canaux, tel une barrette CCD. Ce mode se distingue de la spectrométrie séquentielle par l'absence de balayage et par la simultanéité de la détection des différentes parties du spectre.

**la spectrométrie multi-objets** : (*MOS Multiple Object Spectroscopy*) les spectres de plusieurs sources sélectionnées dans le champ<sup>3</sup> sont simultanément détectés. Cette stratégie a été mise en place avec l'augmentation du nombre de catalogues d'étoiles et de galaxies. La sélection spatiale des objets peut être réalisée par un masque à fente ou par des fibres optiques. Cette méthode a l'avantage de fournir indirectement une information spatiale associée à chaque spectre, en sélectionnant un objet précis. Le fait d'acquérir en même temps le spectre de plusieurs sources permet de les calibrer, vis-à-vis des conditions d'observation. Voici quelques exemples de spectromètres multi-objets utilisés en Astrophysique :

- Le LDSS (*Low Dispersion Survey Spectroscopy*), imaginé par Wynne et Worswick [WYN 88] en 1988, observe jusqu'à 100 objets, dans la gamme spectrale de 370 nm à 750 nm, avec une résolution  $R$  de 70 à 400. Il a été employé à l'*Anglo-Australian Telescope* (AAT) puis remplacé en 1998 par son successeur LDSS++ qui observe jusqu'à 300 objets simultanément.
- Le MEFOS (*Meudon-ESO Fibre Optics Spectroscopy*), présenté par Felenbok en 1997 [FEL 97], peut observer, à partir du télescope de *La Silla*, jusqu'à 30 objets sélectionnés par des fibres optiques robotisées. Il couvre le spectre de 350 nm à 970 nm, avec une résolution de 500 à 1000.

<sup>3</sup>Le champ du télescope est la région du plan focal image à l'intérieur de laquelle la luminosité et le stigmatisme sont satisfaisants [LEN 08].

- Le spectrographe multi-objet FLAMES associé à l’instrument GIRAFFE, est un ensemble hybride, proposé par Pasquini [PAS 02] et mis en place au *Very Large Telescope* en 2006. Capable d’observer 132 galaxies, il est doté de groupes de 20 fibres optiques adjacentes ce qui lui permet d’associer aux spectres détectés une information spatiale. Il est sensible aux ondes électromagnétique de 370 nm à 950 nm avec une résolution spectrale comprise entre 5600 et 46000.
- Le VIMOS (*Visible Multi Object Spectrograph*), imaginé par Le Fevre [Le 03] et installé sur un télescope du VLT en 2003, possède à la fois un mode d’observation de type imagerie classique et un mode spectroscopie multi-objets pouvant détecter jusqu’à 3000 spectres simultanément. Son innovation s’inscrit dans la combinaison de ces deux modes, grâce à l’utilisation d’une grille de micro-lentilles, technique introduite dans le spectrographe TIGER en 1995 [BAC 95] et illustrée par la Fig. I.3. Dans ce troisième mode, ces lentilles sont couplées à des fibres optiques, qui permettent de mesurer le spectre de 6400 positions spatiales afin de couvrir l’intégralité du champ de vue. VIMOS couvre la bande spectrale de 360 nm à 1000 nm, avec une résolution  $R$  de 200 à 2500.

**la spectro-imagerie** Notons à travers les quelques exemples de spectromètres multi-objets, la volonté de multiplier le nombre de sources, afin d’explorer simultanément les domaines spatiaux et spectraux. La spectro-imagerie est l’aboutissement des développements parallèles de l’imagerie multi-spectrale et de la spectroscopie. Elle enregistre le spectre pour un ensemble de positions spatiales consécutives dans le champ. Nous avons déjà évoqué la technique, utilisée par VIMOS, qui consiste à observer le champ à travers une grille de micro-lentilles. Chaque lentille forme alors une source pour un spectromètre de type prisme ou grisme<sup>4</sup>. Cette technique est également appliquée par le spectrographe SAURON (*Spectrographic Areal Unit for Research on Optical Nebulae*) [BAC 01] installé depuis 2002 sur le télescope William Herschel (WHT) sur l’île espagnole de La Palma. Ce dernier fonctionne selon deux modes : le premier fournit, dans la bande spectrale 471-540 nm, une très haute résolution spectrale ( $2.8\text{\AA}$  soit  $R = 1800$ ) (cf. définition B.1), combinée à un bon échantillonnage spatial ( $0.27''$ ) sur un champ de  $11'' \times 9''$ . L’autre mode livre une résolution spectrale légèrement inférieure ( $3.6\text{\AA}$  soit  $R = 1400$  environ), avec un échantillonnage spatial moins bon ( $0.94''$ ) mais sur un champ plus étendu de  $41'' \times 33''$ .

Une autre possibilité réside dans le découpage de l’image par des *slicers* (découpeurs d’image), tel qu’illustré par la Fig. I.4. Chaque partie est ensuite dispersée par une longue fente. C’est la solution choisie dans le cas du spectrographe SPIFFI (*SPectrometer for Infrared Faint Field Imaging*) de l’instrument SINFONI [EIS 03] du VLT. Dédié à l’observation dans l’infrarouge, celui-ci offre une très bonne résolution spatiale ( $0.25''$ ,  $0.10''$  ou  $0.025''$  suivant le mode) sur des champ d’étendue restreinte ( $8'' \times 8''$ ,  $3'' \times 3''$ , ou  $0.8'' \times 0.8''$  respectivement), avec une très bonne résolution spectrale ( $R$  compris entre 1500 et 4000 suivant la bande spectrale).

Ces quatre catégories visent à discriminer les types de spectromètres selon leur mode d’observation dans le domaine spatial, mais elles ne préfigurent pas de la technologie utilisée pour explorer le domaine fréquentiel. Ainsi, le spectre de chaque source peut être obtenu par dispersion par une longue fente, par réseau interférentiel, par grisme, etc.

Dans cette thèse, nous nous sommes concentrés sur les spécificités des données fournies par l’instrument MUSE qui est un spectro-imageur novateur de type intégral de champ.

<sup>4</sup>Un grisme est la combinaison d’un réseau à transmission optique (*gratings*) avec un prisme [KIT 95].

<sup>5</sup>Source : <http://www.strw.leidenuniv.nl/sauron/instrument.html#tiger>.

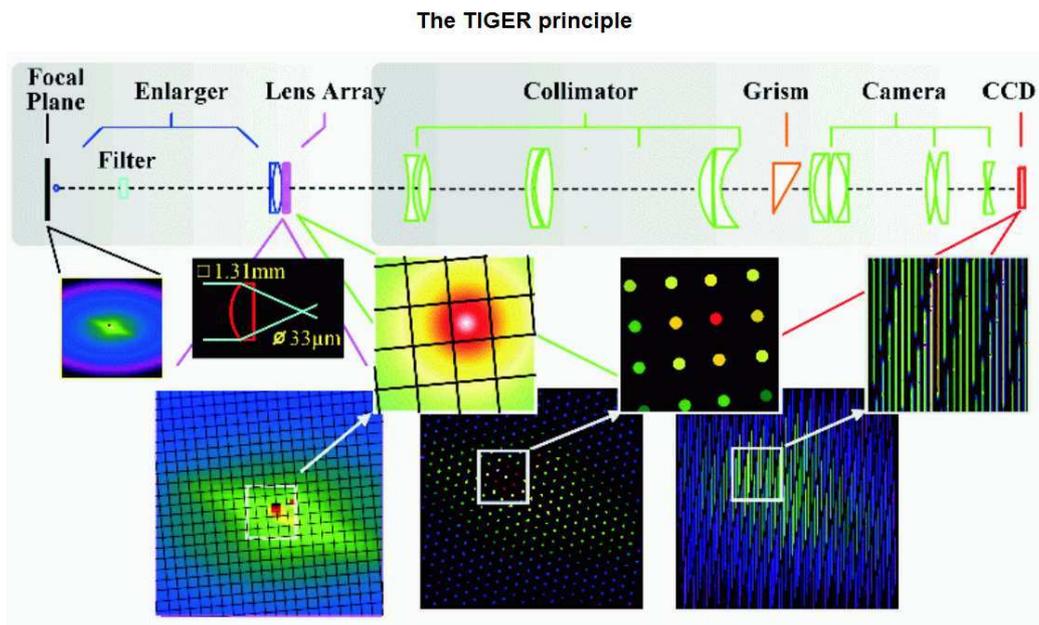


FIG. I.3 – Principe du spectrographe TIGER, doté d'une grille de micro-lentilles. <sup>5</sup>

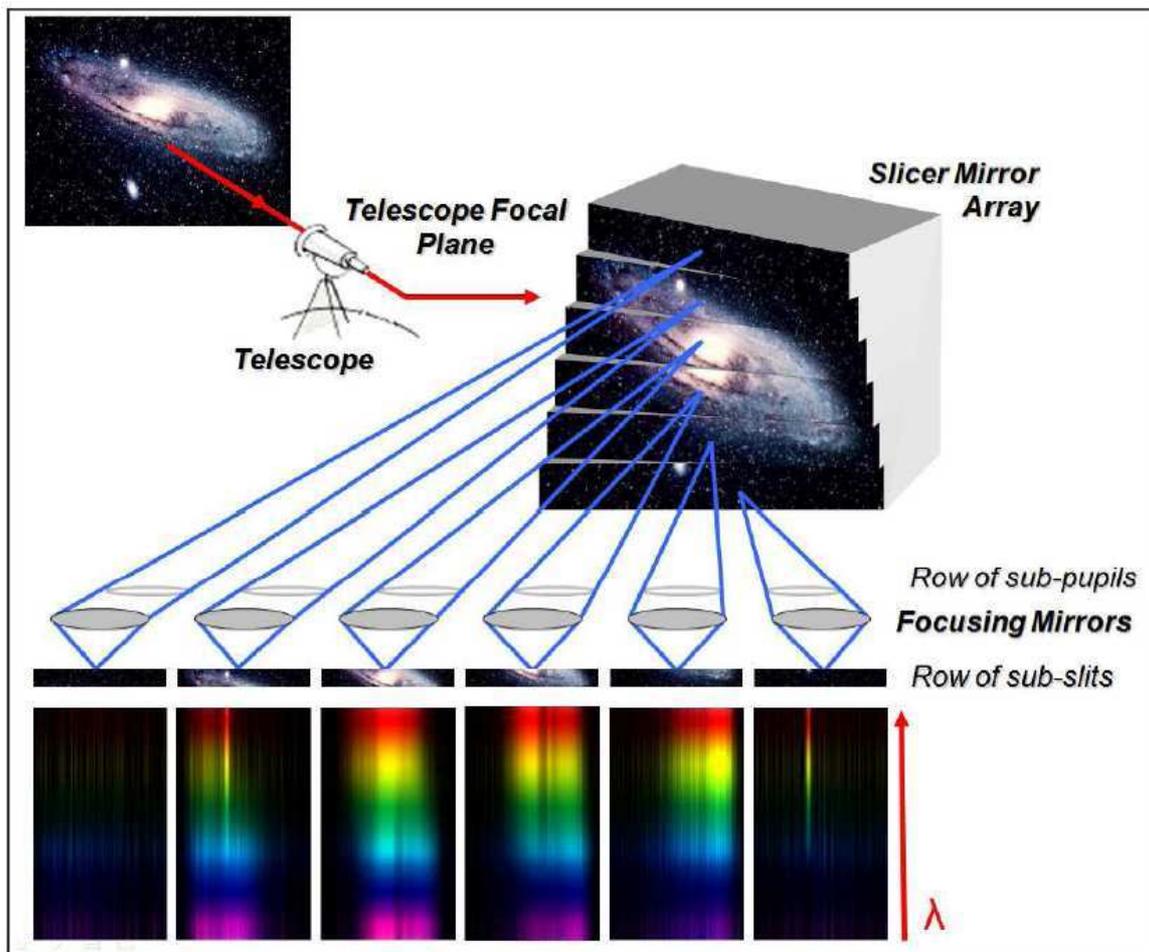


FIG. I.4 – Concept du découpage du champ par des slicers. [VIV 08]

### I.1.3 Le spectro-imageur MUSE

L'instrument MUSE (pour *Multi-Unit Spectroscopic Explorer*) représenté sur la Fig. I.5 est un imageur hyperspectral dont la mise en service au télescope UT4 du *Very Large Telescope* (VLT) est prévue fin 2013.

Il est le fruit de la coopération, sous la direction de Roland Bacon, de sept instituts de recherche : le Centre de Recherche Astrophysique de Lyon (CRAL), le *European Southern Observatory* (ESO), l'Observatoire de Leiden (Pays-Bas), l'Institut de Recherche en Astrophysique et Planétologie (IRAP) de Toulouse, l'Institut d'Astrophysique de Göttingen (Allemagne), le département Astrophysique de l'Institut Polytechnique de Zurich (Suisse) et l'Institut d'Astrophysique de Potsdam (Allemagne).



FIG. I.5 – Vue 3D de l'instrument MUSE.<sup>6</sup>

Comme schématisé Fig. I.6, le spectro-imageur MUSE découpe par un  *slicer*  l'image du plan focal du VLT puis la dirige vers ses 24 *Integral Field Units* (IFU). Chaque IFU découpe à son tour la fraction d'image reçue en 48 blocs placés en entrée d'un spectrographe à longue fente, produisant ainsi 48 *slices*. En fin de chaîne d'acquisition, sur chaque IFU, un capteur CCD de  $4000 \times 4000$  pixels environ, fait l'acquisition d'une image 2D, dont une dimension est spectrale et l'autre est spatiale. Enfin, les données sont réagencées numériquement par le *Data Reduction Software* (DRS) pour former un cube de données (cf. Fig. I.7) de  $300 \times 300$  pixels spatiaux à environ 4000 longueurs d'ondes différentes. La résolution spectrale de ces données est de 0.13 nm sur l'ensemble de la bande spectrale 465-930 nm environ, avec une valeur de  $R$  comprise entre 1770 et 3590.

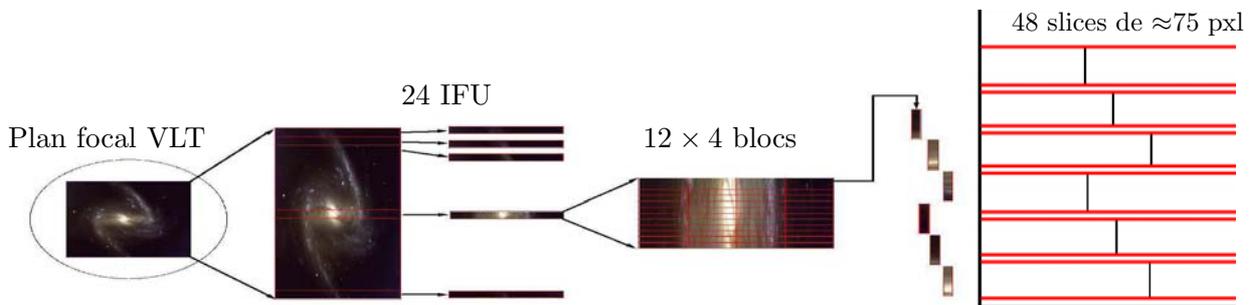


FIG. I.6 – Schéma simplifié d'acquisition sur CCD de l'instrument MUSE

On distingue les 3 modes de fonctionnement de l'instrument :

- *Wide Field Mode* (WFM) sans Optique Adaptative, mode qui sera utilisé durant les deux premières années de fonctionnement de l'instrument ; l'étendue du champ observé est de  $60 \times 60$  (secondes d'arc)<sup>2</sup>, avec un échantillonnage spatial de 0.20".
- WFM avec Optique Adaptative de type GLAO (*Ground Layer Adaptive Optics*)
- *Narrow Field Mode* (NFM) avec Optique Adaptative de type LTAO (*Laser Tomography Adaptive Optics*) ; l'étendue du champ est alors restreinte à  $7.5 \times 7.5$  (secondes d'arc)<sup>2</sup>, avec un échantillonnage spatial de 0.025".

Dans le cadre de cette thèse, nous nous sommes uniquement intéressés au mode grand champ (WFM) qui sera le seul mode utilisé durant les premières années de fonctionnement de MUSE.

<sup>6</sup>Source : <http://muse.univ-lyon1.fr/>.

L'instrument MUSE se démarque des autres instruments par sa dimension hyperspectrale, la taille de son champ, relativement étendue, en regard des instruments précédemment cités (SAURON et SPIFFI) et par sa résolution spatiale.

Illustrons à l'aide de la Fig I.8 la différence entre les données multispectrales et hyperspectrales. Cette distinction peut se faire avant tout à partir du nombre et de la largeur des bandes spectrales observées. En effet, au delà d'une centaine de bandes, on parle souvent d'observation hyperspectrale, voire ultraspectrale. La différence s'inscrit également dans la résolution spectrale : dans le cas des données multispectrales, on ne peut pas à proprement parler de résolution, car les filtres utilisés sont assez larges et se recouvrent les uns les autres. Avec les données hyperspectrales, chaque pixel spatial est associé à un véritable spectre. Celui-ci permet notamment d'étudier la composition chimique et la cinématique des astres.

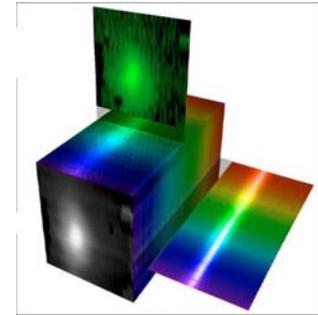


FIG. I.7 – Visualisation d'un cube de données hyperspectral.<sup>7</sup>

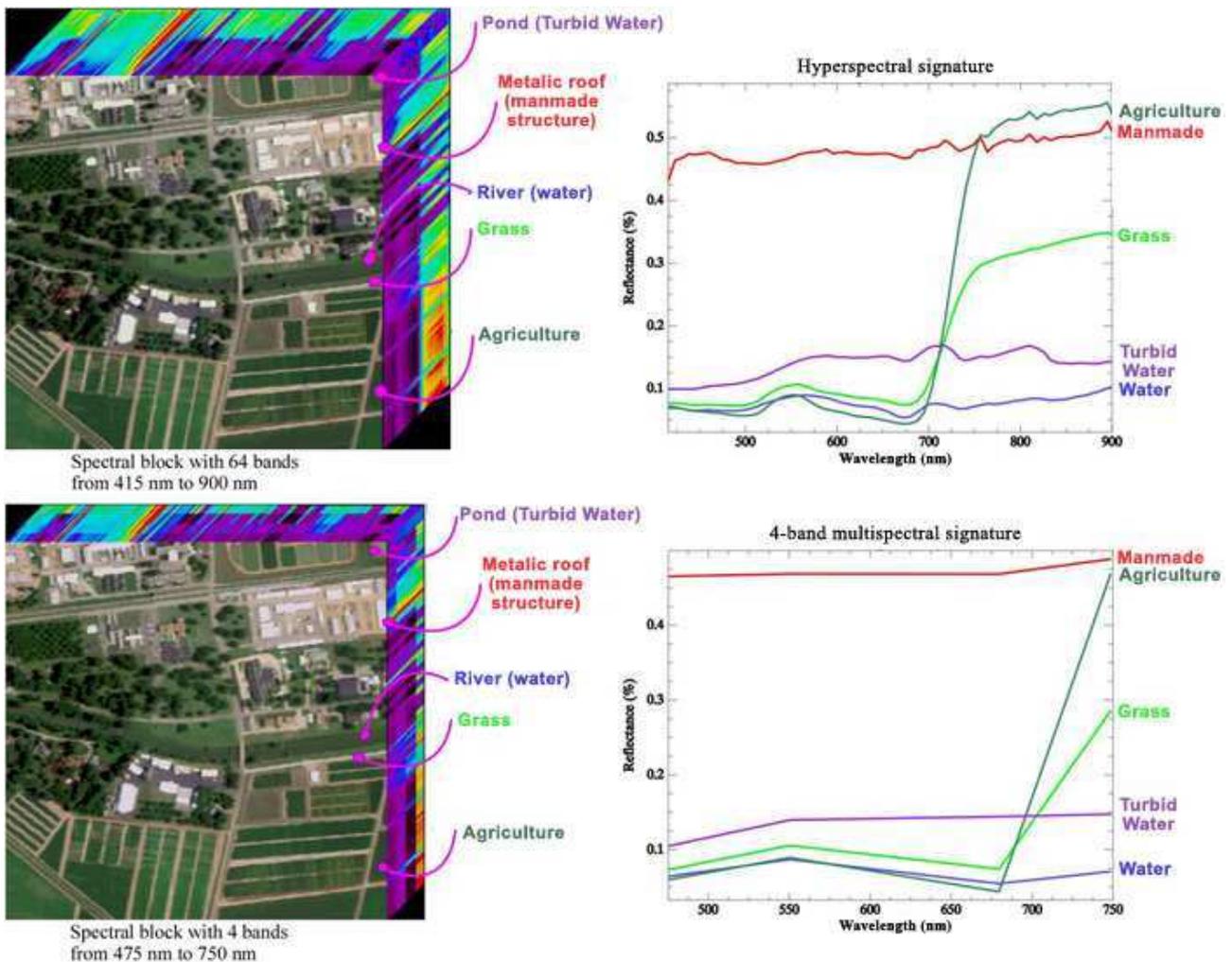


FIG. I.8 – Comparaison de données multispectrales et hyperspectrales sur un exemple de télédétection<sup>8</sup>.

<sup>7</sup>Source : <http://lsiit-miv.u-strasbg.fr/paseo/cubevisualization.php>.

<sup>8</sup>Source : ITRES.

## I.2 L'observation astronomique au sol : MUSE

Lors de l'acquisition d'un cube de données hyperspectrales depuis le sol (Fig. I.9), le signal lumineux traverse l'atmosphère puis parvient dans le plan focal du télescope, éventuellement après correction par un système d'optique adaptative (OA). Cette suite de transformations du signal est commune à tous les instruments au sol. Cependant, l'acquisition de données à partir de l'image reçue par le télescope est propre à chaque instrument. Dans le cas de la construction d'un cube de données hyperspectrales pour l'instrument MUSE, cela nécessite l'étape logicielle du DRS (*Data Reduction System*).

L'image d'une source ponctuelle à travers un système d'acquisition est une tache appelée fonction d'étalement (Point Spread Function, notée PSF) ou réponse impulsionnelle. Nous notons une PSF :

$$h_{\mathbf{z},\mu}(\mathbf{s}, \lambda)$$

avec  $\mathbf{s}$  et  $\mathbf{z}$  des variables spatiales,  $\lambda$  et  $\mu$  des variables spectrales,  $(\mathbf{z}, \mu)$  étant la position de la source. Dans le domaine du traitement du signal, la PSF est généralement considérée invariante par translation (filtre). Elle est alors notée :

$$h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) = h(\mathbf{s} - \mathbf{z}, \lambda - \mu).$$

L'observation  $y(\mathbf{s}, \lambda)$  s'exprime alors comme la convolution de l'image de la source  $x(\mathbf{z}, \mu)$  par la PSF :

$$y(\mathbf{s}, \lambda) = \iint x(\mathbf{z}, \mu) h(\mathbf{s} - \mathbf{z}, \lambda - \mu) d\mathbf{z} d\mu.$$

Afin d'étudier la PSF globale, on peut la décomposer en PSF élémentaires correspondant aux différentes parties de la chaîne d'acquisition. En effet, dans le cas de cette chaîne d'acquisition, certaines PSF élémentaires varient spatialement et spectralement en fonction de la position de la source. L'observation ne s'écrit plus sous la forme d'une convolution mais comme une composition :

$$y(\mathbf{s}, \lambda) = \iint x(\mathbf{z}, \mu) h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) d\mathbf{z} d\mu.$$

On dira qu'une PSF est séparable si elle peut s'écrire sous la forme :

$$h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) = F_{\mathbf{z},\mu}(\mathbf{s}) L_{\mathbf{z},\mu}(\lambda) \quad (\text{I.1})$$

où  $F$  est la fonction d'étalement spatial (*Field Spread Function* - FSF) et  $L$  la fonction d'étalement spectral (*Line Spread Function* - LSF). Nous allons étudier la PSF associée à chacun des éléments de la chaîne afin de modéliser la PSF globale.

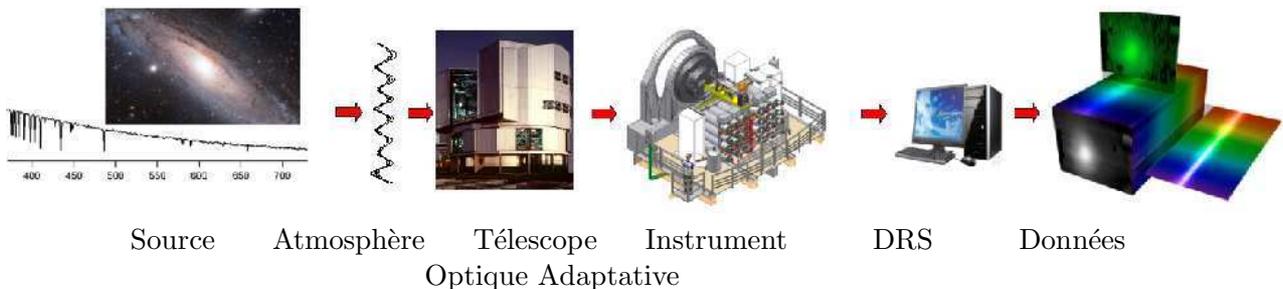


FIG. I.9 – Étapes de formation du cube de données MUSE

### I.2.1 Effets de l'atmosphère

L'atmosphère n'est pas homogène et est constituée d'une superposition de couches qui contiennent des cellules de turbulence. Ses caractéristiques telles que la température et la pression, sont globalement stationnaires dans le temps, avec des variations journalières, annuelles et séculaires [LEN 08]. À la traversée de chaque cellule, les rayons lumineux sont déviés. Donc un rayon lumineux qui traverse toute l'atmosphère est dévié plusieurs fois. L'atmosphère variant dans le temps à l'échelle de la milliseconde, chaque rayon a un trajet qui lui est propre et arrive sur le détecteur avec un angle d'incidence différent. Les caractéristiques de la turbulence atmosphérique sont plus défavorables aux plus courtes longueurs d'onde, donc dans le bleu, par rapport au rouge [ROD 99]. La PSF correspondante est donc plus large dans le bleu que dans le rouge.

La largeur de la PSF de l'atmosphère est couramment appelée « *seeing* ». Précisément, il s'agit de la largeur à mi-hauteur de la PSF de l'atmosphère, que l'on exprime en secondes d'arc (notées "). Ses valeurs courantes sont de 0.5" à 1.0" (à 500 nm) à Paranal au Chili ou à Mauna Kea à Hawaï [LEN 09]. À cela s'ajoutent des phénomènes parasites d'absorption et d'émission. En effet, les constituants de l'atmosphère absorbent une partie du rayonnement, avec un facteur d'absorption qui dépend de la longueur d'onde et de l'altitude. Par la suite, nous ferons abstraction du facteur de transmission de l'atmosphère, celui-ci n'ayant qu'un effet d'atténuation sur les spectres mesurés. D'autre part, l'atmosphère émet des photons, soit par fluorescence, soit par émission thermique [LEN 08]. Ces émissions sont quantifiées par le spectre de l'atmosphère.

L'atmosphère ne produit pas d'étalement spectral ; sa PSF est donc séparable en une FSF et un Dirac pour LSF. Les effets de l'atmosphère étant homogènes compte tenu de la taille réduite du champ, la FSF est donc invariante par translation mais variable en fonction de la longueur d'onde :

$$h_{\mathbf{z},\mu}^{atm}(\mathbf{s}, \lambda) = F_{\mu}^{atm}(\mathbf{s} - \mathbf{z}) \cdot \delta(\lambda - \mu)$$

Pour la durée d'intégration des données MUSE (entre 10 minutes et 1 heure), la FSF peut être considérée circulaire. Pour formaliser la connaissance de la FSF, deux approches se dégagent : la première modélise la FSF via sa transformée de Fourier : l'OTF (*Optical Transfer Function*), la seconde modélise la FSF directement dans le domaine spatial. L'OTF s'exprime à partir de la fonction de structure  $D_{\phi}$  :

$$OTF_{\mu}(\boldsymbol{\nu}) = \exp(-\frac{1}{2}D_{\phi}(\mu|\boldsymbol{\nu}|)),$$

où  $\boldsymbol{\nu}$  est le vecteur de fréquences spatiales. Le modèle le plus courant d'OTF est celui de Kolmogorov [KOL 91] pour lequel :

$$D_{\phi}(r) = 6.88 \left( \frac{r}{r_0} \right)^{\frac{5}{3}}$$

où  $r_0$  est le paramètre de Fried [FRI 66] et représente le diamètre d'un télescope équivalent (voir définition en annexe B.2.3). Dans ce cas, l'atmosphère agit comme un filtre passe-bas de fréquence angulaire de coupure  $\frac{r_0}{\lambda}$ . Le modèle de Von Karman, plus élaboré, fait intervenir l'échelle externe des turbulences, notée  $L_0$  [TOK 02] :

$$D_{\phi}(r) = \frac{\Gamma(11/6)}{2^{11/6}\pi^{8/3}} \left( \frac{24}{5}\Gamma(6/5) \right)^{5/6} \left( \frac{r_0}{L_0} \right)^{-5/3} \left( 2^{-1/6}\Gamma(5/6) - \left( \frac{2\pi r}{L_0} \right)^{5/6} K_{5/6} \left( \frac{2\pi r}{L_0} \right) \right)$$

où  $\Gamma$  est la fonction Gamma et  $K$  la fonction de Bessel modifiée.

La connaissance du paramètre  $r_0$  et éventuellement celle de  $L_0$  permettrait donc de connaître la FSF. L'ESO (*European Southern Observatory*) prévoit de mesurer  $r_0$  par le DIMM (*Differential Image Motion Monitor*), mais les écarts mesurés entre les valeurs obtenues par le DIMM et celles observées par le télescope peuvent atteindre 30%. La précision est donc insuffisante pour en déduire la FSF. Concernant le paramètre  $L_0$ , l'ESO fournit une étude statistique de ces valeurs sur le site [MAR 00], mais ne prévoit pas de mesure simultanée aux observations. Dans [TRU 01], la PSF est décrite selon la

théorie des turbulences atmosphériques, comme la transformée de Fourier de  $\exp[-(|\boldsymbol{\nu}|b)^{5/3}]$  (d'après [FRI 66],[WOO 82]), avec  $\boldsymbol{\nu}$  un vecteur de fréquence spatiale et  $b$  un paramètre d'échelle, et exprime la largeur à mi-hauteur (*Full Width at Half Maximum* - FWHM) par  $\text{FWHM} = 2.9207006 b$ , ce qui s'apparente au modèle de Kolmogorov.

Trujillo et al. [TRU 01] proposent également une expression de la FSF dans le domaine spatial :

$$F^{atm}(r) = \frac{1}{2\pi} \int_0^\infty \nu J_0(\nu r) e^{-(\nu \frac{\text{FWHM}}{2.9207})^{5/3}} d\nu$$

en fonction de la coordonnée radiale  $r$ , avec  $J_0$  fonction de Bessel de première espèce. D'après ce même article, la FSF théorique précédente peut être approchée de manière satisfaisante par une fonction de Moffat avec  $\beta \sim 4.765$ . Toutefois, il est précisé que, si l'on considère la FSF de l'atmosphère et du télescope, les FSF réelles possèdent des ailes plus larges que celles du modèle théorique, car elles prennent en compte également la FSF du télescope. Cette différence peut être corrigée par les fonctions de Moffat avec un paramètre  $\beta$  plus faible.

La FSF de l'atmosphère peut donc être formulée par un modèle physique ou non, lié à des paramètres dépendant des conditions d'observation. Dans le Chapitre II, nous tenterons de modéliser la FSF de l'atmosphère par un modèle non physique mais exprimé directement dans le domaine spatial. Cela limite le coût de calcul par rapport à l'utilisation des expressions citées précédemment, dans le domaine spectral.

### I.2.2 Effet du télescope

Le *Very Large Telescope* (VLT) est un ensemble de 4 télescopes d'un diamètre de 8.2 m et de 4 télescopes mobiles d'un diamètre de 1.8 m, opérationnel depuis 1998. Sa localisation à Cerro Paranal au Chili à 2 635 m d'altitude en fait un instrument privilégié pour l'observation astronomique au sol. L'instrument MUSE doit être installé sur le télescope *Yepun* (UT4) de 8.2 m de diamètre.

Le télescope ne produit pas d'étalement spectral sur l'acquisition des données ; sa PSF 3D est donc séparable en une PSF spatiale invariante par translation mais variable spectralement, et une PSF spectrale, cette dernière étant un Dirac pour modéliser l'absence d'étalement spectral :

$$h_{\mathbf{z},\mu}^{tel}(\mathbf{s}, \lambda) = F_{\mu}^{tel}(\mathbf{s} - \mathbf{z}) \cdot \delta(\lambda - \mu).$$

Dans le cas d'un télescope formé d'une simple pupille de diamètre  $D_t$ , la FSF prend la forme d'une tache dont la forme a été définie par Airy en 1835 [AIR 35] :

$$F_{\lambda}^{tel}(\theta) = \left[ \frac{2J_1\left(\pi \frac{D_t}{\lambda} \theta\right)}{\pi \frac{D_t}{\lambda} \theta} \right]^2$$

où la FSF varie en fonction de la coordonnée angulaire  $\theta$  par rapport à l'axe normal au champ en son centre, avec  $J_1$  la fonction de Bessel de première espèce [LEN 08]. On considère qu'il s'agit d'un filtre passe-bas de fréquence angulaire de coupure  $\frac{D_t}{\lambda}$ , avec  $\frac{D_t}{\lambda} \gg \frac{r_e}{\lambda}$ .

Le télescope du VLT possède, en plus du miroir primaire de 8.2 m de diamètre, un second miroir de 1.1 m de diamètre, maintenu par une « araignée » c'est-à-dire une structure métallique de quatre bras. Il est donc plus complexe que le télescope considéré par Airy, mais sa FSF peut être approchée numériquement.

### I.2.3 Optique Adaptative

Deux ans après la mise en service de l'instrument MUSE, la turbulence atmosphérique sera partiellement corrigée par un système d'optique adaptative (OA). Plus précisément, la fonction de ce système est de corriger le front d'onde<sup>10</sup> qui a été perturbé lors de la traversée de l'atmosphère. Ainsi, le VLT de diamètre 8.2 m, lorsqu'il est équipé de l'optique adaptative, fournit des images dans l'infrarouge d'une résolution angulaire environ 20 fois meilleure que sans OA [LEN 09]. Schématiquement, l'optique adaptative est constituée de trois éléments [LEN 09] illustrés par la Fig. I.2.3 :

- l'analyseur de front d'onde : il utilise une étoile de référence, naturelle ou artificielle pour analyser les perturbations du front d'onde. En l'absence d'étoile naturelle suffisamment brillante dans le champ d'étude, l'analyseur se sert d'une étoile laser (*Laser Guide Star - LGS*). Cette étoile est un faisceau laser, émis depuis le sol, puis rediffusé par une couche atmosphérique dense en sodium située à environ 80 km d'altitude. Elle constitue une source ponctuelle et brillante [LEN 09].
- le calculateur : il calcule les corrections à appliquer au front d'onde,
- le système de correction : il applique les corrections du front d'onde, par le biais d'actuateurs (au nombre de 1170 dans le cas de l'optique adaptative qui est associée à MUSE) qui modifient le profil d'un miroir déformable.

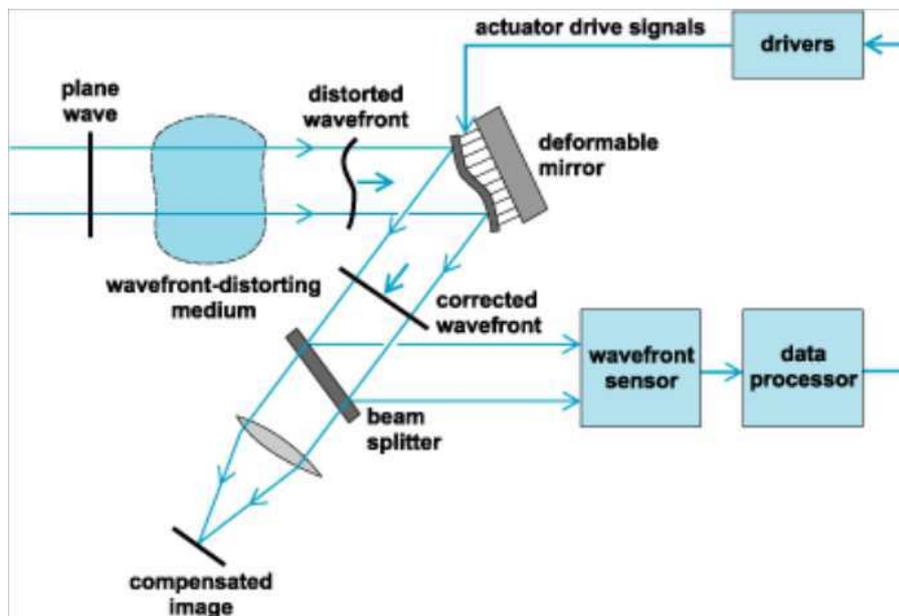


FIG. I.10 – Schéma d'un système d'optique adaptative standard <sup>11</sup>.

L'optique adaptative prévue pour être associée à l'instrument MUSE utilisera quatre étoiles laser et une étoile naturelle proche du champ d'observation pour corriger le *tip-tilt* (basculement du front d'onde dépendant de la longueur d'onde ayant pour effet de décaler spatialement l'image). La correction des effets de l'atmosphère par l'optique adaptative n'est pas toujours uniforme. L'optique adaptative associée au mode grand champ, appelée GLAO (*Ground Layer Adaptive Optics*) corrige essentiellement la couche basse de l'atmosphère [LEN 09]. Elle réalise une correction de la turbulence relativement moyenne vis-à-vis des systèmes d'optique adaptative courants, mais sur un champ très étendu (plusieurs dizaines de secondes d'arc). La FSF résiduelle de l'atmosphère corrigée par optique adaptative varie lentement en fonction de la position de la source dans le champ. En mode petit champ (NFM), l'optique adaptative est de type LTAO (*Laser Tomography Adaptive Optics*). Dans ce cas, la FSF résiduelle varie fortement spatialement, mais la taille du champ est plus restreinte.

<sup>10</sup>Le front d'onde est la surface perpendiculaire aux rayons lumineux émis par une source [LEN 09].

<sup>11</sup>Source : <http://www.answers.com/topic/adaptive-optics>.

L'optique adaptative n'a pas d'effet spectral sur l'acquisition des données mais sa correction est plus performante dans le rouge que dans le bleu [ROD 99]. La FSF résiduelle varie donc spectralement, du fait de la variation de la FSF atmosphérique, mais sa variation pourra être différente de cette dernière. La FSF peut encore être modélisée via son OTF par l'intermédiaire de sa fonction de structure  $D_\phi$  qui dépend toujours des conditions atmosphériques d'observation mais également du système de correction par optique adaptative [JOL 06]. La modélisation analytique de cette fonction de structure, dépendant d'un grand nombre de paramètres, est possible dans le cas asymptotique (c'est-à-dire pour des poses de longues durées par rapport aux variations des turbulences atmosphériques, ce qui sera le cas pour les données MUSE) et a été mise en œuvre dans le code PAOLA [JOL 08], dont nous utilisons des simulations dans le Chapitre II.

Pour l'optique adaptative GLAO de MUSE, l'estimation de la PSF résiduelle à partir de l'analyseur de front d'onde et des commandes du calculateur est en cours d'étude dans le département d'optique théorique et appliquée (DOTA) de l'ONERA.

#### I.2.4 Effet de l'instrument

Étant donné la complexité du système de l'instrument brièvement présenté dans le paragraphe I.1.3, aucun modèle analytique de la PSF de l'instrument n'est disponible, mais un modèle numérique de l'instrument, exploitable pour une simulation réaliste de données, a été développé durant la thèse de A. Jarno [JAR 08]. Comme illustré Fig. I.11 (haut), la PSF 2D de l'instrument sur le CCD, dont une dimension est spectrale et l'autre spatiale, varie fortement dans le champ et en fonction de la longueur d'onde, mais elle reste heureusement de dimension très limitée (de l'ordre de  $2 \times 2$  pixels). L'étalement suivant la dimension spectrale, également illustré Fig. I.11 (bas), principalement dû à la fente du spectrographe, est censé être identique pour tous les *slices*. Mais en pratique, vu le système optique complexe, cet étalement varie lentement sur les CCD dans la dimension spectrale. En revanche, cette variation peut être brutale dans la dimension spatiale car deux pixels voisins dans le cube peuvent être imagés par des systèmes optiques (IFU) différents.

Enfin, il apparaît clairement que la PSF 2D n'est pas séparable en le produit d'une fonction des lignes (dimension spatiale) et d'une fonction des colonnes (dimension spectrale). Notons que l'échantillonnage spectral est à peine suffisant (pas d'échantillonnage supérieur à la limite de Nyquist), ce qui pourrait poser problème lors de l'analyse et des différents traitements de données, en particulier dans une étape d'interpolation [CAR 12].

#### I.2.5 Logiciel de réduction des données

En plus d'étapes de prétraitement (calibrations...), le *Data Reduction Software* (DRS) effectue la conversion d'un ensemble de données 2D issues des CCD vers un cube hyperspectral et nécessite donc une étape d'interpolation [WEI 08]. Cette dernière dépendant directement des données, l'effet du DRS dépend donc des observations. En fonction de la taille de la PSF des étapes en amont, la PSF du DRS peut être considérée comme négligeable. Nous pouvons donc soit négliger la PSF du DRS, dont la modélisation est délicate voire impossible, soit effectuer l'analyse et le traitement à partir de données n'ayant pas subi les étapes critiques. Ainsi, il peut être intéressant pour certaines applications d'analyse et de traitement de données, de travailler directement sur les données acquises sur les CCD ou sur les données pré-traitées par le DRS, avant l'étape d'interpolation.

#### I.2.6 Bilan et approximations pour la PSF globale

En pratique, il n'est pas concevable de calibrer la PSF de l'instrument en chaque pixel du cube, leur nombre s'élevant à environ  $300 \times 300 \times 4000$ . De plus, l'effet du DRS est délicat à maîtriser et

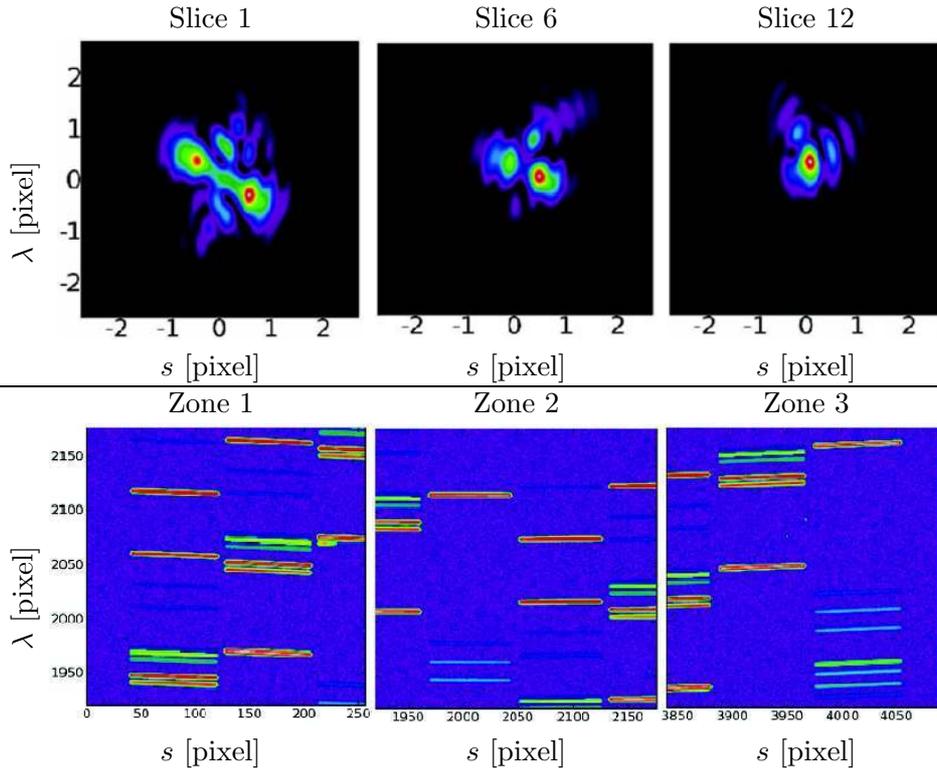


FIG. I.11 – Caractéristiques instrumentales de MUSE en différentes positions du CCD [JAR 08] : simulation à haute résolution d’objets ponctuels (haut, en échelle logarithmique). Simulation d’objet constants spatialement avec un spectre de raies (bas), zoom sur 3 zones d’une image d’un CCD d’un IFU.

dépend des observations. Des hypothèses simplificatrices doivent donc être effectuées afin de pouvoir modéliser et estimer la PSF globale de la chaîne d’acquisition.

Comme nous l’avons vu précédemment, l’étalement spectral de la chaîne d’acquisition des données MUSE est principalement dû à l’instrument MUSE lui même. L’effet du DRS est délicat à modéliser, néanmoins il peut être supposé suffisamment faible et donc négligé. Cet effet ne varie pas en fonction des observations et peut donc être intégré dans la phase de calibration instrumentale. Un moyen d’étudier cet étalement spectral consiste à placer à l’entrée de l’instrument un champ spatialement constant éclairé avec une lampe possédant un spectre de raies (Fig. I.11 bas). Ainsi, on dispose, pour chaque raie de ce spectre, d’une LSF intégrée spatialement (ILSF) qui met bien en avant cet étalement spectral en le supposant faiblement variable dans une slice. Des travaux ont été effectués sur ce sujet dans le cadre du projet DAHLIA, principalement par R. Bacon et H. Carfantan. Il a été montré que cette ILSF était fortement variable d’une slice à l’autre (donc spatialement) et également variable spectralement au sein d’une slice. Un modèle de variation spatiale de cette ILSF a été proposé [BAC 09, CAR 09] ainsi qu’un schéma d’estimation des paramètres de ces modèles à partir de données de calibration [MAI 11]. Ainsi on peut disposer d’un modèle de cette ILSF pour chaque slice de chaque CCD avec une erreur de l’ordre de quelques pourcents. Nous considérerons donc par la suite que la LSF est parfaitement connue en tout point du cube de données.

En mode grand champ sans correction par optique adaptative, cas auquel nous nous sommes le plus intéressés dans cette thèse, nous choisissons de négliger l’étalement spatial dû à l’instrument (et au DRS) d’étendue limitée par rapport à celui du télescope et de l’atmosphère. Cela revient à approximer

la PSF de l'instrument (et du DRS) par une PSF séparable ayant un Dirac comme FSF et une LSF variant spatialement et spectralement.

Nous modélisons donc la PSF globale de la chaîne d'acquisition sans optique adaptative, d'une part par une FSF variable en longueur d'onde, essentiellement due aux effets atmosphériques et au télescope, et d'autre part par une LSF variable spectralement et spatialement propre à l'instrument approximée par son ILSF et estimée une fois pour toutes. Dans ce cas, la FSF de l'atmosphère est invariante par translation, la FSF globale le sera donc également. On peut donc écrire la PSF globale sous la forme :

$$h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) = F_{\mu}(\mathbf{s} - \mathbf{z})L_{\mathbf{z},\mu}(\lambda).$$

L'observation d'un objet  $x$  donne alors :

$$y(\mathbf{s}, \lambda) = \iint x(\mathbf{z}, \mu)F_{\mu}(\mathbf{s} - \mathbf{z})L_{\mathbf{z},\mu}(\lambda)d\mathbf{z}d\mu.$$

Dans le cas de MUSE où l'échantillonnage spectral est de 0.13 nm, il est raisonnable de faire l'approximation que la FSF varie lentement en fonction de  $\lambda$  et donc que  $F_{\mu}(\mathbf{s}) \approx F_{\lambda}(\mathbf{s})$  pour  $\mu$  proche de  $\lambda$ . Comme la LSF est à support limité (quelques pixels spectraux), on peut écrire :

$$\begin{aligned} y(\mathbf{s}, \lambda) &\approx \int \left[ \int x(\mathbf{z}, \mu)L_{\mathbf{z},\mu}(\lambda)d\mu \right] F_{\lambda}(\mathbf{s} - \mathbf{z})d\mathbf{z} \\ &= F_{\lambda} * \left[ \int x(\mathbf{z}, \mu)L_{\mathbf{z},\mu}(\lambda)d\mu \right] \\ &= F_{\lambda} * (L \star x) \end{aligned}$$

Cette relation s'écrit comme une convolution spatiale par la FSF à la longueur d'onde  $\lambda$  du spectre de l'objet étalé par la LSF :

$$\tilde{x}(\mathbf{z}, \lambda) = \int x(\mathbf{z}, \mu)L_{\mathbf{z},\mu}(\lambda)d\mu = L \star x.$$

Notons que cette dernière opération ne correspond pas à une convolution. Certaines méthodes de traitements de données MUSE, pour la séparation aveugle de sources [MEG 11] ou le débruitage [BOU 11a], s'appuient sur cette propriété.

Dans le cas où le télescope est équipé du système d'optique adaptative, la FSF résiduelle de l'atmosphère et du télescope est d'étendue plus réduite. La FSF de l'instrument et du DRS ne peut donc plus être négligée. La FSF résiduelle de la correction par OA variant dans le champ de vue, la PSF globale a donc l'expression générale :

$$h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) = F_{\mathbf{z},\mu}(\mathbf{s})L_{\mathbf{z},\mu}(\lambda).$$

L'observation d'un objet  $x$  n'est donc plus une convolution du fait de la variation spatiale de la FSF, mais la composition successive de la LSF et de la FSF.

Dans l'ensemble de ce document, nous considérons que la PSF 3D est séparable en une FSF spatiale et une LSF spectrale :  $h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) = F_{\mathbf{z},\mu}(\mathbf{s})L_{\mathbf{z},\mu}(\lambda)$ . Il a été démontré récemment [CAR 12] que, pour des raisons liées à la chaîne d'acquisition de MUSE, une écriture de la forme  $h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) = F_{\mathbf{z},\mu}(\mathbf{s})L_{\mathbf{s},\mu}(\lambda)$  est plus correcte, ce qui ne correspond pas exactement à l'hypothèse de séparabilité (indice  $\mathbf{s}$  et non  $\mathbf{z}$  pour la LSF). Si l'on se rappelle que la LSF peut varier très brutalement pour deux pixels voisins, ceux-ci pouvant en effet être imagés sur deux slices différentes et même sur deux IFUs différents, on comprend bien que l'hypothèse de séparabilité n'est pas valable, ou du moins seulement comme une approximation grossière. Cependant, dans le cadre des travaux effectués dans cette thèse, cela n'aura

que peu d'importance au niveau théorique. Éventuellement des modifications seront nécessaires pour la mise en oeuvre pratique sur des données réelles. Notons que toutes les données simulées jusqu'à présent l'ont été avec une PSF séparable. Aussi le fait que cette hypothèse ne soit pas valable en théorie n'aura aucune conséquence pratique sur les résultats obtenus à partir de ces données.

### I.3 Le projet DAHLIA

L'intégralité de mon travail de thèse s'inscrit dans le cadre du projet DAHLIA : *Dedicated Algorithms for HyperspectraL Imaging in Astronomy*. Ce projet d'une durée initiale de 4 ans, renouvelé pour 1 an, est financé par l'Agence Nationale de la Recherche (ANR) 'Programme Blanc', depuis le 1<sup>er</sup> Janvier 2009.



Avec le développement de nouveaux télescopes et instruments devant fournir des données hyperspectrales (dans l'immédiat l'instrument MUSE au VLT et dans un futur proche des instruments de l'ELT (*Extremely Large Telescope*) au sol et le JWST (*James Webb Space Telescope*) dans l'espace), le projet DAHLIA a pour objectif l'élaboration de nouvelles méthodes de traitements spécifiques pour ce type de données. Un premier thème de travail porte sur la simulation de données hyperspectrales réalistes, qui sont par la suite utilisées par les autres thèmes. À cela s'ajoute le développement des méthodes de traitements pour l'analyse, par fusion de données, séparation de sources, déconvolution et détection. Enfin, le projet s'intéresse à la visualisation des données hyperspectrales. Les paragraphes suivants décrivent les différents thèmes de recherche du projet DAHLIA, ainsi que leurs inter-dépendances éventuelles avec les résultats obtenus dans cette thèse.

#### I.3.1 Simulation de données

Les principaux acteurs des travaux sur la simulation sont Roland Bacon, Jeremy Blaizot et Aurélien Jarno du CRAL à Lyon, ainsi que des chercheurs du consortium MUSE.

Afin que l'ensemble des acteurs du projet DAHLIA puisse développer des algorithmes fiables, ce groupe de travail construit des données hyperspectrales de type MUSE aussi réalistes que possible. Ces données peuvent contenir différents objets, tels que des étoiles isolées ou en champ dense, et des galaxies. Elles prennent en compte les déformations dues à la chaîne d'acquisition (atmosphère, optique adaptative, télescope, instrument et logiciel de réduction de données). Pour chaque scène, un premier jeu de données a été simulé par le *Quick Simulator* (QSIM), pour lequel la modélisation de la chaîne d'acquisition est relativement simplifiée. Un autre jeu est simulé à partir de l'INM (*Instrument Numerical Model*) et du logiciel de réduction des données (DRS). Ces différents outils sont issus des travaux du consortium MUSE.

#### I.3.2 Fusion de données

L'acquisition d'un cube de données par l'instrument MUSE est réalisée sur un temps d'intégration relativement court, compris entre quelques minutes et une heure. Si ce temps est suffisant pour observer des étoiles brillantes, il ne l'est pas pour l'exploitation de données visant à étudier des galaxies ou des champs profonds. La solution adoptée pour améliorer le rapport signal sur bruit est l'acquisition de plusieurs cubes d'observation d'une même région du ciel pendant une heure, qui doivent ensuite être combinés.

L'équipe formée de Christophe Collet, Mathieu Petremand, Vincent Mazet, Fabien Salzenstein du LSIIT de Strasbourg ainsi que de André Jalobeanu de l'Université Evora du Portugal a proposé un algorithme de fusion de données ([JAL 11] [PET 11] [PET 12]). Il s'agit de prendre en compte les différentes conditions d'observation des cubes, telles que le *seeing*, le spectre du ciel, etc., afin de construire un cube dont la qualité est meilleure par rapport aux différents cubes fusionnés. Il s'agit

également de lui associer un cube de variance représentant l'incertitude de l'acquisition pour chaque pixel. En particulier, la prise en compte du *seeing* passant par la déconvolution des données, nécessite de connaître la PSF du système des différents cubes à toutes les longueurs d'onde, au moment de leur acquisition.

### I.3.3 Estimation du fond du ciel

Le fond du ciel correspond au spectre d'émission de l'atmosphère et à celui de sources non résolues et vient s'ajouter au spectre des objets que l'on souhaite observer. Il s'agit donc d'un élément parasite qui contribue au bruit d'acquisition et que l'on souhaite soustraire au signal acquis. En tenant compte de la taille restreinte du champ MUSE, on considère que le spectre du fond du ciel est identique dans tous les pixels spatiaux de l'image. La modélisation et l'estimation du fond du ciel est étudiée par Vincent Mazet de LSIIT de Strasbourg. Ce spectre est supposé être constitué d'un continuum et de raies impulsionnelles d'absorption et d'émission dont le nombre et la position sont connus à partir d'un catalogue. L'estimation du spectre est effectuée à partir des données CCD (c'est-à-dire avant réduction par le DRS) et dépend de la PSF de l'instrument. Pour simuler des données réalistes, nous avons utilisé le spectre du ciel illustré par la Fig. I.12.

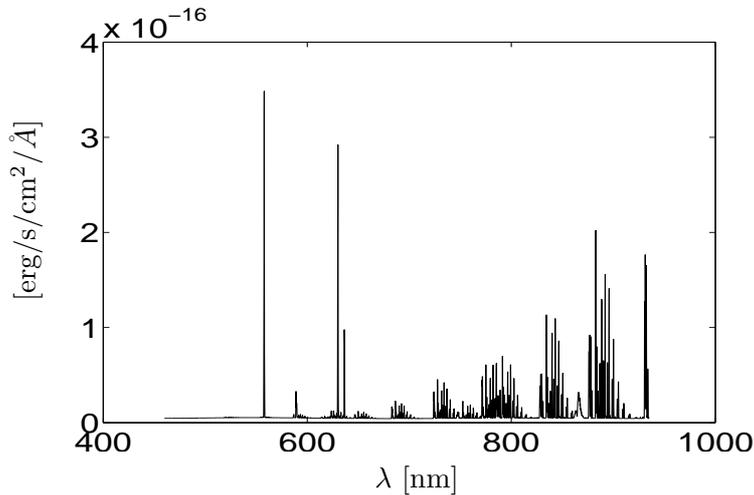


FIG. I.12 – Spectre du ciel à la résolution MUSE.

### I.3.4 Séparation de sources

En raison de l'étalement de la lumière par le système d'acquisition (PSF) et de la résolution spatiale de l'instrument, l'observation d'un champ dense d'étoiles conduit à un cube où chaque pixel spatial contient le spectre de plusieurs étoiles, comme illustré par la Fig I.13. Yannick Deville, Shahram Hosseini, Inès Meganem et Ahmed Selloum de l'IRAP se concentrent sur l'élaboration de nouvelles méthodes de séparation aveugle de sources ponctuelles, afin d'obtenir une estimation de la position et du spectre de chaque étoile [MEG 11]. Cette estimation repose sur des contraintes, telles que la positivité du spectre, le nombre de sources et éventuellement leur position approximative. Elle nécessite également d'estimer la PSF<sup>12</sup> du système qui caractérise les données, car celle-ci détermine les coefficients de mélange en fonction de l'éloignement du pixel par rapport à la source.

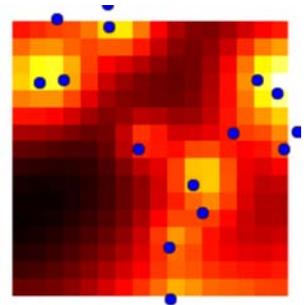


FIG. I.13 – Exemple de superposition de sources (points bleus).

<sup>12</sup>En effet, dans le cas d'observation de champ dense d'étoiles, on ne peut généralement observer d'étoile isolée pouvant servir à l'estimation de la PSF, comme il sera présenté au Chapitre III. C'est pourquoi on parle de séparation aveugle de sources car il faut également estimer cette PSF tout en cherchant à séparer les spectres des étoiles.

### I.3.5 Détection de sources, segmentation et déconvolution

Les données d'observation hyperspectrales sont déformées par la réponse impulsionnelle du système (PSF) et détériorées par le bruit d'acquisition, fortement variables spectralement. Sébastien Bourguignon, David Mary, Eric Slezak de l'OCA ainsi que Hervé Carfantan (IRAP) proposent de nouvelles stratégies de détection, de débruitage et de restauration adaptées aux données hyperspectrales, basées principalement sur la contrainte de parcimonie dans le domaine spectral ([BOU 10] [BOU 11a] [BOU 11b] [BOU 12]). Ces travaux exigent la connaissance de la fonction d'étalement spectral (LSF) pour le débruitage et spatial (FSF) pour la restauration.

Hervé Carfantan et moi-même nous sommes penchés sur un cas particulier de déconvolution de données hyperspectrales de type MUSE, pour l'étude de la cinématique des galaxies, dans un cadre bayésien. Nous en donnons les conclusions dans le Chapitre IV et [VIL 12].

### I.3.6 Visualisation de données hyperspectrales

L'objectif est de pouvoir combiner les aspects spatiaux et spectraux de la visualisation de données hyperspectrales, afin de tirer les informations nécessaires à leur analyse. Il s'agit également de faire apparaître la carte d'incertitude (variance) associée aux données.

Mireille Louys, Christophe Collet et Mathieu Petremand du LSIIT de Strasbourg ont mis en œuvre un outil de visualisation de données hyperspectrales, avec l'aide de François Bonnarel (CDS Strasbourg), Fabien Salzenstein (LSIIT), Thierry Contini (IRAP), Eric Emsellem et Harald Kuntschner (ESO).

Les efforts de ce groupe ont donné naissance à l'outil QuickViz [PET 10], plugin du logiciel Aladin. Ce dernier est un logiciel de visualisation de données astronomiques et de gestion des informations associées. En y associant QuickViz, il devient notamment possible d'extraire les spectres de différents pixels spatiaux et de parcourir les images du cube à différentes longueurs d'onde.

### I.3.7 Modélisation et estimation de la PSF

Plusieurs thèmes de recherche du projet DAHLIA, tels que la fusion de données, la séparation de sources et la déconvolution requièrent la connaissance de la réponse impulsionnelle spatiale (FSF) et spectrale (LSF) du système. Ces deux fonctions peuvent varier en fonction de la position de la source dans le champ (comme illustré par la Fig. I.14), tant spatialement que spectralement. La LSF est constante dans le temps et peut-être calibrée instrumentalement, la FSF dépend des conditions atmosphériques et doit être estimée à chaque observation.

Hervé Carfantan, Laurent Jolissaint (Leiden), Roland Bacon (CRAL), Denis Serre (Leiden), Louis Maire (stagiaire IRAP) et moi-même proposons une modélisation paramétrique des variations spatiales et spectrales de cette PSF au sein du cubes de données. En particulier, H. Carfantan, R. Bacon et L. Maire ont étudié les propriétés de la PSF spectrale (cf. [BAC 09] [CAR 09] [MAI 11]). Dans le cadre de cette thèse, nous nous

sommes intéressés à la modélisation, en tout point d'un cube de données, de la FSF en mode grand champ avec et sans optique adaptative qui fait l'objet du Chapitre II et des communications [SER 10] [VIL 11a]. Cela se traduit par la modélisation des variations des paramètres d'un modèle de Moffat, en fonction de la longueur d'onde, mais aussi en fonction de la position dans le champ dans le cas corrigé par optique adaptative. Puis nous nous sommes confrontés au problème de l'estimation de la FSF en mode grand champ sans optique adaptative en tout point d'un cube de données hyperspectrales, directement à partir des données d'observation d'une étoile isolée. Pour cela, nous nous appuyons sur un modèle simple que nous avons précédemment proposé. Nous présentons ces travaux dans le Chapitre III et dans [VIL 11b] [VIL 11a].

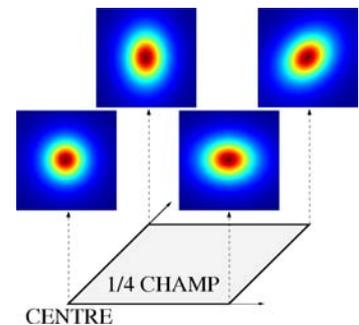


FIG. I.14 – Variation schématique de la FSF dans le champ (avec optique adaptative).



# Chapitre II

## Modélisation de la FSF

### Sommaire

---

<b>Notations</b> . . . . .	<b>30</b>
<b>Introduction</b> . . . . .	<b>31</b>
<b>II.1 Méthodologie</b> . . . . .	<b>31</b>
II.1.1 FSF simulées par PAOLA . . . . .	32
II.1.2 Indices d'erreur . . . . .	32
<b>II.2 Modèles courants pour les PSF spatiales</b> . . . . .	<b>33</b>
II.2.1 Modèles gaussien et Moffat . . . . .	33
II.2.2 Choix de la normalisation . . . . .	34
II.2.3 Fonction de Moffat elliptique . . . . .	35
II.2.4 Paramètres de la fonction de Moffat . . . . .	35
<b>II.3 Modélisation de la FSF en mode WFM sans OA</b> . . . . .	<b>38</b>
II.3.1 Prise en compte de la FSF du télescope . . . . .	38
II.3.2 Modèle elliptique à <i>seeing</i> fixé . . . . .	39
II.3.3 Modèles circulaires à <i>seeing</i> fixé . . . . .	40
II.3.4 Modèles circulaires en fonction du <i>seeing</i> . . . . .	40
II.3.5 Influence du $C_n^2$ . . . . .	42
II.3.6 Influence de l'échelle externe . . . . .	43
II.3.7 Synthèse . . . . .	44
<b>II.4 Modélisation de la FSF en mode WFM avec optique adaptative</b> . . . . .	<b>44</b>
II.4.1 Données et hypothèses . . . . .	44
II.4.2 Modèle circulaire constant dans le champ . . . . .	46
II.4.3 Modèle elliptique variable dans le champ . . . . .	48
<b>Conclusions</b> . . . . .	<b>55</b>

---

## Notations [Unités]

$D(x, y, \lambda)$  données

$L_0$  échelle externe [m]

$M(x, y, \lambda)$  modèle de FSF

$s$  seeing [arcsec]

$x, y$  coordonnées spatiales discrètes [pixels]

$\alpha$  paramètre d'échelle du modèle de Moffat pour la FSF [pixel]

$\alpha_0, \alpha_1, \alpha_2, \alpha_3$  coefficients de la modélisation polynomiale de  $\alpha$

$\beta$  paramètre de forme du modèle de Moffat pour la FSF [sans unité]

$\varepsilon$  indice d'erreur [%]

$\gamma$  paramètre d'ellipticité du modèle de Moffat [sans unité]

$\lambda$  longueur d'onde [nm]

$\phi$  orientation du modèle de Moffat elliptique, par rapport à l'axe des  $x$ ,  
dans le sens trigonométrique [degrés]

$\Phi$  flux d'une image / d'une fonction de Moffat

$\rho, \theta$  coordonnées polaires dans le champ de vue [arcsec, degrés]

## Introduction

Le spectro-imageur MUSE fournit des données 3D composées d'images (2D) à un grand nombre de longueurs d'onde (1D). De par la nature même de ces données, la déformation subie par cet ensemble d'images, à travers la chaîne d'acquisition présentée dans la section I.2, est aussi bien spatiale que spectrale. Ce phénomène d'étalement de la lumière est caractérisé par la *Point Spread Function* (PSF) 3D du système qui correspond à la tache obtenue par l'observation d'un objet ponctuel monochromatique.

Suite à l'analyse de la chaîne d'acquisition, nous supposons que la PSF globale s'exprime (cf. équation I.1) comme la composition d'une PSF spatiale (FSF pour *Field Spread Function*), dépendant des conditions d'observation, par une PSF spectrale (LSF pour *Line Spread Function*), qu'il est possible de calibrer (cf. I.2.6). La PSF due aux turbulences de l'atmosphère, éventuellement corrigée par l'optique adaptative, et du télescope, s'exprime donc selon :

$$h_{\mathbf{z},\mu}^{ATM+TEL+OA}(\mathbf{s}, \lambda) = F_{\mathbf{z},\mu}(\mathbf{s}) \cdot \delta_{\mu}(\lambda) = F_{\mathbf{z},\lambda}(\mathbf{s}) \cdot \delta_{\mu}(\lambda).$$

Nous avons plus particulièrement étudié, dans le cadre de cette thèse, les propriétés de la FSF notée  $F_{\mathbf{z},\lambda}(\mathbf{s})$ .

Rappelons que l'instrument MUSE fonctionnera suivant trois modes d'observation (cf. I.1.3) : le mode grand champ (WFM) d'abord sans, puis avec optique adaptative (OA), ainsi que dans le mode petit champ (NFM) avec OA. Nous nous sommes uniquement intéressés aux modes WFM sans et avec OA, qui seront les seuls utilisés durant les premières années de fonctionnement. Dans chacun de ces cas, la FSF du système doit être connue, que ce soit par un modèle analytique ou par des données numériques, afin de permettre l'utilisation de plusieurs post-traitements du cube de données, tels que la déconvolution, la fusion de cubes et plus généralement l'analyse de données (cf. I.3).

Nous souhaitons donc modéliser analytiquement cette FSF en toute position  $(\mathbf{z}, \lambda)$  spatiale et spectrale d'un cube de données MUSE, dans le mode grand champ, car la FSF dépend des conditions d'observation. Pour cela, nous avons proposé d'établir un modèle analytique de la FSF, avec des paramètres variant :

- en fonction des différentes longueurs d'onde  $\lambda$ ,
- en fonction des différentes positions  $\mathbf{z} = (x, y)$  ou  $\mathbf{z} = (r, \theta)$  dans le champ, si la FSF varie dans le champ (cas avec optique adaptative).

Ces modèles permettront de connaître cette FSF en tout point d'un cube de données réelles à partir de mesures effectuées sur une ou plusieurs étoiles isolées. Dans les cas les plus simples, nous établirons également un modèle de variation des paramètres en fonction des configurations de turbulences de l'atmosphère, pour différentes valeurs du « seeing »  $s$  et pour différents  $C_n^2$ . Enfin, nous quantifierons l'erreur entre les modèles ainsi constitués et les simulations de FSF pour les différentes configurations.

## II.1 Méthodologie

La PSF spatiale de la chaîne d'acquisition, différente pour chaque observation, est constante dans le champ de vue dans le mode grand champ sans optique adaptative mais varie spatialement s'il y a correction par OA. Notre objectif est de disposer d'un modèle de la FSF en tout point du cube de données hyperspectrales, c'est-à-dire pour toute position spatiale et longueur d'onde de la source. Ce modèle doit être suffisamment simple pour pouvoir en estimer les paramètres à partir d'un objet non résolu, éventuellement très faible, isolé dans les données (cf. Chapitre III).

Une alternative consisterait à modéliser la FSF indépendamment à chaque longueur d'onde. Néanmoins, cela n'exploiterait pas la nature hyperspectrale des données et la redondance d'information qui en découle. Cette option sera évoquée dans le Chapitre III dédié à l'estimation de la FSF.

Pour cela, nous disposons de simulations de la FSF de l’atmosphère et du télescope sans et avec optique adaptative, pour plusieurs longueurs d’onde et positions dans le champ (détails dans la section II.1.1). Nous allons donc estimer les paramètres du modèle de FSF que nous aurons choisi associés à chaque simulation, par minimisation d’un critère des moindres carrés. Puis nous étudierons leurs variations et en proposerons une modélisation. Enfin, nous quantifierons les performances des différents modèles proposés par des indices d’erreur, présentés dans la section II.1.2.

### II.1.1 FSF simulées par PAOLA

Les données utilisées pour la modélisation de la FSF sont des simulations effectuées en collaboration avec Denis Serre de l’Université de Leiden (Pays-Bas) en utilisant le logiciel PAOLA de Laurent Jolissaint [JOL 06], en version 6.2.

Nous disposons de FSF de l’ensemble {atmosphère+télescope} avec ou sans optique adaptative en mode grand champ, ainsi que de FSF du télescope seul en mode grand champ, simulées pour différentes valeurs de la longueur d’onde  $\lambda$ , de la position polaire  $(\rho, \theta)$  ou cartésienne  $(x, y)$  dans le champ et différentes valeurs de caractéristiques atmosphériques tels que la *seeing* (cf. annexe A B.2.4) et l’indicateur de turbulence  $C_n^2$ , conformément au Tab. II.1 et à la Fig. II.1.

Le modèle analytique de FSF du télescope implémenté dans PAOLA prend en compte à la fois le diamètre d’ouverture, l’obscuration centrale ainsi que les araignées de support du miroir secondaire [JOL 09]. Néanmoins, la rotation du champ au cours de l’acquisition n’est pas prise en compte.

Variable	Télescope seul	WFM sans OA	WFM avec OA
Longueur d’onde $\lambda$	{465 – 550 – 650 – 750 – 850 – 930} nm		
Coordonnées polaires :	(0,0)	(0,0)	[0 20 30 40]”
$\rho$ $\theta$			[0 15 30 45]°
Coordonnées cartésiennes :			
$x$ et $y$			
Atmosphère :	non pertinent	[0.5 0.7 1.0 1.3 1.7]” « médian »	[0.5 0.7 1.0 1.3]” « médian »
<i>seeing</i> noté $s$ $C_n^2$			
Taille d’un pixel	0.0050” soit 40 fois plus petit qu’un pixel MUSE		

TAB. II.1 – Simulations PAOLA utilisées pour la modélisation de la FSF, à différentes valeurs des variables spatiales, pour différents paramètres atmosphériques.

### II.1.2 Indices d’erreur

Il est capital de quantifier l’erreur commise lors de la modélisation de la FSF par un modèle analytique. Cette erreur correspond à la différence entre une image  $D$  de FSF et le modèle  $M$  correspondant aux paramètres estimés. Il s’agit donc d’une image de même taille que les données. Nous calculons plusieurs statistiques sur l’erreur de modélisation  $D - M$  présentées dans le Tab. II.2, telles que le flux de l’erreur, son énergie et son maximum. En effet, l’erreur commise entre deux images est difficile à quantifier, c’est pourquoi nous utilisons différents types d’erreur qui apportent des informations complémentaires. Notons que ces erreurs sont normalisées par la statistique correspondante.

Chaque indice d’erreur est calculé à partir de la différence entre les données simulées et le modèle pour chaque longueur d’onde, position dans le champ, *seeing*, etc. Sauf cas explicitement précisé, les statistiques telles que la moyenne  $\bar{\varepsilon}$  ou le maximum  $\varepsilon_{max}$  appliquées sur ces termes d’erreur portent sur l’ensemble des erreurs calculées aux différentes valeurs des variables. Les variables considérées pour les

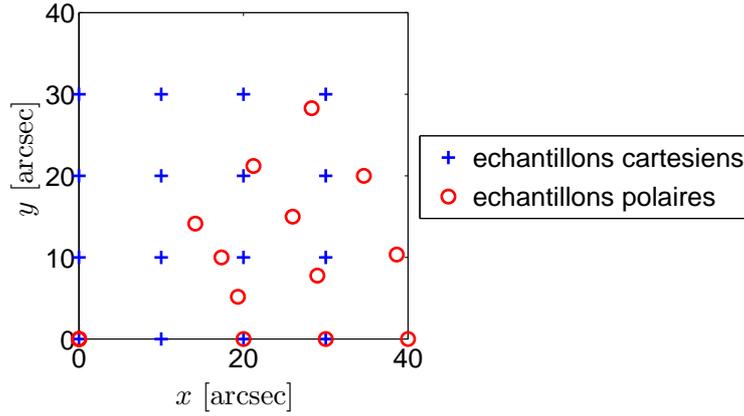


FIG. II.1 – Répartition spatiale des deux séries de simulations de FSF PAOLA, dans le champ de vue, pour le mode grand champ avec correction par optique adaptative.

$$\varepsilon_1 = \frac{\sum_{xy} |D(x, y) - M(x, y)|}{\sum_{xy} |D(x, y)|}$$

$$\varepsilon_2 = \frac{\sqrt{\sum_{xy} (D(x, y) - M(x, y))^2}}{\sqrt{\sum_{xy} D^2(x, y)}}$$

$$\varepsilon_\infty = \frac{\max_{xy} |D(x, y) - M(x, y)|}{\max_{xy} (D(x, y))}$$

TAB. II.2 – Définition des indices d'erreur utilisés pour quantifier la qualité de la modélisation de la FSF.

statistiques sont précisées dans chaque tableau, selon le cas. Par exemple, les indices d'erreur peuvent être moyennés sur les différentes longueurs d'onde, ou positions spatiales.

## II.2 Modèles courants pour les PSF spatiales

### II.2.1 Modèles gaussien et Moffat

Le modèle gaussien est très couramment utilisé pour approximer les FSF, du fait de sa simplicité d'utilisation, en particulier de son invariance par transformée de Fourier [ORI 10]. Ainsi, A.F.J. Moffat choisit en 1969 de modéliser simplement la FSF de l'atmosphère par une distribution gaussienne 2D, qu'il convolue ensuite par les FSF plus complexes du système optique et la plaque photographique [MOF 69]. La FSF résultante est ensuite comparée à des observations d'étoiles, en supposant notamment que toutes les étoiles ont le même profil d'intensité à un facteur d'échelle près. Ce modèle semble coïncider avec les observations faites pour des distances faibles de la position de l'étoile, mais sous-estime les zones éloignées. En d'autres termes, les ailes des FSF sont plus larges que dans le cas gaussien.

Afin d'améliorer la concordance du modèle et des observations, A.F.J. Moffat met en place une nouvelle expression analytique, appelée par la suite fonction de Moffat (MF), de la forme suivante :

$$M_{\alpha, \beta}(r) = \left[ 1 + \left( \frac{r}{\alpha} \right)^2 \right]^{-\beta} \quad (\text{II.1})$$

fonction circulaire du rayon  $r = \sqrt{x^2 + y^2}$ , dans laquelle  $\alpha$  est un paramètre d'échelle et  $\beta$  un paramètre de forme, tous deux fonction du *seeing*. Lorsque celui-ci est mauvais,  $\alpha$  est élevé et  $\beta$  tend vers 1. Ceux-ci sont soumis à certaines contraintes :  $\alpha \in ]0; +\infty[$  et  $\beta \in ]1; +\infty[$ . La largeur à mi-hauteur s'exprime en fonction des paramètres selon la formule :

$$\text{FWHM} = 2\alpha\sqrt{2^{\frac{1}{\beta}} - 1}.$$

La Fig. II.2 compare une FSF simulée par PAOLA, avec son estimation par moindres carrés généralisés ayant pour approximation un modèle de Moffat et un modèle gaussien. Le modèle de Moffat est visiblement plus proche des données réelles que le modèle gaussien.

Historiquement, la fonction de Moffat a été créée afin de modéliser la FSF de l'atmosphère. A.F.J.Moffat précise que le modèle gaussien est un cas limite du modèle de Moffat, lorsque  $\beta \rightarrow \infty$  [TRU 01]. Dans cette étude, nous l'utilisons pour modéliser la FSF de l'ensemble {atmosphère+télescope} dans le cas sans optique adaptative, ou de l'ensemble {atmosphère corrigée par OA +télescope}. La FSF de l'ensemble {atmosphère+télescope} pourrait être modélisée en convoluant un modèle de la FSF de l'atmosphère et un modèle de celle du télescope. Cette option, plus coûteuse en temps de calcul, sera étudiée dans la section II.3.1.

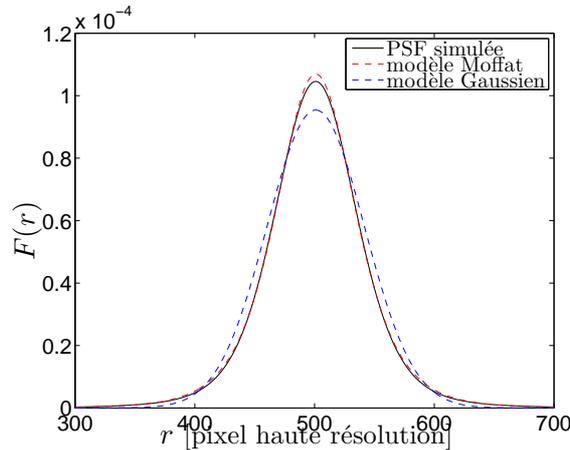


FIG. II.2 – Comparaison d'une FSF simulée à 465 nm et  $s = 0.5''$  en mode WFM sans optique adaptative, avec son approximation des modèles gaussiens et de Moffat à flux normalisé.

## II.2.2 Choix de la normalisation

Avec l'expression (II.1), la fonction est maximale en  $r = 0$  et vaut 1 en ce point. Elle est donc normalisée au maximum. Il est possible de normaliser la fonction :

- soit par son flux théorique  $\Phi = \frac{\pi\alpha^2}{\beta - 1}$  qui correspond à l'intégrale de la fonction sur un champ

$$\text{d'étendue infinie } \Phi = \int_0^\infty \int_0^{2\pi} M_{\alpha,\beta}(r)rdrd\theta$$

- soit par son flux numérique  $\hat{\Phi} = \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} M_{\alpha,\beta}(\sqrt{x^2 + y^2})$  qui est l'estimation du flux théorique

par la somme du champ sur l'image étudiée de taille  $N_x \times N_y$ .

En pratique, si l'on étudie une fenêtre assez grande centrée sur la FSF, le flux numérique approche 99% du flux théorique. Les deux grandeurs sont donc équivalentes, la première tendant vers la seconde lorsque la taille de la fenêtre tend vers l'infini. Par la suite, nous serons amenés à estimer le spectre de l'étoile par le flux de la fenêtre. Cette estimation repose sur l'hypothèse du flux normalisé de la

FSF à chaque longueur d'onde. Pour cette raison, la fonction de Moffat sera normalisée par son flux numérique c'est-à-dire  $M_{\alpha,\beta}(r) = \frac{1}{\Phi} \left[ 1 + \left( \frac{r}{\alpha} \right)^2 \right]^{-\beta}$ .

### II.2.3 Fonction de Moffat elliptique

Sur certaines simulations avec optique adaptative, il apparaît que la FSF ne peut pas être considérée circulaire. Aussi, il est possible d'introduire le paramètre d'ellipticité  $\gamma$  dans le modèle. On obtient alors l'expression générale de la fonction de Moffat elliptique, exprimée dans le repère formé par le centre de l'ellipse, le grand axe et le petit axe, dans le cas de la normalisation par le maximum :

$$M_{\alpha,\beta,\gamma}(x, y) = \left[ 1 + \frac{1}{\alpha^2} \left( x^2 + \frac{y^2}{\gamma^2} \right) \right]^{-\beta}$$

Ainsi, le paramètre  $\gamma$  correspond au rapport entre le grand axe et le petit axe de l'ellipse. Un quatrième paramètre  $\phi$  est ajouté afin d'autoriser la rotation de la fonction de Moffat dans le plan  $(x, y)$ . L'angle  $\phi$  autorise la rotation des isophotes elliptiques dans le plan et correspond à l'opposé de l'angle entre l'axe  $Oy$  et le grand axe de l'ellipse.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

### II.2.4 Paramètres de la fonction de Moffat

$\alpha$  est un paramètre d'échelle qui influe sur l'étalement de la fonction dans le champ. En effet, dans le cas de la fonction de Moffat (MF) circulaire, la FWHM est directement proportionnelle à  $\alpha$ . Dans le cas de la normalisation de la fonction par le flux théorique, la valeur du maximum est inversement proportionnelle au carré de  $\alpha$ .

La Fig. II.3 représente une fonction de Moffat 2D circulaire générée pour différentes valeurs de  $\alpha$ , le paramètre  $\beta$  étant fixé. On vérifie bien que la largeur de la « tache » augmente avec  $\alpha$ , sans affecter la valeur de son maximum.

$\beta$  est un paramètre de forme qui influe également sur l'étalement de la fonction dans le champ. La FWHM diminue lorsque  $\beta$  augmente et tend vers  $2|\alpha|$  lorsque  $\beta$  tend vers 1. Dans le cas de la normalisation de la fonction par le flux théorique, la valeur du maximum est proportionnelle à  $\beta$ . Cette même normalisation impose à  $\beta$  d'être supérieur à 1 ; expérimentalement, nous avons observé que  $\beta$  se trouve dans l'intervalle  $]1; 4]$  pour les FSF simulées.

La Fig. II.4 représente une fonction de Moffat 2D circulaire générée pour différentes valeurs de  $\beta$ , la largeur à mi-hauteur étant fixée. On vérifie bien que la forme est modifiée et que la hauteur des ailes augmente lorsque  $\beta$  tend vers 1.

$\gamma$  et  $\phi$  Le paramètre  $\gamma$  introduit l'ellipticité des isophotes dans la fonction de Moffat. Il correspond au rapport du grand axe sur le petit axe.

Dans la Fig. II.5, le paramètre  $\gamma$  prend différentes valeurs, dont 1, ce qui signifie alors que la fonction de Moffat est circulaire. Les autres paramètres sont fixés.

Dans la Fig. II.6, le grand axe est tourné selon différentes orientations, ses autres paramètres étant fixés.

Si  $\phi = 0$ , le grand axe est parallèle à l'axe  $Oy$  si  $\gamma > 1$ , à l'axe  $Ox$  sinon. On constate une non identifiabilité des paramètres  $\gamma$  et  $\phi$ . En effet, on obtient une fonction de Moffat identique pour un paramètre  $\gamma > 1$  et  $\phi$  quelconque appartenant à l'intervalle  $[0; 180^\circ]$  et pour  $\gamma' = \frac{1}{\gamma} < 1$  et  $\phi' = \phi - 90^\circ$ . Pour lever cette ambiguïté, on peut contraindre les paramètres selon deux choix : soit imposer  $\gamma \geq 1$ , soit imposer  $\phi \in [0; 90^\circ]$ . Par la suite, la première option sera utilisée.

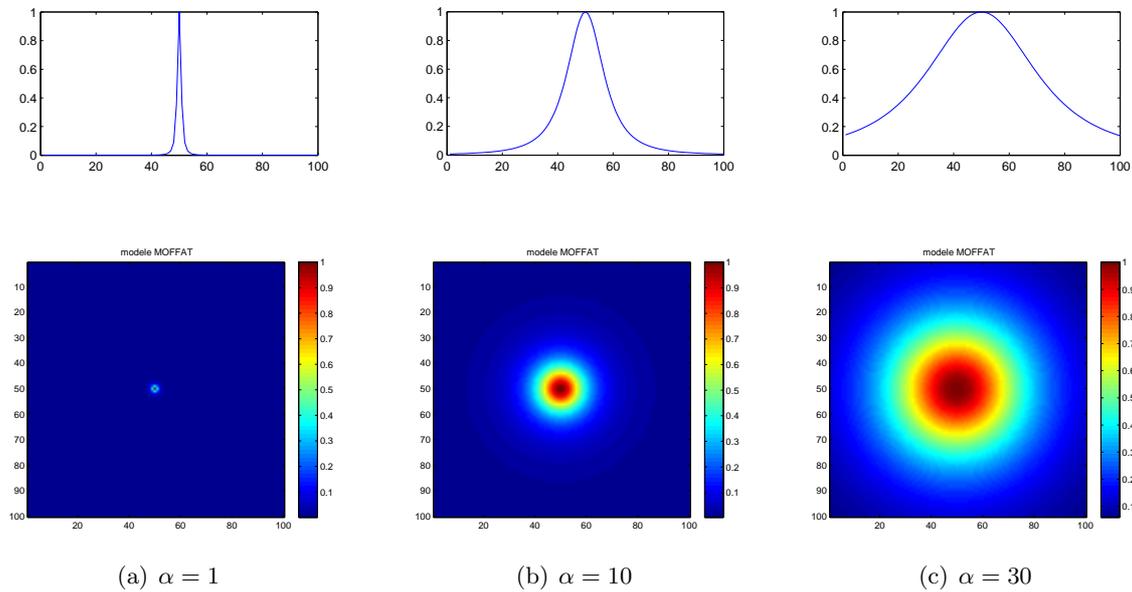


FIG. II.3 – Fonction de Moffat 2D circulaire générée avec  $\alpha \in \{1; 10; 30\}$ ,  $\beta = 1.5$ , centrée sur le pixel (50, 50) d'une image carrée de taille  $100 \times 100$  pixels.

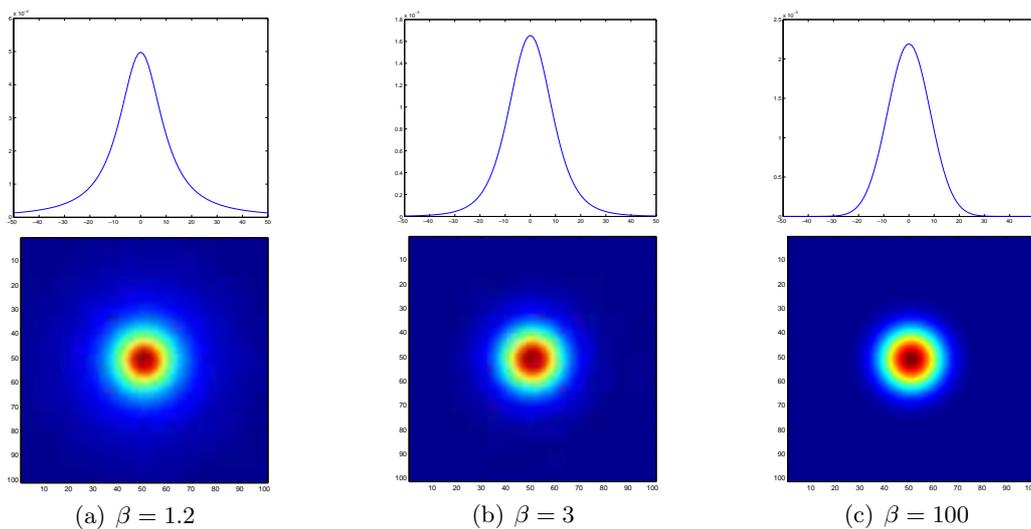


FIG. II.4 – Fonction de Moffat 2D circulaire générée avec  $\beta \in \{1.1; 2; 5\}$ , pour une largeur à mi-hauteur constante de 20 pixels, soit  $\alpha \in \{11.3; 19.6; 120\}$ . Fonction centrée sur le pixel (50, 50) d'une image carrée de taille  $100 \times 100$  pixels.

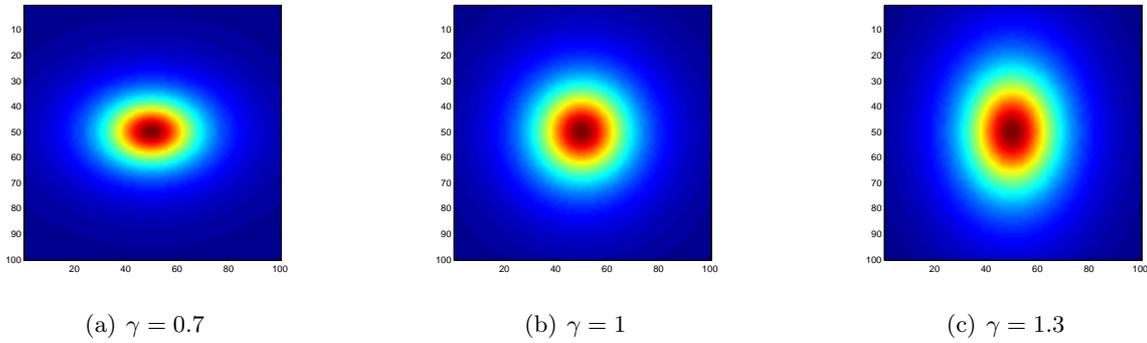


FIG. II.5 – Fonction de Moffat 2D elliptique générée avec  $\alpha = 20$ ,  $\beta = 1.5$  et  $\gamma \in \{0.7; 1; 1.3\}$ , centrée sur le pixel  $(50, 50)$  d'une image carrée de taille  $100 \times 100$  pixels.

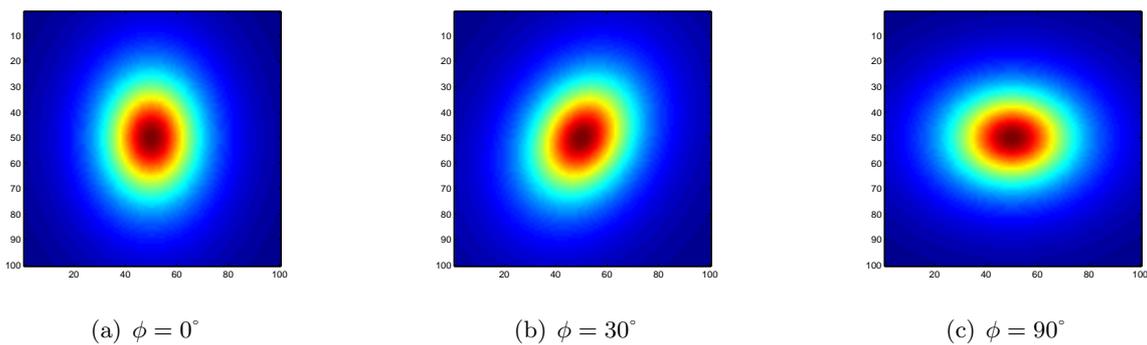


FIG. II.6 – Fonction de Moffat 2D elliptique générée avec  $\alpha = 20$ ,  $\beta = 1.5$ ,  $\gamma = 1.3$  et  $\phi^\circ \in \{0; 30; 90\}$ , centrée sur le pixel  $(50, 50)$  d'une image carrée de taille  $100 \times 100$  pixels.

### II.3 Modélisation de la FSF en mode WFM sans OA

Dans le mode grand champ sans optique adaptative, la FSF de l'ensemble {atmosphère+télescope} est constante dans le champ et varie spectralement. Notre objectif est donc de construire un modèle de FSF permettant de décrire la FSF à toutes les longueurs d'onde.

En premier lieu, nous comparerons la modélisation de la FSF de l'ensemble {atmosphère +télescope} par une seule fonction de Moffat (MF), par rapport à l'utilisation d'une fonction de Moffat modélisant uniquement l'atmosphère, convoluée par une FSF du télescope. En effet, il s'agit de vérifier si le gain de précision potentiellement apporté par la seconde option motive le coût de calcul supplémentaire.

Ensuite, nous estimerons et étudierons les variations des paramètres d'une fonction de Moffat elliptique puis circulaire, en fonction de la longueur d'onde  $\lambda$ . Nous proposerons une modélisation pour chacun de ses paramètres, que nous évaluerons par les indices d'erreur. Enfin, nous évoquerons l'influence du *seeing*, du  $C_n^2$  et de l'échelle externe sur les paramètres du modèle circulaire.

#### II.3.1 Prise en compte de la FSF du télescope

Pour les données MUSE, on peut supposer connue la FSF du télescope aux différentes longueurs d'onde. On doit donc se poser la question de l'intérêt de l'exploitation de ces données pour la modélisation de la FSF {atmosphère +télescope}.

Afin de tirer parti de ces données, on propose de modéliser uniquement la FSF de l'atmosphère par la fonction de Moffat. Pour cela, on considère comme modèle de FSF pour l'ensemble {atmosphère +télescope} la convolution d'une fonction de Moffat circulaire par la FSF du télescope à la longueur d'onde correspondante.

Le Tab. II.3 compare les erreurs de modélisation des FSF par rapport aux simulations PAOLA, à différentes longueurs d'onde et aux *seeings* extrêmes suivant que la fonction de Moffat circulaire ou elliptique modélise l'atmosphère uniquement ou bien l'ensemble {atmosphère + télescope}. Dans chacun des cas, les valeurs optimales des paramètres des modèles sont estimées en minimisant un critère des moindres carrés. On constate qu'à *seeing* et longueurs d'onde donnés, les erreurs de modélisation sont très similaires pour tous les modèles. Ainsi, l'utilisation des données de FSF du télescope seul n'apporte pas d'amélioration significative, tout en augmentant le temps de calcul de l'étape d'optimisation et la complexité de la modélisation. On observe également que les indices d'erreur sont plus élevés pour un haut *seeing*; la fonction de Moffat est donc plus appropriée pour approximer une FSF de bas *seeing*.

La largeur à mi-hauteur de la FSF du télescope (de l'ordre de 0.013" à 500 nm) est largement inférieure à celle de la FSF de l'atmosphère (environ 0.37" à 500 nm soit presque deux pixels). Ainsi, la fonction de Moffat, destinée à l'origine à modéliser la FSF de l'atmosphère, peut inclure celle du télescope, si l'on modifie légèrement ses paramètres.

Pour la suite de nos travaux, nous décidons donc, dans un souci de simplicité, de modéliser par une unique fonction de Moffat à la fois la FSF de l'atmosphère et celle du télescope.

Indices d'erreur à différents <i>seeings</i>							
Modèle MF	<i>seeing</i>	$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
Atmosphère + télescope	0.5"	2.08%	2.11%	5.34%	5.38%	2.18%	2.27%
Atmosphère seule	0.5"	2.08%	2.11%	5.33%	5.36%	2.17%	2.25%
Atmosphère + télescope	1.7"	2.35%	2.45%	5.98%	6.17%	2.44%	2.53%
Atmosphère seule	1.7"	2.35%	2.44%	5.95%	6.14%	2.42%	2.49%

TAB. II.3 – Comparaison de l'erreur entre la FSF de l'ensemble {atmosphère+télescope} simulée par PAOLA et le modèle de FSF, suivant les modélisations et le *seeing*.

### II.3.2 Modèle elliptique à *seeing* fixé

Même si la FSF devrait être circulaire en mode grand champ sans optique adaptative, nous étudions par précaution les variations d'un modèle de Moffat elliptique, plus général, à partir des simulations de FSF comprenant à la fois l'atmosphère et le télescope. Puis, nous en déduisons un modèle de variations de ces paramètres, en fonction de la longueur d'onde, pour un *seeing* donné.

#### II.3.2.1 Paramètres $\alpha$ et $\beta$

L'évolution des paramètres  $\alpha$  et  $\beta$  en fonction de la longueur d'onde est illustrée par la Fig. II.7. Nous pouvons constater que cette variation est continue et peut être approchée par une fonction linéaire ou une fonction puissance de la longueur d'onde. Les indices d'erreur de modélisation de la FSF par une Moffat de paramètres  $\alpha$  et  $\beta$  indépendants à chaque longueur d'onde, puis par une Moffat dont la variation des paramètres est approximée par une droite, sont regroupés dans le Tab. II.4 dans le cas d'un *seeing*  $s = 0.5''$ .

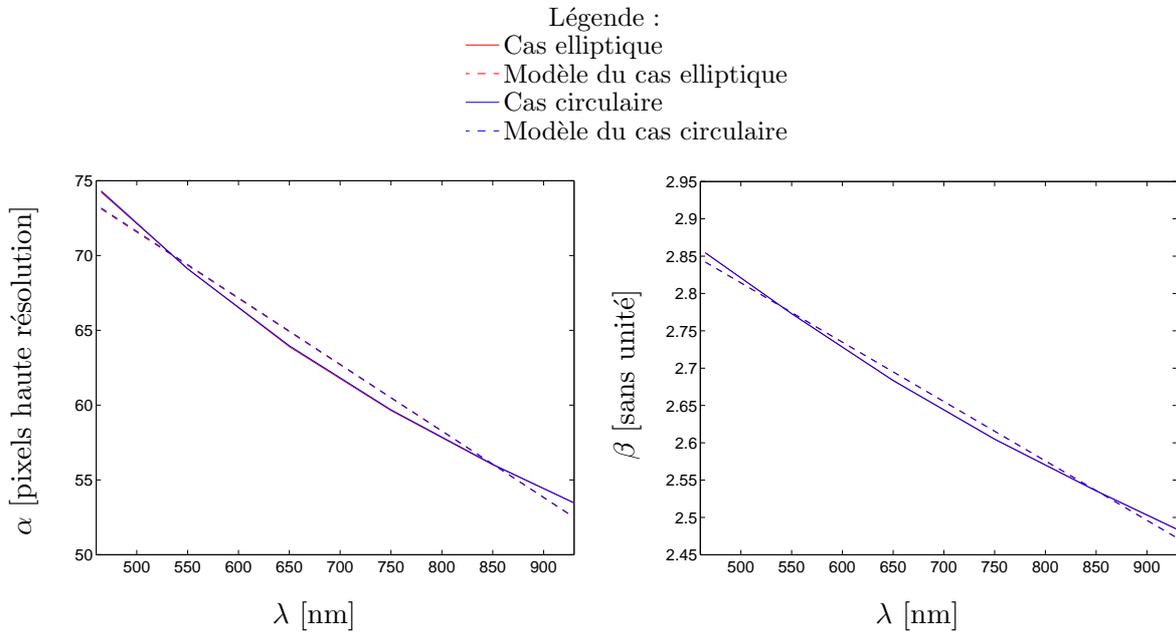


FIG. II.7 – Évolution et modélisation des paramètres  $\alpha$  et  $\beta$  pour une fonction de Moffat elliptique ou circulaire.

#### II.3.2.2 Paramètres $\gamma$ et $\phi$

La Fig. II.8 montre l'évolution des paramètres d'ellipticité et d'orientation en fonction de la longueur d'onde. Comme prévu, on constate que l'ellipticité est très faible ( $\gamma < 1.0025$ ); de ce fait, le paramètre d'orientation  $\phi$  n'est pas significatif.

Si l'on souhaite néanmoins conserver un modèle elliptique, on peut modéliser  $\gamma$  par une constante égale à la moyenne  $\bar{\gamma} = 1.0012$  et l'angle  $\phi$  par la valeur  $135^\circ$ . La fonction de Moffat est alors considérée elliptique avec le même facteur et la même orientation à toute longueur d'onde. Les indices de qualité du Tab. II.4 montrent que la prise en compte de cette faible ellipticité n'améliore pas la qualité de l'approximation.

Il est donc plus simple de supposer  $\gamma$  constant égal à 1,  $\phi$  inexistant, la MF étant alors considérée circulaire à toute longueur d'onde. C'est le modèle que nous avons choisi et qui fait l'objet des sections suivantes.

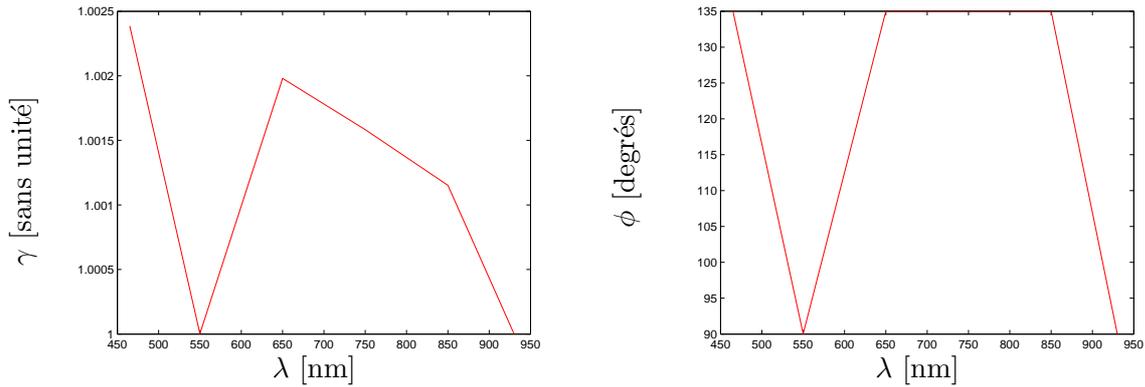


FIG. II.8 – Évolution des paramètres  $\gamma$  et  $\phi$  en fonction de la longueur d'onde.

### II.3.3 Modèles circulaires à *seeing* fixé

Au vu de la Fig. II.7, nous proposons deux modèles des variations des paramètres du modèle de Moffat circulaire en fonction des longueurs d'onde, pour une valeur de *seeing* donnée ( $s = 0.5$ ).

#### II.3.3.1 Modèle circulaire avec $\beta$ variable

L'évolution des paramètres  $\alpha$  et  $\beta$  dans le cas d'un modèle de Moffat circulaire est quasi identique au cas précédent, comme en témoigne la Fig. II.7. Ces deux paramètres peuvent encore être modélisés par une fonction linéaire de la longueur d'onde  $\lambda$ , avec des valeurs de pente et d'ordonnée à l'origine pratiquement identiques à celles du cas elliptique.

Le Tab. II.4 synthétise l'expression du modèle circulaire à  $\beta$  variable et montre que les valeurs des indices de qualité sont relativement faibles, presque identiques à ceux du modèle elliptique. Par conséquent, ce comparatif confirme que ce modèle circulaire est suffisant et satisfaisant pour modéliser la FSF de l'atmosphère et du télescope en mode grand champ sans optique adaptative.

#### II.3.3.2 Modèle circulaire avec $\beta$ constant

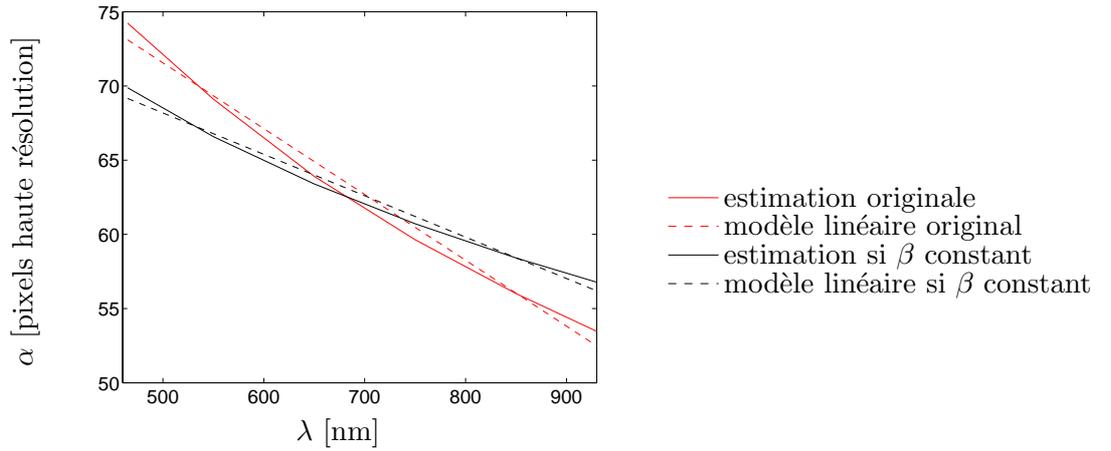
Dans le modèle précédent,  $\alpha$  évolue dans l'intervalle plutôt large [51; 73] alors que  $\beta$  varie sur l'intervalle [2.53; 2.90] plus étroit. Par conséquent, il est envisageable de considérer le paramètre  $\beta$  constant. La moyenne des estimations de  $\beta$  donne  $\bar{\beta} = 2.6559$ .

En fixant  $\beta$  à cette valeur, les valeurs optimales du paramètre  $\alpha$  sont légèrement différentes qu'avec  $\beta$  variable. On constate dans la Fig. II.9 que le paramètre d'échelle est diminué pour les faibles longueurs d'onde et augmenté pour les fortes longueurs d'onde afin de compenser la non variation de forme. Néanmoins, la variation de  $\alpha$  à  $\beta$  fixé reste modélisable par une fonction linéaire de la longueur d'onde.

Les performances de ce modèle, détaillées dans le Tab. II.4, s'avèrent très proches de celles du modèle tenant compte de la variation de  $\beta$ , ces deux séries de statistiques ne différant que de moins de 1%. À ce stade, seuls 3 paramètres suffisent pour modéliser la FSF en tout point du cube de données dans le cas sans optique adaptative. Cependant, les paramètres de variation de  $\alpha$  et la valeur de  $\bar{\beta}$  dépendent des conditions d'observation. On peut donc se demander si l'on peut prédire ces valeurs des conditions d'observation ou si l'on doit absolument les estimer à partir des données.

### II.3.4 Modèles circulaires en fonction du *seeing*

Il serait donc intéressant de disposer d'un tel modèle d'approximation de la FSF pour toute longueur d'onde pour chaque valeur du *seeing*. Ce dernier, grandeur physique liée à la qualité optique


 FIG. II.9 – Évolutions et modélisations de  $\alpha$  avec  $\beta$  variable ou constant.

Modèle	Paramètres $\alpha \beta \gamma \phi$ à $s = 0.5''$	Erreur aux données $\{\lambda_i\}_{i \in [1;6]}$					
		$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
Gaussien circulaire	optimal à chaque longueur d'onde	15.48%	16.88%	25.99%	28.09%	10.07%	11.40%
MF elliptique	optimal à chaque longueur d'onde	2.07%	2.11%	5.33%	5.38%	2.18%	2.27%
Modélisation elliptique	$\begin{cases} \alpha(\lambda) = 93.7342 - 0.0443\lambda \\ \beta(\lambda) = 3.2119 - 7.9531 \cdot 10^{-4}\lambda \\ \gamma = 1.0012 \\ \phi = 135^\circ \end{cases}$	2.42%	2.67%	5.53%	6.02%	2.62%	4.89%
MF circulaire	optimal à chaque longueur d'onde	2.08%	2.11%	5.34%	5.38%	2.18%	2.27%
Modélisation circulaire à $\beta$ variable	$\begin{cases} \alpha(\lambda) = 93.8455 - 0.0445\lambda \\ \beta(\lambda) = 3.2119 - 7.9528 \cdot 10^{-4}\lambda \end{cases}$	2.43%	2.70%	5.53%	6.03%	2.62%	4.96%
Modélisation circulaire à $\beta$ constant	$\begin{cases} \alpha(\lambda) = 82.1278 - 0.0279\lambda \\ \beta(\lambda) = 2.6559 \end{cases}$	2.50%	2.84%	5.79%	6.81%	2.38%	5.50%

 TAB. II.4 – Indices d'erreur des modèles de FSF optimaux à chaque longueur d'onde et des modèles dont les paramètres sont approximés, pour les données en mode grand champ sans optique adaptative à  $s = 0.5''$ .

du ciel, est défini comme la largeur à mi-hauteur de la FSF de l'atmosphère. Il peut être déduit des conditions d'observation par le DIMM (*Differential Image Motion Monitor* cf. I.2.1).

La Fig. II.10 illustre la variation des paramètres  $\alpha$  et  $\beta$  en fonction de la longueur d'onde pour différentes valeurs du *seeing*. On sait que la largeur à mi-hauteur est majoritairement influencée par  $\alpha$ . Or, on observe que le paramètre  $\alpha$  augmente fortement et le paramètre  $\beta$  augmente légèrement avec la valeur du *seeing*. D'après la définition du *seeing*, il est donc parfaitement logique que le paramètre  $\alpha$  du modèle de la FSF de l'ensemble {atmosphère+télescope} augmente fortement avec le *seeing*.

On constate également que le paramètre  $\beta$  a une plus grande variation en fonction de  $\lambda$  pour les faibles *seeings* que pour les forts *seeings*. La forme de la FSF est par conséquent plus susceptible de varier avec la longueur d'onde lorsque que le *seeing* est faible.

Les paramètres  $\alpha$  et  $\beta$  peuvent être modélisés de manière empirique par une fonction linéaire du

*seeing*. On obtient les modèles et les indices d'erreur détaillés dans le Tab. II.5. On constate ainsi qu'une telle modélisation est envisageable pour une erreur d'approximation restant assez faible, mais avec certaines erreurs extrêmes dépassant 15%. Ainsi, on observe une nette dégradation de la qualité de la FSF estimée si la modélisation des paramètres tient compte du *seeing*, par rapport à la qualité du modèle optimal. Il est donc préférable d'estimer les paramètres du modèle de FSF directement à partir des données, plutôt que de les déduire de la valeur du *seeing* mesurée par le DIMM avec une erreur pouvant aller jusqu'à 30%.

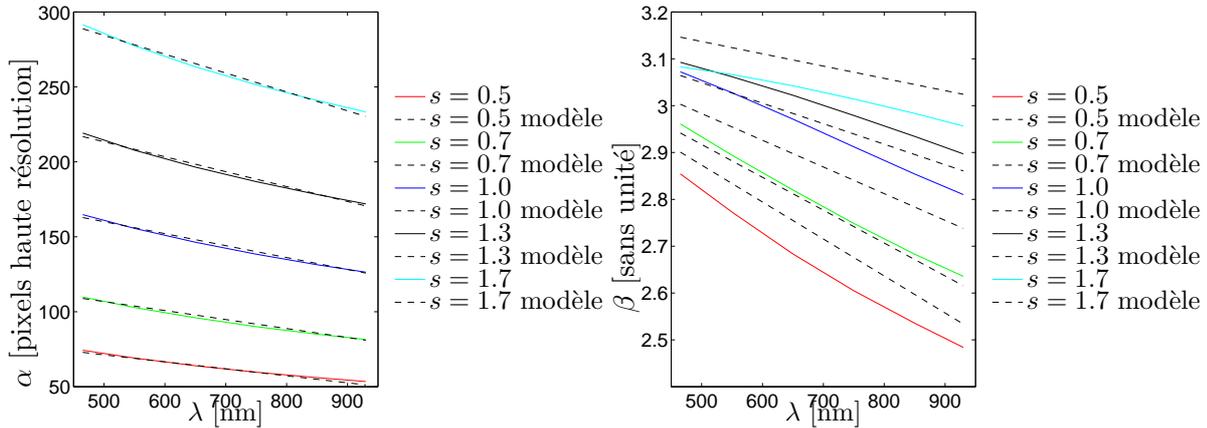


FIG. II.10 – Valeurs optimales de  $\alpha$  et  $\beta$  en fonction de la longueur d'onde  $\lambda$  pour différentes valeurs du *seeing* - Superposition aux modèles linéaires à chaque *seeing*, calculés à partir du modèle global multi-linéaire du Tab. II.5 (- -).

Modèle	Paramètres $\alpha \beta$	Erreur aux données $\{\lambda_i, s_j\}_{i \in [1;6], j \in [1;5]}$					
		$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
MF circulaire	optimal à chaque longueur d'onde et chaque <i>seeing</i>	2.19%	2.45%	5.57%	6.17%	2.30%	2.53%
Modélisation circulaire	$\begin{cases} \alpha = \alpha_1 \cdot \lambda + \alpha_0 \\ \alpha_1 = -0.0660 \cdot s - 0.0137 \\ \alpha_0 = 210.6098 \cdot s - 10.6948 \\ \beta = \beta_1 \cdot \lambda + \beta_0 \\ \beta_1 = 4.3944 * 10^{-4} \cdot s - 0.0010 \\ \beta_0 = 3.2680 \end{cases}$	3.58%	8.53%	6.49%	8.24%	4.34%	15.02%

TAB. II.5 – Indices d'erreur des modèles de FSF optimaux à chaque longueur d'onde et chaque *seeing*, et des modèles dont les paramètres sont approximés, pour les données WFM sans OA.

### II.3.5 Influence du $C_n^2$

Les simulations que nous avons utilisées dans les sections précédentes ont été réalisées à des valeurs fixées de certains paramètres atmosphériques, dont le  $C_n^2$  et l'échelle externe  $L_0$  (cf. Tab. II.1). Nous souhaitons vérifier que les modèles que nous avons proposés sont robustes à la variation d'un de ces paramètres.

Le  $C_n^2$  qualifie la configurations des couches de turbulences dans l'atmosphère. Dans les simulations PAOLA dont nous disposons, il peut être qualifié d'inférieur, de médian ou de supérieur selon l'altitude des turbulences les plus importantes.

Comme décrit dans le Tab. II.6, le  $C_n^2$  a très peu d'influence sur la FWHM de la FSF, quels que soient

la longueur d'onde et le *seeing*. En effet, l'écart-type entre les valeurs de la FWHM pour les 3  $C_n^2$  est inférieur à 0.1%. Le flux des FSF étant normalisé, les paramètres  $\alpha$  et  $\beta$  sont donc indépendants du  $C_n^2$ . Nous pouvons donc en conclure que les évolutions des paramètres  $\alpha$  et  $\beta$  seront du même ordre avec des valeurs différentes du  $C_n^2$  et que les modèles que nous proposons seront encore valables.

Longueur d'onde $\lambda$ [nm]	<i>Seeing</i> $s$			
	0.5''	0.7''	1.0''	1.3''
	Moyenne de la FWHM sur les 3 $C_n^2$ (en pixel haute résolution cf. Tab. II.1)			
500	396.32	572.93	842.28	1116.58
650	363.04	527.76	779.52	1035.93
900	322.63	473.08	704.02	939.39
	Écart-type des FWHM sur les 3 $C_n^2$			
500	0.04%	0.05%	0.05%	0.08%
650	0.02%	0.02%	0.03%	0.06%
900	0.01%	0.01%	0.01%	0.04%

TAB. II.6 – Influence du  $C_n^2$  sur la valeur de la largeur à mi-hauteur à différents ( $s, \lambda$ ).

### II.3.6 Influence de l'échelle externe

L'échelle externe est un paramètre caractérisant l'atmosphère. Elle représente la distance au delà de laquelle « les fluctuations de phase prises entre deux points du front d'onde se décorrèlent »<sup>13</sup>. L'échelle externe de turbulence varie entre quelques mètres et quelques centaines de mètres au voisinage du sol. Elle diffère beaucoup selon les points d'observation et l'état de la turbulence [LEN 08]. La Fig. II.11 illustre la variation des paramètres  $\alpha$  et  $\beta$  en fonction de l'échelle externe, à différentes valeurs de la longueur d'onde. On observe que ces deux paramètres augmentent avec  $L_0$ . L'augmentation de  $\beta$  étant relativement moins significative que celle de  $\alpha$ , il en résulte une augmentation de la largeur à mi-hauteur en fonction de l'échelle externe.

En complétant la modélisation des paramètres  $\alpha$  et  $\beta$  par une fonction linéaire de  $L_0$ , on obtient les modèles et les indices d'erreur détaillés dans le Tab. II.7. On constate que cette modélisation donne des résultats satisfaisants, proches des résultats optimaux, avec une erreur maximale moyenne de 2.86%. Cependant, un tel modèle est de faible intérêt car contrairement au *seeing*, l'échelle externe n'est pas mesurée durant l'acquisition des données.

Modèle circulaire	statistiques sur $\{\lambda_i, L_{0,j}\}_{i,j}$					
	$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
Paramètres optimaux pour chaque $(\lambda, L_0)$	2.16%	2.31%	5.47%	5.81%	2.31%	2.47%
Modèle de variation bi-linéaire pour $\alpha$ et $\beta$ $\begin{cases} \alpha(\lambda, L_0) = 145.58 + 5.23 \cdot 10^{-1} L_0 - 6.82 \cdot 10^{-2} \lambda \\ \beta(\lambda, L_0) = 3.12 + 6.3 \cdot 10^{-3} L_0 - 6.47 \cdot 10^{-4} \lambda \end{cases}$	2.55%	3.47%	5.72%	6.92%	2.86%	6.78%

TAB. II.7 – WFM sans optique adaptative : Performances des modèles de  $\alpha$  et  $\beta$  multi-linéaires en fonction de  $\lambda$  et  $L_0$  avec  $s = 0.8''$ ,  $\lambda \in \{465; 500; 550; 650; 750; 850; 930\}$  nm,  $L_0 \in \{15; 22; 25; 40\}$  m,  $C_n^2$  médian. Comparaison au modèle de Moffat avec les paramètres optimaux.

<sup>13</sup>Source : <http://www-luan.unice.fr/~mekarnia/hivernage/science.html>.

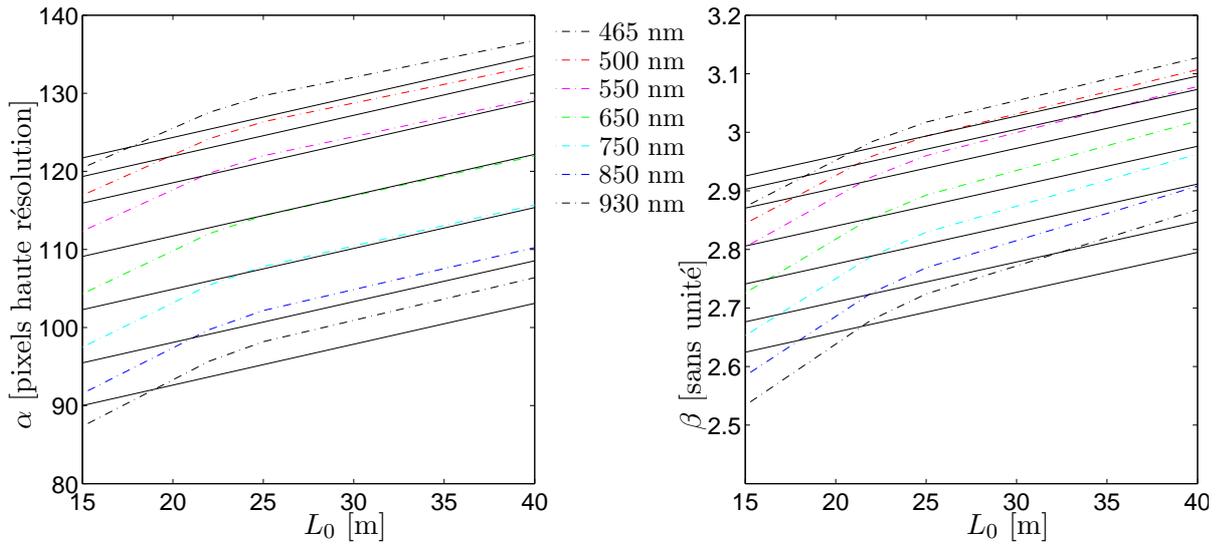


FIG. II.11 – WFM sans optique adaptative : Évolution des paramètres  $\alpha$  et  $\beta$  en fonction de l'échelle externe  $L_0$  pour différentes valeurs de la longueur d'onde  $\lambda$ . Superposition aux modèles de variation linéaire proposés.

### II.3.7 Synthèse

En mode WFM sans optique adaptative, nous avons montré que la FSF de l'atmosphère et du télescope, variable en longueur d'onde, peut être approximée de façon satisfaisante par une fonction de Moffat circulaire, dont les paramètres sont modélisés par une fonction simple. Ainsi, le paramètre d'échelle  $\alpha$  peut être modélisé par une fonction linéaire de la longueur d'onde, voire une fonction quadratique ou une loi de puissance, si l'on souhaite davantage de précision. Le paramètre de forme  $\beta$  peut être approximé par une fonction linéaire ou par une constante si l'on préfère un modèle simple. Nous pouvons donc fournir une approximation satisfaisante de la FSF sans optique adaptative en tout point du cube de données avec seulement trois paramètres. L'estimation de ces paramètres est l'objet du chapitre III. Enfin, nous avons vérifié la robustesse de cette modélisation vis-à-vis de la variation des caractéristiques de l'atmosphère qui sont quantifiées par le *seeing*, le  $C_n^2$  et l'échelle externe.

## II.4 Modélisation de la FSF en mode WFM avec optique adaptative

Un système de correction de la turbulence atmosphérique basse couche par optique adaptative (GLAO) devrait être disponible à partir de la 3<sup>ème</sup> année d'exploitation de l'instrument MUSE. Dans ce cas, contrairement au cas sans optique adaptative étudié précédemment, la FSF de la chaîne d'acquisition peut varier en fonction de la position de la source dans le champ de vue. En effet, la correction fournie par ce système sera plus performante au centre du champ que sur les bords. La modélisation des paramètres de la fonction de Moffat devient alors plus complexe car ceux-ci dépendent en outre de variables spatiales. Notre objectif est de quantifier les variations spatiales et spectrales de la FSF et de proposer si possible un modèle permettant de les approximer.

### II.4.1 Données et hypothèses

Comme dans le cas sans optique adaptative, nous avons étudié les variations des paramètres du modèle de Moffat elliptique en fonction de la longueur d'onde et de la position, et nous proposons plusieurs hypothèses simplificatrices, issues de la géométrie du problème.

Notons que la simulation des FSF avec PAOLA ne fait pas intervenir l'utilisation d'une étoile naturelle,

située en dehors du champ, pour corriger le *tip-tilt*. Cette correction n'étant pas parfaite, elle devrait introduire une ellipticité de la FSF, d'orientation constante dans le champ dans la direction de l'étoile [CAR 12]. En estimant cette ellipticité, il serait alors possible de corriger notre modèle en le convoluant par un modèle elliptique.

#### II.4.1.1 Symétrie des quadrants

Nous émettons l'hypothèse que les paramètres de la FSF sont symétriques par rapport aux axes  $Ox$  et  $Oy$  du champ de vue, du fait de la position des quatre étoiles guides laser sur ces deux axes. Par conséquent, nous réduisons l'étude de leur variation au quadrant ( $x \geq 0, y \geq 0$ ) ou ( $\rho \geq 0, \theta \in [0^\circ; 90^\circ]$ ).

#### II.4.1.2 Symétrie par rapport à la 1<sup>ère</sup> diagonale

Chaque quadrant est composé de deux octants. Afin de réduire encore la zone du champ de vue étudiée, nous cherchons à savoir si l'approximation de symétrie de deux octants, dans ce cas précis par rapport à l'axe faisant un angle de  $+45^\circ$  par rapport à l'axe  $Ox$ , est plausible. Afin d'étudier cette symétrie, nous comparons les symétriques des FSF du deuxième octant situées à  $\theta \in \{60^\circ, 75^\circ, 90^\circ\}$  par rapport à la première diagonale, aux FSF du premier octant situées à  $\theta \in \{30^\circ, 15^\circ, 0^\circ\}$ , aux différents rayons  $\rho \in \{20, 30, 40\}$ . Les erreurs entre ces FSF appariées sont évaluées par les moyennes et les maxima de indices d'erreur dans la première ligne du Tab. II.8. Si les FSF du deuxième octant sont symétriques à celle du premier, alors les FSF comparées sont identiques et les indices d'erreur doivent être nuls. Dans cette même hypothèse, les FSF situées sur l'axe  $+45^\circ$  doivent être elles-mêmes symétriques par rapport à cet axe. La deuxième ligne du Tab. II.8 contient les normes d'erreur entre ces FSF et leur symétrique. Ces deux séries d'erreur n'étant pas négligeable, on ne peut pas recourir à l'hypothèse de la symétrie des FSF dans le champ par rapport à l'axe  $+45^\circ$ .

	Statistiques sur les paires d'octants					
	$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
Différence au symétrique	3.38%	5.63%	2.60%	3.98%	5.15%	8.68%
Erreur de symétrie à $45^\circ$	0.84%	1.16%	1.00%	1.40%	0.43%	0.58%
Différence à $\theta = 0$ et même $\rho$	5.75%	10.37%	6.69%	11.18%	5.68%	11.22%
Erreur due à la rotation $\times 10^{14}$ (%)	3.80	3.83	10.5	10.6	7.37	9.20

TAB. II.8 – Étude des symétries dans le 1<sup>er</sup> quadrant à  $\lambda = 650nm$

#### II.4.1.3 Invariance par rotation

Nous souhaitons évaluer la variation des FSF selon l'angle dans le champ, à l'intérieur du quadrant, afin de vérifier la validité de l'hypothèse d'invariance par rotation. Pour cela, nous effectuons une rotation de toutes les FSF du quadrant d'un angle  $-\theta$  avec  $\theta$  correspondant à leur coordonnée polaire, pour les comparer ensuite avec la FSF située à  $(\rho, \theta = 0)$ .

D'après la 4<sup>ème</sup> ligne du Tab. II.8, l'erreur numérique due à l'opération de rotation des FSF est négligeable ( $< 10^{-13}$ ). Les normes d'erreur entre une FSF et sa FSF correspondante sur l'axe  $Ox$ , dans la troisième ligne du Tab. II.8 sont relativement importantes et permettent d'écarter pour l'instant l'hypothèse de la symétrie de rotation dans le champ des FSF.

#### II.4.1.4 Influence de l'intégration

Les données réelles devant être acquises par l'imageur MUSE le seront pendant un temps d'intégration d'une heure. Pendant cette période, le champ de vue tourne d'un angle de  $15^\circ$  maximum, dans le plan

de l'image afin de garder les objets observés fixes dans le champ. Les simulations de FSF sur lesquelles est basée cette étude correspondent à des FSF instantanées ne prenant en compte ni la rotation du champ ni l'intégration.

La question se pose alors de l'influence de ces deux phénomènes sur la valeur des paramètres de la fonction de Moffat. Pour l'évaluer, une FSF « intégrée » est calculée comme la somme de toutes les FSF situées à  $\rho = 30''$  aux angles compris entre  $0^\circ$  et  $15^\circ$  tous les  $1^\circ$ . Ce mode de calcul suppose que la vitesse de rotation du champ est constante, hypothèse peu probable mais retenue en première approximation. Toutes les FSF considérées sont normalisées par leur flux et appartiennent à la dernière longueur d'onde, à laquelle la FSF est la plus petite et où les variations sont le plus susceptible d'être perçues. On estime ensuite l'ensemble de ces FSF instantanées et la FSF intégrée par une fonction de Moffat, dont les paramètres sont rassemblés dans le Tab. II.9.

On constate que le paramètre de forme  $\beta$  et d'ellipticité  $\gamma$  restent inchangés, même si l'on s'attendait à voir décroître l'ellipticité du fait de la rotation du champ. Le paramètre  $\alpha$  augmente par rapport à la moyenne de l'intervalle d'intégration, ce qui représente logiquement un élargissement de la FSF. De plus, l'orientation des isophotes de la FSF intégrée est égale à la moyenne des orientations des FSF instantanées sur l'intervalle considéré. La rotation et l'intégration augmentent donc la FWHM de la FSF via le paramètre  $\alpha$  et modifient son orientation via le paramètre  $\phi$ . Enfin, on constate que l'erreur de modélisation de la FSF intégrée par la fonction de Moffat est très proche de celles des FSF instantanées. La fonction de Moffat est donc autant appropriée pour approximer les FSF instantanées que pour les FSF intégrées.

Paramètres de la MF de FSF instantanées/intégrée $\lambda = 930 \text{ nm} - s = 0.5'' - C_n^2 \text{ médian} - \rho = 30''$					
	$\alpha(\text{px})$	$\beta$	$\gamma$	$\phi(^{\circ})$	$\varepsilon_2$
FSF instantanées : Statistiques sur $\{\theta(^{\circ})\} = [0:1:15]$					
Minimum	15.96	1.706	1.077	71.00	1.92%
Moyenne	16.79	1.707	1.078	79.99	1.93%
Maximum	17.27	1.709	1.079	89.27	1.94%
FSF intégrée	17.17	1.707	1.076	79.94	1.92%

TAB. II.9 – Influence de la rotation et de l'intégration sur les paramètres de la fonction de Moffat. Erreur  $\varepsilon_2$  entre la FSF simulée instantanée / intégrée et son approximation par le modèle de Moffat.

## II.4.2 Modèle circulaire constant dans le champ

En mode grand champ sans correction par optique adaptative, le modèle adopté est la fonction de Moffat circulaire constante dans le champ. Avec cette hypothèse, l'observation est bien le résultat de la convolution de l'image source par la FSF, ce qui n'est pas le cas lorsque la FSF varie dans le champ. En outre, les paramètres sont uniquement variables en fonction de la longueur d'onde, ce qui réduit le nombre d'hyper-paramètres à seulement 4 (3 si  $\beta$  constant).

Afin de déterminer dans quelle mesure ce modèle pourrait être utilisé dans le mode avec GLAO, nous avons estimé les valeurs optimales des paramètres  $\alpha$  et  $\beta$  puis construit leur propre modèle, comme précédemment.

Pour commencer, il est crucial d'évaluer la pertinence de la fonction de Moffat circulaire pour approximer la FSF du mode grand champ avec correction par optique adaptative, en différents points du champ. Pour cela, nous pouvons comparer les performances de la fonction de Moffat circulaire dont les paramètres sont optimaux à chaque longueur d'onde, dans ce cas avec optique adaptative (cf. 1<sup>ère</sup> ligne du Tab. II.10) par rapport au cas précédent sans OA (cf. 4<sup>ème</sup> ligne du Tab. II.4). Ainsi, l'erreur quadratique moyenne  $\bar{\varepsilon}_2$  passe de 2.08% dans le cas sans OA, à 4.65% dans l'autre cas. En outre, l'er-

reur maximale sur l'ensemble des simulations de FSF passe de 2.27% dans le cas sans OA, à 16.21% dans le cas présent. On constate donc une nette augmentation, vis-à-vis de l'erreur  $\varepsilon_{\infty, max}$  dans le cas sans OA mais également de l'erreur quadratique moyenne avec OA. Il s'avère que les erreurs les plus grandes sont observées dans les bords du champ. Ceci révèle la variabilité de la FSF dans le champ qui n'est pas prise en compte par le modèle. Nous pouvons donc en déduire que le modèle de Moffat constant dans le champ est moins approprié pour le cas avec OA que dans le cas sans OA.

Par ailleurs, si l'on regarde les performances du modèle dont le paramètre  $\alpha$  est approximé par une fonction linéaire de la longueur d'onde et  $\beta$  par une constante (cf. 2<sup>nd</sup>e ligne du Tab. II.10), on remarque que l'erreur maximale atteint en moyenne 6.04% ce qui reste bas dans l'absolu. Néanmoins, on constate une forte augmentation de tous les indices d'erreur par rapport au cas de l'approximation de la FSF par le même modèle dans le cas sans OA (cf. 6<sup>ème</sup> ligne du Tab. II.4).

Enfin, la Fig. II.12, dont la partie droite présente l'erreur horizontale et verticale entre le modèle de Moffat circulaire et la FSF simulée à une longueur d'onde et une position données, met en lumière la dissymétrie de la FSF. Nous pouvons donc en conclure que l'ajout d'un paramètre d'ellipticité au modèle pourrait permettre d'améliorer la précision de la modélisation.

Modèle MF circulaire	Statistiques sur $\{\rho_i, \theta_j, \lambda_k\}_{i,j,k}$					
	$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
Paramètres optimaux en chaque $\lambda$	4.65%	9.67%	7.26%	11.68%	5.07%	16.21%
Paramètres modélisés $\begin{cases} \alpha(\lambda) = 34.85 - 1.91 \cdot 10^{-2} \lambda \\ \beta = 1.76 \end{cases}$	5.45%	11.49%	7.64%	12.29%	6.04%	15.43%

TAB. II.10 – WFM avec optique adaptative : Performances du modèle circulaire constant dans le champ à  $s = 0.5''$ ,  $C_n^2$  médian - Statistiques des indices d'erreur sur  $\rho \in \{0; 20; 30; 40\}''$ ,  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}$  degrés,  $\lambda \in \{465; 550; 650; 750; 850; 930\}$  nm - Comparaison aux performances du modèle de Moffat basé sur les paramètres estimés.

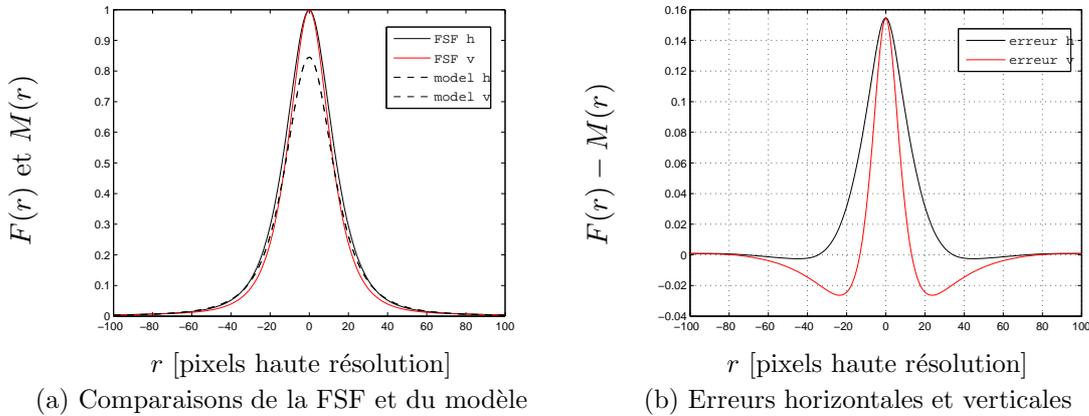


FIG. II.12 – Comparaison entre la FSF simulée et le modèle de Moffat elliptique à  $\lambda = 750$  nm,  $\rho = 40''$ ,  $\theta = 0^\circ$  - Extraits 1D horizontaux et verticaux - Zoom sur les 200 pixels centraux.

### II.4.3 Modèle elliptique variable dans le champ

La correction par optique adaptative induit une variation des FSF dans le champ qui s'ajoute à la variation avec la longueur d'onde. En particulier, les FSF tendent à devenir elliptiques vers les bords du champ.

Nous analyserons le comportement des paramètres du modèle de Moffat en fonction de la position dans le champ et de la longueur d'onde  $\lambda$  puis nous proposerons un modèle simple de variation. Ainsi, nous pourrions évaluer dans quelle mesure la prise en compte de l'ellipticité de la FSF et de sa variabilité dans le champ apporte une amélioration à la modélisation.

#### II.4.3.1 Variations de l'orientation $\phi$

Les variations du paramètre d'orientation  $\phi$  sont illustrées dans la Fig. II.13. Pour les sources situées à une distance  $\rho$  supérieure ou égale à  $30''$  du centre du champ, on observe une variation de l'orientation  $\phi$  en fonction de la longueur d'onde, pouvant aller jusqu'à  $20^\circ$  (Fig. II.13 (a)). Néanmoins, certaines de ces sources sont à l'extérieur du quadrant de taille  $30'' \times 30''$ . Pour les autres points situés à l'intérieur, il apparaît clairement que  $\phi$  est indépendant de la longueur d'onde. Ensuite, si on écarte le point situé au centre du champ où la fonction de Moffat est quasi circulaire ( $\gamma \simeq 1$ ), on peut considérer que les variations de  $\phi$  par rapport à la distance au centre  $\rho$  sont négligeables (Fig. II.13 (b)).

En conservant cette restriction, la variation de  $\phi$  en fonction de l'angle  $\theta$  (Fig. II.13 (c)) est monotone et décroissante et peut être modélisée sans grande approximation par la fonction linéaire :

$$\phi = 90 - \theta.$$

La 2<sup>ème</sup> ligne du Tab. II.11 montre que les erreurs aux données obtenues avec ce modèle sont particulièrement proches de celles obtenues avec les valeurs optimales de  $\phi$ . Le choix de cette fonction linéaire est donc approprié.

#### II.4.3.2 Variations de l'ellipticité $\gamma$

Les variations du paramètre d'ellipticité  $\gamma$  sont illustrées par la Fig. II.14. D'après la troisième série de figures, la variation de  $\gamma$  en fonction de l'angle  $\theta$  est relativement faible et sera négligée. Par ailleurs, l'ellipticité  $\gamma$  semble dépendre des deux autres variables  $\lambda$  et  $\rho$  de façon croissante et monotone, de sorte que ces variations peuvent être approximées par une fonction linéaire. En pratique, on constate donc que la FSF est plus elliptique aux bords du champ qu'au centre et plus elliptique dans le rouge que dans le bleu.

Pour conserver la cohérence du modèle construit avec la modélisation du paramètre  $\phi$ , la fonction de Moffat est considérée circulaire au centre. On peut envisager le modèle de variation de  $\gamma$  suivant :

$$\gamma = 1 + (\gamma_0 + \gamma_1 \cdot \lambda) \rho.$$

Les performances de ce modèle, données dans la 3<sup>ème</sup> ligne du Tab. II.11, sont comparables à celles obtenues avec les paramètres optimaux et à peine dégradées par rapport à celles obtenues avec le modèle de variation de  $\phi$  uniquement (avec  $\alpha$  et  $\beta$  optimaux). Ceci valide la qualité du modèle de variation de  $\gamma$ . Par la suite, ce modèle sera utilisé pour étudier la variation des autres paramètres.

#### II.4.3.3 Variations du paramètre de forme $\beta$

Les variations du paramètre de forme  $\beta$  sont illustrées par la Fig. II.15. Celles-ci sont plutôt limitées par rapport à la valeur moyenne de  $\beta$  et difficilement modélisables. Cela signifie que la forme de la FSF est globalement homogène dans le champ de vue pour toutes les longueurs d'onde.

En considérant  $\beta$  constant, on obtient le modèle et les indices de qualité détaillés dans la 4<sup>ème</sup> ligne du Tab. II.11. On constate une légère augmentation des statistiques d'erreur, augmentation plus importante que lors de la modélisation des deux paramètres précédents. L'approximation des FSF par un modèle de Moffat de forme constante doit donc être utilisée avec précaution et la valeur de  $\beta$  doit être ajustée judicieusement.

Toutefois, les indices d'erreur obtenus avec les trois premiers paramètres modélisés sont assez faibles, avec par exemple une erreur maximale valant en moyenne 2.27%. Par la suite,  $\beta$  sera considéré constant comme dans le cas sans optique adaptative :

$$\beta = \beta_0.$$

#### II.4.3.4 Variations du paramètre d'échelle $\alpha$

Les variations du paramètre d'échelle  $\alpha$  sont illustrées par la Fig. II.16. Les variations en fonction de  $\lambda$  et  $\rho$  sont monotones décroissantes, celles en fonction de  $\theta$  négligeables. Lorsque  $\beta$  est considéré constant, le paramètre  $\alpha$  est proportionnel à la FWHM. Ainsi, il apparaît que la FSF est plus large dans le bleu que dans le rouge et plus large au centre du champ que sur les bords. Si cette dernière remarque est surprenante, elle peut être expliquée par le positionnement des étoiles guides laser à l'extérieur du champ.

Nous avons choisi de modéliser les variations de  $\alpha$  par une fonction linéaire de  $\rho$  et une fonction quadratique de  $\lambda$  :

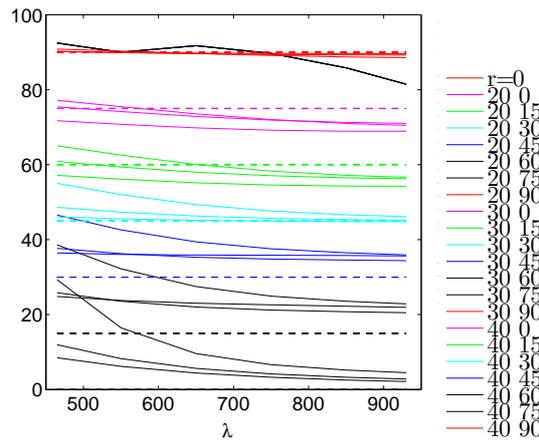
$$\alpha = \alpha_0 + \alpha_1\rho + \alpha_2\lambda + \alpha_3\lambda^2.$$

En première approximation, une fonction linéaire de  $\lambda$  pourrait suffire. Il faut cependant se rappeler que le *seeing* fixé lors de l'étude est relativement optimiste et que l'aspect quadratique de cette variation augmente avec le *seeing*, d'après les résultats obtenus en WFM sans optique adaptative.

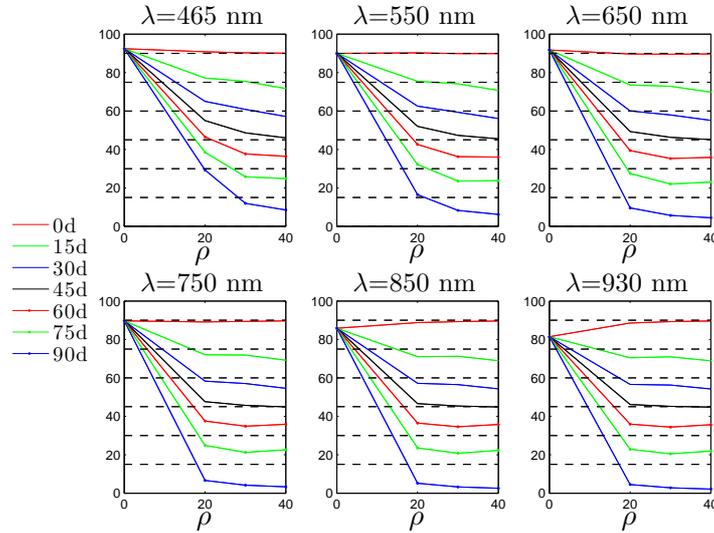
La dernière ligne du Tab. II.11 compare les performances de la fonction de Moffat dont les paramètres sont tous modélisés, avec celle dont les paramètres sont optimaux. Les premières erreurs sont légèrement supérieures par rapport aux secondes mais restent relativement basses. En effet, l'erreur maximale entre les FSF simulées et leur modèle atteint en moyenne seulement 3.00%. Elle peut atteindre 10.39% dans le cas illustré par la Fig. II.17 mais on notera qu'il s'agit d'un modèle calculé en dehors du champ. Enfin, nous pouvons remarquer que les indices d'erreur associés à ce modèle de variation des quatre paramètres  $[\alpha, \beta, \gamma, \phi]$ , sont nettement inférieurs d'un facteur 1.5 à 2 à ceux obtenus avec le modèle circulaire constant dans le champ (cf. Tab II.10). La prise en compte de l'ellipticité et de la variabilité spatiale de la FSF apporte donc une réelle amélioration de la modélisation.

#### II.4.3.5 Conclusion

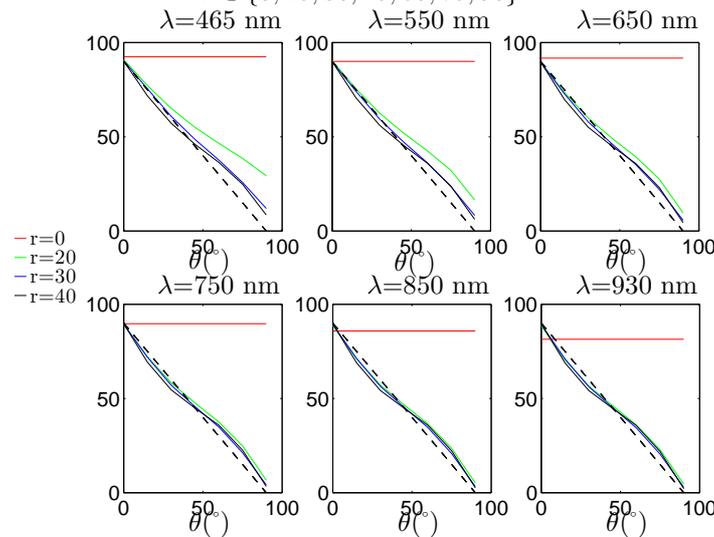
Pour le mode grand champ avec correction par optique adaptative, nous avons proposé un modèle de variations des quatre paramètres  $[\alpha, \beta, \gamma, \phi]$  de la fonction de Moffat elliptique en fonction de la longueur d'onde et de la position dans le champ. Il découle de la modélisation de ces quatre paramètres que la FSF peut être approximée par un modèle à symétrie de révolution. Ce modèle de variations s'exprime en fonction de 7 hyper-paramètres  $[\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \gamma_0, \gamma_1]$ . Ainsi, lorsque ces paramètres sont connus, il est possible de connaître la FSF en tout point du cube de données avec une erreur relativement faible (10% maximum).



(a)  $\phi$  (°) en fonction de la longueur d'onde  $\lambda$  [nm], pour différentes positions  $(\rho, \theta)$  dans le champ avec  $\rho \in \{0; 20; 30; 40\}$  et  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}$ .

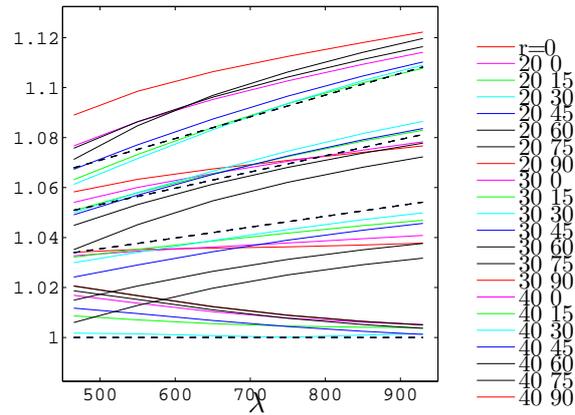


(b)  $\phi$  en fonction de la distance au centre du champ  $\rho$ , pour différentes valeurs de  $\lambda$  et pour  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}$ .

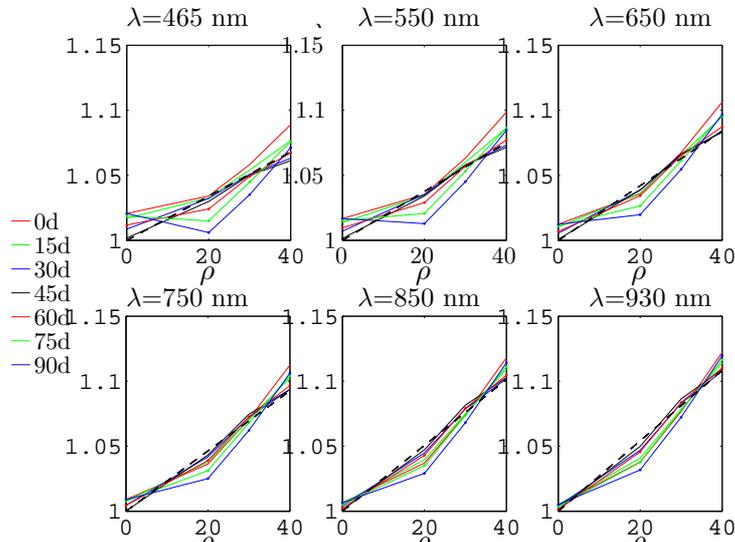


(c)  $\phi$  en fonction de l'angle  $\theta$ , pour différentes valeurs de  $\lambda$  et pour  $\rho \in \{0; 20; 30; 50\}$ .

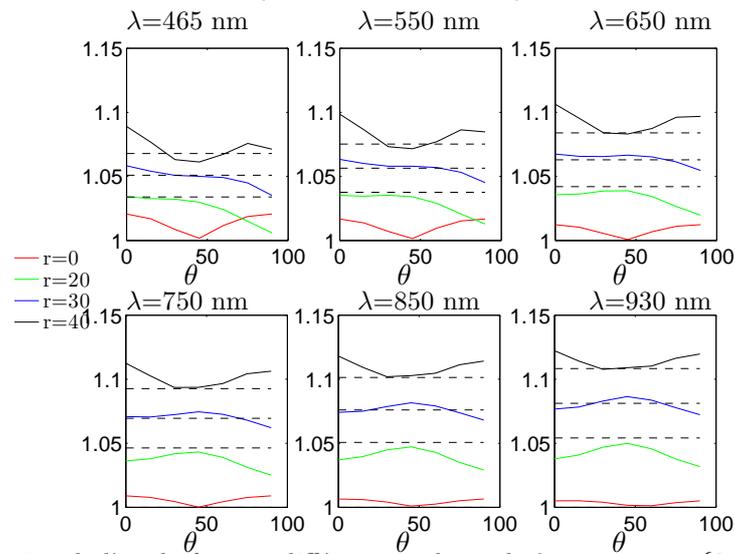
FIG. II.13 – Évolutions et modèles [- -] de variation de  $\phi$  en fonction de différentes variables.



(a)  $\gamma$  en fonction de la longueur d'onde  $\lambda$  [nm], pour différentes positions  $(\rho, \theta)$  dans le champ avec  $\rho \in \{0; 20; 30; 40\}$  et  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}^\circ$ .

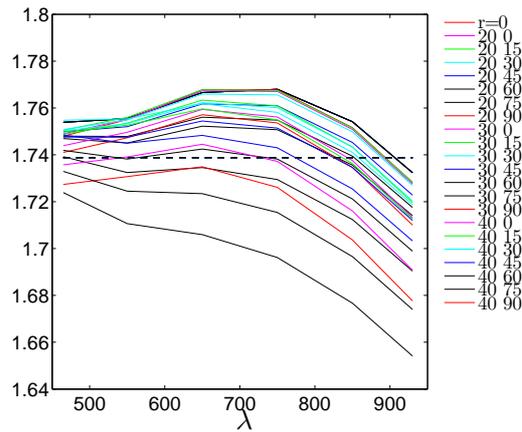


(b)  $\gamma$  en fonction de la distance au centre du champ  $\rho$ , pour différentes valeurs de  $\lambda$  et pour  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}^\circ$ .

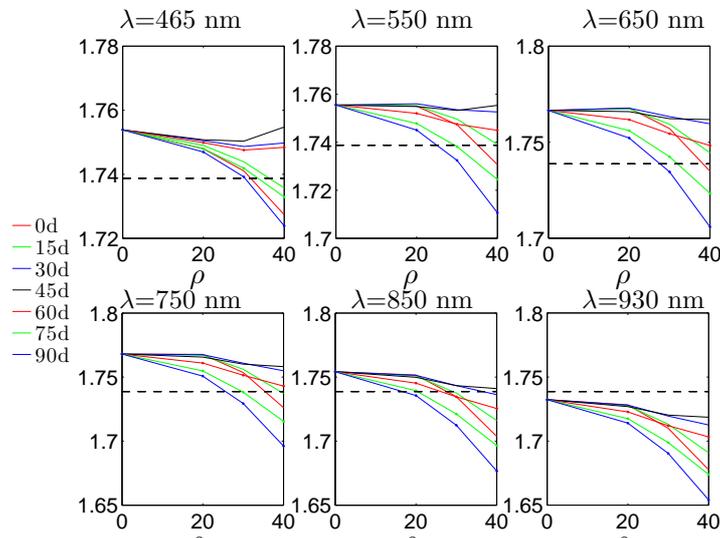


(c)  $\gamma$  en fonction de l'angle  $\theta$ , pour différentes valeurs de  $\lambda$  et pour  $\rho \in \{0; 20; 30; 50\}$ .

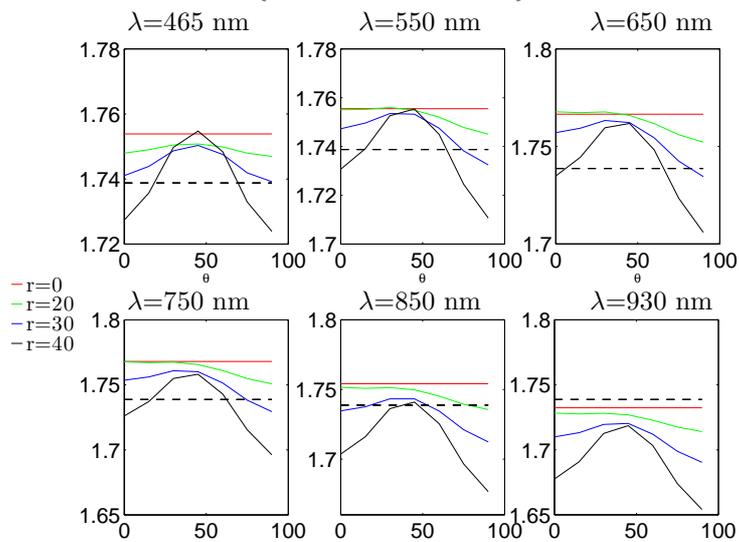
FIG. II.14 – Évolutions et modèles [- -] de variation de  $\gamma$  en fonction de différentes variables.



(a)  $\beta$  en fonction de la longueur d'onde  $\lambda$  [nm], pour différentes positions  $(\rho, \theta)$  dans le champ avec  $\rho \in \{0; 20; 30; 40\}$  et  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}^\circ$ .

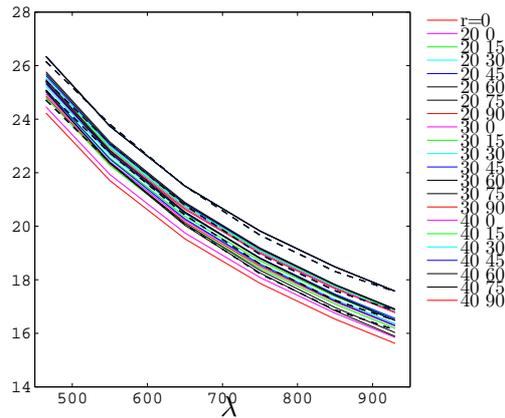


(b)  $\beta$  en fonction de la distance au centre du champ  $\rho$ , pour différentes valeurs de  $\lambda$  et pour  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}^\circ$ .

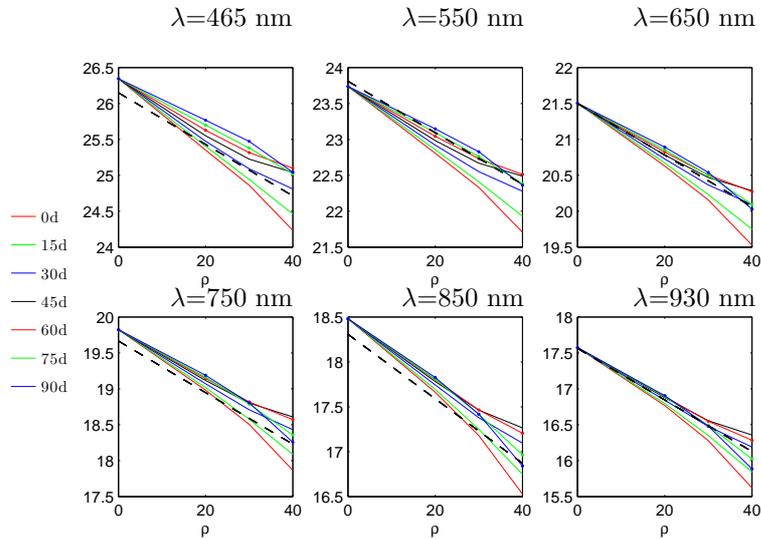


(c)  $\beta$  en fonction de l'angle  $\theta$ , pour différentes valeurs de  $\lambda$  et pour  $\rho \in \{0; 20; 30; 50\}$ .

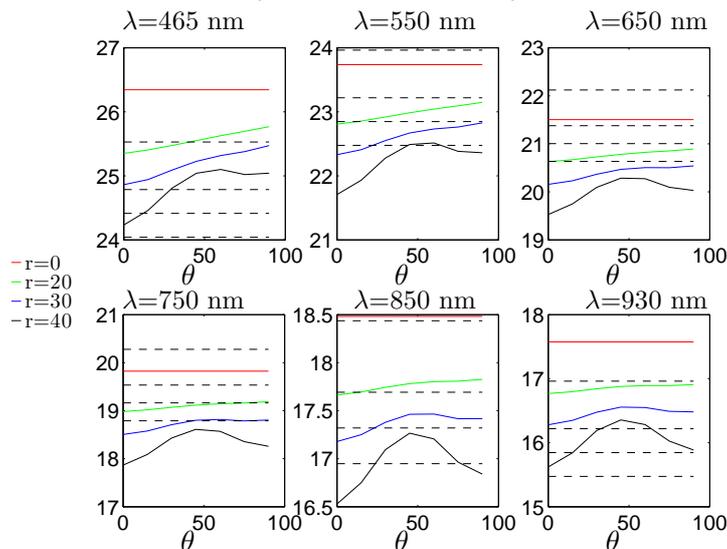
FIG. II.15 – Évolutions et modèles [- -] de variation de  $\beta$  en fonction de différentes variables.



(a)  $\alpha$  en fonction de la longueur d'onde  $\lambda$  [nm], pour différentes positions  $(\rho, \theta)$  dans le champ avec  $\rho \in \{0; 20; 30; 40\}$  et  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}^\circ$ .



(b)  $\alpha$  en fonction de la distance au centre du champ  $\rho$ , pour différentes valeurs de  $\lambda$  et pour  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}^\circ$ .



(b)  $\alpha$  en fonction de l'angle  $\theta$ , pour différentes valeurs de  $\lambda$  et pour  $\rho \in \{0; 20; 30; 50\}$ .

FIG. II.16 – Évolutions et modèles [- -] de variation de  $\alpha$  en fonction de différentes variables.

Modèle MF elliptique	Statistiques sur $\{\rho_i, \theta_j, \lambda_k, \}_{i,j,k}$					
	$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
Paramètres optimaux	2.22%	2.73%	5.38%	6.61%	1.99%	3.00%
$\alpha, \beta$ et $\gamma$ optimaux $\phi(\theta) = 90 - \theta$	2.27%	2.77%	5.39%	6.69%	1.97%	3.03%
$\alpha$ et $\beta$ optimaux $\begin{cases} \phi(\theta) = 90 - \theta \\ \gamma(\rho, \lambda) = 1 + (6.86 \cdot 10^{-4} + 2.17 \cdot 10^{-6} \lambda) \rho \end{cases}$	2.31%	2.77%	5.43%	6.72%	1.98%	3.04%
$\alpha$ optimal $\begin{cases} \phi(\theta) = 90 - \theta \\ \gamma(\rho, \lambda) = 1 + (6.86 \cdot 10^{-4} + 2.17 \cdot 10^{-6} \lambda) \rho \\ \beta = 1.74 \end{cases}$	2.39%	3.13%	5.50%	7.21%	2.27%	5.04%
Paramètres modélisés $\begin{cases} \phi(\theta) = 90 - \theta \\ \gamma(\rho, \lambda) = 1 + (6.86 \cdot 10^{-4} + 2.17 \cdot 10^{-6} \lambda) \rho \\ \beta = 1.74 \\ \alpha = 45.03 - 3.59 \cdot 10^{-2} \rho \\ \quad - 5.17 \cdot 10^{-2} \lambda + 2.38 \cdot 10^{-5} \lambda^2 \end{cases}$	2.69%	4.71%	5.63%	7.80%	3.14%	10.39%

TAB. II.11 – WFM avec optique adaptative : Performances du modèle elliptique variable dans le champ avec différentes valeurs des paramètres  $[\alpha, \beta, \gamma, \phi]$  : optimales ou modélisées. Valeur du seeing :  $s = 0.5''$ ,  $C_n^2$  médian. Moyennes et maxima des indices d'erreur sur  $\rho \in \{0; 20; 30; 40\}''$ ,  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}$  degrés,  $\lambda \in \{465; 550; 650; 750; 850; 930\}$  nm.

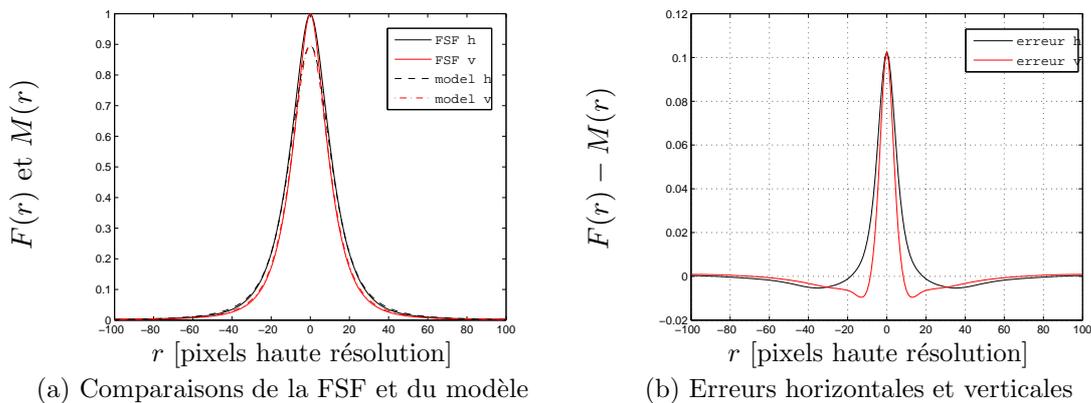


FIG. II.17 – Comparaison entre la FSF simulée et le modèle de Moffat elliptique à  $\lambda = 930$  nm,  $\rho = 40$  arcsec,  $\theta = 0^\circ$  correspondant au cas où l'erreur  $\varepsilon_\infty$  est maximale. Extraits 1D horizontaux et verticaux - Zoom sur les 200 pixels centraux.

## Conclusions

Nous avons proposé plusieurs modèles, basés sur la fonction de Moffat, de la FSF du mode d'observation grand champ de MUSE, principalement due à la turbulence atmosphérique et au télescope. En mode grand champ sans correction par OA, la FSF est constante dans le champ de vue et varie en longueur d'onde. La variation du paramètre d'échelle  $\alpha$  peut être approximée par une fonction linéaire de la longueur d'onde et le paramètre de forme  $\beta$  peut être considéré constant.

Dans le cadre du WFM avec correction par optique adaptative, la FSF varie en tout point du cube de données. Aussi, nous avons proposé un modèle de Moffat elliptique, dont les paramètres varient à la fois en longueur d'onde et en fonction des coordonnées polaires dans le champ, à symétrie de révolution.

D'après les erreurs mesurées entre les FSF simulées et les modèles associés, seuls quelques paramètres suffisent pour décrire la FSF en tout point d'un cube de données hyperspectrales. Il est alors possible, en utilisant les différentes mesures d'erreurs, de choisir la complexité du schéma de variation des paramètres du modèle de FSF, suivant l'usage que l'on souhaite en faire.

Si nous nous sommes concentrés sur certains modèles en particulier, tel que le modèle circulaire constant dans le champ, avec  $\alpha$  linéaire en longueur d'onde et  $\beta$  fixé, nous avons également envisagé d'opter pour d'autres formes de variations, telles que des fonctions puissances, exponentielles, etc. L'ensemble des modèles et des valeurs numériques rassemblés dans ce chapitre et en annexe C constituent essentiellement une base de travail pour permettre la simulation de données astrophysiques hyperspectrales, et faciliter le développement de méthodes de traitement de données, tels que la séparation de sources et la fusion.

Enfin, un modèle hyperspectral de FSF apporte davantage de robustesse, vis-à-vis de l'estimation des paramètres, par rapport à l'estimation de la FSF indépendamment à chaque longueur d'onde, ce qui fait l'objet du chapitre suivant.



# Chapitre III

## Estimation de la FSF

### Sommaire

---

<b>Notations</b> . . . . .	<b>58</b>
<b>Introduction</b> . . . . .	<b>59</b>
<b>III.1 Modèle de données</b> . . . . .	<b>59</b>
III.1.1 FSF . . . . .	59
III.1.2 FSF simulées . . . . .	60
III.1.3 Simulations de données réalistes . . . . .	60
III.1.4 Magnitude et rapport signal sur bruit . . . . .	62
III.1.5 Méthodologie . . . . .	62
<b>III.2 Estimation par maximum de vraisemblance</b> . . . . .	<b>64</b>
III.2.1 Estimation des hyper-paramètres . . . . .	64
III.2.2 Estimations des paramètres de nuisance . . . . .	65
III.2.3 Résultats d'expérimentations . . . . .	67
<b>III.3 Amélioration de l'estimation des paramètres de nuisance</b> . . . . .	<b>71</b>
III.3.1 Normalisation des données par le spectre . . . . .	72
III.3.2 Moyenne non pondérée . . . . .	72
III.3.3 Moyenne pondérée par la variance du bruit . . . . .	74
III.3.4 Résultats d'expérimentations . . . . .	75
III.3.5 Choix de l'intervalle moyenné . . . . .	76
<b>III.4 Estimation jointe</b> . . . . .	<b>77</b>
<b>III.5 Estimation par espérance <i>a posteriori</i></b> . . . . .	<b>79</b>
III.5.1 Modèle bayésien . . . . .	80
III.5.2 Méthode d'échantillonnage de Metropolis-Hastings à marche aléatoire . . . . .	81
III.5.3 Estimation conjointe du spectre de l'étoile et de la FSF . . . . .	82
III.5.4 Résultats d'expérimentations . . . . .	83
<b>III.6 Synthèse des méthodes d'estimation du spectre</b> . . . . .	<b>83</b>
<b>III.7 Estimation à partir d'une étoile non isolée</b> . . . . .	<b>84</b>
III.7.1 Données . . . . .	84
III.7.2 FSF de référence . . . . .	84
III.7.3 Masques . . . . .	88
III.7.4 Résultats d'expérimentations . . . . .	89
<b>Conclusions</b> . . . . .	<b>91</b>

---

## Notations [Unités]

- $c$  spectre du ciel convolué par la LSF [photons ou erg/s/cm<sup>2</sup>/pixel spatial]
- $D$  données bruitées après soustraction du spectre du ciel
- $\mathbf{D}$  ensemble des données bruitées, sous forme vectorielle
- $d_c$  données bruitées contenant l'étoile et le fond du ciel
- $F_\lambda$  FSF d'une source située à la longueur d'onde  $\lambda$
- $L_{x,y,\lambda}$  LSF d'une source située en  $(x, y, \lambda)$
- $m$  modèle paramétrique de Moffat pour la FSF, avec  $\alpha$  et  $\beta$  exprimés en fonction de la longueur d'onde et d'hyper-paramètres
- $N_x, N_y, N_\lambda$  taille du cube hyperspectral dans les dimensions  $x$ ,  $y$ , et  $\lambda$  [pixels]
- $o$  modèle d'observation idéale non bruitée d'une étoile et du ciel
- $O$  modèle d'observation non bruitée d'une étoile et du ciel, convolués par la PSF
- $s$  spectre d'une étoile convolué par la LSF [photons ou erg/s/cm<sup>2</sup>/pixel spatial]
- $x, y$  variables spatiales (généralement discrètes) [pixel spatial]
- \* convolution spectrale
- ★ convolution spatiale
- composition spatiale / spectrale (opérateur non convolutif)
- $\hat{\theta}$  estimateur de la variable  $\theta$
- $\alpha$  paramètre d'échelle de la fonction de Moffat [pixel spatial]
- $\alpha_0$  ordonnée à l'origine du modèle de variation linéaire de  $\alpha$  [pixel spatial]
- $\alpha_1$  pente du modèle de variation linéaire de  $\alpha$  [pixel spatial / nm]
- $\beta$  paramètre de forme de la fonction de Moffat [sans unité]
- $\beta_0$  valeur fixe de  $\beta$  [sans unité]
- $\mathbf{\Gamma}$  matrice diagonale contenant la variance du bruit
- $\boldsymbol{\theta}$  vecteur d'hyper-paramètres à estimer
- $\lambda$  longueur d'onde [nm]
- $\Lambda$  nombre de longueurs d'onde consécutives utilisées pour le moyennage spectral
- $\sigma(x, y, \lambda)$  écart-type du bruit gaussien additif des données

## Introduction

Nous venons de voir dans le Chapitre II que la fonction d'étalement spatial (FSF pour *Field Spread Function*) du télescope et de l'atmosphère peut être approximée en tout point du cube de données par une fonction de Moffat dépendant d'un faible nombre de paramètres. Cette étude se place dans le mode de fonctionnement grand champ (*Wide Field Mode* : WFM) du télescope sans correction par optique adaptative, pour lequel la FSF est constante dans le champ. L'objectif de ce chapitre réside dans la conception et la validation d'une méthodologie d'estimation de ces paramètres à partir de données réelles afin d'obtenir une approximation de la FSF en tout point de ce cube de données. Pour cela, nous utilisons le fait que les étoiles peuvent être raisonnablement considérées comme des sources ponctuelles spatiales [MAG 98]. L'observation idéale d'une étoile isolée correspond donc à la FSF pondérée par le spectre de l'étoile à chaque longueur d'onde.

Dans une première partie, nous modélisons le problème de l'observation d'une étoile isolée et présentons la méthodologie que nous allons adopter ainsi que les données simulées. Puis nous proposons deux approches d'estimation de la FSF, la première par maximum de vraisemblance, la seconde par espérance *a posteriori* par simulations de Monte-Carlo. Nous nous confrontons également au problème de l'estimation des paramètres de nuisance, dont nous tentons de réduire la variance. Enfin, nous présentons quelques résultats d'expérimentations à la fois sur des données idéales et sur des données plus réalistes contenant plusieurs sources.

### III.1 Modèle de données

Pour commencer, nous rappelons le modèle de FSF ainsi que le modèle de variation de ses paramètres. Puis nous présentons les données que nous avons simulées ainsi que la méthodologie que nous adoptons pour vérifier les performances de nos algorithmes.

#### III.1.1 FSF

D'après les résultats du Chapitre II, la FSF est approximée par la fonction de Moffat (MF) [MOF 69], de la forme suivante :

$$M_{\alpha,\beta}(r) = \left[ 1 + \left( \frac{r}{\alpha} \right)^2 \right]^{-\beta}$$

dans laquelle  $\alpha$  est un paramètre d'échelle et  $\beta$  un paramètre de forme, tous deux fonction des conditions d'observations. Ceux-ci sont soumis à certaines contraintes :  $\alpha \in ]0; +\infty[$  et  $\beta \in ]1; +\infty[$ . Avec cette expression, la fonction est maximale en  $r = 0$  et vaut 1 en ce point. Elle est donc normalisée au maximum. Afin que ce modèle corresponde à une fonction d'étalement, cette fonction sera normalisée par son flux numérique :

$$\hat{\Phi} = \sum_x \sum_y M_{\alpha,\beta}(x, y).$$

Enfin, la largeur à mi-hauteur (ou FWHM : *Full Width at Half-Maximum*) s'exprime en fonction des paramètres selon la formule :

$$\text{FWHM} = 2\alpha \sqrt{2^{\frac{1}{\beta}} - 1}.$$

Dans le cas du mode WFM sans correction par optique adaptative, la FSF est constante dans le champ de vue et varie lentement en longueur d'onde. Cette variation se traduit par celle des paramètres  $\alpha$  et  $\beta$  du modèle de Moffat. Notons que nous reprenons ici le modèle de variation des paramètres proposé et validé au paragraphe II.3.3.2 mais que la méthodologie d'estimation proposée par la suite peut s'adapter à tout modèle de variation de  $\alpha$  et  $\beta$ . Dès lors que ceux-ci sont exprimés en fonction de la

longueur d'onde et dépendent d'un nombre limité d'hyper-paramètres, alors l'estimation de la FSF se résume à celle des hyper-paramètres :

- le paramètre d'échelle  $\alpha$  qui influe sur l'étalement de la fonction dans le champ est modélisé par une fonction linéaire de la longueur d'onde :

$$\alpha(\lambda) = \alpha_0 + \alpha_1 \cdot \lambda$$

où  $\alpha_0$  et  $\alpha_1$  sont deux hyper-paramètres à estimer à partir du cube de données ;

- le paramètre de forme  $\beta$  qui influe également sur l'étalement de la fonction dans le champ est considéré constant en 1<sup>ère</sup> approximation :

$$\beta = \beta_0,$$

dans ce cas, la forme de la FSF est la même pour toutes les longueurs d'onde, seule sa taille varie.

L'estimation de la FSF se fait donc par le biais de l'estimation des trois hyper-paramètres  $\alpha_0, \alpha_1, \beta_0$ .

### III.1.2 FSF simulées

Il est crucial de bien différencier le modèle de FSF décrit ci-dessus, utilisé pour l'estimation, de la FSF utilisée pour générer les simulations. Pour valider la méthode, nous utiliserons plusieurs jeux de données simulées.

Dans un premier temps, nous travaillons avec un cube idéal, dont la FSF est générée en utilisant le modèle précédent. Ainsi, la méthode développée doit permettre d'estimer exactement la FSF à partir de données non bruitées, puisqu'il n'y a pas d'erreur de modèle.

Dans un deuxième temps, nous estimons la FSF à partir de données plus réalistes (voir paragraphe III.7) fournies par le consortium MUSE. Celles-ci contiennent des FSF de type Moffat, dont le paramètre  $\alpha$  varie suivant un modèle non linéaire en fonction de la longueur d'onde. Il y a donc une erreur de modèle entre la FSF simulée et son approximation.

### III.1.3 Simulations de données réalistes

Dans un souci de simplicité, nous simulons des données de taille  $12 \times 12 \times 3577$  contenant une seule étoile caractérisée par son spectre à laquelle s'ajoute le fond du ciel. Ces données sont ensuite convoluées par la PSF du système puis bruitées lors de l'acquisition par le capteur photométrique.

#### Données non bruitées

L'observation idéale non bruitée est celle d'une étoile positionnée en  $(x_0, y_0)$  de spectre  $\tilde{s}(\lambda)$  à laquelle s'ajoute le fond du ciel  $\tilde{c}(\lambda)$ , illustrés par la Fig. III.1 :

$$o(x, y, \lambda) = \tilde{s}(\lambda) \cdot \delta(x - x_0, y - y_0) + \tilde{c}(\lambda)$$

Notons que les variables  $x, y, \lambda$  sont ici discrètes. En raison du chemin parcouru par la lumière à travers la chaîne d'acquisition (atmosphère, télescope, instrument et DRS), l'observation est déformée par la PSF du système. Cela se traduit par la composition<sup>14</sup>, notée  $\circ$ , par la LSF variable spatialement et spectralement, et par la convolution spatiale, notée  $*$ , par la FSF à chaque longueur d'onde  $\lambda$  :

$$O(x, y, \lambda) = (\tilde{s}(\lambda) \cdot \delta(x - x_0, y - y_0) + \tilde{c}(\lambda)) \circ L_{x,y,\lambda} * F_\lambda$$

<sup>14</sup>Dans ce chapitre, nous préférons utiliser le terme « composition » à « convolution » lorsque la LSF ou FSF est variable en fonction de la position de la source, car l'opérateur n'est plus une véritable convolution.

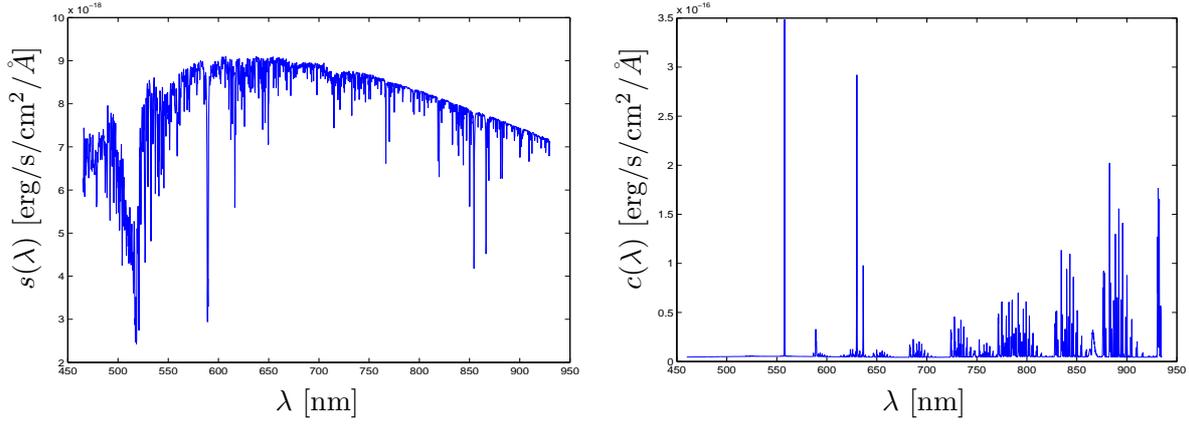


FIG. III.1 – Spectres de l'étoile et du ciel utilisés pour générer des cubes de données réalistes.

$$O(x, y, \lambda) = (\tilde{s}(\lambda) \circ L_{x_0, y_0, \lambda}) \cdot F_\lambda(x - x_0, y - y_0) + \tilde{c}(\lambda) \circ L_{x, y, \lambda}.$$

Par la suite, nous notons  $s(\lambda) = \tilde{s}(\lambda) \circ L_{x_0, y_0, \lambda}$  et  $c(\lambda) = \tilde{c}(\lambda) \circ L_{x, y, \lambda}$  les spectres après la composition spectrale par la LSF  $L$ , ce qui donne :

$$O(x, y, \lambda) = s(\lambda) \cdot F_\lambda(x - x_0, y - y_0) + c(\lambda).$$

### Données bruitées

Les données MUSE sont perturbées par le bruit de comptage des photons (de type poissonien [MCL 08]) et le bruit de lecture du capteur (de type gaussien), ce dernier étant supposé négligeable. Le cube de données acquis suit alors une loi de poisson de moyenne et de variance l'observation :

$$d_c(x, y, \lambda) \sim \mathcal{P}(O(x, y, \lambda))$$

Pour cela, nous supposons que les électrons générés par l'absorption d'un photon par un pixel du détecteur sont tous comptabilisés sur ce même pixel, sans phénomène de perte ou de débordement vers les pixels voisins [JAR 08]. Tous les pixels du cube, donc les variables associées, sont considérés indépendants.

Le temps d'intégration étant compris entre 10 minutes et plusieurs heures, ce qui fournit un flux de photons très important si la source est lumineuse, le bruit poissonien peut être approximé par un bruit gaussien de mêmes moyenne et variance [MUG 04] :

$$d_c(x, y, \lambda) \sim \mathcal{N}(O(x, y, \lambda), O(x, y, \lambda))$$

soit :

$$d_c(x, y, \lambda) = O(x, y, \lambda) + b(x, y, \lambda) \quad \text{avec } b \sim \mathcal{N}(0, O(x, y, \lambda)).$$

La première étape du traitement des données est la soustraction du fond du ciel, supposée parfaitement effectuée. En pratique, seule une estimation imparfaite du fond du ciel sera soustraite, ce qui engendrera un bruit supplémentaire non modélisable. À l'issue de cette soustraction, les données bruitées ont pour expression :

$$d(x, y, \lambda) = s(\lambda) \cdot F_\lambda(x - x_0, y - y_0) + b(x, y, \lambda) \quad \text{avec } b(x, y, \lambda) \sim \mathcal{N}(0, O(x, y, \lambda)). \quad (\text{III.1})$$

En pratique, on ne connaît pas la variance de  $b$  qui est  $\sigma^2(x, y, \lambda) = O(x, y, \lambda)$ . Néanmoins, on peut prendre comme estimateur de la variance les données bruitées  $d_c(x, y, \lambda)$  qui en sont une réalisation.

### III.1.4 Magnitude et rapport signal sur bruit

Afin de générer des données réalistes, nous utilisons le spectre d'une étoile, illustré par la Fig. III.1a, fourni par le consortium MUSE. La luminosité de cette étoile est caractérisée par sa magnitude.

La magnitude est une mesure relative logarithmique de la luminosité d'une étoile. Elle est généralement associée à une bande spectrale qui correspond au filtre à travers lequel l'objet est observé. Il existe alors autant de systèmes de magnitudes que d'ensembles de filtres. La magnitude AB, utilisée ici, a pour expression :

$$m_\nu = -2.5 \log_{10} e_\nu - 56.1$$

avec  $\nu = \frac{c}{\lambda}$  la fréquence d'onde et  $e_\nu$  l'éclairement monochromatique à cette fréquence, exprimé en  $W.m^{-2}.nm^{-1}$  [LEN 08].

Le spectre d'étoile utilisé correspond à une magnitude de 21.5 dans le vert notée  $a_0$ . Afin de générer plusieurs jeux de données contenant des étoiles plus ou moins lumineuses, nous appliquons un facteur d'amplification / de réduction à ce spectre, ce qui se traduit par un changement de sa magnitude  $a$  :

$$s_a(\lambda) = 10^{-0.4(a-a_0)} s_{a_0}(\lambda)$$

Notons que la magnitude augmente lorsque l'étoile perd en luminosité. Si la magnitude  $a_0$  concerne à l'origine la bande spectrale verte, nous appliquons cette opération sans distinction sur tout l'intervalle spectral. Par la suite, nous ne faisons donc plus référence à des bandes spectrales. L'évolution logarithmique de l'intensité des étoiles que nous avons simulées est illustrée dans la Fig. III.2.

La magnitude peut être vue, dans notre cas, comme un indicateur du rapport signal sur bruit (RSB). En effet, plus la magnitude est faible, plus l'étoile est brillante et plus le RSB est élevé.

Pour quantifier le bruit contenu dans nos données, nous définissons un RSB global sur l'intégralité du cube hyperspectral, défini comme le rapport entre l'énergie des données et la puissance moyenne du bruit :

$$\text{RSB} = \frac{\sum_{x,y,\lambda} d^2(x,y,\lambda)}{\sum_{x,y,\lambda} \sigma^2(x,y,\lambda)} \quad \text{et} \quad \text{RSB [dB]} = 10 \log_{10}(\text{RSB}).$$

Les RSB pour les simulations utilisées sont indiqués dans la Fig. III.2 et vont de RSB élevés (20 dB) à très faibles (-50dB).

### III.1.5 Méthodologie

Nous simulons des données réalistes à la résolution MUSE, conformément au modèle établi au paragraphe III.1.3, avec pour FSF une fonction de Moffat dont les hyper-paramètres sont précisés dans le Tab. III.1. L'ensemble formé par la fonction de Moffat et l'expression de ses paramètres est appelé par la suite Modèle de Moffat et noté  $m$ .

L'estimation de la FSF s'effectue par l'estimation des hyper-paramètres  $\alpha_0$ ,  $\alpha_1$  et  $\beta_0$  ainsi que la position sub-pixellaire de l'étoile, que nous rassemblons dans le vecteur  $\theta = [\alpha_0, \alpha_1, \beta_0, x_0, y_0]$ .

La comparaison directe des hyper-paramètres estimés avec ceux du modèle de Moffat simulé permet de quantifier la qualité de la FSF obtenue, mais n'est pas très parlante physiquement. En conséquence, nous comparerons également la FSF calculée à partir de  $\hat{\theta}$ , c'est-à-dire  $m(x,y,\lambda,\hat{\theta})$  avec celle de référence  $F_\lambda$ . Cette comparaison sera effectuée à toutes les longueurs d'onde de MUSE, à une résolution spatiale 40 fois plus fine que celle de MUSE, en utilisant les paramètres haute résolution du Tab. III.1. La différence entre les deux modèles de FSF à chaque longueur d'onde sera quantifiée par des indices d'erreur normalisés, de type  $L_1$  (erreur en module),  $L_2$  (erreur quadratique),  $L_\infty$  (erreur maximale), notés respectivement  $\varepsilon_1(\lambda)$ ,  $\varepsilon_2(\lambda)$ ,  $\varepsilon_\infty(\lambda)$  comme défini dans le Tab. III.2. Les performances d'estima-

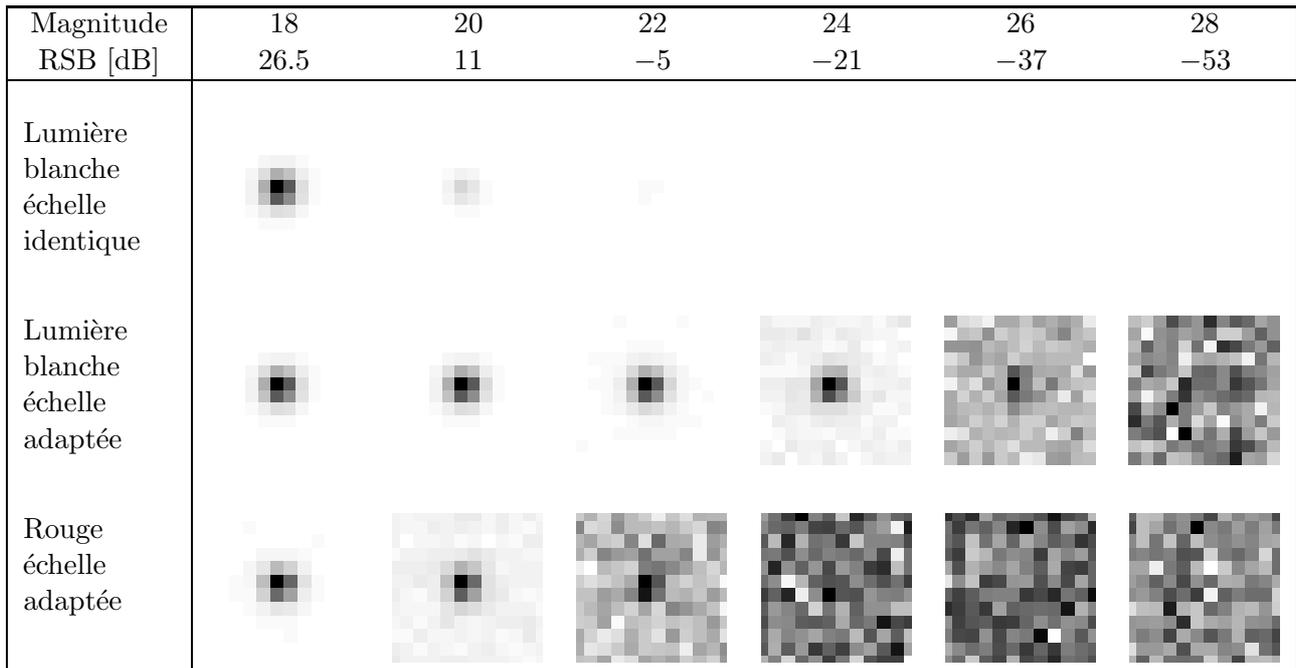


FIG. III.2 – Extraits de cubes de données bruitées de taille  $12 \times 12$  pour une étoile isolée. Affichage en lumière blanche (somme de toutes les longueurs d'onde) et dans le rouge (dernière longueur d'onde), en fonction de la magnitude de l'étoile. Rapport signal sur bruit (RSB) correspondant, exprimé en décibels (dB), pour un temps d'intégration de 20 minutes.

tion de la FSF sur l'intégralité du cube sont évaluées à partir de la moyenne  $\bar{\varepsilon}_i$  et du maximum  $\varepsilon_{i,max}$  de ces critères, par rapport à la longueur d'onde.

Dans nos simulations, l'étoile est positionnée en trois points différents, à savoir au centre du pixel, sur un quart de pixel et à la limite entre deux pixels. Dans les résultats d'estimation de ce chapitre, si la position de l'étoile n'est pas indiquée, le résultat correspond à la moyenne des résultats obtenus pour chaque position.

$\theta$	Résolution MUSE 1 pxl MUSE = 0.20"		Haute résolution (/40) 1 pxl HR = 0.005" arcsec	
	valeur	unité	valeur	unité
$\alpha_0$	2.50	pxl MUSE	100	pxl HR
$\alpha_1$	$-10^{-3}$	pxl MUSE / nm	-0.04	pxl HR / nm
$\beta_0$	3.00	sans unité	3.00	sans unité
$(x_0, y_0)$	(6.00; 6.00)	pxl MUSE	(240; 240)	pxl HR
	(6.25; 6.25)		(250; 250)	
	(6.50; 6.50)		(260; 260)	

TAB. III.1 – Valeurs et unités des hyper-paramètres de référence utilisés pour simuler les FSF contenues dans les données pseudo-réelles, à la résolution MUSE et à haute résolution (HR).

$$\begin{aligned}\varepsilon_1(\lambda) &= \frac{\sum_{xy} |F_\lambda(x, y) - m(x, y, \lambda)|}{\sum_{xy} |F_\lambda(x, y)|} \\ \varepsilon_2(\lambda) &= \frac{\sqrt{\sum_{xy} (F_\lambda(x, y) - m(x, y, \lambda))^2}}{\sqrt{\sum_{xy} F_\lambda^2(x, y)}} \\ \varepsilon_\infty(\lambda) &= \frac{\max_{xy} (|F_\lambda(x, y) - m(x, y, \lambda)|)}{\max_{xy} (F_\lambda(x, y))}\end{aligned}$$

TAB. III.2 – Définition des indices d'erreur utilisés, pour quantifier l'erreur entre la vraie FSF recherchée  $F_\lambda$  et le modèle de FSF  $m$ .

## III.2 Estimation par maximum de vraisemblance

L'objectif de cette étude est de développer et valider une méthode d'estimation de la FSF en tout point d'un cube de données réelles.

Cette estimation s'effectue via celle d'hyper-paramètres en comparant le cube de données  $D$  à un cube modèle  $d$  contenant le spectre estimé de l'étoile convolué par le modèle de FSF en chaque longueur d'onde.

### III.2.1 Estimation des hyper-paramètres

L'estimation de la FSF  $F$  nécessite celle des trois hyper-paramètres  $(\alpha_0, \alpha_1, \beta_0)$  et des coordonnées du centre de l'étoile  $(x_0, y_0)$ , rassemblés dans le vecteur  $\boldsymbol{\theta} = [\alpha_0, \alpha_1, \beta_0, x_0, y_0]$ . Vu le nombre de données disponibles dans le cube hyperspectral, et la connaissance de la statistique du bruit, nous pouvons espérer obtenir de bons résultats d'estimation en utilisant l'estimateur du maximum de vraisemblance.

#### III.2.1.1 Vraisemblance poissonnienne

D'un point de vue théorique comme à la section III.1, on devrait considérer pour modèle de données celui du cube bruité poissonnien de moyenne et de variance  $O(x, y, \lambda) : d_c(x, y, \lambda) \sim \mathcal{P}(O(x, y, \lambda))$ . Chaque pixel de ce cube est une variable aléatoire poissonnienne de probabilité :

$$p(d_c(x, y, \lambda) = k) = \frac{O(x, y, \lambda)^k}{k!} e^{-O(x, y, \lambda)}.$$

La vraisemblance du modèle de paramètres  $\boldsymbol{\theta}$  par rapport aux données  $d_c$  s'écrit :

$$L(\boldsymbol{\theta}, \mathbf{D}_c) = p(\mathbf{D}_c | \boldsymbol{\theta}) = \prod_{x, y, \lambda} \frac{O(x, y, \lambda, \boldsymbol{\theta})^{d_c(x, y, \lambda)}}{d_c(x, y, \lambda)!} e^{-O(x, y, \lambda, \boldsymbol{\theta})}$$

avec  $\mathbf{D}_c$  le vecteur rassemblant les données bruitées  $d_c(x, y, \lambda)$ ,

$$O(x, y, \lambda, \boldsymbol{\theta}) = s(\lambda) \cdot m(x, y, \lambda, \boldsymbol{\theta}) + c(\lambda)$$

$$\text{et } m(x, y, \lambda, \boldsymbol{\theta}) = \frac{1}{\widehat{\Phi}_m(\lambda)} \left[ 1 + \frac{(x - x_0)^2 + (y - y_0)^2}{(\alpha_0 + \alpha_1 \cdot \lambda)^2} \right]^{-\beta_0}$$

où  $\widehat{\Phi}_m$  est le flux du modèle de FSF sur un support restreint de taille  $N$ , défini au paragraphe III.1.1. L'anti-log-vraisemblance est définie à un facteur additif près par :

$$NLL(\boldsymbol{\theta}, \mathbf{D}_c) = \sum_{x, y, \lambda} O(x, y, \lambda, \boldsymbol{\theta}) - d_c(x, y, \lambda) \ln(O(x, y, \lambda, \boldsymbol{\theta})).$$

L'estimateur du maximum de vraisemblance  $\hat{\theta}$  est donc la valeur de  $\theta$  qui minimise l'anti-log-vraisemblance du modèle de paramètres  $\theta$  par rapport aux données  $d_c$  :  $\hat{\theta} = \arg \min_{\theta} NLL(\theta, \mathbf{D}_c)$ . Cette estimation nécessite de connaître le spectre de l'étoile  $s(\lambda)$  et celui du ciel  $c(\lambda)$ .

En pratique, les données pré-traitées par le DRS seront fournies après soustraction du ciel. En outre, on ne disposera pas du spectre du ciel en chaque pixel du cube de données. Aussi, il sera plus simple de travailler à partir du modèle de bruit additif gaussien centré, assez réaliste après soustraction du ciel pour les observations considérées.

### III.2.1.2 Vraisemblance gaussienne

Le bruit peut donc être supposé additif gaussien en raison du grand nombre de photons reçus :  $d(x, y, \lambda) = s(\lambda) \cdot m(x, y, \lambda, \theta) + b(x, y, \lambda)$  avec  $b(x, y, \lambda) \sim \mathcal{N}(0, \sigma^2(x, y, \lambda))$ . La vraisemblance du modèle de paramètres  $\theta$  par rapport aux données a pour expression :

$$\begin{aligned} L(\theta, \mathbf{D}) = p(\mathbf{D}|\theta) &= \frac{1}{\sqrt{2\pi}^{N_x N_y N_\lambda}} \frac{1}{\sqrt{\det \Gamma}} \prod_{x,y,\lambda} \exp\left(-\frac{1}{2} \frac{(D(x, y, \lambda) - s(\lambda) \cdot m(x, y, \lambda, \theta))^2}{\sigma_{x,y,\lambda}^2}\right) \\ &= \frac{1}{\sqrt{2\pi}^{N_x N_y N_\lambda} \sqrt{\det \Gamma}} \exp\left(-\frac{1}{2} \|\mathbf{D} - s(\lambda) \cdot m(x, y, \lambda, \theta)\|_{\Gamma^{-1}}^2\right) \end{aligned} \quad (\text{III.2})$$

avec  $\mathbf{D}$  le cube de données bruitées  $d(x, y, \lambda)$  vectorisé,  $N_x, N_y, N_\lambda$  les tailles du cube suivant  $x, y, \lambda$  et  $\Gamma = \text{diag}(\sigma^2(x, y, \lambda))$ . L'anti-log-vraisemblance du modèle de paramètres  $\theta$  par rapport aux données  $\mathbf{D}$  est définie à une constante additive près par :

$$NLL(\theta, \mathbf{D}) = \sum_{xy\lambda} \frac{(D(x, y, \lambda) - s(\lambda) \cdot m(x, y, \lambda, \theta))^2}{\sigma_{x,y,\lambda}^2} \quad (\text{III.3})$$

L'estimateur  $\hat{\theta}$  minimise donc un critère des moindres carrés pondéré par la variance des données :  $\hat{\theta} = \arg \min_{\theta} NLL(\theta, \mathbf{D})$ .

Contrairement à l'estimation par vraisemblance poissonnienne, cette estimation peut s'effectuer sur les données après soustraction du fond du ciel et nécessite de connaître, en plus du spectre du ciel et de l'étoile, la variance du bruit. Ces deux dernières variables sont des paramètres de nuisance qu'il est nécessaire d'estimer avant l'estimation même de la FSF.

## III.2.2 Estimations des paramètres de nuisance

Les paramètres de nuisance que nous devons estimer afin de recourir à l'estimateur du maximum de vraisemblance gaussienne sont donc le spectre de l'étoile et la variance du bruit. Dans un premier temps, nous procédons en deux étapes : l'estimation des paramètres de nuisance, puis l'estimation des paramètres  $\theta$ . La qualité de l'estimation des paramètres de nuisance est cruciale pour obtenir des performances satisfaisantes d'estimation de la FSF.

### III.2.2.1 Spectre de l'étoile

Le spectre de l'étoile  $s(\lambda)$  doit être connu à toutes les longueurs d'onde du cube de données hyperspectrales. La FSF ayant un support infini, il faudrait également des données de support infini pour mesurer parfaitement, dans un cas idéal non bruité, le spectre de l'étoile.

**Cas idéal à support infini :** Le cube d'observation idéale d'une étoile isolée s'écrit comme le produit du spectre de l'étoile par la FSF<sup>15</sup> :

$$\forall(x, y, \lambda) \in \mathbb{N}^2 \times \mathbb{R}_+, \quad O(x, y, \lambda) = s(\lambda) \cdot m(x, y, \lambda).$$

Sachant que le flux d'une FSF est normalisé par définition à chaque longueur d'onde :

$$\forall \lambda \quad \Phi_m(\lambda) = \sum_{x=-\infty}^{+\infty} \sum_{y=-\infty}^{+\infty} m(x, y, \lambda) = 1,$$

alors le spectre de l'étoile peut être mesuré par le flux de l'observation idéale :

$$s(\lambda) = \sum_{x=-\infty}^{+\infty} \sum_{y=-\infty}^{+\infty} O(x, y, \lambda) = s(\lambda) \sum_{x=-\infty}^{+\infty} \sum_{y=-\infty}^{+\infty} m(x, y, \lambda).$$

**Cas idéal à support fini :** En pratique, la taille du champ observé par l'instrument est évidemment finie, et le cube contenant l'étoile est un sous-cube de taille  $N_x \times N_y$  extrait des données d'observation. De la même manière que dans le cas d'un cube à support infini, nous pouvons mesurer le spectre de l'étoile par le flux spatial des données :

$$\forall(x, y, \lambda) \in \llbracket 1; N_x \rrbracket \times \llbracket 1; N_y \rrbracket \times \mathbb{R}_+ \quad \hat{s}(\lambda) = \hat{\Phi}(\lambda) = s(\lambda) \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} m(x, y, \lambda).$$

Le spectre est alors multiplié par le flux de la FSF sur l'intervalle spectral du cube, qui tend vers 1 lorsque la taille du cube tend vers l'infini. En choisissant un cube assez étendu par rapport à la FSF, nous pouvons obtenir une mesure fiable du spectre de l'étoile.

**Cas bruité à support fini :** L'observation est en réalité détériorée par un bruit que nous considérons centré additif gaussien (voir équation III.1).

L'estimateur du spectre de l'étoile par le flux des données a alors pour formulation :

$$\hat{s}(\lambda) = \hat{\Phi}(\lambda) = s(\lambda) \cdot \hat{\Phi}_m(\lambda) + \sum_{x=1}^N \sum_{y=1}^N b(x, y, \lambda).$$

L'espérance de cet estimateur est  $E\{\hat{s}(\lambda)\} = s(\lambda) \cdot \hat{\Phi}_m(\lambda)$ ; l'estimation n'est donc biaisée que par la restriction spatiale des données sur un support fini. Plus la taille spatiale du cube sera grande, moins l'estimation du spectre de l'étoile étudiée sera biaisée. Néanmoins, l'étoile doit pouvoir être considérée isolée, ce qui restreint la taille du cube.

La variance de l'estimateur est due à celle du bruit :

$$Var\{\hat{s}(\lambda)\} = \sum_{x=1}^N \sum_{y=1}^N \sigma^2(x, y, \lambda),$$

les bruits des différents pixels étant des variables indépendantes. Cette variance peut donc être très importante, ce qui pourrait être un handicap pour l'estimation de la FSF.

<sup>15</sup>Nous notons ici la longueur d'onde comme une variable continue, elle est en réalité échantillonnée à la résolution MUSE de 0.13 nm. Néanmoins, cela n'a pas d'influence sur les formules de ce paragraphe.

Par la suite, nous normalisons toutes les FSF simulées à basse ou haute résolution par leur flux numérique. Cela est réaliste étant donnée la taille du cube par rapport à celle de la FSF, typiquement des cubes de taille  $12 \times 12$  pour un *seeing* de 0.5. Cela permet de s'affranchir du biais d'estimation du spectre, car celle-ci n'est pas l'objet premier de cette étude.

### III.2.2.2 Variance du bruit

Nous pouvons estimer la variance du bruit  $\sigma^2(x, y, \lambda) = O(x, y, \lambda)$ , nécessaire à l'estimation par vraisemblance gaussienne, par les données bruitées avant soustraction du spectre du ciel  $d_c(x, y, \lambda)$ , qui suivent une loi de Poisson de moyenne et de variance  $O(x, y, \lambda)$ .

Cet estimateur d'espérance  $E \left\{ \widehat{\sigma^2}(x, y, \lambda) \right\} = E \{ d_c(x, y, \lambda) \} = O(x, y, \lambda)$ , est non biaisé, mais sa variance  $\text{Var} \left\{ \widehat{\sigma^2}(x, y, \lambda) \right\} = \text{Var} \{ d_c(x, y, \lambda) \} = O(x, y, \lambda)$  peut être très grande, en particulier aux longueurs d'onde où le spectre de l'étoile ou du ciel est important.

## III.2.3 Résultats d'expérimentations

### III.2.3.1 Algorithme de minimisation

Afin de minimiser l'anti-log-vraisemblance, nous utilisons l'algorithme de Nelder-Mead [LAG 98] mis en œuvre dans la fonction *fminsearch* de Matlab. Avant cela, nous devons initialiser le vecteur d'hyper-paramètres  $\theta$ . À cette fin, nous sélectionnons les images des première et dernière longueurs d'onde du cube de données bruitées, normalisées par le spectre de l'étoile estimé. Sur chacune de ces deux images, nous mesurons la position du pixel d'amplitude maximale  $m_{\max}$ , que nous moyennons pour obtenir l'initialisation des variables  $x_0$  et  $y_0$  à  $(x_0^i, y_0^i)$ . Le paramètre de forme  $\beta_0$  étant initialisé à 3 (valeur moyenne au vu de la modélisation effectuée au Chapitre II), nous en déduisons la valeur de  $\alpha = \sqrt{\frac{\beta_0 - 1}{\pi m_{\max}}}$  aux deux longueurs d'onde. Ceci permet d'initialiser la pente  $\alpha_1^i = \frac{\alpha^{(2)} - \alpha^{(1)}}{\lambda^{(2)} - \lambda^{(1)}}$  et l'ordonnée à l'origine  $\alpha_0^i = \frac{1}{2} (\alpha^{(1)} + \alpha^{(2)} - \alpha_1^i \cdot (\lambda^{(1)} + \lambda^{(2)}))$ . Il s'agit d'une initialisation assez simpliste qui pourrait être améliorée.

### III.2.3.2 Tests sur un cube non bruité

Un premier objectif est de vérifier le fonctionnement de l'estimation des hyper-paramètres par maximum de vraisemblance sur un cube idéal non bruité, à différentes positions sub-pixellaires de l'étoile. Ce cube correspond à la FSF sur  $12 \times 12$  pixels à 94 longueurs d'onde. Il s'agit ensuite de quantifier l'erreur au cube de référence commise en basse et haute résolution.

La variance de ces données est nulle. Pour estimer les hyper-paramètres, nous minimisons donc un critère légèrement différent, dans lequel l'erreur est normalisée par la valeur de la FSF, en tout point du cube. Notons que ce critère ne fait pas apparaître le spectre de l'étoile, car en l'absence de bruit, son estimation est exacte.

$$\hat{\theta} = \arg \min_{\theta} \sum_{xy\lambda} \frac{(F(x, y, \lambda) - m(x, y, \lambda, \theta))^2}{F(x, y, \lambda)^2}$$

D'après les résultats donnés dans le Tab. III.3, les hyper-paramètres estimés sont très proches de ceux de référence : l'erreur est inférieure à 0.2%. L'erreur maximale entre le cube de référence à la résolution de MUSE et le cube reconstitué ne dépasse pas 0.18% avec une moyenne inférieure à 0.01%. Ces erreurs sont logiquement très basses car le modèle d'estimation utilisé, à savoir une fonction de Moffat, de paramètres  $\alpha$  linéaire en longueur d'onde et  $\beta$  constant, est identique à celui du cube de référence. Les performances ne semblent pas dépendre de la position sub-pixellaire de l'étoile, et sont

comparables pour les deux résolutions utilisées. Par conséquent, nous nous limiterons, sauf indication contraire, à des indices d'erreur moyennés sur les trois positions sub-pixellaires de l'étoile.

Décalage simulé ( $x_0, y_0$ )	Hyper-paramètres estimés	Erreur aux données - Statistiques sur les $\lambda$					
		$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
(0,0)	$\left\{ \begin{array}{l} \alpha(\lambda) = 2.50 - 1.00 \cdot 10^{-3} \lambda \\ \beta(\lambda) = 3.0000 \\ \text{shift} (7.10^{-6}, -3.10^{-6}) \end{array} \right.$	échantillonnage MUSE					
		4.12 $10^{-5}$	140 $10^{-5}$	3.19 $10^{-5}$	105 $10^{-5}$	5.36 $10^{-5}$	176 $10^{-5}$
$(\frac{1}{4}, \frac{1}{4})$	$\left\{ \begin{array}{l} \alpha(\lambda) = 2.50 - 1.00 \cdot 10^{-3} \lambda \\ \beta(\lambda) = 3.0000 \\ \text{shift} (0.2500, 0.2500) \end{array} \right.$	40×échantillonnage MUSE					
		3.57 $10^{-5}$	118 $10^{-5}$	3.23 $10^{-5}$	106 $10^{-5}$	5.45 $10^{-5}$	180 $10^{-5}$
$(\frac{1}{2}, \frac{1}{2})$	$\left\{ \begin{array}{l} \alpha(\lambda) = 2.50 - 1.00 \cdot 10^{-3} \lambda \\ \beta(\lambda) = 3.0000 \\ \text{shift} (0.5000, 0.5000) \end{array} \right.$	échantillonnage MUSE					
		4.28 $10^{-5}$	117 $10^{-5}$	3.84 $10^{-5}$	107 $10^{-5}$	5.60 $10^{-5}$	152 $10^{-5}$
$(\frac{1}{2}, \frac{1}{2})$	$\left\{ \begin{array}{l} \alpha(\lambda) = 2.50 - 1.00 \cdot 10^{-3} \lambda \\ \beta(\lambda) = 3.0000 \\ \text{shift} (0.5000, 0.5000) \end{array} \right.$	40×échantillonnage MUSE					
		4.24 $10^{-5}$	117 $10^{-5}$	3.86 $10^{-5}$	106 $10^{-5}$	6.40 $10^{-5}$	179 $10^{-5}$
$(\frac{1}{2}, \frac{1}{2})$	$\left\{ \begin{array}{l} \alpha(\lambda) = 2.50 - 1.00 \cdot 10^{-3} \lambda \\ \beta(\lambda) = 3.0000 \\ \text{shift} (0.5000, 0.5000) \end{array} \right.$	échantillonnage MUSE					
		3.13 $10^{-5}$	93 $10^{-5}$	3.47 $10^{-5}$	110 $10^{-5}$	3.39 $10^{-5}$	93 $10^{-5}$
$(\frac{1}{2}, \frac{1}{2})$	$\left\{ \begin{array}{l} \alpha(\lambda) = 2.50 - 1.00 \cdot 10^{-3} \lambda \\ \beta(\lambda) = 3.0000 \\ \text{shift} (0.5000, 0.5000) \end{array} \right.$	40×échantillonnage MUSE					
		3.73 $10^{-5}$	118 $10^{-5}$	3.37 $10^{-5}$	107 $10^{-5}$	5.70 $10^{-5}$	180 $10^{-5}$

TAB. III.3 – Performances d'estimation du modèle de référence sans bruit

### III.2.3.3 Estimation indépendante à chaque longueur d'onde

La solution la plus simple et la plus intuitive pour estimer la FSF serait d'estimer les paramètres  $\alpha$  et  $\beta$  du modèle, indépendamment à chaque longueur d'onde. Afin d'évaluer la validité de cette méthode, nous avons estimé ces paramètres par maximum de vraisemblance<sup>16</sup> sur un cube de données faiblement bruité.

La Fig. III.3 (a) et (b) compare les valeurs des paramètres  $\alpha$  et  $\beta$  estimées aux valeurs théoriques. Nous pouvons constater que les valeurs estimées peuvent s'avérer très éloignées des valeurs théoriques pour certaines longueur d'onde et que de façon générale elles sont très variables. En effet, l'optimisation réduit, à chaque longueur, l'écart aux données qui peuvent être très bruitées aux longueurs d'onde où le spectre de l'étoile est très faible.

La Fig. III.3 (d) montre les variations de l'indice d'erreur  $\varepsilon_2$  calculé à chaque longueur d'onde. Sa moyenne de 3.22% sur toutes les longueurs d'onde montre que cette méthode d'estimation est globalement satisfaisante pour une étoile relativement brillante, de RSB = 18.8 dB. Néanmoins, cette qualité d'estimation varie beaucoup suivant la longueur d'onde. La Fig. III.4 donne quelques exemples de FSF simulées et de modèles estimés, à différentes longueurs d'onde, ce qui permet de visualiser ce que représentent les indices d'erreur  $\varepsilon_2$ . On observe qu'en dessous d'une erreur de 5% les deux courbes ne peuvent être visuellement différenciées.

Nous avons voulu comparer l'évolution de  $\bar{\varepsilon}_2$  avec celle du rapport signal sur bruit en fonction de la longueur d'onde. Celui-ci dépend du spectre de l'étoile et du fond du ciel. En effet, lorsque l'intensité du spectre de l'étoile est faible, le signal et donc le RSB le sont aussi ; à l'inverse, lorsque le spectre du fond du ciel est élevé, la variance du bruit l'est aussi donc le RSB est faible. Par conséquent, le courbe RSB( $\lambda$ ) fait apparaître à la fois les raies d'absorption de l'étoile et les raies d'émission du fond du ciel. En sélectionnant la longueur d'onde des raies les plus faibles de cette courbe de la Fig. III.3 (c), nous pouvons constater que nous retrouvons sur la figure (d) les pics où l'erreur d'estimation est particulièrement élevée, pouvant aller jusqu'à 27%. Du point de vue de l'interprétation des résultats,

<sup>16</sup>L'estimateur est donné par  $\widehat{\theta}_\lambda = \arg \min_{\theta_\lambda} NLL(\theta_\lambda, D_\lambda)$  en définissant le vecteur  $\theta_\lambda = (\alpha_\lambda, \beta_\lambda)$  et l'anti-log-vraisemblance par  $NLL(\theta_\lambda, D_\lambda) = \sum_{x,y} \frac{(D_\lambda(x,y) - s(\lambda) \cdot m_\lambda(x,y, \theta_\lambda))^2}{\sigma_{x,y,\lambda}^2}$ .

il n'est pas satisfaisant que la fiabilité des données soit variable à ce point en fonction de la longueur d'onde, et qu'elle dépende autant de l'étoile.

La Fig. III.5 montre en gris les performances d'estimation indépendante à chaque longueur d'onde, pour différentes valeurs du RSB, et donc pour différents niveaux de l'intensité de l'étoile. Bien sûr, les performances d'estimation se dégradent lorsque le RSB diminue. Ces courbes serviront de référence pour vérifier que les méthodes que nous proposons apportent une amélioration.

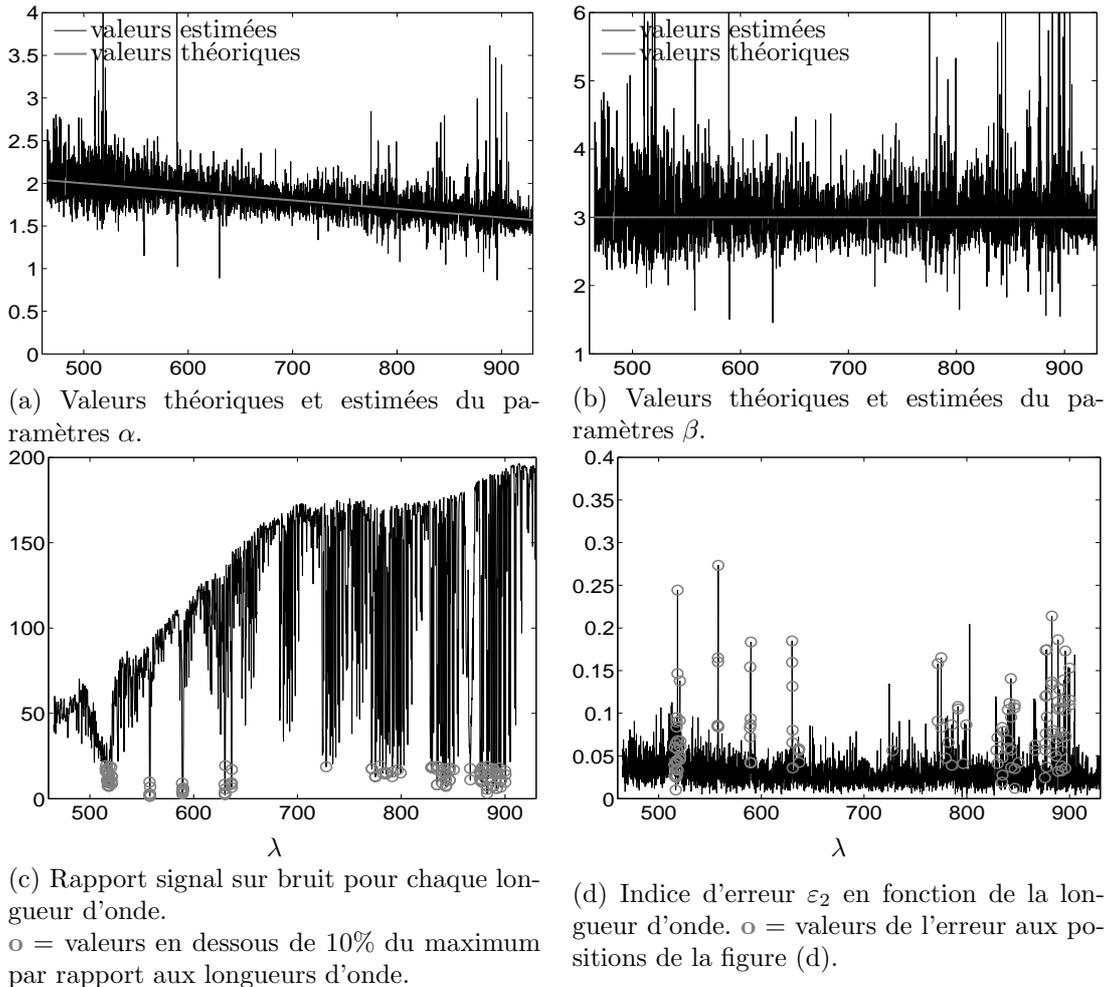


FIG. III.3 – Résultats de l'estimation de la FSF indépendamment à toutes les longueurs d'onde. Rapport signal sur bruit<sup>17</sup> pour chaque longueur d'onde, indice d'erreur  $\varepsilon_2$  obtenu (valeur moyenne de l'erreur sur l'ensemble des longueurs d'onde :  $\bar{\varepsilon}_2 = 3.22\%$ , valeur maximale :  $\varepsilon_{2,max} = 27.3\%$ ), valeurs de  $\alpha$  et  $\beta$  théoriques et estimées. Pour une étoile de magnitude 19, soit un RSB équivalent de 18.80 dB.

<sup>17</sup> $RSB(\lambda) = \frac{\sum_{x,y} o^2(x,y,\lambda)}{\sum_{x,y} \sigma^2(x,y,\lambda)}$ .

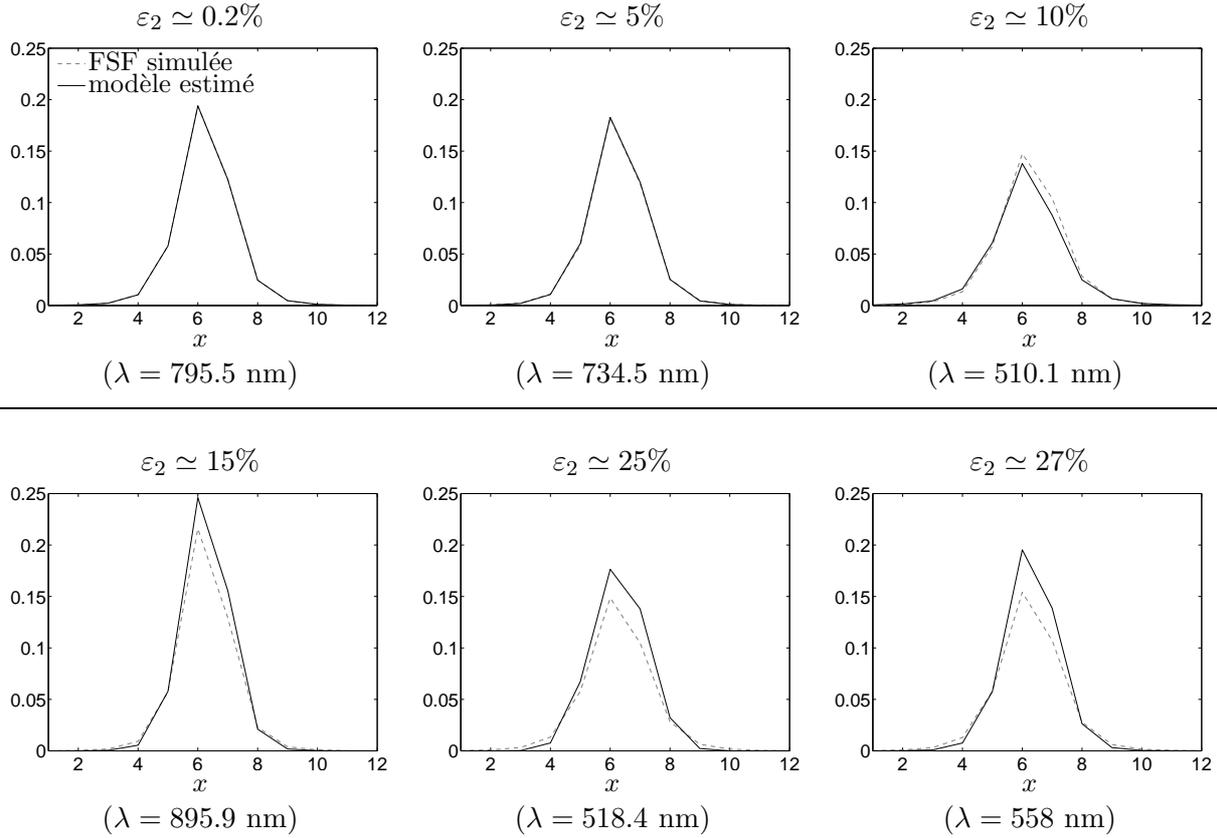


FIG. III.4 – Comparaison de FSF simulées de type Moffat, à différentes longueurs d’onde, et des modèles estimés correspondants. Coupes 1D de chaque FSF / modèle à  $y = 6$  (position approximative de l’étoile). Erreurs calculées sur les images d’erreur 2D.

### III.2.3.4 Tests sur des cubes bruités

Nous avons estimé les hyper-paramètres du modèle de FSF proposé par maximum de vraisemblance pour différents RSB. La Fig. III.5 présente l’évolution des indices d’erreur en fonction du rapport signal sur bruit (RSB) des données, les données les plus bruitées ayant le RSB le plus faible, et la compare à celle obtenue par l’estimation de la FSF indépendamment à chaque longueur d’onde.

Tout d’abord, nous pouvons remarquer que l’estimation à partir du cube semble plus performante au vu des indices d’erreur à haute résolution, mais pas à la résolution MUSE. Cette différence est due au fait qu’à la résolution MUSE, l’estimation indépendante est plus apte à coller à chaque image, car elle fournit une valeur des paramètres pour chaque longueur d’onde. Inversement, l’estimation sur le cube a pour contrainte de réduire l’erreur à l’ensemble du cube. Elle est donc moins sensible au bruit et fournit donc un modèle plus proche de la vraie FSF, ce que traduit la haute résolution. Pour cette raison, nous nous limiterons par la suite à l’analyse des performances par l’étude des indices d’erreur en haute résolution. Ces résultats nous confirment clairement qu’il est préférable d’estimer les hyper-paramètres d’un modèle de FSF pour l’ensemble du cube que d’estimer la FSF indépendamment à chaque longueur d’onde.

Dans le cas d’un cube idéal très peu bruité, les performances d’estimation de la FSF par maximum de vraisemblance, répertoriées dans la Fig. III.5 sont relativement satisfaisantes, avec une erreur quadratique moyenne  $\bar{\varepsilon}_2$  d’environ 11.3% (à la résolution MUSE) à 18.8 dB. Le fait que l’erreur maximale moyenne soit assez basse (environ 5.4% au même niveau de bruit) suggère l’absence d’écart significatif entre la vraie FSF et son modèle sur une zone ponctuelle de l’image. À l’opposé, le flux moyen de l’erreur  $\bar{\varepsilon}_1$ , légèrement plus élevé à 24.8% environ, montre que l’écart entre les 2 FSF est diffus sur

une zone étendue de l'image. Cette répartition de l'erreur est logique du fait de l'estimation de la FSF par un modèle paramétrique. Il s'agit en effet d'une erreur d'estimation de la taille ou de la forme de la FSF, et non sur une zone ponctuelle de la FSF.

On constate que les performances d'estimation de la FSF se dégradent fortement lorsque la puissance du bruit augmente, c'est-à-dire lorsque le RSB diminue, et ceci quels que soient l'indice d'erreur et la résolution.

L'estimation à partir des cubes bruités contenant une étoile de magnitude supérieure à 21, dont le RSB est inférieur à 3 dB, entraîne une erreur quadratique moyenne supérieure à 50%. Au delà de cette limite, on peut considérer que la FSF estimée n'est pas fiable. Il est donc crucial de non seulement améliorer les performances d'estimation de la FSF sur les cubes peu bruités, mais également de repousser cette limite jusqu'à laquelle l'estimation reste raisonnablement possible.

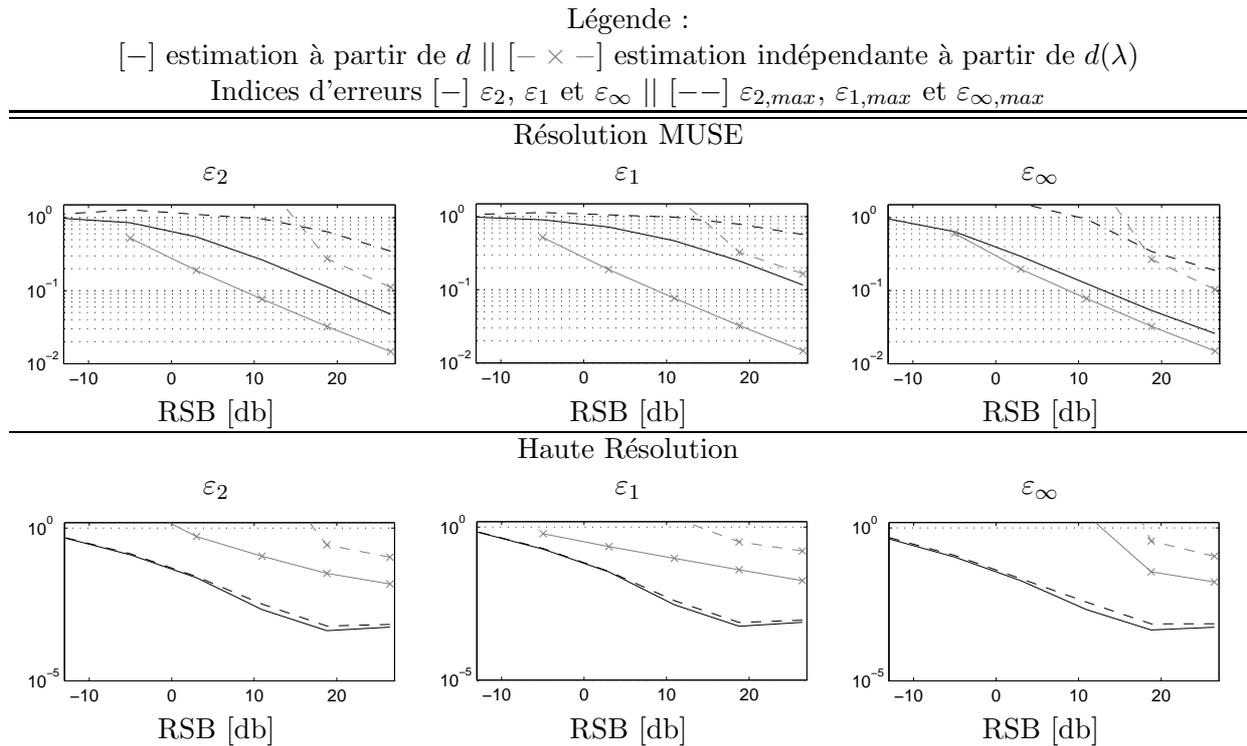


FIG. III.5 – Performances d'estimation de la FSF sur un cube idéal bruité dont la FSF est un modèle de Moffat. Indices d'erreurs  $\bar{\varepsilon}_2, \bar{\varepsilon}_1$  et  $\bar{\varepsilon}_\infty$  en [ligne continue] et  $\varepsilon_{2,max}, \varepsilon_{1,max}$  et  $\varepsilon_{\infty,max}$  en [tirets] en fonction du rapport signal sur bruit des données bruitées. Moyenne des résultats obtenus sur les 3 positions sub-pixellaires de l'étoile.

### III.3 Amélioration de l'estimation des paramètres de nuisance

Lorsque le rapport signal sur bruit est faible, la variance de l'estimation du spectre de l'étoile et de l'estimation de la variance du bruit est très grande. Or, l'estimation des hyper-paramètres de la FSF s'appuie sur ces valeurs estimées. Pour réduire cette variance d'estimation, nous proposons d'effectuer un moyennage spectral, c'est-à-dire de découper le cube de données  $d(x, y, \lambda)$  en groupes de  $\Lambda$  longueurs d'ondes consécutives autour de  $\bar{\lambda}$  (noté par la suite  $\lambda \sim \bar{\lambda}$ ), chaque groupe étant ensuite moyenné spectralement. Il existe plusieurs possibilités de combiner les données au sein de chaque groupe. Aussi, il est important d'étudier le biais et la variance des estimateurs du spectre ainsi que la variance du bruit afin de choisir la méthode adéquate.

### III.3.1 Normalisation des données par le spectre

Dans le cas idéal où l'on connaît le spectre de l'étoile et s'il est non nul à toutes les longueurs d'onde, il est possible de normaliser les données par ce spectre. Le cube de données normalisé  $d'(x, y, \lambda)$  a alors pour moyenne la FSF  $m$ , ce qui permet de s'affranchir de la pondération du spectre, toutes les images du cube ayant le même poids. En outre, sa variance est réduite d'un facteur  $s(\lambda)^2$  :

$$d'(x, y, \lambda) = m(x, y, \lambda) + b'(x, y, \lambda) \text{ avec } b'(x, y, \lambda) \sim \mathcal{N}\left(0, \sigma^{2'} = \frac{\sigma^2(x, y, \lambda)}{s(\lambda)^2}\right)$$

Néanmoins, en pratique, le spectre n'est pas connu mais seulement estimé. Chaque image du cube est alors le produit de la FSF par le rapport entre le spectre et son estimation :

$$d'(x, y, \lambda) = m(x, y, \lambda) \cdot \frac{s(\lambda)}{\widehat{s(\lambda)}} + b'(x, y, \lambda) \text{ avec } b'(x, y, \lambda) \sim \mathcal{N}\left(0, \sigma^{2'} = \frac{\sigma^2(x, y, \lambda)}{\widehat{s(\lambda)}^2}\right)$$

La variance d'estimation du spectre de l'étoile étant très grande (cf. III.2.2), nous pouvons présumer que le ratio  $\frac{s(\lambda)}{\widehat{s(\lambda)}}$  sera loin de la valeur 1 attendue, même s'il est impossible de calculer théoriquement l'espérance et la variance de ce cube normalisé. Par conséquent, nous préférons travailler par la suite avec le cube non normalisé.

### III.3.2 Moyenne non pondérée

Nous supposons que la FSF varie spectralement de façon suffisamment lente pour être considérée constante sur un court intervalle spectral. Cette hypothèse est réaliste dans le cas WFM sans OA, pour lequel nous avons validé le modèle linéaire de variation spectrale du paramètre d'échelle du modèle de Moffat.

D'après cette hypothèse, le résultat de la moyenne spectrale  $\bar{d}(x, y, \bar{\lambda})$ , par groupes du cube de données bruitées  $d$  a pour expression :

$$\bar{d}(x, y, \bar{\lambda}) = \frac{1}{\Lambda} \sum_{\lambda \sim \bar{\lambda}} d(x, y, \lambda) = \bar{s}(\bar{\lambda}) \cdot m(x, y, \bar{\lambda}) + \bar{b}(x, y, \bar{\lambda})$$

où  $\bar{s}(\bar{\lambda})$  est le spectre de l'étoile moyenné et  $\bar{b}$  le bruit moyenné.

#### III.3.2.1 Estimation de la variance du bruit moyen

Le bruit moyen  $\bar{b}$  est le résultat de la moyenne spectrale groupée du bruit  $b$  additif gaussien, de moyenne nulle et de variance  $\sigma^2(x, y, \lambda)$ . Il est donc également additif gaussien, de moyenne nulle et de variance  $\bar{\sigma}^2(x, y, \bar{\lambda})$ . Cette nouvelle variance s'exprime en fonction de la première :

$$\begin{aligned} \bar{\sigma}^2(x, y, \bar{\lambda}) &= \text{Var} \{ \bar{b}(x, y, \bar{\lambda}) \} \\ &= \text{Var} \left\{ \frac{1}{\Lambda} \sum_{\lambda \sim \bar{\lambda}} b(x, y, \lambda) \right\} \\ &= \frac{1}{\Lambda^2} \sum_{\lambda \sim \bar{\lambda}} \text{Var} \{ b(x, y, \lambda) \} \\ &= \frac{1}{\Lambda^2} \sum_{\lambda \sim \bar{\lambda}} \sigma^2(x, y, \lambda) \end{aligned}$$

La variance du bruit moyen est  $\Lambda$  fois plus petite que la variance moyenne du bruit d'origine.

La variance  $\sigma^2$  du bruit  $b$ , dont la vraie valeur est l'observation  $O(x, y, \lambda)$ , est estimée par les données bruitées avec ciel  $d_c(x, y, \lambda)$ , qui sont une réalisation de l'observation. Pour estimer la variance  $\bar{\sigma}^2$  du

bruit moyenné, il suffit de remplacer, dans son expression,  $\sigma^2$  par son estimateur :

$$\widehat{\sigma^2}(x, y, \bar{\lambda}) = \frac{1}{\Lambda^2} \sum_{\lambda \sim \bar{\lambda}} d_c(x, y, \lambda)$$

Cet estimateur est non biaisé :

$$\begin{aligned} E \left\{ \widehat{\sigma^2}(x, y, \bar{\lambda}) \right\} &= E \left\{ \frac{1}{\Lambda^2} \sum_{\lambda \sim \bar{\lambda}} d_c(x, y, \lambda) \right\} \\ &= \frac{1}{\Lambda^2} \sum_{\lambda \sim \bar{\lambda}} E \{ d_c(x, y, \lambda) \} \\ &= \frac{1}{\Lambda^2} \sum_{\lambda \sim \bar{\lambda}} \sigma^2(x, y, \lambda) \\ &= \bar{\sigma^2}(x, y, \bar{\lambda}) \end{aligned}$$

et sa variance est environ  $\Lambda^3$  fois plus petite que celle de l'estimateur de la variance du bruit d'origine :

$$\begin{aligned} \text{Var} \left\{ \widehat{\sigma^2}(x, y, \bar{\lambda}) \right\} &= \text{Var} \left\{ \frac{1}{\Lambda^2} \sum_{\lambda \sim \bar{\lambda}} d_c(x, y, \lambda) \right\} \\ &= \frac{1}{\Lambda^4} \sum_{\lambda \sim \bar{\lambda}} \text{Var} \{ d_c(x, y, \lambda) \} \\ &= \frac{1}{\Lambda^4} \sum_{\lambda \sim \bar{\lambda}} \sigma^2(x, y, \lambda) \\ &= \frac{1}{\Lambda^4} \sum_{\lambda \sim \bar{\lambda}} \text{Var} \left\{ \widehat{\sigma^2}(x, y, \bar{\lambda}) \right\}. \end{aligned}$$

### III.3.2.2 Estimation du spectre moyenné

Lorsque l'on estime les paramètres du modèle de Moffat à partir du cube moyenné, il n'est pas nécessaire de connaître le spectre de l'étoile à la résolution MUSE mais seulement le spectre moyenné  $\bar{s}(\bar{\lambda}) = \sum_{\lambda \sim \bar{\lambda}} s(\lambda)$ . Si l'on suppose que la taille du cube étudié est suffisamment grande de telle sorte

que le flux numérique  $\widehat{\Phi}_m$  de la FSF vaut 1, alors le spectre moyenné est estimé de la même façon que précédemment, en remplaçant le cube d'origine par le cube moyenné :

$$\widehat{\bar{s}}(\bar{\lambda}) = \sum_{x,y} \bar{d}(x, y, \bar{\lambda})$$

Cet estimateur est non biaisé :

$$\begin{aligned}
E \left\{ \widehat{\bar{s}(\bar{\lambda})} \right\} &= E \left\{ \sum_{x,y} \bar{d}(x, y, \bar{\lambda}) \right\} \\
&= E \left\{ \sum_{x,y} \bar{s}(\bar{\lambda}) \cdot m(x, y, \bar{\lambda}) + \bar{b}(x, y, \bar{\lambda}) \right\} \\
&= \bar{s}(\bar{\lambda}) \sum_{x,y} m(x, y, \bar{\lambda}) + \sum_{x,y} E \left\{ \bar{b}(x, y, \bar{\lambda}) \right\} \\
&= \bar{s}(\bar{\lambda})
\end{aligned}$$

et sa variance est approximativement  $\Lambda$  fois plus petite que celle de l'estimateur du spectre complet :

$$\begin{aligned}
\text{Var} \left\{ \widehat{\bar{s}(\bar{\lambda})} \right\} &= \text{Var} \left\{ \sum_{x,y} \bar{d}(x, y, \bar{\lambda}) \right\} \\
&= \sum_{x,y} \text{Var} \left\{ \bar{d}(x, y, \bar{\lambda}) \right\} \\
&= \sum_{x,y} \sigma^2(x, y, \bar{\lambda}) \\
&= \frac{1}{\Lambda^2} \sum_{x,y} \sum_{\lambda \sim \bar{\lambda}} \sigma^2(x, y, \lambda) \\
&= \frac{1}{\Lambda^2} \sum_{\lambda \sim \bar{\lambda}} \text{Var} \left\{ \widehat{s(\lambda)} \right\}.
\end{aligned}$$

### III.3.3 Moyenne pondérée par la variance du bruit

Dans le cas de bruit non identiquement distribué, il est souvent utile, lorsque l'on estime une moyenne, de pondérer les données par leur variance, afin de réduire la variance d'estimation. Dans notre cas, nous pouvons effectuer la moyenne spectrale  $\underline{d}(x, y, \bar{\lambda})$  groupée du cube en pondérant chaque pixel par la variance du bruit :

$$\begin{aligned}
\underline{d}(x, y, \bar{\lambda}) &= \left( \sum_{\lambda \sim \bar{\lambda}} \frac{1}{\sigma^2(x, y, \lambda)} \right)^{-1} \sum_{\lambda \sim \bar{\lambda}} \frac{d(x, y, \lambda)}{\sigma^2(x, y, \lambda)} \\
&= \left( \sum_{\lambda \sim \bar{\lambda}} \frac{1}{\sigma^2(x, y, \lambda)} \right)^{-1} \left[ \sum_{\lambda \sim \bar{\lambda}} \frac{s(\lambda) \cdot m(x, y, \lambda)}{\sigma^2(x, y, \lambda)} + \sum_{\lambda \sim \bar{\lambda}} \frac{b(x, y, \lambda)}{\sigma^2(x, y, \lambda)} \right] \\
&= \left( \sum_{\lambda \sim \bar{\lambda}} \frac{1}{\sigma^2(x, y, \lambda)} \right)^{-1} \left[ m(x, y, \bar{\lambda}) \sum_{\lambda \sim \bar{\lambda}} \frac{s(\lambda)}{\sigma^2(x, y, \lambda)} + \sum_{\lambda \sim \bar{\lambda}} \frac{b(x, y, \lambda)}{\sigma^2(x, y, \lambda)} \right] \\
&= \underline{s}(x, y, \bar{\lambda}) \cdot m(x, y, \bar{\lambda}) + \underline{b}(x, y, \bar{\lambda})
\end{aligned}$$

avec le « spectre » moyenné

$$\underline{s}(x, y, \bar{\lambda}) = \left( \sum_{\lambda \sim \bar{\lambda}} \frac{1}{\sigma^2(x, y, \lambda)} \right)^{-1} \sum_{\lambda \sim \bar{\lambda}} \frac{s(\lambda)}{\sigma^2(x, y, \lambda)}$$

et le bruit moyenné

$$\underline{b}(x, y, \bar{\lambda}) = \left( \sum_{\lambda \sim \bar{\lambda}} \frac{1}{\sigma^2(x, y, \lambda)} \right)^{-1} \sum_{\lambda \sim \bar{\lambda}} \frac{b(x, y, \lambda)}{\sigma^2(x, y, \lambda)}.$$

Le bruit  $\underline{b}(x, y, \bar{\lambda})$  est de moyenne nulle et de variance

$$\underline{\sigma}^2(x, y, \bar{\lambda}) = \sum_{\lambda \sim \bar{\lambda}} \frac{1}{\sigma^2(x, y, \lambda)}.$$

On peut montrer que  $\underline{\sigma}^2(x, y, \bar{\lambda}) \leq \sum_{\lambda \sim \bar{\lambda}} \sigma(x, y, \lambda)$ ; une telle moyenne permet effectivement de réduire la variance du bruit par rapport à la moyenne non pondérée de la section III.3.2. La variance du bruit peut être estimée par

$$\underline{\sigma}^2(x, y, \bar{\lambda}) = \sum_{\lambda \sim \bar{\lambda}} \frac{1}{\widehat{\sigma}^2(x, y, \lambda)}$$

dont il n'est pas aisé d'estimer le biais et la variance.

Cependant, le principal problème de ce moyennage pondéré vient du fait que l'on ne peut pas exploiter la propriété  $\sum_{x,y} m(x, y, \bar{\lambda}) = 1$  pour estimer le spectre de l'étoile (éventuellement moyenné). Nous n'avons donc pas retenu cette méthode de moyennage par la suite.

### III.3.4 Résultats d'expérimentations

L'objectif est de vérifier si le moyennage spectral non pondéré du cube apporte une amélioration de l'estimation des variables de nuisance, à savoir le spectre de l'étoile et la variance du bruit. Pour cela, nous comparons les performances d'estimation de la FSF à toutes les longueurs d'onde, à partir d'un cube idéal bruité non moyenné et à partir d'un cube moyenné, dont la FSF est un modèle de Moffat.

#### III.3.4.1 Influence du spectre et de la variance

Le Tab. III.4 témoigne de l'influence de l'estimation du spectre de l'étoile et de la variance du bruit sur les performances d'estimation de la FSF.

Si ces deux paramètres de nuisance étaient parfaitement connus, il serait possible d'atteindre des indices d'erreur  $\bar{\varepsilon}_2$  particulièrement bas, même pour des RSB négatifs, grâce à la nature hyperspectrale des données.

En estimant ces variables à partir des données, on constate que l'estimateur empirique de la variance  $\hat{\sigma}(x, y, \lambda)$  présenté au paragraphe III.3.2 donne toujours de meilleurs résultats qu'en supposant un bruit stationnaire de variance  $\sigma_0^2$ , mais ces résultats sont très insuffisants pour un RSB faible.

Le moyennage spectral des données améliore de façon significative les performances d'estimation, en particulier pour les RSB négatifs où l'indice  $\bar{\varepsilon}_2$  est divisé par un facteur 10 environ. Seul le cas d'un RSB très élevé est dégradé par le moyennage spectral. Cela est dû à une erreur de modèle car on suppose, en effectuant le moyennage spectral, que la FSF est constante pour toutes les longueurs d'onde  $\lambda \sim \bar{\lambda}$ , ce qui n'est qu'une approximation.

#### III.3.4.2 Impact du moyennage spectral

Nous comparons les indices d'erreur d'estimation de la FSF par maximum de vraisemblance (ML) (§ III.2) obtenus à partir de deux cubes de données, l'un étant un cube de données bruité original, l'autre ayant subi un moyennage spectral non pondéré sur  $\Lambda = 40$  longueurs d'onde.

La Fig. III.6 révèle que les performances obtenues avec moyennage spectral sont incontestablement meilleures que les précédentes, en particulier pour les RSB négatifs. Dans le cas particulier d'un RSB relativement élevé, supérieur à 10, les résultats des deux méthodes sont très similaires; ainsi

	Magnitude RSB [dB]	19	21	23
spectre et variance connus	$s(\lambda) \& \sigma(x, y, \lambda)$	0.04%	0.11%	0.73%
variance constante (i.i.d.)	$\widehat{s}(\lambda) \& \sigma_0$	0.11%	3.63%	70.67%
estimateurs empiriques	$\widehat{s}(\lambda) \& \widehat{\sigma}(x, y, \lambda)$	0.04%	2.35%	47.15%
moyennage spectral $\Lambda = 20$	$\widehat{s}(\bar{\lambda}) \& \widehat{\sigma}(x, y, \bar{\lambda})$	0.073%	0.38%	6.21%
moyennage spectral $\Lambda = 60$	$\widehat{s}(\bar{\lambda}) \& \widehat{\sigma}(x, y, \bar{\lambda})$	0.067%	0.39%	3.49%

TAB. III.4 – Influence de l’estimation du spectre de l’étoile et de la variance du bruit sur l’estimation de la FSF. Indices d’erreur  $\varepsilon_2(\widehat{\theta})$  calculés à Haute Résolution spatiale, pour différentes magnitudes d’étoiles / différents RSB, pour différentes valeurs du spectre et de la variance du bruit (valeurs vraies / estimées).

le moyennage spectral n’est pas utile dans ce cas, mais ne dégrade pas de manière significative les performances.

Un élément crucial est la limite en puissance du bruit ou en brillance de l’étoile au delà de laquelle l’estimation n’est plus raisonnablement fiable. En prenant comme critère le RSB en dessous duquel l’indice  $\varepsilon_2$  dépasse 20%, alors l’estimation avec moyennage spectral est possible jusqu’à  $-21$  dB environ, soit une magnitude d’étoile de 24, contre  $-5$  dB / 22 précédemment. Il s’agit d’un gain considérable en performances, qui dépend néanmoins de la valeur du paramètre de moyenne  $\Lambda$ .

### III.3.5 Choix de l’intervalle moyenné

Comme cela a déjà été évoqué par le Tab III.4, le paramètre  $\Lambda$  correspond au nombre de longueurs d’onde consécutives sur lesquelles s’effectue le moyennage spectral. Il influence les performances de l’estimation de la FSF. En effet, plus  $\Lambda$  est élevé et plus la variance d’estimation du spectre et de la variance du bruit est réduite par rapport à celle des estimateurs du cube original. Par contre, l’hypothèse selon laquelle la FSF est constante spectralement sur l’intervalle moyenné devient de moins en moins valable lorsque  $\Lambda$  augmente. Il s’agit donc de trouver un compromis entre la validité du modèle et le gain en terme d’estimation des paramètres de nuisance.

Le Tab. III.5 rassemble les indices d’erreur  $\varepsilon_2$  à haute résolution, obtenus pour l’estimation à partir d’un cube de données non moyenné (1<sup>ère</sup> ligne) ou moyenné pour différentes valeurs du paramètre  $\Lambda$ , en fonction du rapport signal sur bruit.

Dans le cas de données très faiblement bruitées ( $\text{RSB} > 10$  dB), les indices d’erreur sont faibles et à peu près équivalents pour toutes les valeurs du paramètre  $\Lambda$ . Les meilleurs résultats sont obtenus sans moyennage, car l’estimation des paramètres de nuisance est de bonne qualité.

En dessous d’un RSB de 10 dB, le moyennage spectral est utile et améliore les performances. Pour les données qui restent relativement peu bruitées, les meilleurs résultats sont obtenus pour  $\Lambda \in [20 : 200]$ , les performances restant assez stables sur cette large plage.

Pour les données fortement bruitées, à RSB négatif, il est préférable d’utiliser les plus grandes valeurs de  $\Lambda$  possibles, entre 400 et 1000, le gain en précision étant très important.

Pour conclure, si l’on ne connaît pas la qualité des données que l’on utilise, la plage optimale de  $\Lambda$  est de 200 à 400, sauf en cas d’étoile extrêmement faible où il faut choisir un  $\Lambda$  supérieur à 400. En effet, l’amélioration des résultats obtenue pour les données très bruitées est supérieure à la dégradation de celles moins bruitées, tant en terme de moyenne des indices  $\varepsilon_2$  sur l’ensemble des longueurs d’onde que sur les valeurs maximales. Le temps de calcul est par ailleurs considérablement réduit.

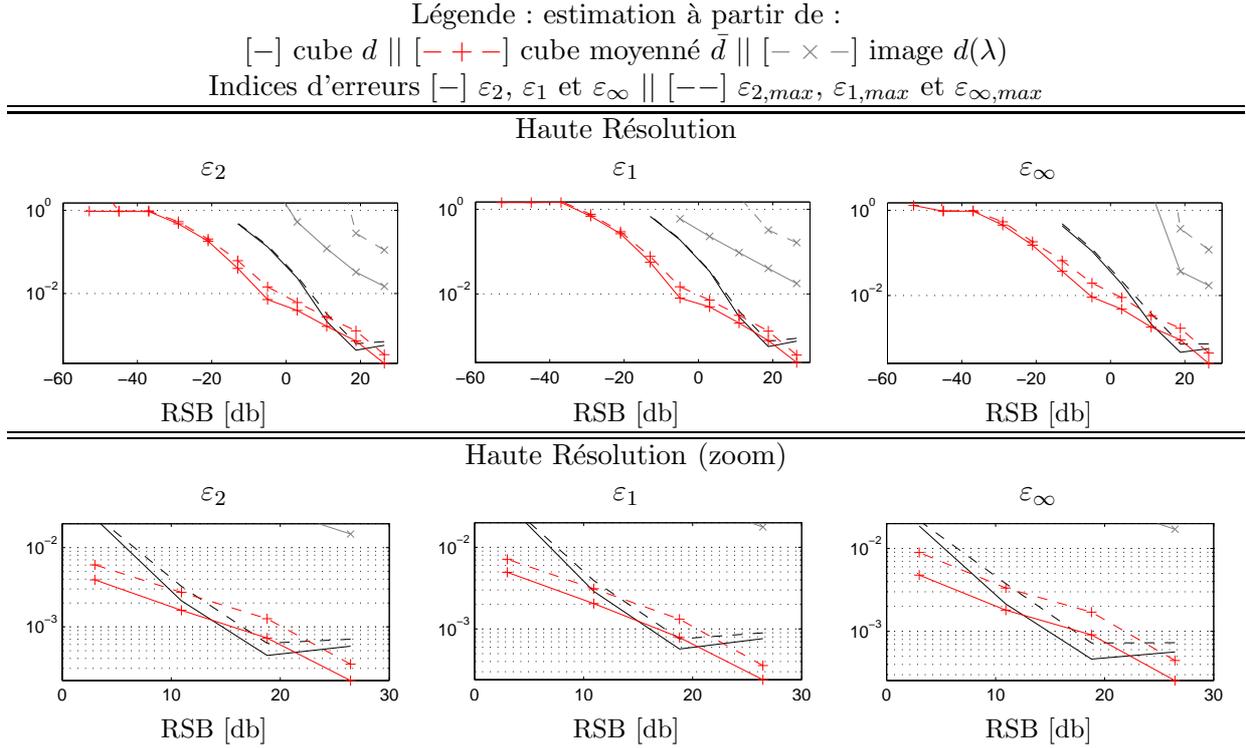


FIG. III.6 – Comparaison des performances d'estimation de la FSF à partir d'un cube bruité MUSE [noir] et d'un cube moyenné [rouge] par un facteur  $\Lambda = 40$ . Indices d'erreurs en fonction du rapport signal sur bruit des données bruitées. Moyenne des résultats obtenus sur les 3 positions sub-pixellaires de l'étoile.

### III.4 Estimation jointe

Dans les paragraphes III.2 et III.3, nous avons estimé le spectre de l'étoile  $s(\lambda)$  à partir des données puis le vecteur d'hyper-paramètres  $\theta$  pour cette valeur de  $s(\lambda)$ . Dans cette partie, nous allons estimer conjointement  $\theta$  et  $s(\lambda)$ .

Nous avons vu précédemment qu'en connaissant ou en estimant  $s(\lambda)$  à partir des données bruitées par  $\widehat{s(\lambda)} = \sum_{x,y} d(x,y,\lambda)$ , il est possible d'estimer  $\theta$  en minimisant l'anti-log-vraisemblance par rapport à  $\theta$  (équation III.3). Cependant, il est légitime de chercher à estimer  $s(\lambda)$  par maximum de vraisemblance. Cela revient à estimer conjointement les paramètres  $\theta$  et  $s(\lambda)$  par maximum de vraisemblance jointe, soit à minimiser l'anti-log-vraisemblance jointe ( $NLJL$ ) :

$$NLJL(s(\lambda), \theta | \mathbf{D}) = \sum_{xy} \frac{(d(x,y,\lambda) - s(\lambda) \cdot m(x,y,\lambda, \theta))^2}{\sigma_{x,y,\lambda}^2} \quad (\text{III.4})$$

On constate que cette fonction est quadratique en  $s(\lambda)$  pour  $\theta$  fixé. On peut donc, pour  $\theta$  fixé, obtenir une solution explicite du spectre en minimisant cette fonction :

$$s(\lambda)_{MC} = \arg \min_{s(\lambda)} NLJL(s(\lambda), \theta | \mathbf{D}) = \frac{1}{\sum_{x,y} \frac{m^2(x,y,\lambda, \theta)}{\sigma^2(x,y,\lambda)}} \sum_{x,y} \frac{m(x,y,\lambda, \theta) d(x,y,\lambda)}{\sigma^2(x,y,\lambda)} \quad (\text{III.5})$$

Une première manière de minimiser la  $NLJL$  conjointement par rapport à  $s(\lambda)$  et  $\theta$  consiste à faire une minimisation alternée, dont la structure est décrite par le Tab. III.6.

Magnitude SNR (dB)	18 26.46	19 18.80	21 3.00	23 -12.96	24 -20.96	25 -28.96
$\Lambda = 1$	0.052%	<b>0.043%</b>	<i>2.35%</i>	<i>47.15%</i>	> 3200%	> 4900%
$\Lambda = 20$	<b>0.021%</b>	0.073%	<b>0.38%</b>	<i>6.21%</i>	<i>26.71%</i>	59.83%
$\Lambda = 40$	<b>0.021%</b>	0.072%	0.39%	4.02%	18.04%	47.37%
$\Lambda = 60$	0.022%	0.067%	0.39%	3.49%	12.51%	41.20%
$\Lambda = 80$	0.022%	0.064%	0.42%	2.53%	10.84%	69.38%
$\Lambda = 100$	0.022%	0.057%	<b>0.38%</b>	2.44%	9.02%	89.91%
$\Lambda = 200$	0.025%	<b>0.055%</b>	0.40%	2.13%	6.99%	114.22%
$\Lambda = 400$	0.040%	0.059%	0.48%	<b>1.85%</b>	6.41%	98.11%
$\Lambda = 1000$	<i>0.077%</i>	<i>0.074%</i>	<i>0.81%</i>	2.53%	<b>5.64%</b>	<b>20.20%</b>

TAB. III.5 – Erreur d'estimation  $\varepsilon_2$  en haute résolution, pour différentes valeurs du paramètre  $\Lambda$  de moyennage spectral et différentes magnitudes d'étoiles. [*italique* / **gras**] valeur de  $\bar{\varepsilon}_2$  la plus élevée/basse à un RSB donné.

1. Estimation de la variance $\widehat{\sigma}^2(x, y, \lambda)$ à partir du cube de données bruitées $d_c$ contenant le ciel.
2. Initialisation du spectre $s(\lambda)^{(0)}$ par calcul du flux des données bruitées $d$ à chaque longueur d'onde.
3. Initialisation des paramètres $\theta^{(0)}$ à partir du cube normalisé par le spectre $s(\lambda)^{(0)}$ (voir III.2.3.1).
4. Itération k :
– Calcul des paramètres $\theta^{(k)}$ par minimisation de la <i>NLJL</i> (équation III.4) pour $s(\lambda) = s(\lambda)^{(k-1)}$ et $\widehat{\sigma}^2(x, y, \lambda)$ fixés.
– Calcul du spectre de l'étoile $s(\lambda)$ minimisant la <i>NLJL</i> de l'équation III.4 en utilisant l'expression III.5, pour $\theta = \theta^{(k)}$ .

TAB. III.6 – Algorithme d'estimation jointe de la FSF et du spectre de l'étoile par minimisation alternée.

Notons que cet algorithme peut être appliqué sur un cube moyenné spectralement. Le spectre mis à jour par maximum de vraisemblance est alors le spectre moyenné  $\bar{s}(\bar{\lambda})$ . En pratique, ce schéma d'estimation converge en peu d'itérations (typiquement moins de 10).

La Fig. III.7 compare, par l'indice d'erreur  $\bar{\varepsilon}_2$  calculé en haute résolution, les performances d'estimation jointe par processus itératif à celles du processus sans itération, pour différentes valeurs du facteur  $\Lambda$  de moyennage spectral. Les résultats, quantifiés par les autres indices d'erreur, sont tout à fait similaires.

L'estimation par maximum de vraisemblance jointe évite les résultats aberrants que l'on obtient avec l'estimation par maximum de vraisemblance simple en dessous de  $-37$  dB (l'absence de point à certains RSB signifie que le résultat n'a pas pu être exploité). Néanmoins, sur cet intervalle, l'indice d'erreur  $\bar{\varepsilon}_2$  obtenu avec cette méthode avoisine 100%. La FSF estimée est alors inexploitable.

Pour un RSB supérieur à  $-37$  dB, l'estimation jointe fournit des résultats légèrement meilleurs que l'estimation simple. En effet, les indices d'erreur  $\bar{\varepsilon}_2$  obtenus avec 10 itérations sont presque toujours inférieurs à ceux obtenus avec la méthode précédente. L'amélioration est plus perceptible lorsque le bruit est important car celui-ci perturbe davantage l'estimation du spectre. Donc l'amélioration progressive de l'estimation du spectre est d'autant plus utile qu'il y a beaucoup de bruit.

Si cette méthode d'estimation jointe par processus itératif ne semble pas repousser la limite en magnitude jusqu'à laquelle il est possible d'estimer la FSF, elle améliore la qualité des résultats qui étaient déjà fiables sans itération ainsi que la connaissance du spectre.

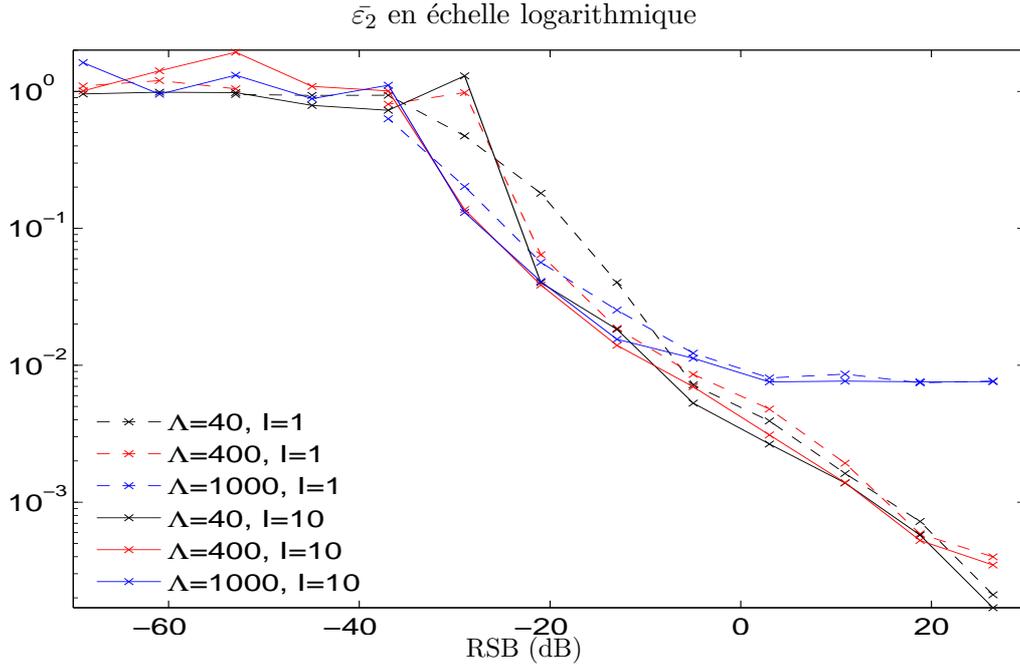


FIG. III.7 – Performances d’estimation jointe de la FSF et du spectre de l’étoile par maximum de vraisemblance avec algorithme itératif pour  $I = 10$  itérations, sur des cubes moyennés d’un facteur  $\Lambda = 40, 400$  ou  $1000$ , en fonction du RSB. Comparaison aux performances de l’estimation par ML simple avec les mêmes facteurs.

Une seconde méthode pour estimer conjointement le spectre et la FSF serait d’intégrer le critère par rapport à  $s(\lambda)$ . En effet, puisque l’estimateur du spectre a une expression explicite en fonction de  $\theta$ , il est possible de remplacer le spectre par son estimateur dans l’expression de l’anti-log-vraisemblance. Ainsi, plutôt que de minimiser alternativement  $NLJL(\theta, s(\lambda)|\mathbf{D})$  par rapport à  $s(\lambda)$  et  $\theta$ , on préfère minimiser :

$$K(\theta) = \arg \min_{s(\lambda)} NLJL(\theta, s(\lambda)|\mathbf{D}) = \sum_{xy} \frac{(d(x, y, \lambda) - s(\lambda)_{MC} \cdot m(x, y, \lambda, \theta))^2}{\sigma_{x,y,\lambda}^2}$$

avec  $s(\lambda)_{MC}$  donné par l’équation III.5.

La faiblesse de ces deux méthodes d’estimation par maximum de vraisemblance réside néanmoins dans l’existence potentielle de minima locaux pour les critères à minimiser, dus à la non-linéarité en les paramètres du modèle d’observation. Ainsi, les résultats obtenus sont dépendants de la valeur initiale fournie à l’algorithme de minimisation. Un travail supplémentaire consisterait à améliorer cette initialisation, mais rien ne permettrait cependant de garantir que le minimum global serait atteint.

### III.5 Estimation par espérance *a posteriori*

Pour estimer les paramètres  $\theta = [\alpha_0, \alpha_1, \beta_0, x_0, y_0]$ , nous proposons de recourir aux méthodes de simulation de Monte-Carlo, qui constituent une alternative efficace aux méthodes d’optimisation, pour pallier le problème des minima locaux. Pour cela, nous allons définir un cadre bayésien en définissant les lois *a priori* des paramètres. Le principe des simulations de Monte-Carlo est de simuler un grand nombre d’échantillons des paramètres  $\theta$  suivant la loi *a posteriori*, puis d’estimer ces paramètres par l’espérance de la loi *a posteriori* (notée PM). Dans un premier temps, nous supposons le spectre de l’étoile et la variance du bruit connus ou estimés à partir des données.

### III.5.1 Modèle bayésien

Définir un cadre bayésien consiste à ajouter aux informations disponibles à travers la vraisemblance, certaines contraintes parfois d'origine physique aux paramètres que l'on étudie, sous la forme de lois de probabilité *a priori*.

#### III.5.1.1 Vraisemblance

Nous nous plaçons comme précédemment dans le cadre de la vraisemblance gaussienne. Comme le détaille le paragraphe III.1, le modèle des données bruitées s'écrit sous la forme d'un spectre d'étoile convolué par la FSF, auquel s'ajoute un bruit gaussien. La vraisemblance gaussienne des données a alors l'expression (III.2).

#### III.5.1.2 Lois *a priori*

Les lois *a priori* contiennent des informations qui permettent de contraindre les valeurs des variables en fonction du problème étudié. Nous avons choisi d'associer aux variables des lois peu informatives, ne prenant en compte que des limites physiques :

- paramètre  $\alpha_0$  : D'après l'évolution du paramètre  $\alpha$  étudiée dans le Chapitre II et le modèle que nous utilisons ici, l'ordonnée à l'origine  $\alpha_0$  du paramètre de taille est positif. Aussi, nous avons choisi de le décrire suivant une loi uniforme positive :  $\alpha_0 \sim \mathcal{U}_{[0;10 \alpha_0^i]}$  (cf. III.2.3.1).

Nous pourrions également le définir par une loi normale  $\alpha_0 \sim \mathcal{N}^+(0, r)$  ou une loi inverse-gamma  $\alpha_0 \sim \mathcal{IG}(a, b)$  mais il est plus délicat de régler les paramètres  $r$  ou  $(a, b)$  que dans le cas d'une loi uniforme.

- paramètre  $\alpha_1$  : La traversée de l'atmosphère étant moins favorable aux faibles longueurs d'onde, la pente  $\alpha_1$  du paramètre de taille en fonction de la longueur d'onde est négative. Nous le décrivons par une loi uniforme pour des valeurs négatives :  $\alpha_1 \sim \mathcal{U}_{[10 \alpha_1^i; 0]}$ .
- paramètre  $\beta_0$  : Pour que la fonction de Moffat définissant le modèle soit intégrable, il faut que le paramètre de forme  $\beta_0$  soit strictement supérieur à 1. Aussi, nous avons choisi une loi inverse gamma décalée :  $\beta_0 = 1 + \beta'$  avec  $\beta' \sim \mathcal{IG}(\alpha_b, \beta_b)$ .

Une autre possibilité serait une loi normale positive de moyenne 1 :  $\beta_0 \sim \mathcal{N}^+(1, r_b)$ .

- positions  $x_0, y_0$  : Nous supposons que la position de l'étoile  $(x_0, y_0)$  est proche de la position déterminée par l'initialisation. Nous la décrivons donc par une loi normale de moyenne la valeur initiale et de faible variance :  $x_0$  et  $y_0 \sim \mathcal{N}(x_0^i$  ou  $y_0^i, r_{xy})$ .

Par ailleurs, nous travaillons avec des données extraites d'un cube hyperspectral dans lequel se trouve une étoile. La position de l'étoile  $(x_0, y_0)$  appartient donc à l'intervalle de la taille du cube en pixels. Nous pourrions donc choisir une loi uniforme  $x_0$  et  $y_0 \sim \mathcal{U}_{[1;N]}$  ou une loi triangulaire  $x_0 \sim \Lambda_{(1, x_0^i, N)}(x_0)$  centrée en  $x_0^i$  ou  $y_0^i$  et définie sur l'intervalle  $[1; N]$ .

Notons qu'en raison des lois *a priori* que nous avons choisies, les paramètres à régler sont  $\Theta = [\alpha_0^i, \alpha_1^i, x_0^i, y_0^i, \alpha_b, \beta_b, r_x, r_y]$ . Pour cela, nous nous sommes servis des valeurs initiales de chaque paramètre  $\theta_i = [\alpha_0^i, \alpha_1^i, \beta_0^i, x_0^i, y_0^i]$  que nous avons calculé dans le cas de l'estimation par maximum de vraisemblance. Les paramètres  $(\alpha_b, \beta_b)$  ont été réglés de manière à favoriser les valeurs entre 1 et 4. Nous émettons l'hypothèse que tous les paramètres du vecteur  $\theta$  sont indépendants. En conséquence, la loi de probabilité *a priori* du vecteur  $\theta$  s'exprime comme le produit de la loi de probabilité de chaque paramètre :  $p(\theta) = \prod_k p(\theta_k)$ .

### III.5.1.3 Loi *a posteriori*

La loi de Bayes [IDI 01] permet d'exprimer la loi *a posteriori*  $p(\boldsymbol{\theta}|\mathbf{D})$  en fonction de la vraisemblance  $p(\mathbf{D}|\boldsymbol{\theta})$  et des lois *a priori*  $p(\boldsymbol{\theta})$  et  $p(\mathbf{D})$  :

$$p(\boldsymbol{\theta}|\mathbf{D}) = \frac{p(\mathbf{D}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{D})} \quad (\text{III.6})$$

En remplaçant la vraisemblance gaussienne et les probabilités *a priori* par leurs expressions, on obtient<sup>17</sup> :

$$\begin{aligned} p(\boldsymbol{\theta}|\mathbf{D}) &= p(\mathbf{D}|\boldsymbol{\theta}) \cdot p(\alpha_0)p(\alpha_1)p(\beta_0)p(x_0)p(y_0) \\ &= \frac{1}{\sqrt{2\pi}^{N_x N_y N_\lambda} \sqrt{\det \mathbf{\Gamma}}} \exp\left(-\frac{1}{2}\|\mathbf{D} - s(\lambda) \cdot m(x, y, \lambda, \boldsymbol{\theta})\|_{\mathbf{\Gamma}^{-1}}^2\right) \\ &\cdot \frac{1}{10 \alpha_0^i} \mathbf{1}_{[0;10 \alpha_0^i]}(\alpha_0) \\ &\cdot \frac{1}{10 \alpha_1^i} \mathbf{1}_{[10 \alpha_1^i;0]}(\alpha_1) \\ &\cdot \frac{\beta_b^{\alpha_b-1}}{\Gamma(\alpha_b)} (\beta - 1)^{\alpha_b-1} \exp\left(\frac{-\beta_b}{(\beta - 1)}\right) \\ &\cdot \frac{1}{\sqrt{2\pi} r_x} \exp\left(-\frac{(x_0 - x_0^i)^2}{2r_x}\right) \\ &\cdot \frac{1}{\sqrt{2\pi} r_y} \exp\left(-\frac{(y_0 - y_0^i)^2}{2r_y}\right). \end{aligned}$$

Le principe de l'estimation par espérance *a posteriori* par méthode de Monte-Carlo consiste à générer des réalisations de  $\boldsymbol{\theta}$  suivant cette loi *a posteriori* et à calculer la moyenne empirique de ces  $P$  réalisations :

$$\hat{\boldsymbol{\theta}} = \frac{1}{P} \sum_{k=1}^P \boldsymbol{\theta}^{(k)}.$$

Par ailleurs, la variance de ces échantillons fournit des barres d'erreur, utiles pour évaluer la fiabilité des résultats.

Cette loi *a posteriori* n'a pas de structure particulière à exploiter pour générer les réalisations des paramètres. En effet, les lois *a posteriori* conditionnelles de chaque paramètre ne correspondent pas à des lois classiques, ce qui pourrait être exploité par un échantillonneur de Gibbs, comme ce sera le cas dans le Chapitre IV. Nous nous tournons plutôt vers l'algorithme de Metropolis-Hastings, dans sa version vectorielle.

### III.5.2 Méthode d'échantillonnage de Metropolis-Hastings à marche aléatoire

L'algorithme de Metropolis-Hastings, dont le principe est présenté dans le Tab. III.7, permet de générer des échantillons  $\boldsymbol{\theta}$  distribués suivant une loi cible  $\pi$ , à partir d'échantillons générés suivant une loi de proposition  $Q$ . Cette loi, également appelée loi instrumentale, est souvent symétrique, comme dans le cas d'une loi normale, ce qui simplifie le calcul de la valeur de test. Notons que les lois  $\pi$  et  $Q$  (si cette dernière est asymétrique) doivent être simplement connues à un facteur près. Dans le cas où la loi de proposition dépend de l'échantillon, cas que nous utilisons ici, alors les échantillons générés suivant  $Q$  forment une chaîne de Markov, que l'on appelle une marche aléatoire.

<sup>17</sup> $\Gamma(\alpha_b)$  désigne ici la fonction Gamma en  $\alpha_b$  et non la matrice de variance du bruit.

Pour générer des échantillons distribués suivant la loi cible  $\pi$  :

- générer aléatoirement ou de façon déterministe  $\boldsymbol{\theta}^{(0)}$ ,
- pour chaque itération  $n$  :
  - générer un échantillon  $\boldsymbol{\theta}^*$  suivant la loi de proposition  $Q(\cdot|\boldsymbol{\theta}^{n-1})$
  - générer un échantillon  $u$  suivant la loi uniforme  $\mathcal{U}_{[0;1]}$
  - si  $u \leq \min\left(1, \frac{\pi(\boldsymbol{\theta}^*)Q(\boldsymbol{\theta}^{n-1}|\boldsymbol{\theta}^*)}{\pi(\boldsymbol{\theta}^{n-1})Q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{n-1})}\right)$   
 alors  $\boldsymbol{\theta}^n = \boldsymbol{\theta}^*$   
 sinon  $\boldsymbol{\theta}^n = \boldsymbol{\theta}^{n-1}$ .

TAB. III.7 – Algorithme de Metropolis-Hastings, d'après [DOU 05].

Nous avons choisi de générer le vecteur candidat suivant la loi de proposition gaussienne  $\boldsymbol{\theta}^* \sim \mathcal{N}(\boldsymbol{\theta}^{(n-1)}, \mathbf{s})$  avec  $\mathbf{s}$  les variances respectives des lois de proposition des cinq paramètres à générer. Une autre possibilité est de générer les échantillons par une loi triangulaire  $\Lambda(\boldsymbol{\theta}^{(n-1)}, r)$  centrée en  $\boldsymbol{\theta}^{(n-1)}$  avec  $r$  la largeur (triangle défini sur  $[\mu - r, \mu + r]$ ).

Afin d'explorer d'avantage l'espace des variables, on peut générer à fréquence connue, un échantillon des paramètres sur une loi plus large, et dans le cas d'une loi gaussienne, avec un écart-type beaucoup plus grand. Ce processus n'altère pas la qualité de l'estimation car les échantillons peu probables sont rejetés.

### III.5.3 Estimation conjointe du spectre de l'étoile et de la FSF

Comme dans le cas de l'estimation par maximum de vraisemblance, on peut s'attendre à une amélioration de l'estimation si l'on estime les paramètres de nuisance (en particulier le spectre de l'étoile) conjointement aux paramètres de la FSF.

Il est possible d'adapter le processus d'estimation itérative de la FSF vu au paragraphe III.4 au cas d'estimation par espérance *a posteriori* (PM). A l'issue de l'estimation des paramètres par PM, le spectre approximativement calculé lors de l'initialisation, est calculé à nouveau par moindres carrés. Cet enchaînement est répété plusieurs fois, ce qui permet d'améliorer successivement l'estimation des paramètres et du spectre. Même si ce processus n'est pas cohérent dans un cadre d'estimation par espérance *a posteriori*, il est simple à mettre en œuvre.

Pour conserver une estimation par espérance *a posteriori*, il faudrait échantillonner conjointement le spectre et les paramètres du modèle de FSF. Le principal inconvénient d'une telle approche vient de la dimension de l'espace des paramètres qui est largement augmenté. Notons que pour une loi *a priori* gaussienne du spectre, la loi *a posteriori* du spectre conditionnellement aux autres paramètres est également gaussienne donc simple à échantillonner. Cette propriété pourrait être exploitée pour un tel schéma d'échantillonnage.

Une autre solution consisterait à intégrer le spectre de l'étoile hors du problème, comme cela a été fait dans [DOU 01]. Ainsi, pour un *a priori* gaussien du spectre, il est aisé de calculer la loi *a posteriori*  $p(\boldsymbol{\theta}|\mathbf{D})$  par marginalisation de  $p(\boldsymbol{\theta}, \{s(\lambda)\}_\lambda | \mathbf{D})$ . On pourrait alors échantillonner suivant cette loi *a posteriori* marginale, ce qui permettrait à la fois d'avoir un espace de paramètre de faible dimension et de prendre en compte de façon cohérente les paramètres  $\{s(\lambda)\}_\lambda$  dans cette estimation (cf. annexe D.2). Notons que cet estimateur est asymptotiquement identique à celui obtenu en échantillonnant conjointement le spectre et les paramètres du modèle. Toutefois, la vitesse de convergence peut différer entre les deux méthodes.

### III.5.4 Résultats d'expérimentations

L'objectif est de comparer les performances d'estimation de la FSF par espérance *a posteriori* (PM) avec une ou plusieurs itérations, par rapport aux résultats obtenus par ML et ML itératif.

Les données sont les mêmes que celles des paragraphes III.2.3 et III.3.4.2 ; ce sont des simulations de cubes idéaux bruités d'observation d'une étoile isolée, suivant le modèle précédemment établi III.1.3.

Les valeurs des paramètres  $\theta$  sont initialisées de la même façon que précédemment. On génère ensuite un grand nombre d'échantillons (10 000) selon l'algorithme de Metropolis-Hastings (MH) par marche aléatoire afin d'approcher la loi *a posteriori*. On en conserve uniquement la seconde moitié (période de *burn-in*) afin d'assurer l'indépendance des échantillons vis-à-vis de l'initialisation. Les paramètres sont estimés par l'espérance de cette loi en calculant la moyenne des échantillons.

Comme dans les sections précédentes, on calcule l'erreur quadratique moyenne  $\bar{\varepsilon}_2$  entre la vraie FSF en haute résolution et la FSF estimée pour quantifier la qualité des résultats.

D'après la figure III.8 (a), l'estimation par espérance *a posteriori* donne des indices d'erreur similaires, très légèrement supérieurs à ceux obtenus par estimation itérative par maximum de vraisemblance. Ces résultats sont plutôt satisfaisants. Ils dépendent des réglages des loi *a priori*, de la loi de proposition de la marche aléatoire ainsi que de la durée du *burn-in* et du nombre total d'échantillons.

Ici, les paramètres de la marche aléatoire ont été réglés de façon à accepter environ un quart des échantillons. La période de *burn-in* correspond à la première moitié des échantillons (conformément aux recommandations de Gelman and Shirley ([BRO 11]-chap.6)). Le nombre total d'échantillons a été fixé visuellement afin d'atteindre la convergence.

L'estimation jointe de la FSF et du spectre donne des résultats moins bons lorsque il y a peu de bruit, ceci pouvant être expliqué par le fait que les paramètres de *burn-in* et du nombre d'échantillons n'ont pas été ajustés. Par ailleurs, l'erreur ne semble pas diminuer au cours des itérations, rendant ce processus inutile. Ce processus mixte semble uniquement utile lorsque le bruit est important.

## III.6 Synthèse des méthodes d'estimation du spectre

Dans la section III.3.4, nous avons constaté l'impact de la méconnaissance du spectre de l'étoile sur l'estimation de la FSF. En effet, celle-ci nécessite ce spectre, alors estimé à partir du flux des données bruitées. Nous avons indiqué, dans les paragraphes précédents, plusieurs méthodes pour estimer le spectre de l'étoile et pour réduire sa variance d'estimation.

- Dans la section III.2, le spectre est estimé à partir du cube de données par le flux total reçu à chaque longueur d'onde  $\widehat{s}(\lambda) = \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} f(x, y, \lambda)$  (spectre servant pour l'initialisation de l'estimation de la FSF par ML).
- Dans la section III.3, nous proposons une méthode pour minimiser l'impact de la variance d'estimation du spectre, en travaillant non plus avec le spectre de l'étoile  $s(\lambda)$  à la résolution spectrale MUSE, mais avec un spectre moyenné  $\bar{s}(\bar{\lambda})$  dont la variance d'estimation est réduite. Le spectre pour toute longueur d'onde peut alors être estimé par moindres carrés pondérés à la fin du processus d'estimation de la FSF, à partir du cube de données et du modèle estimé.
- Enfin, le spectre peut être estimé conjointement aux paramètres du modèle de la FSF, par maximum de vraisemblance, en utilisant le processus itératif présenté dans la section III.4. Cette méthode peut être appliquée aussi bien sur le cube de données  $d$  que sur le cube moyenné  $\bar{d}$ .

Ainsi, nous pouvons nous demander quelles sont les performances d'estimation du spectre à l'échantillonnage spectral MUSE en utilisant ces différentes méthodes. La Fig. III.9 compare les spectres estimés pour différentes magnitudes d'étoiles, grâce aux méthodes rappelées ci-dessus. Les résultats obtenus par estimation conjointe du spectre de l'étoile et de la FSF sont meilleurs que ceux obtenus lors

de l'initialisation par simple calcul du flux, car il y a un apport d'information supplémentaire via la FSF. Cette méthode d'estimation conjointe est uniquement efficace si le RSB est élevé. Dans le cas contraire, l'estimation du spectre de l'étoile est très bruitée, en particulier à cause du spectre du ciel qui participe au bruit. On voit de fortes erreurs d'estimation du spectre de l'étoile à certaines longueurs d'onde correspondant à des raies du ciel. Néanmoins, si l'on souhaite utiliser cette méthode, on constate qu'il n'est pas nécessaire d'estimer la FSF à partir du cube entier (Fig. III.9 – 2<sup>ème</sup> ligne). Au contraire, l'estimation de la FSF à partir d'un cube moyenné, suivie de l'estimation du spectre de l'étoile par MC sur le cube entier, donne de meilleurs résultats (Fig. III.9 – 3<sup>ème</sup> ligne).

### III.7 Estimation à partir d'une étoile non isolée

Les méthodes et les tests présentés précédemment supposent qu'il est possible d'extraire des données le sous-cube d'une étoile isolée. Cette hypothèse, nécessaire à l'écriture du modèle direct, est réaliste mais n'est pas forcément vérifiée dans le cas d'un champ d'observation, par exemple. Le cube extrait peut contenir une autre étoile, ou un objet étendu moins lumineux. Il est donc crucial de mesurer la capacité à estimer la FSF à partir d'une étoile qui n'est pas parfaitement isolée.

#### III.7.1 Données

Nous disposons de données issues de QSIM, le simulateur simplifié de données MUSE, contenant plusieurs types d'objet, que l'on peut observer dans la Fig. III.10 et dont les magnitudes sont précisées dans le Tab. III.8.

	Obj	mag(V)	mag(I)
STAR-01	2	22.94	21.12
STAR-02	3	24.85	23.03
LYA-OVERLAP-02	16	28.22	28.15

TAB. III.8 – Magnitude des objets extraits des données de référence QSIM, dans les bandes V (visible) et I (infrarouge) du spectre électromagnétique.

L'étoile n°3 appelée 'STAR-02' est relativement isolée car elle est assez éloignée de ses plus proches voisins n°4 et n°10. Par contre, l'étoile 'STAR-01' est très proche d'un émetteur Lyman- $\alpha$  (objet n°16), tous deux illustrés par la Fig. III.12, ce qui peut parasiter l'estimation de la FSF. En comparant l'amplitude des spectres de ces deux objets sur la Fig. III.11, on constate que le spectre de l'émetteur Lyman- $\alpha$  est plus faible d'un facteur 100 environ que le spectre de l'étoile 'STAR-01' sur la quasi-totalité des longueurs d'onde, sauf autour d'un pic d'émission situé entre 500 nm et 550 nm. Sur cet intervalle, les deux spectres ont le même ordre de grandeur. L'intensité des deux spectres, quantifiée par leur magnitude (cf. Tab. III.8), diffère de 5 magnitudes environ dans la bande V.

#### III.7.2 FSF de référence

La FSF utilisée pour simuler ces données est une fonction de Moffat circulaire, constante dans le champ, de paramètre  $\beta$  constant de valeur connue  $\beta = 2.5$  et de paramètre  $\alpha$  variable en longueur d'onde suivant une loi de puissance  $\lambda^{-\frac{1}{5}}$ . Il y a donc forcément une erreur de modèle entre la variation de  $\alpha$  et notre modèle linéaire.

Ces valeurs de  $\alpha$  constituent notre référence pour évaluer les performances d'estimation de la FSF à partir d'une étoile non isolée (cf. Fig. III.14).

---

<sup>18</sup>L'indice d'erreur utilisé ici est défini par  $\epsilon_2 = \frac{\sum_{\lambda} (s(\lambda) - \widehat{s(\lambda)})^2}{\sum_{\lambda} s^2(\lambda)}$ .

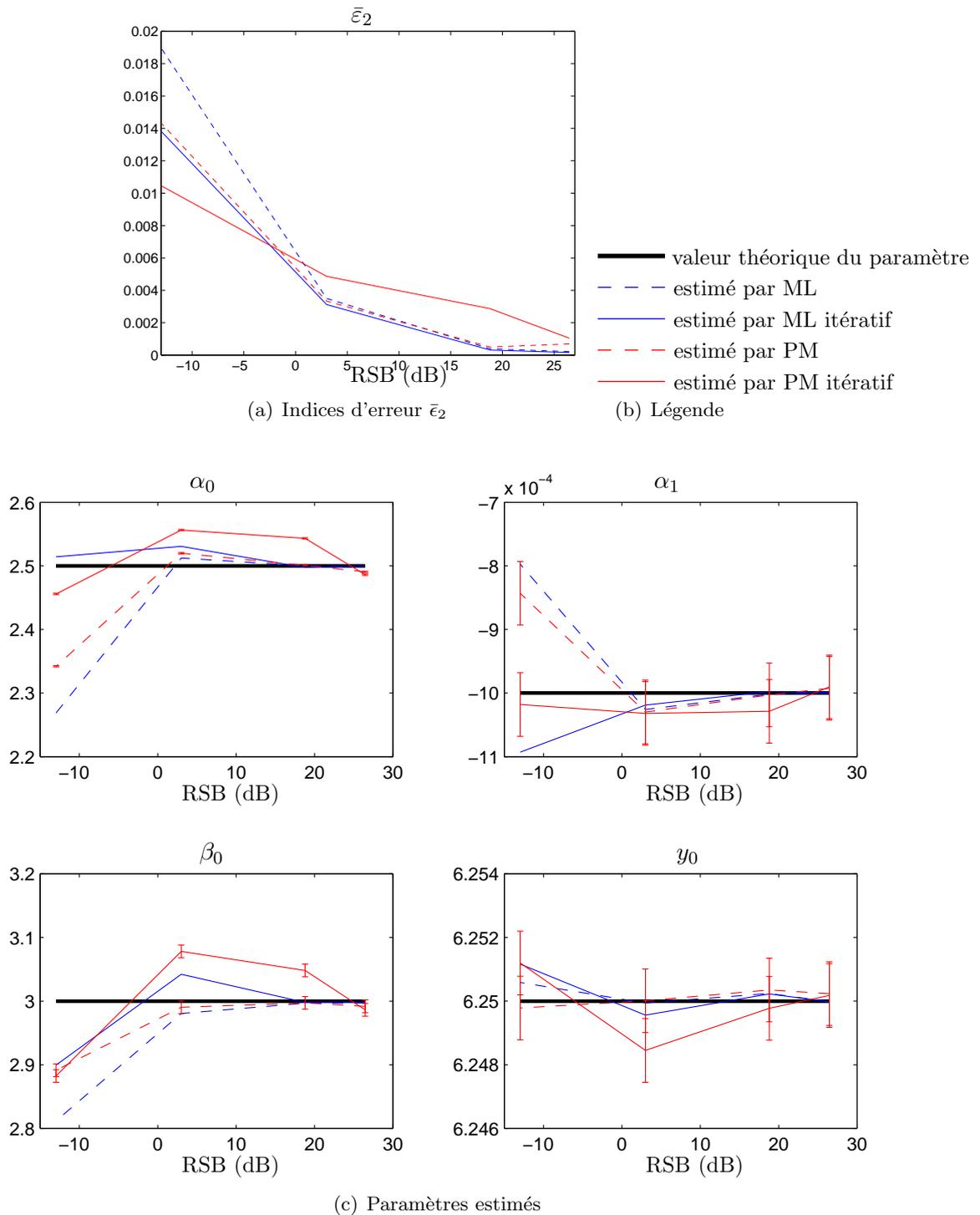


FIG. III.8 – Performances d'estimation en fonction du rapport signal sur bruit du cube exprimé en dB. Comparaison des 4 méthodes : estimation de la FSF par maximum de vraisemblance (ML), ML jointe par algorithme itératif, par espérance de la loi *a posteriori* (PM), par PM itérative. Paramètres théoriques et estimés (excepté  $x_0$  : résultats similaires à ceux de  $y_0$ ), associés aux barres d'erreur évaluées. L'étoile est positionnée sur le quart d'un pixel. Pour chaque niveau de RSB testé, le cube est moyenné par groupes de  $\Lambda = 100$  dans les 3 RSB positifs, par groupe de  $\Lambda = 400$  pour le RSB négatif, ce qui correspond dans tous les cas à un  $\Lambda$  optimal.

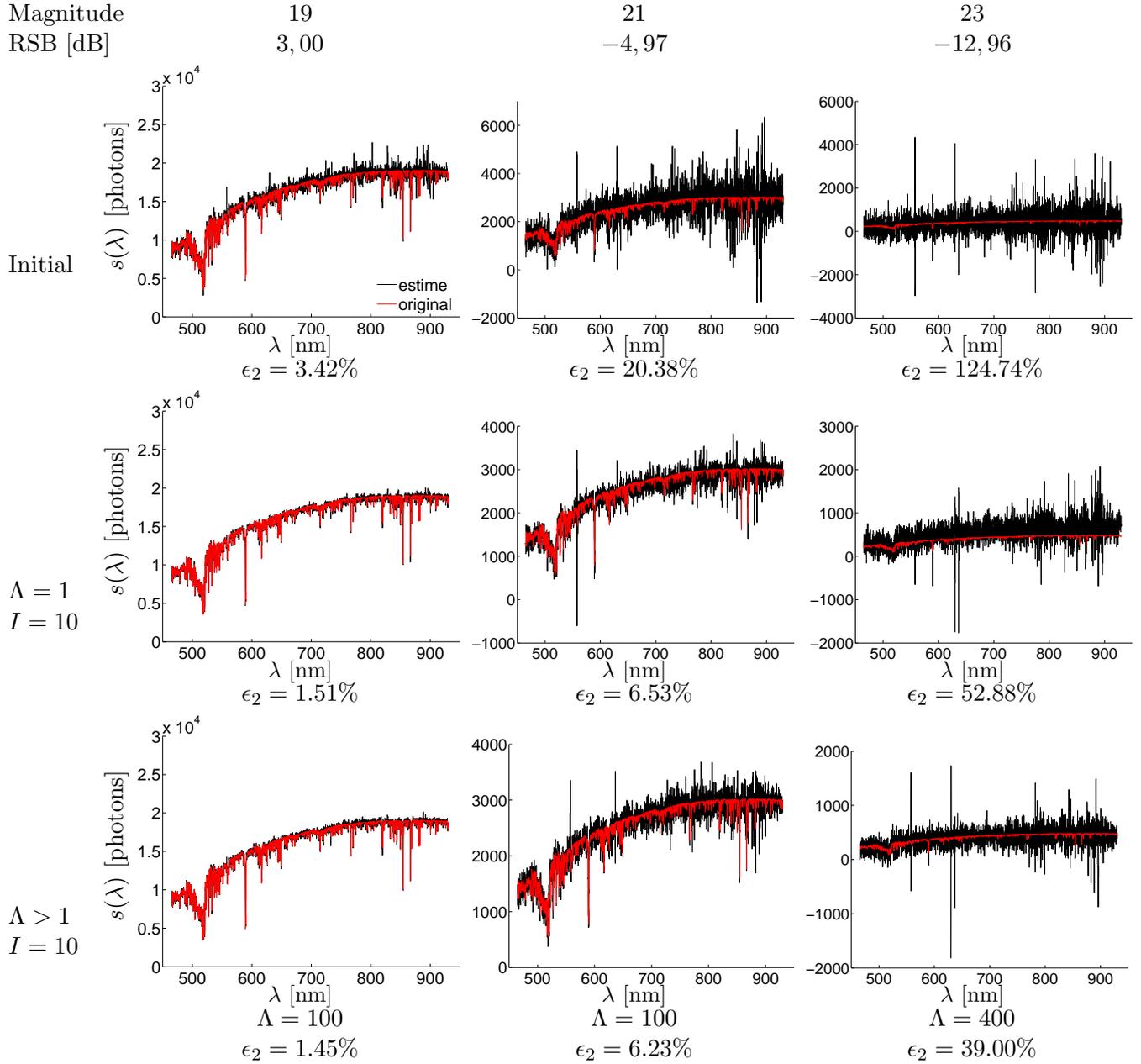


FIG. III.9 – Performances d'estimation<sup>18</sup> du spectre d'une étoile isolée. Spectre original [rouge], spectre estimé [noir]. En fonction du RSB des données, par différentes méthodes. 1<sup>ère</sup> ligne : estimation par calcul du flux des données bruitées, spectre utilisé pour l'initialisation de l'estimation de la FSF par ML. 2<sup>ème</sup> ligne : spectre estimé conjointement à la FSF par maximum de vraisemblance jointe sur 10 itérations. 3<sup>ème</sup> ligne : idem, à partir d'un cube moyenné spectralement.

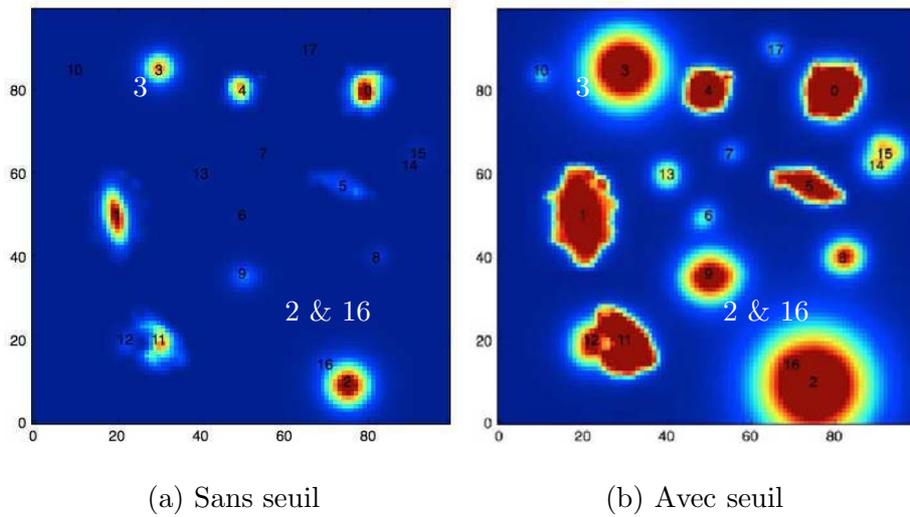


FIG. III.10 – Données QSIM de référence utilisée pour l'estimation de la FSF à partir d'une étoile non isolée. Image en lumière blanche du cube hyperspectral simulé, contenant 17 objets numérotés. Simulations dans le cas du WFM sans AO, pour une durée d'intégration de 10 heures, avec un *seeing*  $s = 0.8''$ .

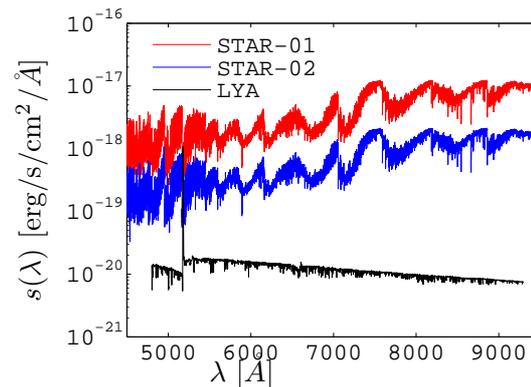


FIG. III.11 – Spectres des objets extraits des données de référence QSIM, en fonction de la longueur d'onde, en échelle logarithmique.

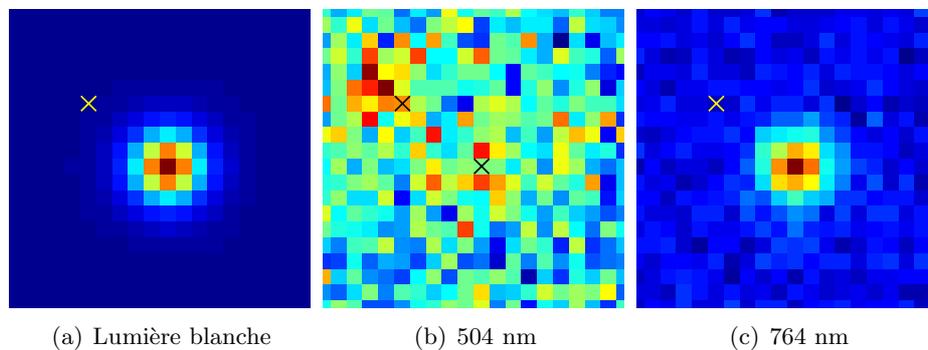


FIG. III.12 – Illustration du cube de données extrait des données QSIM, en lumière blanche (a), à 504 nm (b) qui correspond au pic d'émission de l'émetteur Lyman- $\alpha$ , et à une raie d'émission de l'étoile : 764 nm. Ce cube contient une étoile (au centre) et un émetteur Lyman- $\alpha$  marqué par une croix.

### III.7.3 Masques

L'influence des objets parasites sur les performances d'estimation de la FSF peut être atténuée en appliquant des masques spatiaux et/ou spectraux sur le cube d'observation ou sur la variance associée. Pour cela, nous proposons de construire un masque binaire 3D de la même taille que les données, dans lequel les 1 indiquent les valeurs à prendre en compte dans l'estimation et les 0 celles à omettre. La répartition des données écartées peut être spatiale, spectrale ou mixte. L'exploitation de ce masque pour l'estimation des paramètres  $s(\lambda)$  et  $\theta$  est rendue possible en remplaçant la valeur de la variance des pixels marqués par un 0, par une valeur infinie. Ainsi, ces pixels ne sont pas comptabilisés dans la vraisemblance ou dans la loi *a posteriori*.

D'après les caractéristiques de nos données présentées au paragraphe III.7.1, nous proposons un masque spatial et un autre spectral. Connaissant la position de l'étoile à partir de laquelle on souhaite estimer la FSF et celle de l'objet voisin, on peut séparer le champ de vue en deux par la médiatrice des deux points. Alors nous suggérons de masquer le demi plan contenant l'objet parasite, à toutes les longueurs d'onde, comme illustré par la Fig. III.13 (a). Le choix de la médiatrice comme élément découpeur du plan a pour but d'optimiser la perte d'informations associées à l'étoile cible vis-à-vis de la suppression d'informations relatives à l'objet parasite. On peut envisager de déplacer cette droite le long du segment [cible-parasite], en fonction du ratio entre leurs magnitudes. En outre, l'objet parasite a un flux du même ordre de grandeur que le flux de l'étoile cible, sur une partie du spectre et un flux négligeable sur le reste du spectre. Alors l'opération de masquage spectral consiste à associer la valeur 0 dans le cube binaire aux images de l'intervalle spectral parasite (cf. Fig. III.13 (b)). Ce masque peut être appliqué sur le cube de données entier ou sur le cube ayant subi un moyennage spectral.

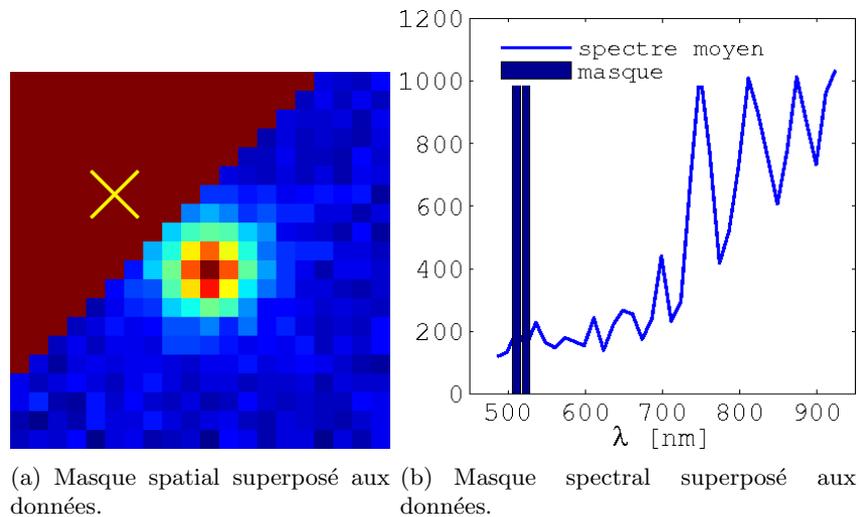


FIG. III.13 – Illustrations de masques spatiaux et spectraux pour l'estimation de la FSF à partir d'un objet non isolé. (a) Variance en lumière blanche associée aux données : la zone rouge correspond à la zone spatiale masquée ; la croix jaune indique l'emplacement de l'objet parasite. (b) Spectre de l'étoile cible et positionnement du masque spectral.

Notons que ces deux masques peuvent être appliqués aussi bien dans le cadre de l'estimation par maximum de vraisemblance, que dans celui du calcul de la moyenne *a posteriori*. Dans le cas où un objet voisin gêne l'estimation sur un intervalle restreint de longueurs d'onde, son flux étant négligeable sur le reste du spectre, les deux masques peuvent être combinés, en masquant spatialement l'objet parasite uniquement aux longueurs d'onde de l'intervalle critique. Cela permet de préserver les informations reliées à l'étoile cible qui sont prédominantes sur le reste du spectre.

### III.7.4 Résultats d'expérimentations

Un cube d'observation est extrait autour de l'étoile 'STAR-01' qui est proche d'un émetteur Lyman- $\alpha$ , dont les positions sont indiquées sur la Fig. III.12. Que ce soit en lumière blanche ou à une longueur d'onde quelconque (ici 764 nm), on vérifie, conformément à la magnitude des objets annoncés par le Tab. III.8, que le flux de l'objet voisin est nettement inférieur à celui de l'étoile parce qu'il n'apparaît pas sur ces deux images. À 504 nm, longueur d'onde à laquelle le flux de l'objet voisin est le plus élevé, les deux objets sont à peine distinguables dans le bruit.

La Fig. III.14 illustre les résultats de l'estimation de la FSF à partir de cette étoile non isolée, par maximum de vraisemblance jointe pour  $I = 10$  itérations et avec  $\Lambda = 100$  longueurs d'onde. Les paramètres  $\alpha_0$  et  $\alpha_1$  estimés fournissent des valeurs de  $\alpha$  très proches de celles de référence. Le paramètre de forme estimé  $\hat{\beta} = 2.48$  est également proche de la valeur théorique  $\beta = 2.50$ . Les plus grands écarts en terme de  $\alpha$  sont obtenus aux extrémités de l'intervalle spectral, c'est-à-dire dans le rouge et dans le bleu. En particulier, la plus grande erreur est commise dans le bleu à 465 nm, et la différence entre la vraie FSF et l'estimée à cette longueur d'onde (2<sup>nd</sup>e ligne de la Fig. III.14) est faible.

L'évolution de l'erreur est confirmée par celle de l'indice d'erreur  $\varepsilon_2$  en fonction de la longueur d'onde qui décroît depuis 3.5%. L'indice moyen sur toutes les longueurs d'onde est  $\bar{\varepsilon}_2 = 0.71\%$ , ce qui est très satisfaisant. La présence d'un objet voisin dont le flux est négligeable par rapport à celui de l'étoile n'est donc pas un frein à l'estimation de la FSF.

En terme de méthodologie, on constate sur la Fig. III.14 (a) que les valeurs de  $\alpha$  estimées se stabilisent à partir de la 2<sup>ème</sup> itération, car les valeurs de la 2<sup>ème</sup> et de la 3<sup>ème</sup> itérations sont superposées. Ainsi, dans ce cas, 2 itérations du processus d'estimation jointe décrit au paragraphe III.4 suffisent.

Le Tab. III.9 rassemble les valeurs de  $\beta$  estimées et les performances en terme de  $\bar{\varepsilon}_2$  pour différents paramètres algorithmiques. On vérifie l'utilité du moyennage spectral : l'indice d'erreur  $\bar{\varepsilon}_2 = 0.71\%$  obtenu avec un facteur  $\Lambda = 100$  est nettement plus satisfaisant que celui obtenu sans moyennage, de valeur  $\bar{\varepsilon}_2 = 2.13\%$ .

$I = 10$ itérations				
$\Lambda = 1$		$\Lambda = 100$		
	sans masque	sans masque	masque spatial	masque spectral
$\hat{\beta}$	2.60	2.48	2.49	2.48
$\bar{\varepsilon}_2$	2.13%	0.71%	0.72%	0.75%

TAB. III.9 – Performances d'estimation de la FSF et valeurs de  $\beta$  estimées à partir de l'étoile 'STAR-01' non isolée, par maximum de vraisemblance jointe, avec/sans moyennage spectral, pour 10 itérations, avec/sans masque.

En outre, l'estimation a été testée à partir de cubes ayant subi un masque spatial ou spectral. Les performances, mesurées par la valeur de  $\beta$  et l'indice  $\bar{\varepsilon}_2$ , ne sont pas améliorées par rapport à celle de l'estimation sans masque. Les masques spatiaux ou spectraux n'ont donc pas d'intérêt dans cet exemple. Ceci s'explique par la faiblesse du ratio entre le flux de l'objet parasite et celui de l'étoile. Néanmoins, on montre que la perte d'information induite par les masques ne provoque pas une détérioration significative des performances de l'estimation. L'information contenue dans le reste du cube masqué est suffisante pour estimer correctement les paramètres du modèle. L'utilisation de masques pourrait être plus intéressante pour l'estimation de la FSF à partir d'objets parasités de manière plus significative.

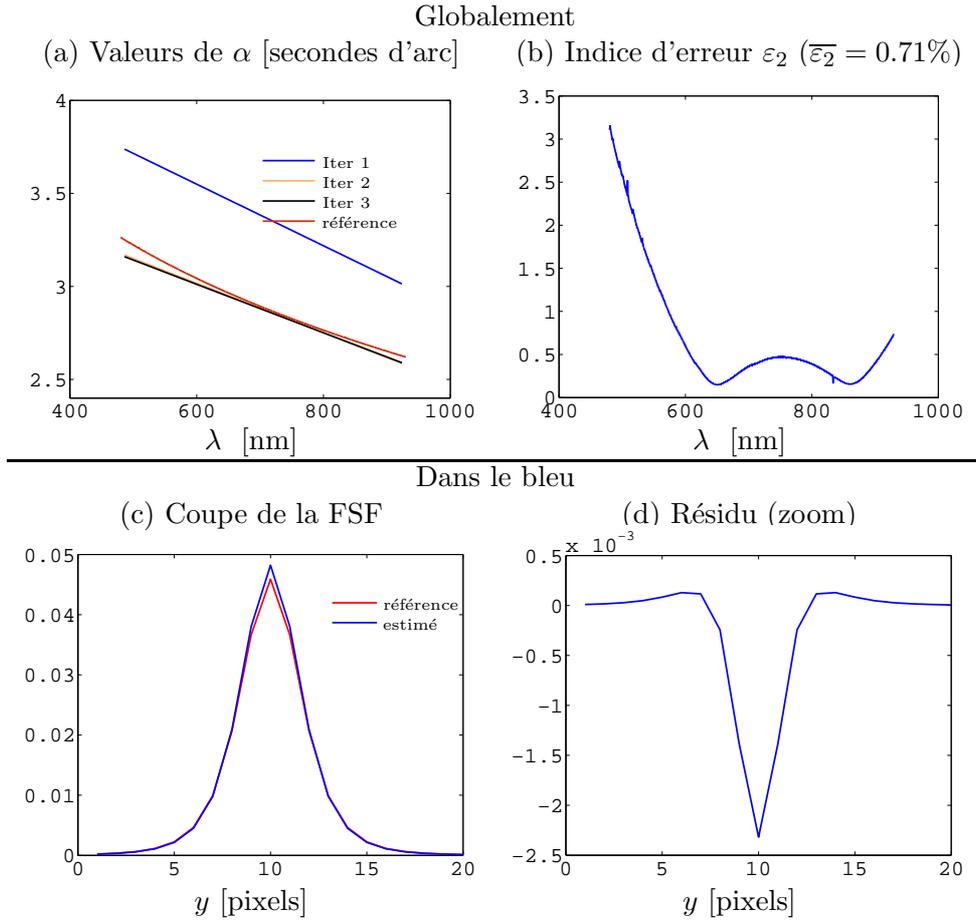


FIG. III.14 – Valeurs de  $\alpha$  estimées à partir de l'étoile 'STAR-01' non isolée, par maximum de vraisemblance, après moyennage spectral du cube avec  $\Lambda = 100$  et  $I = 10$  itérations (seules les 3 premières itérations sont affichées, les itérations suivantes donnant des résultats identiques). Quantification des performances par l'indice  $\varepsilon_2$  à la résolution MUSE en fonction de la longueur d'onde. Comparaison de la FSF estimée et de la FSF de référence dans le bleu ( $\lambda = 465$  nm), à un  $x$  fixé au centre du cube.

## Conclusions

La connaissance de la FSF spatiale de la chaîne instrumentale à toutes les longueurs d'onde est un enjeu crucial, non seulement pour évaluer la qualité des données, mais également pour appliquer plusieurs types de traitements, tels que la fusion, la séparation de sources et la déconvolution.

Dans le chapitre précédent, nous avons montré que dans le mode grand champ de MUSE sans correction par optique adaptative, il est possible de modéliser la FSF en toute position du cube par un modèle de Moffat dépendant d'un faible nombre de paramètres. Nous nous sommes intéressés dans ce chapitre à l'estimation de ces paramètres à partir de l'observation d'une étoile isolée.

Plutôt que d'estimer la FSF séparément à chaque longueur d'onde, nous avons proposé d'estimer les paramètres du modèle de FSF variant en longueur d'onde par maximum de vraisemblance sur l'intégralité d'un cube d'observation bruité. La redondance d'information sur l'ensemble des longueurs d'onde assure des performances d'estimation satisfaisantes même pour des étoiles faiblement brillantes. Une des difficultés du problème réside dans l'estimation de paramètres de nuisance que sont le spectre de l'étoile et la variance du bruit. Aussi, nous avons proposé une amélioration de l'estimation de ces paramètres de nuisance, s'appuyant sur le moyennage spectral des données. Dans ce cas, l'estimation de la FSF utilise des paramètres de nuisance moyennés dont la variance d'estimation est réduite. Nous avons pu constater que le moyennage spectral réduit significativement les indices d'erreur d'estimation de la FSF, notamment pour les données de RSB négatifs, et n'altère pas significativement les résultats dans les autres cas.

Nous suggérons également d'estimer plus finement le spectre de l'étoile en calculant sa valeur à chaque longueur d'onde par estimation jointe du spectre et de la FSF, par un processus itératif. Celui-ci améliore les performances de l'estimation de la FSF à partir d'étoiles peu brillantes.

En outre, dans le cadre d'un modèle bayésien des données, les paramètres du modèle de FSF peuvent être estimés par espérance *a posteriori*. Pour cela, il est nécessaire de générer des échantillons de la loi *a posteriori*, ce qui est fait asymptotiquement par un algorithme de Metropolis-Hastings. Cette méthode a pour principal intérêt de s'affranchir du problème de l'existence de minima locaux.

Par ailleurs, nous avons quantifié les performances d'estimation du spectre de l'étoile par les différentes méthodes présentées. Les deux stratégies proposées permettent une nette amélioration de l'estimation du spectre. Néanmoins, contrairement à l'estimation du modèle de FSF, celle du spectre ne profite pas d'une corrélation entre les différentes longueurs d'onde, ce qui la limite aux étoiles suffisamment brillantes.

Enfin, nous avons envisagé le cas de l'estimation de la FSF à partir d'une étoile non isolée, grâce à un masquage spatial et spectral des données. Les tests ont démontré l'utilité d'un modèle hyperspectral qui utilise conjointement toutes les longueurs d'onde et qui permet de minimiser l'influence de l'objet parasite, si celui-ci a un faible flux. Les stratégies envisagées pour réduire le biais induit par le voisin ne semblent pas détériorer de façon notable les performances par rapport à celles de l'estimation à partir d'une étoile isolée. Il faudrait explorer les différentes configurations possibles de voisinage de l'étoile et quantifier les performances de l'estimation en fonction de la magnitude des objets voisins.



# Chapitre IV

## Déconvolution de données hyperspectrales pour la cinématique des galaxies

### Sommaire

---

Notations . . . . .	94
Introduction . . . . .	95
<b>IV.1 Données . . . . .</b>	<b>95</b>
IV.1.1 Modèle d'objet . . . . .	95
IV.1.2 Modèle de PSF . . . . .	97
IV.1.3 Modèle de données . . . . .	97
IV.1.4 Estimation et déconvolution . . . . .	98
<b>IV.2 Méthodes usuelles d'estimation de cartes de cinématique . . . . .</b>	<b>98</b>
IV.2.1 Méthode des moments . . . . .	99
IV.2.2 Estimation par maximum de vraisemblance . . . . .	100
<b>IV.3 Estimation par espérance <i>a posteriori</i> . . . . .</b>	<b>101</b>
IV.3.1 Lois <i>a priori</i> . . . . .	101
IV.3.2 Échantillonnage stochastique de la loi <i>a posteriori</i> . . . . .	102
IV.3.3 Échantillonneur de Gibbs et lois conditionnelles . . . . .	102
IV.3.4 Choix des lois de proposition pour l'échantillonnage de Metropolis-Hastings . . . . .	104
<b>IV.4 Déconvolution de cartes de cinématique . . . . .</b>	<b>105</b>
IV.4.1 Maximum de vraisemblance conjointe . . . . .	105
IV.4.2 Échantillonnage stochastique de la loi <i>a posteriori</i> . . . . .	106
IV.4.3 Lois <i>a posteriori</i> conditionnelles . . . . .	106
IV.4.4 Déconvolution par espérance <i>a posteriori</i> conjointe . . . . .	108
<b>IV.5 Expérimentations sur données 1D1D . . . . .</b>	<b>108</b>
IV.5.1 Description des données . . . . .	108
IV.5.2 Méthodologie . . . . .	109
IV.5.3 Indicateurs de convergence pour PM et JPM . . . . .	109
IV.5.4 Résultats d'estimation . . . . .	112
IV.5.5 Résultats de déconvolution . . . . .	112
<b>IV.6 Expérimentations sur données 3D . . . . .</b>	<b>113</b>
IV.6.1 Description des données . . . . .	113
IV.6.2 Méthodologie . . . . .	114
IV.6.3 Résultats d'estimation . . . . .	114
IV.6.4 Résultats de déconvolution . . . . .	115
<b>Conclusions . . . . .</b>	<b>125</b>

---

## Notations

$a, c, w$  flux, centre et largeur d'une raie d'émission.

$C$  continuum.

$D$  données bruitées contenant les raies spectrales.

$D_{cc}$  données bruitées, contenant les raies, le continuum et le spectre du ciel.

$e_k$  ( $e_k(c_k, w_k, \lambda)$ ) contribution 3D d'une raie convoluée par la PSF 3D.

$f_k$  ( $f_k(c_k, w_k, \lambda)$ ) raie de flux 1D convoluée par la LSF.

$F_{z,\mu}$  FSF d'une source située en  $(z, \mu)$ .

$F$  ensemble des FSF utilisées.

$L_{z,\mu}$  LSF d'une source située en  $(z, \mu)$ .

$L$  ensemble des LSF utilisées.

**MM** méthode des moments.

**ML** (*maximum likelihood*) maximum de vraisemblance (cas de l'estimation).

**JML** (*joint maximum likelihood*) maximum de vraisemblance conjoint (cas de la déconvolution).

$N = N_x \times N_y \times N_\lambda$  taille des données suivant les dimensions spectrale et spatiales.

$O$  observation non bruitée.

**PM** (*posteriori mean*) espérance *a posteriori* (cas de l'estimation).

**JPM** (*joint posteriori mean*) espérance *a posteriori* conjointe (cas de la déconvolution).

$r_a$  variance de la loi de probabilité *a priori* gaussienne de  $a$ .

$\mathcal{R}$  modèle de raie spectrale.

$\hat{R}$  indicateur de convergence de la chaîne de Markov simulée par MCMC : *potential scale reduction factor*.

$x, y, \mathbf{s}, z$  variables spatiales.

$\Gamma$  matrice diagonale de la variance du bruit :  $\Gamma = \text{diag}(\sigma^2(x, y, \lambda)) = \text{diag}(\sigma^2(\mathbf{s}, \lambda))$ .

$\Gamma_k$  matrice diagonale de la variance du bruit en un pixel spatial :  $\mathbf{s}_k \Gamma_k = \text{diag}(\sigma^2(\mathbf{s}_k, \lambda))$ .

$\theta_k$  vecteur des trois paramètres d'une raie d'émission à la position  $\mathbf{s}_k$  :  $\theta_k = [a_k, c_k, w_k]$ .

$\theta$  vecteur des paramètres de l'ensemble des raies pour toutes les positions  $\mathbf{s}$  :

$$\theta = [\theta_1, \dots, \theta_{N_{xy}}] = [a_1, c_1, w_1, a_2, \dots, w_{N_{xy}}].$$

$\lambda, \mu$  longueur d'onde.

$\lambda_1, \lambda_2$  limites de l'intervalle de la loi a priori de  $c$ .

$\sigma(x, y, \lambda)$  écart-type du bruit.

$\omega$  la largeur maximale a priori de la raie.

$\mathbf{1}_I$  fonction indicielle : vaut 1 sur l'intervalle  $I$ , 0 ailleurs.

$*_\lambda$  convolution spectrale.

$\star_{\mathbf{s}}$  convolution spatiale.

$\circ$  opérateur spatial / spectral non convolutif.

$\hat{\theta}$  estimateur de la variable  $\theta$ .

## Introduction

L'un des thèmes de recherche astrophysique du consortium MUSE est l'étude de la formation et de l'évolution des galaxies. Pour cela, le spectro-imageur MUSE est un formidable instrument permettant d'observer à la fois la distribution spatiale d'une galaxie proche ainsi que son spectre en différentes positions spatiales. Il permet d'étudier, grâce à l'effet Doppler, la cinématique des galaxies. Dans le cadre de l'observation astronomique au sol et donc de l'instrument MUSE, les données sont convoluées par une fonction d'étalement (PSF) à la fois spatiale et spectrale, ce qui dégrade la résolution que l'on peut obtenir sur cette étude cinématique.

Dans ce chapitre, nous nous concentrerons sur la déconvolution de cartes de cinématique de galaxies, dont une carte de distribution spatiale, une carte de vitesse relative de la galaxie et une carte de dispersion en vitesse le long de la ligne de visée. Notre objectif sera de développer une méthode d'estimation de ces cartes en prenant en compte la convolution 3D des données par la PSF 3D du système. En se plaçant autour d'une raie d'émission de la galaxie, ce problème se traduit, grâce à l'effet Doppler, en l'estimation de cartes de flux, de position et de largeur de raies spectrales.

En premier lieu, nous présenterons le modèle de données que nous avons adopté en fonction des raies spectrales, de la PSF de la chaîne d'acquisition et du bruit pour le mode grand champ de MUSE sans optique adaptative. Puis nous rappellerons les méthodes qui sont généralement utilisées par la communauté astrophysique, à savoir la méthode des moments et l'estimation par maximum de vraisemblance. Pour pallier les faiblesses de ces deux méthodes face au bruit et aux minima locaux, nous proposerons d'estimer les cartes par des simulations de Monte-Carlo par Chaînes de Markov (MCMC), dans un cadre bayésien. En particulier, nous utiliserons l'estimateur de l'espérance *a posteriori*. Par la suite, nous montrerons que les méthodes d'estimation par maximum de vraisemblance et par espérance *a posteriori* peuvent être étendues de manière à prendre en compte la convolution des données par la PSF spectrale et spatiale, ce qui permet d'améliorer la résolution spatiale des cartes estimées. Enfin, nous évaluerons les performances de l'ensemble de ces données à l'aide de simulations à une dimension spectrale et une dimension spatiale dans un premier temps, puis avec des simulations 3D réalistes.

## IV.1 Données

Dans cette première partie, nous présentons le modèle 3D de raies gaussiennes que nous avons adopté et son lien avec les cartes de cinématique. Nous précisons également les modèles de PSF spectrale et spatiale utilisés ainsi que la nature du bruit.

### IV.1.1 Modèle d'objet

On considère une raie d'émission, de longueur d'onde au repos  $\lambda_0$ , du spectre d'une galaxie. Cette raie peut être modélisée par une fonction de Dirac de flux  $a_0$  :

$$\mathcal{R}_0(\mathbf{s}, \lambda) = a_0(\mathbf{s}) \cdot \delta_{\lambda_0}(\lambda)$$

En raison de l'expansion de l'univers, la galaxie s'éloigne vis-à-vis de l'observateur. Ainsi, la longueur d'onde observée de la raie d'émission diffère de celle dite au repos  $\lambda_0$  (galaxie immobile) par un décalage vers le rouge (red-shift) dû à l'effet Doppler-Fizeau (cf. annexe A). Ce décalage est identique, quelle que soit l'origine spatiale de la raie d'émission à l'intérieur de la galaxie, parce qu'il est dû à l'éloignement global de la galaxie :

$$\mathcal{R}_1(\mathbf{s}, \lambda) = a_0(\mathbf{s}) \cdot \delta_{(\text{vitesse de la galaxie})}(\lambda)$$

Par ailleurs, la galaxie tourne sur elle-même, ce qui provoque le décalage local de la raie d'émission. Ce décalage est alors fonction de la vitesse relative locale de chaque source de la galaxie : les sources qui se rapprochent de l'observateur, relativement à l'éloignement global de la galaxie, émettent une raie observée décalée vers le bleu par rapport à celle de la galaxie en déplacement, alors que celles qui s'éloignent sont observées davantage vers le rouge :

$$\mathcal{R}_2(\mathbf{s}, \lambda) = a_0(\mathbf{s}) \cdot \delta_{(\text{vitesse relative de la source})}(\lambda)$$

Le signal perçu par un pixel du capteur CCD est la somme des signaux de toutes les sources de la galaxie dans l'angle solide correspondant sur la ligne de visée. Toutes ces sources ont une vitesse relative différente et produisent donc une raie observée à une longueur d'onde différente. L'intégration de toutes ces émissions ponctuelles, dénommée par simplicité raie du spectre et illustrée par la Fig. IV.1, peut être approximée [EPI 08] par une fonction gaussienne notée  $\ell$  :

$$\begin{aligned} \mathcal{R}(\mathbf{s}, \lambda) &= \frac{a(\mathbf{s})}{\sqrt{2\pi}w(\mathbf{s})} \cdot \exp\left(-\frac{1}{2} \frac{(\lambda - c(\mathbf{s}))^2}{w(\mathbf{s})^2}\right) \\ &= a(\mathbf{s}) \cdot \ell(c(\mathbf{s}), w(\mathbf{s}), \lambda) \end{aligned} \quad (\text{IV.1})$$

de flux  $a(\mathbf{s})$ , de position centrale  $c(\mathbf{s})$  liée à la longueur d'onde au repos  $\lambda_0$  et à la vitesse de la galaxie, et de largeur  $w(\mathbf{s})$  liée à la dispersion en vitesse des sources au sein de la galaxie, au pixel de coordonnées  $\mathbf{s} = (x, y)$ . Notons que cette raie n'est pas forcément symétrique, mais qu'en raison du pas d'échantillonnage spectral de MUSE, ce modèle est raisonnable. Dans le cas de l'étude d'un doublet de raies comme le doublet O<sub>II</sub>, il est possible de modéliser l'objet comme l'addition de deux raies identiques dont l'écart spectral peut être supposé connu.

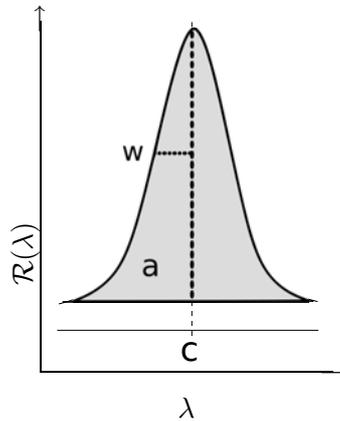


FIG. IV.1 – Raie gaussienne de flux  $a$ , centrée en  $c$ , d'écart-type  $w$ .

L'étude de la cinématique d'une galaxie requiert donc l'estimation de cartes de vitesse et de dispersion de vitesse illustrées par la Fig. IV.2, lesquelles sont directement liées au centre et à la largeur de la raie du spectre. Il s'agit alors d'estimer les cartes spatiales de  $a(\mathbf{s})$ ,  $c(\mathbf{s})$  et  $w(\mathbf{s})$  pour chaque position  $\mathbf{s}$ .

Le spectre d'une galaxie est composé d'un continuum, auquel s'ajoute des raies d'émission et d'absorption :

$$S(\mathbf{s}, \lambda) = \mathcal{R}(\mathbf{s}, \lambda) + C(\mathbf{s}, \lambda)$$

Ce continuum également appelé ligne de base, noté  $C(\mathbf{s}, \lambda)$ , est lentement variable spatialement et spectralement.

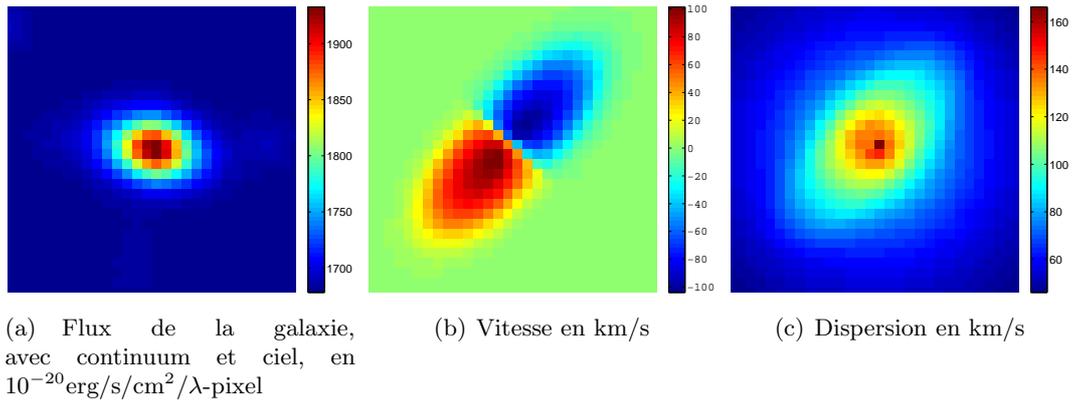


FIG. IV.2 – Exemple de cartes de flux, de vitesse et de dispersion de vitesse à la résolution MUSE.

### IV.1.2 Modèle de PSF

L'objet observé est étalé spatialement et spectralement par les effets de l'atmosphère, du télescope et de l'instrument. Dans le cas de la chaîne d'acquisition de données MUSE, rappelons que la PSF est tridimensionnelle et peut s'exprimer, sous réserve de certaines approximations, comme la composition d'une PSF spatiale (FSF) et d'une PSF spectrale (LSF) [VIL 11a].

Pour le problème visé ici,

- la LSF, notée  $L_{z,\mu}(\lambda)$ , varie lentement en longueur d'onde et plus significativement dans le champ de vue. Elle est considérée connue pour chaque pixel et chaque longueur d'onde sous une forme analytique. L'estimation de cette LSF peut être réalisée dans les phases de calibration de l'instrument [BAC 09] [CAR 09] [MAI 11]. Par simplicité, nous supposons la LSF gaussienne, mais les méthodes proposées pourront s'adapter sans difficulté à d'autres modèles de LSF. En outre, du fait que nous travaillons avec un intervalle spectral réduit autour d'une longueur d'onde, nous considérons que la LSF est constante par translation spectrale :  $L_{z,\mu}(\lambda) = L_z(\lambda - \mu)$ .
- la FSF, notée  $F_\mu(\mathbf{s} - \mathbf{z})$ , est constante dans le champ de vue dans le cadre du mode grand champ (WFM) sans correction. Elle varie lentement spectralement et peut donc être considérée constante autour d'une raie (cf. Chapitre II). Nous la supposons connue ici, estimée à partir des données d'une étoile isolée (cf. Chapitre III).

La PSF de la chaîne d'acquisition s'écrit donc, dans ce cas, sous la forme :

$$h_{z,\mu}(\mathbf{s}, \lambda) = L_z(\lambda - \mu) \cdot F_\mu(\mathbf{s} - \mathbf{z}) \quad (\text{IV.2})$$

Les formes, analytique ou numérique, de la LSF et de la FSF n'ont pas d'impact sur les méthodes que nous allons proposer, seul le coût de calcul peut s'en trouver modifié. L'essentiel, d'un point de vue théorique, est d'être capable de calculer la convolution ou composition tridimensionnelle des données par la PSF.

### IV.1.3 Modèle de données

En supposant que la PSF de la chaîne d'acquisition soit séparable, on peut écrire que le signal est composé par le produit de LSF par la FSF. Par linéarité de l'opération de composition, le spectre composé est la somme de la raie composée et du continuum composé.

$$\begin{aligned} \mathcal{S}_{LF}(\mathbf{s}, \lambda) &= \iint (\mathcal{R}(\mathbf{z}, \mu) + C(\mathbf{z}, \mu)) L_z(\lambda - \mu) F_\mu(\mathbf{s} - \mathbf{z}) d\mathbf{z} d\mu \\ &= \mathcal{R}_{LF}(\mathbf{s}, \lambda) + C_{LF}(\mathbf{s}, \lambda). \end{aligned}$$

Par la suite, nous supposons que le continuum composé est connu et parfaitement soustrait aux données ; nous n'étudions que la raie spectrale composée.

Dans ce chapitre, nous pouvons considérer que la FSF, variable en longueur d'onde, est constante sur l'intervalle spectral réduit de la raie étudiée :  $F_\mu(\mathbf{s} - \mathbf{z}) \simeq F_\lambda(\mathbf{s} - \mathbf{z})$ . Par conséquent, on peut écrire le modèle de raie convoluée comme le résultat de la convolution successive par la LSF (notée  $*$ ) puis par la FSF (notée  $\star$ ) :

$$\begin{aligned} \mathcal{R}_{LF}(\mathbf{s}, \lambda) &= \iint \mathcal{R}(\mathbf{z}, \mu) L_{\mathbf{z}}(\lambda - \mu) F_\lambda(\mathbf{s} - \mathbf{z}) d\mathbf{z} d\mu \\ &= \int \left( \int \mathcal{R}(\mathbf{z}, \mu) L_{\mathbf{z}}(\lambda - \mu) d\mu \right) F_\lambda(\mathbf{s} - \mathbf{z}) d\mathbf{z} \\ &= (\mathcal{R}(\mathbf{s}, \lambda) * L_{\mathbf{s}}) \star F_\lambda. \end{aligned}$$

Remarquons que les convolutions par la LSF et par la FSF ne sont pas commutatives. L'observation non bruitée est la somme du spectre de l'objet convolué et du spectre du ciel convolué :

$$O(\mathbf{s}, \lambda) = \mathcal{S}_{LF}(\mathbf{s}, \lambda) + \text{ciel}_{LF}(\mathbf{s}, \lambda).$$

Nous supposons que le spectre du ciel convolué est connu et parfaitement soustrait aux données, car son estimation fait l'objet de recherches au sein du projet DAHLIA et du consortium MUSE. Il intervient néanmoins dans la variance du bruit. En raison du temps d'intégration élevé (supérieur à 1h pour l'observation de galaxies), nous approximons, comme à la section III.1, le bruit d'acquisition par un bruit additif gaussien de variance l'observation idéale  $\sigma^2(\mathbf{s}, \lambda) = O(\mathbf{s}, \lambda)$ . Le cube bruité contenant les raies, le continuum et le spectre du ciel s'écrit alors :

$$D_{cc}(\mathbf{s}, \lambda) = O(\mathbf{s}, \lambda) + b(\mathbf{s}, \lambda) \text{ avec } b(\mathbf{s}, \lambda) \sim \mathcal{N}(0, O(\mathbf{s}, \lambda)).$$

Après soustraction du spectre du ciel et du continuum, les données sont modélisées par l'expression :

$$D(\mathbf{s}, \lambda) = \mathcal{R}_{LF}(\mathbf{s}, \lambda) + b(\mathbf{s}, \lambda) \text{ avec } b(\mathbf{s}, \lambda) \sim \mathcal{N}(0, O(\mathbf{s}, \lambda)). \quad (\text{IV.3})$$

Il est important de noter que la variance du bruit dépend non seulement du spectre de l'objet mais également du continuum et du spectre du ciel.

#### IV.1.4 Estimation et déconvolution

L'objectif de notre travail consiste donc à estimer les cartes de flux  $a(\mathbf{s})$ , de vitesse moyenne  $c(\mathbf{s})$  et de dispersion de vitesse  $w(\mathbf{s})$  à l'origine de la raie modélisée par l'équation (IV.2), à partir des données formées selon le modèle (IV.3).

Tout d'abord, nous évoquerons dans les sections IV.2 et IV.3, les méthodes que l'on qualifie d'estimation qui ne prennent pas en compte la convolution par la PSF 3D et qui s'appliquent donc indépendamment sur chaque pixel. Puis nous les distinguerons des méthodes dites de déconvolution, présentées dans la section IV.4, qui prennent en compte cette convolution 3D et estiment les cartes à partir de l'ensemble des pixels.

## IV.2 Méthodes usuelles d'estimation de cartes de cinématique

Nous rappelons deux méthodes classiquement utilisées pour l'estimation de cartes de cinématique : la méthode des moments et l'optimisation d'un critère quadratique que l'on peut interpréter comme une estimation par maximum de vraisemblance.

### IV.2.1 Méthode des moments

La méthode des moments est généralement utilisée par la communauté astrophysique (cf. [EPI 08]) pour estimer les caractéristiques d'une raie d'émission. Nous proposons de l'étendre au cas d'un doublet de raies, sous certaines hypothèses simplificatrices, afin de pouvoir étudier les données que le consortium nous a fournies (cf. section IV.5). Notons que cette méthode ne tient pas compte de la convolution spatiale par la FSF et que, de fait, elle est appliquée indépendamment sur chaque pixel spatial. En outre, elle ne prend pas en compte la variation du niveau de bruit sur chaque pixel.

#### IV.2.1.1 Méthode pour une raie d'émission

En premier lieu, le flux est estimé par la somme de l'intensité des pixels des données sans le continuum, suivant la dimension spectrale. Le centre et la largeur sont alors estimés par les moments spectraux du premier et second ordre [EPI 08] :

$$\hat{a} = \sum_{\lambda} D(\lambda) \quad \hat{c} = \frac{\sum_{\lambda} \lambda D(\lambda)}{\hat{a}} \quad \hat{w} = \sqrt{\frac{\sum_{\lambda} (\lambda - \hat{c})^2 D(\lambda)}{\hat{a}}}$$

Cette méthode ne suppose pas de forme particulière de la raie. Avec une telle méthode, la convolution spectrale des données par la LSF n'est pas prise en compte. Les quantités estimées ne correspondent donc pas aux caractéristiques de la raie  $\mathcal{R}(\lambda)$ . Néanmoins, si la LSF est symétrique, la position de la raie n'est pas décalée par la convolution spectrale, donc  $\hat{c}$  donne une estimation non biaisée de  $c$ . Par ailleurs, si la LSF et la raie sont toutes les deux de forme gaussienne, alors la raie peut être déconvoluée spectralement en soustrayant la variance de la LSF au carré de la largeur estimée  $\hat{w}$ .

L'inconvénient de cette méthode est d'être très sensible au bruit, en particulier pour l'estimation de la largeur, dont la qualité dépend directement de celle de l'estimation du flux et de la position.

#### IV.2.1.2 Extension de la méthode pour un doublet de raies

Pour certains éléments chimiques, les raies sont émises par paire, comme par exemple pour le doublet de  $O_{II}$ . Il est donc intéressant d'étendre la méthode des moments, initialement appliquée à une seule raie d'émission, au cas d'un doublet de raies. Le modèle d'objet est alors défini comme la somme de deux raies. En première approximation, on peut considérer que les deux raies sont identiques, de même flux  $a$ , de même largeur  $w$  et que la distance  $\Delta$  entre leur position respective est connue. Soit le modèle de doublet en une position  $s$  donnée :

$$\mathcal{R}_d(\lambda) = \mathcal{R}_1(\lambda) + \mathcal{R}_2(\lambda) \tag{IV.4}$$

avec  $a$  et  $w$  correspondant au flux et à la largeur des deux raies et  $c$  correspondant à la position du centre du doublet.

On peut alors relier les paramètres de chaque raie aux caractéristiques estimées à partir du doublet dans son ensemble. En effet, le moment d'ordre 0 donne le flux total du doublet  $a_d$ ; le flux d'une raie peut donc être estimé comme la moitié du flux total estimé :  $\hat{a} = \frac{\hat{a}_d}{2}$ .

Les raies étant symétriques, le moment d'ordre 1 estimé  $c_d$  donne un estimateur du barycentre des raies. Connaissant  $\Delta$ , il est aisé d'en déduire un estimateur de la position de chaque raie :  $\hat{c}_d - \frac{\Delta}{2}$  et  $\hat{c}_d + \frac{\Delta}{2}$ .

Pour la largeur, on peut montrer en raisonnant sur les données non bruitées en exploitant des pa-

paramètres  $a_d$  et  $c_d$  que :

$$\begin{aligned}
w_d^2 &= \frac{1}{a_d} \sum_{\lambda} (\lambda - c_d)^2 \mathcal{R}_d(\lambda) = \frac{1}{a_d} \sum_{\lambda} \lambda^2 \mathcal{R}_d(\lambda) - c_d^2 \\
&= \frac{1}{2a} \sum_{\lambda} \lambda^2 \mathcal{R}_1(\lambda) + \frac{1}{2a} \sum_{\lambda} \lambda^2 \mathcal{R}_2(\lambda) - \left( \frac{(c - \frac{\Delta}{2}) + (c + \frac{\Delta}{2})}{2} \right)^2 \\
&= \frac{1}{2} \left( \frac{1}{a} \sum_{\lambda} \lambda^2 \mathcal{R}_1(\lambda) - (c - \frac{\Delta}{2})^2 \right) + \frac{1}{2} \left( \frac{1}{a} \sum_{\lambda} \lambda^2 \mathcal{R}_2(\lambda) - (c + \frac{\Delta}{2})^2 \right) + \left( \frac{(c - \frac{\Delta}{2}) - (c + \frac{\Delta}{2})}{2} \right)^2 \\
&= \frac{1}{2} (w^2 + w^2) + \left( \frac{\Delta}{2} \right)^2 = w^2 + \left( \frac{\Delta}{2} \right)^2
\end{aligned}$$

On obtient donc l'expression d'estimateurs des trois paramètres de chaque raie à partir d'estimateurs calculés sur le doublet :

$$\begin{array}{c} \text{Flux} \\ \hat{a} = \frac{\hat{a}_d}{2} \end{array} \left| \begin{array}{c} \text{Positions} \\ \hat{c} = \hat{c}_d - \frac{\Delta}{2} \quad \text{et} \quad \hat{c} = \hat{c}_d + \frac{\Delta}{2} \end{array} \right| \begin{array}{c} \text{Largeur} \\ \hat{w} = \sqrt{\hat{w}_d^2 - \left( \frac{\Delta}{2} \right)^2} \end{array}$$

L'hypothèse selon laquelle les deux raies ont le même flux et la même largeur, peut être assouplie à condition que le rapport des flux et des largeurs soient connus (cf. annexe D.1).

#### IV.2.2 Estimation par maximum de vraisemblance

La seconde méthode utilisée pour l'estimation de cartes est l'ajustement d'un modèle aux données par minimisation d'une distance quadratique aux données, ce qui peut être interprété comme une estimation par maximum de vraisemblance (ML).

Notons que même si nous ne prenons pas en compte ici l'effet de la PSF spatiale et que nous traitons tous les pixels  $\mathbf{s}_k$  indépendamment, nous pouvons prendre en compte l'effet de la PSF spectrale en  $\mathbf{s}_k$  notée pour simplifier  $L_k(\lambda)$ . Nous pouvons alors chercher les paramètres  $\boldsymbol{\theta}_k = [a_k, c_k, w_k]$  tels que le modèle de raie correspondant convolué par la LSF, noté  $\mathcal{R}(\boldsymbol{\theta}_k) * L_k$ , colle le plus aux données.

Conformément au modèle de données détaillé à la section IV.1, le bruit est supposé additif gaussien centré, de variance  $\sigma^2(\mathbf{s}, \lambda)$  connue. La vraisemblance des données s'écrit alors comme la densité de probabilité des données conditionnellement aux paramètres :

$$p(\mathbf{D}_k | \boldsymbol{\theta}_k) = \frac{1}{\sqrt{2\pi}^{N_\lambda} \sqrt{\det \boldsymbol{\Gamma}_k}} \exp \left( -\frac{1}{2} \|\mathbf{D}_k - \mathcal{R}(\boldsymbol{\theta}_k) * L_k\|_{\boldsymbol{\Gamma}_k^{-1}}^2 \right)$$

avec :

- $N_\lambda$  la taille des données dans la dimension spectrale,
- $\mathbf{D}_k = \{\mathbf{D}(\mathbf{s}_k, \lambda_i)\}_{i=1 \dots N}$  vecteur de données bruitées à une position spatiale,
- $\boldsymbol{\Gamma}_k = \text{diag}\{\sigma^2(\mathbf{s}_k, \lambda_i)\}_i$ .

L'estimateur du maximum de vraisemblance est donc le vecteur de paramètres qui minimise l'anti-log-vraisemblance :

$$\hat{\boldsymbol{\theta}}_k = \arg \min_{\boldsymbol{\theta}_k} \|\mathbf{D}_k - \mathcal{R}(\boldsymbol{\theta}_k) * L_k\|_{\boldsymbol{\Gamma}_k^{-1}}^2.$$

L'avantage de cette méthode par rapport à la méthode des moments est qu'elle prend en compte la statistique du bruit et l'étalement spectral dû à la LSF, quelle que soit l'expression de cette dernière. Cependant, le critère qui doit être minimisé n'est pas convexe à cause de la non-linéarité du modèle vis-à-vis des paramètres  $c_k$  et  $w_k$ , et peut alors avoir des minima locaux. Donc, suivant l'algorithme de minimisation utilisé, cette méthode peut être très sensible à l'initialisation.

En pratique, nous avons initialisé l'algorithme par les résultats de la méthode des moments, puis nous avons réalisé une minimisation locale.

### IV.3 Estimation par espérance *a posteriori*

Afin de compenser la sensibilité au bruit ou à l'initialisation des deux méthodes précédentes, nous proposons d'estimer les cartes par des simulations de Monte-Carlo par Chaînes de Markov (MCMC), théoriquement insensibles à l'initialisation. Cette approche s'applique aisément dans un cadre bayésien. D'après le théorème de Bayes, la probabilité *a posteriori* du vecteur  $\boldsymbol{\theta}_k$  conditionnellement aux données est proportionnelle au produit de la vraisemblance par la loi *a priori* de ce même vecteur suivant l'expression :

$$p(\boldsymbol{\theta}_k | \mathbf{D}_k) = \frac{p(\mathbf{D}_k | \boldsymbol{\theta}_k) \cdot p(\boldsymbol{\theta}_k)}{p(\mathbf{D}_k)}$$

À partir d'une telle densité de probabilité, plusieurs estimateurs peuvent être proposés. Dans notre cas, l'estimateur du maximum *a posteriori* aura les mêmes problèmes que la vraisemblance en terme d'optima locaux. Aussi, nous avons choisi l'estimateur de l'espérance *a posteriori* (ou de la moyenne *a posteriori*) notée PM (pour *posterior mean*).

Cet estimateur ne peut pas être calculé directement d'après l'expression de cette loi *a posteriori* mais des méthodes de simulations de Monte-Carlo peuvent être utilisées. Ces méthodes permettent de fournir, en plus des paramètres estimés, des barres d'erreur en calculant l'écart-type *a posteriori*.

#### IV.3.1 Lois *a priori*

Le cadre bayésien de notre approche nécessite l'apport de lois *a priori* sur les paramètres. L'idée est d'apporter des informations simples relatives à chaque paramètre, tout en s'abstenant de privilégier certaines valeurs afin de ne pas biaiser l'estimation. Aussi, nous avons opté pour des lois *a priori* peu informatives, liées aux contraintes physiques du problème.

Nous avons supposé que les composantes du vecteur  $\boldsymbol{\theta}_k$ , c'est-à-dire le flux, la position et la largeur de la raie, sont des paramètres indépendants. La loi *a priori* du vecteur est donc le produit de la loi de chaque composante :

$$p(\boldsymbol{\theta}_k) = p(a_k)p(c_k)p(w_k).$$

La position approximative de la raie doit être détectée, visuellement ou automatiquement, afin de pouvoir extraire un sous-cube de données du cube de type MUSE autour de la raie. Il est donc possible de déterminer visuellement un intervalle  $[\lambda_1; \lambda_2]$  dans lequel se trouve le centre de la raie. Nous avons choisi d'associer au paramètre  $c$  une loi uniforme sur cet intervalle :  $c_k \sim \mathcal{U}_{[\lambda_1; \lambda_2]}$  soit une densité de probabilité :

$$p(c_k) = \frac{1}{\lambda_2 - \lambda_1} \mathbf{1}_{[\lambda_1; \lambda_2]}(c_k).$$

Par ailleurs, nous considérons que la largeur de la raie est obligatoirement positive et qu'il est possible de déterminer visuellement la largeur maximale  $\omega$  des raies à partir des données. Nous décrivons donc le paramètre  $w$  par une variable aléatoire suivant une loi uniforme sur un intervalle  $[0; \omega]$  :  $w_k \sim \mathcal{U}_{[0; \omega]}$  de densité de probabilité :

$$p(w_k) = \frac{1}{\omega} \mathbf{1}_{[0; \omega]}(w_k).$$

Enfin, le flux de la raie est naturellement positif. Nous considérons qu'il est possible de déterminer visuellement l'ordre de grandeur  $r_a$  du flux de la raie. Notons que nous pourrions, comme pour les autres paramètres, prendre pour  $a$  une loi *a priori* uniforme sur un intervalle. Mais la loi gaussienne positive sera plus simple à manipuler par la suite, car c'est une loi conjuguée à la vraisemblance pour

l'amplitude. La densité de probabilité s'exprime alors par :

$$p(a_k) = 2 \frac{1}{\sqrt{2\pi r_a}} \exp\left(\frac{-a_k^2}{2r_a}\right) \mathbf{1}_{[0;+\infty[}(a_k).$$

Notons que les paramètres de ces lois *a priori* sont considérés constants spatialement. On pourrait envisager de les faire varier pour les différents pixels, ce qui nécessiterait un réglage automatique délicat ou des informations *a priori* supplémentaires sur la galaxie.

### IV.3.2 Échantillonnage stochastique de la loi *a posteriori*

En multipliant la vraisemblance gaussienne par les lois *a priori* choisies, on obtient l'expression de la loi *a posteriori* :

$$\begin{aligned} p(\boldsymbol{\theta}_k | \mathbf{D}_k) &= \frac{1}{\sqrt{2\pi}^{N_k} \sqrt{\det \boldsymbol{\Gamma}_k}} \exp\left(\frac{1}{2} \|\mathbf{D}_k - \mathcal{R}(\boldsymbol{\theta}_k) * L_k\|_{\boldsymbol{\Gamma}_k^{-1}}^2\right) / p(\mathbf{D}_k) \\ &\cdot 2 \frac{1}{\sqrt{2\pi r_a}} \exp\left(\frac{-a_k^2}{2r_a}\right) \mathbf{1}_{[0;+\infty[}(a_k) \\ &\cdot \frac{1}{\lambda_2 - \lambda_1} \mathbf{1}_{[\lambda_1; \lambda_2]}(c_k) \\ &\cdot \frac{1}{\omega} \mathbf{1}_{[0; \omega]}(w_k). \end{aligned}$$

Le problème décisif réside dans le fait qu'il est impossible de générer directement des échantillons suivant la loi *a posteriori*. Il faut recourir à des méthodes de simulations de Monte-Carlo par chaînes de Markov (MCMC) pour y parvenir. Le principe des MCMC [ROB 96] consiste à générer des échantillons à partir d'une loi de proposition, pour produire une chaîne de Markov  $\{\boldsymbol{\theta}_k^{(i)}\}_{i=1\dots P}$ , dont la distribution asymptotique est la loi cible  $\pi(\boldsymbol{\theta}_k)$ .

Dans le cas de l'estimation des paramètres, la variable à échantillonner est le vecteur  $\boldsymbol{\theta}_k$  suivant la loi *a posteriori*  $\pi(\boldsymbol{\theta}_k) = p(\boldsymbol{\theta}_k | \mathbf{D}_k)$ . Nous avons choisi de recourir à la méthode de Gibbs qui échantillonne successivement les différentes composantes d'un vecteur suivant la loi cible conditionnellement aux autres paramètres, de telle sorte que la distribution du vecteur résultant tend asymptotiquement vers la loi cible. La méthode de Gibbs permet de réaliser l'échantillonnage d'un vecteur en effectuant l'échantillonnage de variables scalaires. Il est particulièrement adapté aux cas où les lois *a posteriori* conditionnelles sont simples à échantillonner. Nous exploiterons cette propriété pour le paramètre  $a$  puisque sa loi *a priori* est conjuguée à la vraisemblance.

Il est également envisageable de générer directement le vecteur  $\boldsymbol{\theta}_k$  entier à partir d'une loi de proposition vectorielle, en appliquant la méthode de Metropolis-Hastings. Cependant, il nous semble préférable d'exploiter la particularité du paramètre  $a$  et nous pensons que cela sera profitable pour la généralisation au cas de la déconvolution qui sera évoqué à la section IV.4.

### IV.3.3 Échantillonneur de Gibbs et lois conditionnelles

Rappelons que l'objectif est de générer des échantillons du vecteur  $\boldsymbol{\theta} = [a_k, c_k, w_k]$  suivant la loi *a posteriori*  $p(\boldsymbol{\theta}_k | \mathbf{D}_k)$ . Le principe de l'échantillonneur de Gibbs, dont l'algorithme est décrit par le Tab. IV.1, est de générer des échantillons de chaque composante du vecteur suivant leur loi conditionnelle [ROB 96].

Notons que, si la génération d'échantillons des différentes composantes du vecteur  $\boldsymbol{\theta}$  est réalisée dans l'ordre dans le Tab. IV.1, elle peut être réalisée dans un ordre aléatoire en tenant compte des composantes qui ont déjà été échantillonnées à l'itération courante.

Le recours à la méthode d'échantillonnage de Gibbs nécessite la connaissance des lois *a posteriori* conditionnelles des trois paramètres. Dans notre cas, ces lois s'expriment par :

Pour échantillonner le vecteur  $\boldsymbol{\theta}$  suivant la loi  $\pi(\boldsymbol{\theta})$  :

initialiser aléatoirement ou de façon déterministe la chaîne par  $\boldsymbol{\theta}^{(0)} = [\theta_1^{(0)}, \dots, \theta_J^{(0)}]$ ,

pour chaque itération  $n \geq 1$ , générer :

1.  $\theta_1^{(n)} \sim \pi(\cdot | \theta_{j \neq 1}^{(n-1)})$
2.  $\theta_2^{(n)} \sim \pi(\cdot | \theta_1^{(n)}, \theta_3^{(n-1)}, \dots, \theta_J^{(n-1)})$
- ...
- J.  $\theta_J^{(n)} \sim \pi(\cdot | \theta_1^{(n)}, \dots, \theta_{J-1}^{(n)})$

L'ensemble des vecteurs générés  $\{\boldsymbol{\theta}^{(n)}\}_{n=0:P}$  constitue une chaîne de Markov, dont la distribution tend asymptotiquement vers  $\pi(\boldsymbol{\theta})$ .

TAB. IV.1 – Échantillonneur de Gibbs.

**Flux** Notre modèle étant linéaire par rapport au paramètre  $a_k$  et les perturbations gaussiennes étant additives, on sait, du fait que la loi a priori sur  $a_k$  est gaussienne positive, que la loi a posteriori conditionnelle sur  $a_k$  l'est également (principe de lois conjuguées) :  $p(a_k | c_k, w_k, \mathbf{D}_k) \sim \mathcal{N}^+(\mu, \rho)$  de moyenne  $\mu = \rho \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{D}_k$  et de variance  $\rho = \frac{r_a}{1 + r_a \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{f}_k}$ , avec  $\mathbf{f}_k = \ell(c_k, w_k, \lambda) * L_k(\lambda)$ . En effet, en adaptant la démonstration de Vincent Mazet [MAZ 05], on peut montrer que :

$$p(a_k | c_k, w_k, \mathbf{D}_k) \propto \exp\left(-\frac{1}{2} \|\mathbf{D}_k - \mathcal{R}(\boldsymbol{\theta}_k) * L_k\|_{\boldsymbol{\Gamma}_k^{-1}}^2\right) \exp\left(\frac{-a_k^2}{2r_a}\right) \mathbf{1}_{[0;+\infty[}(a_k)$$

Or, on peut écrire :

$$\begin{aligned} \|\mathbf{D}_k - \mathcal{R}(\boldsymbol{\theta}_k) * L_k\|_{\boldsymbol{\Gamma}_k^{-1}}^2 &= (\mathbf{D}_k - a_k \mathbf{f}_k)^T \boldsymbol{\Gamma}_k^{-1} (\mathbf{D}_k - a_k \mathbf{f}_k) \\ &= \mathbf{D}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{D}_k - 2a_k \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{D}_k + a_k^2 \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{f}_k \end{aligned}$$

En incluant l'expression ci-dessus dans la loi conditionnelle, on obtient :

$$\begin{aligned} p(a_k | c_k, w_k, \mathbf{D}_k) &\propto \exp\left(-\frac{1}{2} (\mathbf{D}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{D}_k - 2a_k \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{D}_k + a_k^2 \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{f}_k) - \frac{a_k^2}{2r_a}\right) \mathbf{1}_{[0;+\infty[}(a_k) \\ &\propto \exp\left(-\frac{1}{2} a_k^2 \left(\frac{1}{r_a} + \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{f}_k\right) + a_k \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{D}_k\right) \mathbf{1}_{[0;+\infty[}(a_k) \\ &\propto \exp\left(-\frac{1}{2} \frac{a_k^2}{\rho} + \frac{a_k \mu}{\rho}\right) \mathbf{1}_{[0;+\infty[}(a_k) \propto \exp\left(-\frac{1}{2} \frac{(a_k - \mu)^2}{\rho}\right) \mathbf{1}_{[0;+\infty[}(a_k) \\ &\sim \mathcal{N}^+(\mu, \rho) \text{ avec } \rho = \frac{r_a}{1 + r_a \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{f}_k} \text{ et } \mu = \rho \mathbf{f}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{D}_k. \end{aligned}$$

Remarquons que des termes constants vis-à-vis du paramètre du flux  $a_k$ , tels que le terme  $\exp(-\frac{1}{2} \mathbf{D}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{D}_k)$  ont été supprimés ou ajoutés au terme exponentiel, puisque l'expression de la loi est donnée à un coefficient multiplicatif près.

La loi *a posteriori* conditionnelle de  $a_k$  peut donc être échantillonnée directement, en calculant sa moyenne et sa variance à partir des données et des autres échantillons.

C'est pour cette raison technique, souvent utilisée en échantillonnage stochastique [ROB 96], que nous avons choisi une loi a priori gaussienne positive pour  $a_k$ .

**Position**  $p(c_k | a_k, w_k, \mathbf{D}_k) \propto \exp\left(\frac{1}{2} \|\mathbf{D}_k - \mathcal{R}(\boldsymbol{\theta}_k) * L_k\|_{\boldsymbol{\Gamma}_k^{-1}}\right) \mathbf{1}_{[\lambda_1; \lambda_2]}(c_k)$ .

Étant donnée l'expression de la loi *a posteriori* conditionnelle de  $c_k$ , il est impossible de générer

directement des échantillons de la position à partir de cette loi.

La solution que nous avons adoptée consiste à générer des échantillons de  $c_k$  à partir de lois de propositions standards et à les accepter ou les rejeter selon la méthode de Metropolis-Hastings à marche aléatoire (cf. section III.7). Les échantillons ainsi produits sont distribués suivant la loi conditionnelle. Cette option nécessite de régler la variance de la loi normale que nous avons choisie pour loi de proposition, de manière à ajuster le taux d'acceptation à une valeur proche de 50% (cf. § IV.3.4).

**Largeur**  $p(w_k | a_k, c_k, \mathbf{D}_k) \propto \exp\left(\frac{1}{2}\|\mathbf{D}_k - \mathcal{R}(\boldsymbol{\theta}_k) * L_k\|_{\Gamma_k^{-1}}\right) \mathbf{1}_{[0;\omega]}(w_k)$ .

Pour les mêmes raisons que précédemment, nous échantillons le paramètre de largeur par un algorithme de Metropolis-Hastings à marche aléatoire (cf. § IV.3.4).

Enfin, nous écartons la première moitié des échantillons de la chaîne, pour ne garder que la seconde moitié afin de s'affranchir de l'initialisation. Cette proportion d'échantillons conservés est conseillée par [BRO 11], mais elle peut être inférieure pour s'assurer davantage de la convergence. L'estimation du vecteur de paramètres  $\boldsymbol{\theta}_k$  est ensuite obtenu en calculant la moyenne des échantillons de la chaîne de Markov  $\{\boldsymbol{\theta}_k^{(i)}\}_{i=1\dots P}$ . La variance empirique des échantillons donne alors un estimateur de la variance des paramètres estimés.

La méthode d'estimation des caractéristiques des raies spectrales que nous présentons a pour avantage de fournir des résultats asymptotiquement indépendants de l'initialisation et de leur associer des barres d'erreur grâce à l'échantillonnage par MCMC des variables.

Néanmoins, comme la méthode des moments et l'estimation par maximum de vraisemblance, elle ne prend pas en compte la convolution des données par la PSF du système et ne permet donc pas leur déconvolution.

#### IV.3.4 Choix des lois de proposition pour l'échantillonnage de Metropolis-Hastings

Les paramètres de position et de largeur sont échantillonnés par une étape de Metropolis-Hastings (cf. Tab. III.7) suivant leur loi conditionnelle *a posteriori*  $\pi(\theta_k^i) = p(\theta_k^i | \theta_k^{rest}, \mathbf{D}_k)$  avec  $\theta_k^i$  un des trois paramètres de  $\boldsymbol{\theta}_k = [a_k, c_k, w_k]$ .

Avant toute chose, il faut s'assurer que la loi de proposition permet d'explorer tout le support de la loi cible.

Les lois de proposition que nous avons choisies pour générer les paramètres de position et de largeur sont des lois normales. Chaque chaîne de Markov suit donc une marche aléatoire, dans le sens où la création d'un nouvel échantillon est un pas de direction aléatoire partant de l'échantillon précédent. Ces lois dépendent de paramètres, en l'occurrence de variances, dans le cas de lois normales, dont la valeur doit être fixée. Dans un souci de simplicité, chaque type de paramètre (position / largeur) est associé à une valeur unique de variance, ne variant pas en fonction de la position spatiale.

De cette valeur de la variance dépend l'efficacité de la génération des échantillons, à la fois en terme d'exploration de l'espace des paramètres et de coût de calcul. En effet, si la variance est trop grande, les candidats sont en majorité rejetés car peu probables, remplacés par le dernier échantillon accepté. La chaîne apparaît donc comme des marches d'escalier. À l'opposé, si la variance est trop petite, les échantillons sont tous acceptés. Dans ce cas, l'exploration de l'espace des paramètres est très lente et le nombre d'échantillons requis pour atteindre la convergence très important. Il s'agit donc de trouver un compromis entre l'efficacité de l'exploration de l'espace des paramètres et le coût de calcul.

Un premier indicateur de l'efficacité de la loi de proposition est le taux d'acceptation  $\tau$  qui est la proportion d'échantillons acceptés définie par [ROB 01] :

$$\tau = \int A(\theta, \Theta) \pi(\theta) Q(\theta, \Theta) d\theta d\Theta$$

avec  $\pi$  la loi cible,  $Q$  la loi de proposition et  $A(\theta, \Theta) = \min\left(1, \frac{\pi(\Theta)Q(\theta, \Theta)}{\pi(\theta)Q(\Theta, \theta)}\right)$  la probabilité d'acceptation utilisée par la méthode de Metropolis-Hastings. En pratique, il peut être facilement estimé par le ratio entre le nombre d'échantillons acceptés sur le nombre d'échantillons candidats.

Müller a proposé de calibrer automatiquement l'échelle de proposition afin que le taux d'acceptation tende vers 0.25 [MUL 91]. Gelman, Gilks et Roberts recommandent la valeur  $\tau = 0.25$  si la dimension du vecteur échantillonné est élevée, et  $\tau = 0.50$  si cette dimension vaut 1 ou 2 [GEL 96] [ROB 97]. Le problème du réglage de l'échelle de la loi de proposition d'un algorithme de Metropolis-Hastings a été largement étudié par Roberts et Rosenthal [ROB 01], qui explorent le cas du réglage de lois de propositions à plusieurs dimensions séparables. Ceux-ci expliquent que l'efficacité de l'échantillonnage par marche aléatoire est maximale pour un taux d'acceptation asymptotique de valeur  $\tau = 0.234$ , lorsque la dimension du problème tend vers l'infini. Ils ont également montré que la vitesse d'exploration de l'espace des paramètres est satisfaisante pour un taux d'acceptation compris entre 0.15 et 0.50, ce qui est beaucoup moins contraignant. Du fait que nous échantillonnons le vecteur  $\theta$  par la méthode de Gibbs, chaque étape de Metropolis-Hastings ne génère que des paramètres scalaires. Par conséquent, le taux d'acceptation recommandé par [GEL 96] et [ROB 97] est  $\tau = 0.50$  dans notre cas.

Un autre indicateur de l'efficacité de la loi de proposition est le temps inter-saut  $t_\sigma$  [ROB 96]. Si l'on sait que la loi cible possède plusieurs modes que l'on connaît, alors le temps inter-saut correspond au nombre d'itérations entre deux passages d'un voisin de  $\theta_i$  au voisinage de  $\theta_{i(i \neq j)}$ , soit entre deux modes de la distribution. Si le temps inter-saut est grand, alors l'exploration des différents modes est lente, ce qui signifie que la variance de la loi de proposition est trop petite. À l'inverse, si le temps inter-saut est petit, alors cela signifie que la chaîne saute d'un mode à un autre, mais que leur exploration n'est pas assez fine à cause d'une variance trop élevée. Dans le cas présent, nous ne connaissons pas le nombre ni la position des modes de la loi cible, nous ne pouvons donc pas ajuster l'échelle des lois de proposition par ce temps caractéristique.

Pour régler la variance de chaque loi de proposition, nous avons choisi d'utiliser uniquement le taux d'acceptation. En effet, il est l'indice le plus simple à calculer et le principal critère d'évaluation pour l'échantillonnage d'une variable de dimension 1.

## IV.4 Déconvolution de cartes de cinématique

Pour corriger le principal défaut des méthodes d'estimation précédentes, nous nous intéressons maintenant à la prise en compte de la PSF 3D dans l'estimation des cartes de cinématique, ce que nous appelons problème de déconvolution.

Nous proposons d'estimer conjointement tous les paramètres de flux, de position et de largeur (supposés indépendants) :  $\theta = [\theta_1, \dots, \theta_{Nxy}]$  en tenant compte non seulement de la convolution spectrale (notée  $*$ ) par la LSF  $L$  mais aussi de la convolution spatiale (notée  $\star$ ) par la FSF  $F$ . Pour cela, nous suggérons deux méthodes qui sont directement issues des méthodes précédentes, à savoir l'estimation par maximum de vraisemblance conjointe et par espérance *a posteriori* conjointe.

### IV.4.1 Maximum de vraisemblance conjointe

Le bruit sur chaque pixel du cube est indépendant, donc la vraisemblance du cube  $\mathbf{D}$  est le produit de la vraisemblance de chaque vecteur spectral  $\mathbf{D}_k$  conditionnellement au vecteur  $\theta$  et peut s'écrire sous la forme :

$$p(\mathbf{D}|\theta) = \frac{1}{\sqrt{2\pi}^N \sqrt{\det \mathbf{\Gamma}}} \exp\left(-\frac{1}{2} \|\mathbf{D} - \mathcal{R}(\theta) * L \star F\|_{\mathbf{\Gamma}^{-1}}^2\right)$$

avec :

- $N = N_{xy} \times N_\lambda$  la taille des données toutes dimensions confondues,
- $\mathbf{\Gamma}$  la matrice diagonale de variance du bruit  $\mathbf{\Gamma} = \text{diag}(\sigma^2(\mathbf{s}, \lambda))$ , de taille  $N \times N$
- $\mathbf{D}$  le vecteur de données bruitées de dimension  $N$ ,
- $\mathcal{R}(\boldsymbol{\theta})$  le modèle de raies 3D, dépendant des paramètres  $\boldsymbol{\theta} = \{\boldsymbol{\theta}_k\}_{k=1:N_{xy}}$ , vectorisé de taille  $N$ ,
- la notation  $\mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} \star \mathbf{F}$  symbolise les convolutions spectrales et spatiales de la raie 3D, le résultat étant rangé dans un vecteur de dimension  $N$ . L'estimateur du maximum de vraisemblance conjointe s'obtient alors par la minimisation de l'anti-log-vraisemblance :

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\mathbf{D} - \mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} \star \mathbf{F}\|_{\mathbf{\Gamma}^{-1}}^2.$$

L'opération de minimisation de ce critère est d'autant plus délicate que le nombre de paramètres par rapport auxquels le critère n'est pas linéaire a été multiplié par le nombre de pixels spatiaux  $N_{xy}$ . La minimisation par un algorithme d'optimisation standard, tel que la méthode de Nelder-Mead mise en œuvre dans la fonction *fminsearch* de Matlab que nous utilisons, a donc « de grandes chances » de fournir des paramètres correspondant à un minimum local.

#### IV.4.2 Échantillonnage stochastique de la loi *a posteriori*

Pour pallier ce défaut caractéristique de l'estimation par optimisation d'une fonction non convexe, nous souhaitons généraliser la méthode d'estimation par espérance *a posteriori* présentée à la section IV.3.

Nous supposons que tous les paramètres du vecteur  $\boldsymbol{\theta}$  sont indépendants, ce qui implique que la loi *a priori*  $p(\boldsymbol{\theta})$  s'exprime comme le produit des lois *a priori* de chaque paramètre que nous conservons inchangées par rapport à la section IV.3. De ce fait, et d'après le théorème de Bayes, la loi *a posteriori* a pour expression :

$$p(\boldsymbol{\theta}|\mathbf{D}) = \frac{p(\mathbf{D}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{D})} \text{ avec } p(\boldsymbol{\theta}) = \prod_{k=1}^{N_{xy}} p(a_k)p(c_k)p(w_k).$$

L'espérance *a posteriori* n'est pas calculable directement à partir de l'expression de cette loi. Pour l'estimer, nous proposons d'échantillonner le vecteur  $\boldsymbol{\theta}$  suivant la loi *a posteriori*  $p(\boldsymbol{\theta}|\mathbf{D})$  en étendant la méthode de Gibbs que nous avons présentée dans la même section IV.3 et qui nécessite de connaître les lois *a posteriori* conditionnelles.

#### IV.4.3 Lois *a posteriori* conditionnelles

Remarquons qu'en raison de l'indépendance *a priori* des paramètres, la loi *a posteriori* « conjointe » conditionnelle diffère de la loi *a posteriori* conditionnelle au pixel  $s_k$ , uniquement par l'expression de la vraisemblance.

Introduisons les notations qui vont être utiles pour présenter la méthode. Rappelons que  $\ell(c_k, w_k, \lambda)$  est le modèle de raie de flux unitaire en une position spatiale  $s_k$ . Le vecteur

$$\mathbf{f}_k = \ell(c_k, w_k, \lambda) * L_k(\lambda) \tag{IV.5}$$

de dimension  $N_\lambda$  représente le modèle de raie de flux unitaire à la position  $s_k$ , convolué spectralement par la LSF  $L_k$ ;  $\mathbf{f}_k$  dépend donc des paramètres  $c_k$  et  $w_k$  et chacun de ces éléments correspond à une longueur d'onde  $\lambda$ .

La contribution 3D de cette même raie convoluée à chaque longueur d'onde  $\lambda$  par la FSF  $\mathbf{F}_\lambda$ , c'est-à-dire une raie convoluée par la PSF 3D, est notée par le vecteur de dimension  $N$  :

$$\mathbf{e}_k = \mathbf{f}_k \delta(s_k) \star \mathbf{F}, \tag{IV.6}$$

cette contribution dépend des paramètres  $c_k$  et  $w_k$  et les éléments de ce vecteur correspondent aux différentes positions  $\mathbf{s}_k$  et longueurs d'onde  $\lambda$ . Alors, on peut écrire le modèle  $\mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} * \mathbf{F}$  comme la somme des contributions des raies pour chaque pixel :

$$\mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} * \mathbf{F} = \sum_{k=1}^{N_{xy}} a_k \mathbf{f}_k \delta(\mathbf{s}_k) * \mathbf{F} = \sum_{k=1}^{N_{xy}} a_k \mathbf{e}_k.$$

Nous souhaitons isoler la contribution d'une raie dans l'expression de la vraisemblance, ce qui nous permettra de marginaliser certains paramètres pour exprimer les lois conditionnelles. On peut pour cela écrire :

$$\mathbf{D} - \mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} * \mathbf{F} = \mathbf{D} - \sum_{k=1}^{N_{xy}} a_k \mathbf{e}_k = \mathbf{D} - \sum_{j \neq k} a_j \mathbf{e}_j - a_k \mathbf{e}_k = \mathbf{U}_k - a_k \mathbf{e}_k$$

où  $\mathbf{U}_k$ , appelé résidu, est le vecteur correspondant au cube de données auquel on a soustrait l'ensemble des contributions des raies pour la valeur actuelle des paramètres  $\boldsymbol{\theta}$ , sauf celle de la raie positionnée au pixel  $\mathbf{s}_k$ . Alors l'anti-log-vraisemblance des données  $\mathbf{D}$  s'écrit :

$$\|\mathbf{D} - \mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} * \mathbf{F}\|_{\Gamma^{-1}}^2 = \|\mathbf{U}_k - a_k \mathbf{e}_k\|_{\Gamma^{-1}}^2.$$

Nous pouvons remarquer que l'anti-log-vraisemblance a la même forme que dans le cas de l'estimation indépendante à chaque  $\mathbf{s}_k$  (cf. section IV.2.2), en remplaçant d'une part les données  $\mathbf{D}$  par le résidu  $\mathbf{U}_k$  et d'autre part le modèle de raie convoluée par la LSF  $\mathcal{R}(\boldsymbol{\theta}_k) * L_k$  par la contribution 3D de la raie convoluée  $a_k \mathbf{e}_k$ .

En exploitant ces notations, on peut aisément écrire l'expression des lois *a posteriori* conditionnelles de chaque variable connaissant l'ensemble des autres variables :

**Flux** La loi *a posteriori* conditionnelle du flux est une loi normale positive  $p(a_k | \boldsymbol{\theta}_{\text{rest}}, \mathbf{D}) \sim \mathcal{N}^+(\mu; \rho)$  de moyenne  $\mu = \rho \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{U}_k$  et de variance  $\rho = \frac{r_a}{1 + r_a \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{e}_k}$ .

En effet, on peut montrer que l'erreur aux données se décompose sous la forme suivante :

$$\begin{aligned} \|\mathbf{D} - \mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} * \mathbf{F}\|_{\Gamma^{-1}}^2 &= (\mathbf{U}_k - a_k \mathbf{e}_k)^T \boldsymbol{\Gamma}^{-1} (\mathbf{U}_k - a_k \mathbf{e}_k) \\ &= \mathbf{U}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{U}_k - 2a_k \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{U}_k + a_k^2 \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{e}_k \end{aligned}$$

Alors, par des calculs similaires à ceux effectués à la section IV.3.3, la loi *a posteriori* conditionnelle du flux s'exprime :

$$\begin{aligned} p(a_k | \boldsymbol{\theta}_{\text{rest}}, \mathbf{D}) &\propto \exp\left(-\frac{1}{2}(\mathbf{U}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{U}_k - 2a_k \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{U}_k + a_k^2 \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{e}_k) - \frac{a_k^2}{2r_a}\right) \mathbf{1}_{[0; +\infty[}(a_k) \\ &\propto \exp\left(-\frac{1}{2}a_k^2 \frac{1 + r_a \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{e}_k}{r_a} + a_k \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{U}_k\right) \mathbf{1}_{[0; +\infty[}(a_k) \\ &\propto \exp\left(-\frac{1}{2} \frac{a_k^2}{\rho} + \frac{a_k \mu}{\rho}\right) \mathbf{1}_{[0; +\infty[}(a_k) \\ &\sim \mathcal{N}^+(\mu, \rho) \end{aligned}$$

avec  $\rho = \frac{r_a}{1 + r_a \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{e}_k}$  et  $\mu = \rho \mathbf{e}_k^T \boldsymbol{\Gamma}^{-1} \mathbf{U}_k$ . Il est donc possible d'échantillonner directement le paramètre  $a_k$  suivant sa loi *a posteriori* conditionnelle.

**Position**

$$\begin{aligned}
p(c_k | \boldsymbol{\theta}_{\text{rest}}, \mathbf{D}) &\propto \exp\left(-\frac{1}{2}\|\mathbf{D} - \mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} \star \mathbf{F}\|_{\Gamma^{-1}}^2\right) \mathbf{1}_{[\lambda_1; \lambda_2]}(c_k) \\
&\propto \exp\left(-\frac{1}{2}\|\mathbf{U}_k - a_k \mathbf{e}_k\|_{\Gamma^{-1}}^2\right) \mathbf{1}_{[\lambda_1; \lambda_2]}(c_k)
\end{aligned}$$

**Largeur**

$$\begin{aligned}
p(w_k | \boldsymbol{\theta}_{\text{rest}}, \mathbf{D}) &\propto \exp\left(-\frac{1}{2}\|\mathbf{D} - \mathcal{R}(\boldsymbol{\theta}) * \mathbf{L} \star \mathbf{F}\|_{\Gamma^{-1}}^2\right) \mathbf{1}_{[0; \omega]}(w_k) \\
&\propto \exp\left(-\frac{1}{2}\|\mathbf{U}_k - a_k \mathbf{e}_k\|_{\Gamma^{-1}}^2\right) \mathbf{1}_{[0; \omega]}(w_k)
\end{aligned}$$

Ces deux dernières lois *a posteriori* conditionnelles dépendent respectivement des paramètres  $c_k$  et  $w_k$  par l'intermédiaire de  $\mathbf{e}_k$  (équations IV.5 et IV.6).

**IV.4.4 Déconvolution par espérance *a posteriori* conjointe**

En résumé, nous généralisons la méthode d'estimation des caractéristiques des raies spectrales présentée au paragraphe IV.3, en échantillonnant le vecteur  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_{Nxy}]$ , suivant la loi *a posteriori*  $p(\boldsymbol{\theta} | \mathbf{D})$ .

Pour cela, nous appliquons l'algorithme de l'échantillonneur de Gibbs (présenté § IV.1) : nous balayons dans un ordre aléatoire et échantillonnons les différentes composantes du vecteur  $\boldsymbol{\theta} = [a_1, c_1, w_1, \dots, a_{Nxy}, c_{Nxy}, w_{Nxy}]$ . Les variables  $a_k$  sont échantillonnées à partir de la loi normale positive  $\mathcal{N}^+(\mu; \rho)$ , de la même manière que dans le paragraphe IV.3.2. Pour les mêmes raisons que dans le cas de l'estimation par espérance *a posteriori*, nous échantillonnons les paramètres de position et de largeur par un algorithme de Metropolis-Hastings à marche aléatoire, dont nous réglons les variances de manière à obtenir un taux d'acceptation proche de 50%. Les variables  $c_k$  et  $w_k$  sont échantillonnées suivant leur loi conditionnelle par un algorithme de Metropolis-Hastings.

Ce balayage est effectué  $P$  fois afin de construire la chaîne de Markov  $\{\boldsymbol{\theta}^{(i)}\}_{i=1\dots P}$ . Enfin, la moyenne des échantillons donne une estimation du vecteur de paramètres  $\boldsymbol{\theta}$  ; la variance des échantillons fournit un estimateur de la variance des paramètres estimés, ce qui peut être interprété comme des barres d'erreur sur les paramètres estimés.

**IV.5 Expérimentations sur données 1D1D**

Dans un souci de simplicité et d'économie de temps de calcul, les méthodes présentées précédemment ont tout d'abord été testées sur des données simulées dites 1D1D, c'est-à-dire à une dimension spatiale et une dimension spectrale. Il s'agit donc d'une tranche de cube MUSE, non pas dans le sens instinctif à une longueur d'onde donnée (2D), mais à une position spatiale  $x$  donnée.

**IV.5.1 Description des données**

Les données sont simulées conformément au modèle présenté à la section IV.1, à la différence près qu'il s'agit d'un doublet de raies. Elles sont échantillonnées à la résolution MUSE, c'est-à-dire avec un pas spatial  $\Delta s = 0.2''$  et spectral  $\Delta \lambda = 0.13$  nm.

Dans le cadre de données 1D1D, les signaux de cinématique ne sont pas des cartes 2D mais des courbes 1D en fonction de  $y$ . Les courbes de flux, position et largeur que nous utilisons sont extraites de cartes de cinématique standard et sont affichées en rouge sur la Fig. IV.3.

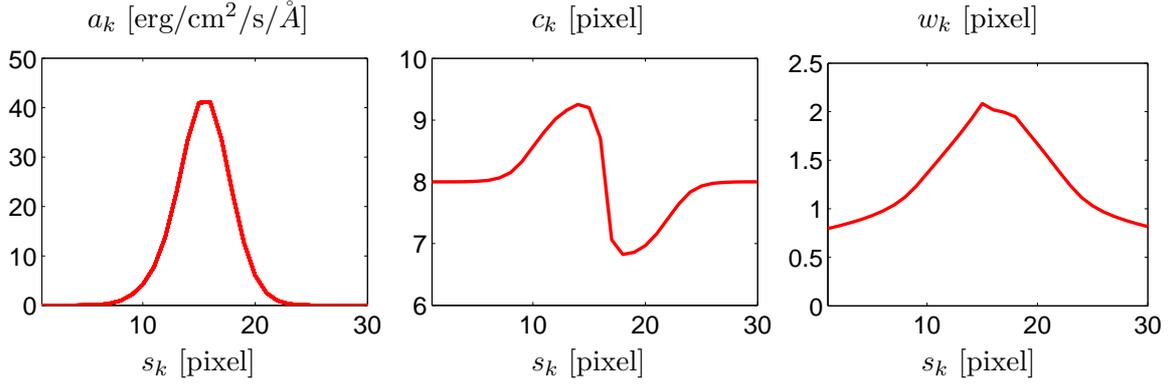


FIG. IV.3 – Courbes de flux, de position et de largeur des simulations de données 1D1D.

La PSF spectrale est de forme gaussienne, constante dans le champ et en longueur d'onde, donc d'écart-type constant. La PSF spatiale est également de forme gaussienne, d'écart-type constant dans le champ et variable en longueur d'onde. Notons que la PSF spatiale est ici 1D car les simulations que nous réalisons ont une seule dimension spatiale. Sa taille correspond à une observation dans des conditions de mauvais *seeing* de  $1.1''$ , afin de mettre en lumière l'utilité de la déconvolution. Nous supposons connaître les PSF spatiales et spectrales pour les opérations d'estimation et de déconvolution. Les données sont simulées à deux niveaux de bruit différents. Le premier possédant un flux équivalent à 80 heures d'intégration, équivalent à un rapport signal sur bruit (RSB) (défini p. 9) élevé de valeur  $\text{RSB} = 27.2$  dB, est utilisé essentiellement pour valider les méthodes étudiées. Le deuxième, correspondant à 1 heure d'intégration soit un  $\text{RSB} = 8.2$  dB, correspond à des données plus réalistes. Enfin, le continuum et le spectre du ciel sont supposés parfaitement soustraits. Le résultat de cette simulation est illustré par la Fig. IV.4.

### IV.5.2 Méthodologie

Afin de comparer les performances des différentes méthodes d'estimation (méthode des moments notée MM, maximum de vraisemblance ML et espérance *a posteriori* PM) et de déconvolution (maximum de vraisemblance conjointe JML et espérance *a posteriori* conjointe JPM), nous quantifions l'écart entre les valeurs théoriques des paramètres et leurs estimées. Dans le cas du vecteur de flux  $\mathbf{a} = [a_1, \dots, a_{N_y}]$  rassemblant les valeurs du flux aux différents pixels spatiaux, nous mesurons cet écart sur l'ensemble des pixels spatiaux par l'erreur quadratique moyenne :

$$\text{EQM} = \frac{1}{N_y} \sum_{k=1}^{N_y} (a_k - \hat{a}_k)^2.$$

Pour les paramètres de position  $\mathbf{c}$  et de largeur  $\mathbf{w}$ , cette EQM est calculée uniquement sur la zone centrale des données dans laquelle le flux est significatif. Cette zone correspond à l'intervalle spatial  $[8; 22]$  qui contient 99.6% du flux de la galaxie contenu dans l'intervalle original des données  $[1; 30]$ . Cette restriction permet de ne pas prendre en compte les valeurs aberrantes estimées dans les pixels de flux très faibles.

### IV.5.3 Indicateurs de convergence pour PM et JPM

Rappelons que nous souhaitons générer des échantillons du vecteur  $\boldsymbol{\theta}_k$  pour PM,  $\boldsymbol{\theta}$  pour JPM suivant la loi *a posteriori*. Pour cela, nous échantillonnons chaque paramètre du vecteur suivant sa loi *a posteriori* conditionnelle, suivant l'algorithme de Gibbs. Celui-ci fournit des valeurs du vecteur  $\boldsymbol{\theta}$  asymptotiquement distribuées suivant la loi *a posteriori*. Nous devons donc générer un nombre suf-

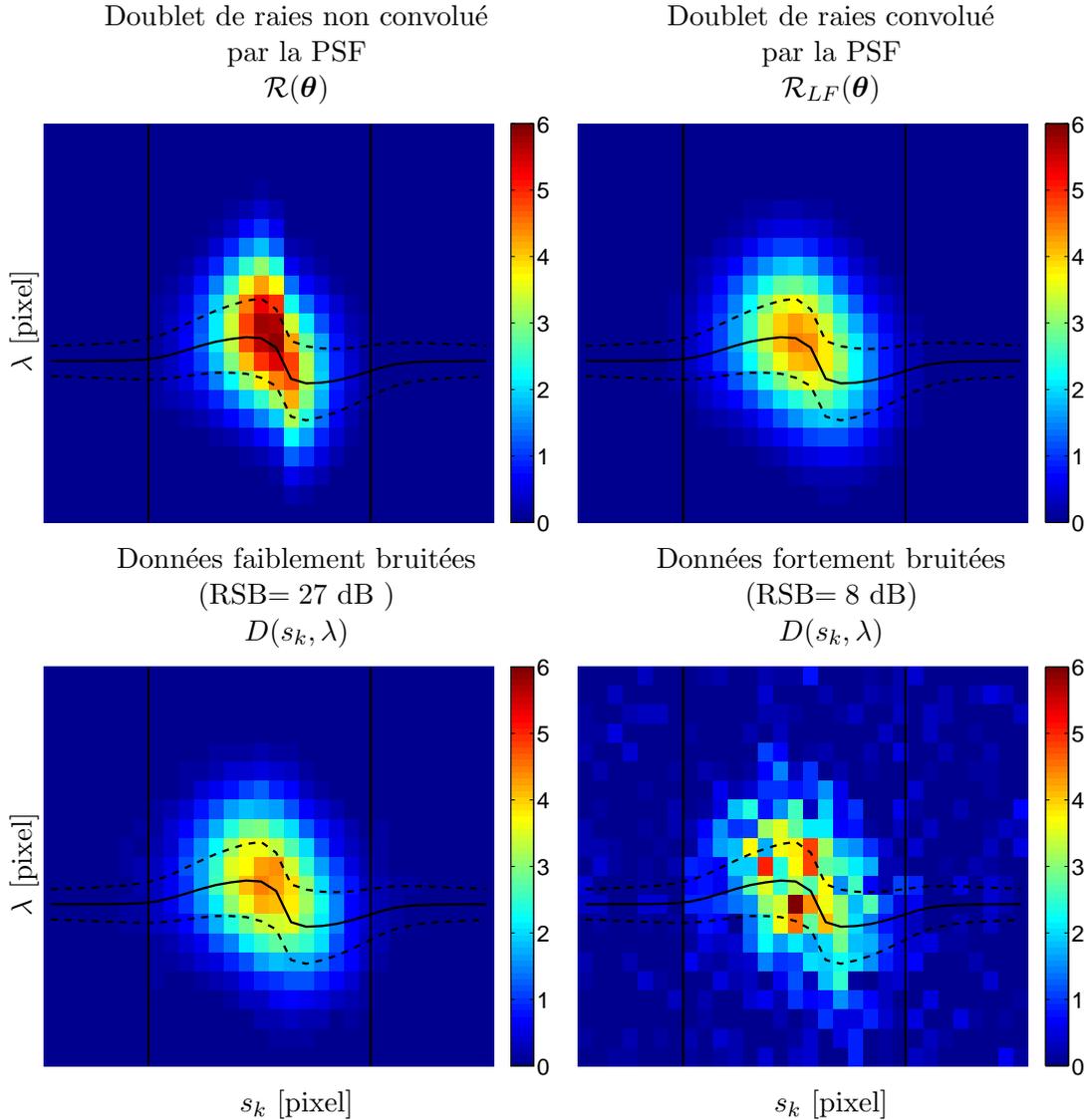


FIG. IV.4 – Données simulées 1D1D :  $21 \times 30$  pixels dans la dimension spectrale (verticale) et spatiale (horizontale). Unité de l'intensité de chaque pixel :  $10^{20}$  erg/cm<sup>2</sup>/s/Å. La zone délimitée par les barres noires correspond à l'intervalle spatial [8; 22] et contient 99.6% du flux de la galaxie calculé sur les données non convoluées. La courbe noire centrale [ligne] indique la position  $c_k$  théorique, les deux autres courbes [tirets] représentent la position théorique  $c_k \pm$  la largeur théorique  $w$ , en chaque position  $s_k$ .

fisant d'échantillons pour pouvoir considérer qu'ils sont distribués suivant la loi cible. Au final, seule la seconde moitié des échantillons de la chaîne est utilisée pour calculer l'espérance et la variance *a posteriori*, les autres sont écartés et constituent une période de préchauffage (*burn-in*).

Il existe plusieurs méthodes pour évaluer la convergence de la chaîne. Dans le cas de variables scalaires et de lois simples, la méthode la plus simple consiste à vérifier, soit par le tracé des échantillons, soit par leur histogramme, que la chaîne est bien distribuée suivant la loi cible.

Dans le cas de l'échantillonnage du vecteur  $\theta$ , il est préférable d'utiliser des tests plus formalisés, appelés diagnostics de convergence. Ainsi, une méthode consiste à générer plusieurs chaînes d'échantillons. En effet, chaque chaîne dépend de son initialisation et la comparaison des caractéristiques statistiques de chaque chaîne permet de vérifier leur convergence. Néanmoins, le nombre et la longueur des chaînes est sujet à débat et les avis sont contradictoires [GIL 96]. Ainsi, il est souvent suggéré de générer en parallèle plusieurs longues chaînes [GEL 92] [BRO 11] ayant chacune leur propre initialisa-

tion. Il est également conseillé de générer une très longue chaîne afin de trouver d'éventuels nouveaux modes [GEY 92].

Pour chaque expérimentation, nous avons généré trois chaînes indépendantes ayant chacune une initialisation aléatoire différente, selon les recommandations de Gelman et Shirley [BRO 11]. Nous avons ensuite calculé pour chaque paramètre  $a_k$ ,  $c_k$  et  $w_k$  le *potential scale reduction factor* [GEL 92] :

$$\widehat{R} = \sqrt{\frac{\sigma_{all}^2}{\sigma_c^2}}$$

où  $\sigma_{all}^2$  est la variance des échantillons des chaînes réunies et  $\sigma_c^2$  la moyenne des variances des échantillons de chaque chaîne. Dans l'idéal,  $\widehat{R}$  tend vers 1 lorsque les chaînes convergent ; en pratique, nous considérons que la convergence a été atteinte lorsque  $\widehat{R}$  passe en dessous de 1.1, d'après [BRO 11].

La Fig. IV.5 montre l'évolution de  $\widehat{R}$  pour les différents paramètres  $a_k$ ,  $c_k$  et  $w_k$  dans le cas de l'estimation par espérance *a posteriori* conjointe, en fonction de groupes de 1000 itérations. L'ensemble des courbes de  $\widehat{R}$  se situe en dessous du seuil après environ 40 000 itérations. Prévenus du fait que cette méthode a tendance à déclarer la convergence prématurément, nous avons choisi de générer 100 000 échantillons, dont 50 000 sont rejetés en guise de *burn-in*.

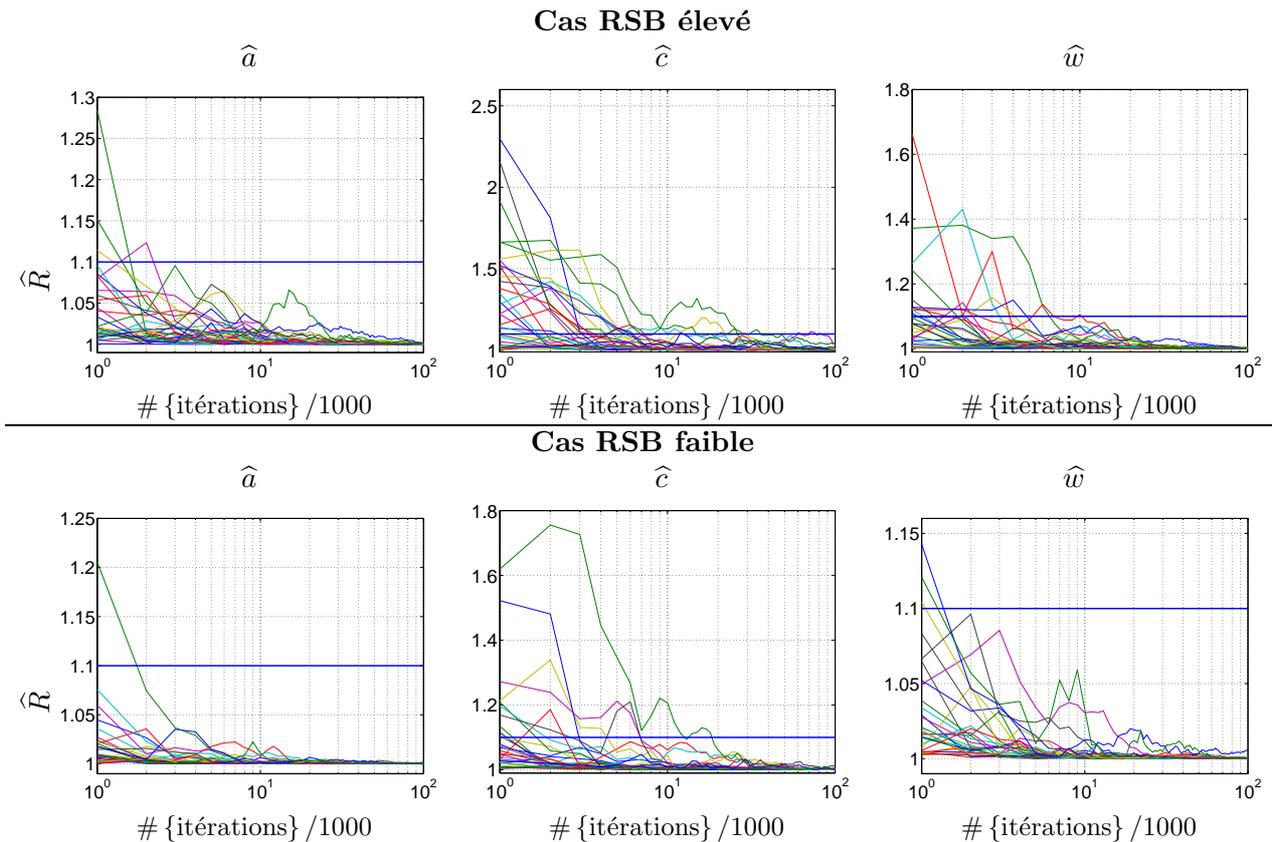


FIG. IV.5 – Indicateur de convergence  $\widehat{R}$  des 3 chaînes générées pour la déconvolution JPM de données 1D1D, en fonction du logarithme des groupes de 1000 échantillons, pour les 30 valeurs de  $x$ . La droite horizontale indique le seuil de 1.1 en dessous duquel on peut considérer, avec précaution, que la chaîne a convergé vers la distribution cible.

#### IV.5.4 Résultats d'estimation

L'estimateur du maximum de vraisemblance (ML) a été calculé par l'algorithme de minimisation du simplexe de Nelder-Mead [LAG 98], mis en œuvre dans la fonction *fminsearch* de Matlab, initialisé par les résultats de la méthode des moments (MM).

Dans le cas des données à RSB élevé, les estimateurs par MM, ML et PM donnent des résultats assez similaires, tel que le montrent les paramètres estimés sur la Fig. IV.6 et les indices d'erreur du Tab. IV.2. En particulier, la méthode des moments semble plus appropriée pour estimer le flux, avec une erreur quadratique moyenne de 2.46 contre 2.47 pour l'estimateur du maximum de vraisemblance et 2.49 pour l'espérance *a posteriori*. Néanmoins, il faudrait confirmer ce léger avantage sur plusieurs réalisations des données. En terme d'erreur quadratique, les résultats des estimateurs ML et PM sont identiques, ce qui est confirmé visuellement avec les deux premières lignes de la Fig. IV.6, du moins sur la zone de flux [8; 22].

Dans le cas des données plus réalistes dont le RSB est faible, les estimations effectuées par les trois méthodes sont significativement détériorées. En effet, les erreurs quadratiques moyennes obtenues pour l'estimation du flux augmentent de 40% minimum. Sur ce point, la méthode des moments et la méthode ML confirment leur robustesse pour l'estimation du flux. Concernant l'estimation de la position, la méthode PM semble moins sensible au bruit car l'erreur quadratique de la position est multipliée par 2 par rapport au cas moins bruité, contre une multiplication par 2.5 avec la méthode ML et 3 avec la méthode des moments. Enfin, l'estimation de la largeur ne semble pas détériorée par le bruit avec la méthode PM, alors que l'erreur quadratique obtenue avec la méthode MM (resp. ML) est multipliée par 6 (resp. 3).

Dans les deux cas, il apparaît clairement que les paramètres estimés (en noir) ne correspondent pas aux valeurs théoriques (en rouge). Le flux et la position de la raie semblent plus étalés qu'ils ne le sont en réalité, ce qui est logique parce que l'étalement spatial par la FSF n'a pas été pris en compte. Ainsi, les modèles de raie reconstitués à partir des cartes estimées par les méthodes ML et PM (dernière colonne de la Fig. IV.6) sont plus proches du modèle convolué  $\mathcal{R}_{LF}(\theta)$  de la Fig. IV.4 que du modèle non convolué  $\mathcal{R}(\theta)$ .

	Cas RSB élevé			Cas RSB faible		
	MM	ML	PM	MM	ML	PM
$\hat{a}$	2.46	2.47	2.49	3.48	3.53	3.79
$\hat{c}$	0.26	0.22	0.22	0.77	0.51	0.44
$\hat{w}$	0.14	0.11	0.11	0.84	0.32	0.087

TAB. IV.2 – **Résultats d'estimation** : Erreurs quadratiques moyennes par rapport aux valeurs théoriques.

#### IV.5.5 Résultats de déconvolution

L'algorithme d'optimisation que nous utilisons pour calculer l'estimateur JML pouvant être sensible à l'initialisation, nous utilisons deux jeux de valeurs initiales : les résultats de l'estimation ML, donc ne prenant pas en compte la FSF, et les vraies valeurs qui constituent un cas idéal de référence mais non exploitable en pratique. Les résultats correspondants sont notés respectivement JML et JML<sup>+</sup>.

La Fig. IV.7 et le Tab. IV.3 illustrent les résultats de déconvolution obtenus par les estimateurs JML, JML<sup>+</sup> et JPM, à la fois dans les cas peu et assez bruités. L'estimation du flux est clairement améliorée par les méthodes de déconvolution, ce qui se traduit par une réduction de l'erreur quadratique moyenne, passant de 2.5 environ pour les méthodes d'estimation à 1.93 par maximum de vraisemblance et 0.18 par espérance *a posteriori*. Toutefois, l'évolution des performances d'estimation de la position et de

la largeur diffère suivant la méthode employée.

D'une part, les erreurs quadratiques d'estimation de  $c$  et  $w$  des méthodes de maximum de vraisemblance conjointe (JML) sont nettement supérieures à celles par maximum de vraisemblance indépendant à chaque pixel (ML). Le phénomène d'alternance entre des valeurs hautes et basses de la position estimée  $\hat{c}_k$  pour des positions spatiales  $s_k$  voisines, y compris dans le cas où l'algorithme est initialisé par les vraies valeurs, est dû au fait que le modèle est ensuite convolué par la FSF, ce qui annule cet effet. Cette mauvaise estimation de la position de la raie affecte directement celle de sa largeur. Ces deux difficultés montrent la réelle complexité du problème de déconvolution.

Notons que les deux initialisations de l'estimateur JML donnent des résultats légèrement différents, ce qui était prévisible à cause de l'existence de minima locaux. Toutefois, même si l'estimateur JML est initialisé avec les vraies valeurs, il fournit également des résultats en dents de scie.

D'autre part, les performances pour l'estimation de la position et de la largeur sont meilleures par la méthode de déconvolution par espérance *a posteriori* conjointe que celle de la méthode d'estimation par PM indépendante à chaque pixel. Cela se vérifie visuellement sur la Fig. IV.7, car les valeurs estimées sont plus proches des valeurs théoriques que précédemment, ainsi que sur les erreurs quadratiques du Tab. IV.3. Nous pouvons le constater également car le modèle déconvolué  $\mathcal{R}(\hat{\theta})$  est plus proche du modèle de raies d'origine  $\mathcal{R}(\theta)$  que du modèle convolué  $\mathcal{R}_{LF}(\theta)$ , ce qui différencie l'approche de déconvolution de celle de l'estimation.

Pour conclure, les performances de l'estimateur de l'espérance *a posteriori* conjointe sont meilleures que celles de celui du maximum de vraisemblance conjointe. Le flux et la position sont bien estimés, même si la position est surestimée dans les zones où le flux est très faible, donc non significatif. Néanmoins, la largeur n'est pas correctement reconstruite dans le cas fortement bruité.

	Cas RSB élevé			Cas RSB faible		
	JML	JML <sup>+</sup>	JPM	JML	JML <sup>+</sup>	JPM
$\hat{a}$	1.93	0.38	0.18	7.92	5.32	2.15
$\hat{c}$	1.35	0.89	0.14	2.15	2.20	0.28
$\hat{w}$	0.33	0.40	0.052	0.90	1.48	0.18

TAB. IV.3 – **Résultats de déconvolution** : Erreurs quadratiques moyennes par rapport aux valeurs théoriques.

## IV.6 Expérimentations sur données 3D

Dans la section IV.5, nous avons appliqué les différentes méthodes d'estimation et de déconvolution sur des données 1D1D, dans un souci de simplicité et de gain en temps de calcul. Les données de type MUSE et plus généralement les données hyperspectrales pour l'observation des galaxies possèdent deux dimensions spatiales et une dimension spectrale.

Dans cette section, nous présentons des simulations de données 3D réalistes. Puis nous évaluons les performances des méthodes d'estimation et de déconvolution de ces données 3D, par maximum de vraisemblance (ML, JML) et par espérance *a posteriori* (PM, JPM).

### IV.6.1 Description des données

Nous avons généré des cubes de données à partir du modèle de raies développé à la section IV.1 et des cartes de flux, de position et de largeur, illustrées par la Fig. IV.2. Il s'agit ici de simples raies,

alors que les données 1D1D contenaient des doublets de raies, cela n'affectant en rien le nombre de variables à estimer. La PSF spatiale correspond à un *seeing* relativement mauvais de valeur  $s = 1.1''$ . Afin d'obtenir plusieurs jeux de données, nous y avons ajouté du bruit gaussien avec deux variances différentes. La première donne un jeu de données faiblement bruité de  $\text{RSB} = 18.1$  dB, correspondant à dix heures d'observation ; l'autre est plus réaliste, avec un  $\text{RSB} = 8.7$  dB, et correspond à une heure d'observation (cf. Fig IV.8).

### IV.6.2 Méthodologie

Nous comparons les performances des méthodes d'estimation par maximum de vraisemblance indépendantes à chaque pixel spatial (ML) et par espérance *a posteriori* indépendante (PM), ainsi que les méthodes de déconvolution par maximum de vraisemblance jointe (JML) et par espérance *a posteriori* jointe (JPM).

Concernant les méthodes d'estimation par maximum de vraisemblance (ML), l'algorithme d'optimisation a été initialisé par les résultats de la méthode des moments dans le cas indépendant (MM). Dans le cas de la déconvolution par maximum de vraisemblance, il a été initialisé par les résultats de l'estimation par la méthode ML (JML), puis par les vraies valeurs (JML<sup>+</sup>), et enfin par les valeurs obtenues par la méthode de déconvolution par JPM (JML<sup>\*</sup>) dans le cas faiblement bruité.

Concernant les méthodes d'estimation et de déconvolution par espérance *a posteriori* (PM et JPM), nous associons aux paramètres les mêmes lois *a priori* que pour l'étude des données 1D1D. Ainsi, la composante du flux suit une loi gaussienne positive de variance  $r_a = 100$  (PM, JPM<sub>1</sub>) et  $r_a = 1000$  (JPM<sub>2</sub>) ; le paramètre de position suit une loi uniforme sur l'intervalle  $[5; 15]$  que l'on définit visuellement à partir des données ; enfin, la composante de largeur suit une loi uniforme sur l'intervalle  $[0; 3]$ , intervalle qui peut être également mesuré à partir des données. En outre, nous avons choisi des lois de propositions gaussiennes d'écart-type  $\sigma_c = 1$  pour la position et  $\sigma_w = 1$  pour la largeur. Nous avons ajusté la valeur de ces paramètres de manière à ce que le taux d'acceptation des échantillons candidats soit proche de 50%. Le calcul des paramètres estimés a été effectué à partir de trois chaînes de  $P = 50\,000$  échantillons, après avoir écarté les  $P$  premiers échantillons de la chaîne.

Afin de quantifier les performances des cinq méthodes, nous quantifions l'écart entre les cartes théoriques de cinématiques et les cartes estimées par l'erreur quadratique moyenne (EQM) qui s'écrit dans le cas du paramètre de flux :

$$\text{EQM} = \frac{1}{N_{xy}} \sum_{k=1}^{N_{xy}} (a_k - \hat{a}_k)^2.$$

Pour les paramètres de position  $\mathbf{c}$  et de largeur  $\mathbf{w}$ , cette erreur quadratique moyenne est calculée uniquement sur la zone centrale des données dans laquelle le flux est significatif. Cette zone correspond à l'intervalle spatial  $[4; 22] \times [1; 23]$  qui contient 99.4% du flux de la galaxie contenu dans l'intervalle original des données  $[1; 24] \times [1; 30]$ . Cette restriction permet de ne pas prendre en compte les valeurs aberrantes estimées dans les pixels de flux très faibles. Notons que, dans le cas de l'estimation par maximum de vraisemblance, cette erreur quadratique moyenne est calculée sur les cartes seuillées aux limites des lois *a priori*.

### IV.6.3 Résultats d'estimation

De manière identique au cas 1D1D, l'estimation des cartes à partir des données 3D par la méthode ML a été réalisée en appliquant la méthode d'optimisation de Nelder-Mead [LAG 98], mise en œuvre dans la fonction *fminsearch* de Matlab. Rappelons que les valeurs d'initialisation sont les résultats de la méthode des moments appliquée indépendamment sur chaque pixel spatial. Les résultats d'estimation de ces deux méthodes sont illustrés par la Fig. IV.9, dont la qualité est mesurée en terme d'erreur

quadratique moyenne dans le Tab. IV.4.

	Cas RSB élevé		Cas RSB faible	
	ML	PM	ML	PM
$\hat{a}$	51.44	11.71	53.36	16.09
$\hat{c}$	5.39	0.84	8.40	1.82
$\hat{w}$	1.12	0.21	1.72	0.28

TAB. IV.4 – **Résultats d'estimation 3D** : Erreurs quadratiques moyennes.

Les cartes de flux estimées par la méthode PM sont sous-estimées, celles par ML le sont davantage. Cela est dû au fait qu'elles ne sont pas déconvoluées spatialement, de même que ne l'étaient pas les cartes 1D estimées de la Fig. IV.6. Si l'on fait abstraction du fait qu'elles sont déformées par la FSF, les cartes de position et de largeur estimées par la méthode PM sont relativement satisfaisantes, et ce malgré un bruit important. Les cartes de position estimées par ML sont acceptables, même si elles sont visiblement moins précises que celles de PM, d'après les EQM du Tab. IV.4. La carte de largeur est relativement bien estimée par ML à partir des données de RSB élevé, mais son estimation est très sensible au bruit, et paraît donc difficile avec les données fortement bruitées.

Notons que les valeurs estimées sont uniquement significatives dans la zone où le flux des raies est remarquable. Les valeurs estimées par PM dans les zones de faible flux sont proches du centre de l'intervalle de loi *a priori*, en raison du manque d'information. Nous pouvons observer que la zone spatiale dans laquelle l'estimation est réaliste est plus large avec la méthode PM qu'avec la méthode ML. En considérant que le RSB est localement inférieur dans les zones de faible flux, alors cela confirme le fait que l'estimation par PM est plus robuste au bruit que celle par ML.

En conclusion, l'estimation des cartes par espérance *a posteriori* est bien meilleure que celle par maximum de vraisemblance, actuellement utilisée par la communauté astrophysique.

#### IV.6.4 Résultats de déconvolution

La déconvolution de cartes a été réalisée par la méthode JPM avec deux valeurs de la variance de la loi *a priori* du flux, ainsi que par la méthode JML avec différentes initialisations. Les résultats de déconvolution 3D sont illustrés par les Fig. IV.10 et IV.11, leur performance est quantifiée par le Tab. IV.5.

	Cas RSB élevé					Cas RSB faible				
	JML	JML <sup>+</sup>	JML*	JPM <sub>1</sub>	JPM <sub>2</sub>	JML	JML <sup>+</sup>	JML*	JPM <sub>1</sub>	JPM <sub>2</sub>
$\hat{a}$	45.62	0.057	1.23	1.22	2.84	50.31	0.068	4.78	5.11	3.83
$\hat{c}$	5.08	0.021	1.83	1.41	1.42	8.10	0.038	2.98	1.80	1.84
$\hat{w}$	1.04	0.032	0.24	0.25	0.23	1.65	0.034	0.24	0.24	0.21

TAB. IV.5 – **Résultats de déconvolution 3D** : Erreurs quadratiques moyennes.

#### Déconvolution par maximum de vraisemblance jointe

Quel que soit le niveau de bruit, les cartes de position et de largeur déconvoluées par JML, avec initialisation réaliste par les résultats de ML, sont visuellement très bruitées. On retrouve en effet

le phénomène d’alternance spatiale entre des valeurs hautes et d’autres basses, qui avait été mis en évidence sur la Fig. IV.7 en 1D1D, et qui est annihilé par la convolution spatiale. Le gain en précision entre les résultats ML et les cartes déconvoluées JML est minime en terme d’EQM, puisque les valeurs correspondantes des Tab. IV.4 et Tab. IV.5 sont très proches.

En initialisant l’algorithme avec les vraies valeurs (JML<sup>+</sup>), on obtient des cartes extrêmement précises, ponctuées par quelques pixels parasites. Il est donc encourageant de pouvoir déconvoluer des données très bruitées. Néanmoins, cette initialisation et par conséquent son résultat sont inaccessibles en pratique. Ils mettent en évidence l’existence de minima locaux de la vraisemblance jointe. En l’initialisant par les résultats de JPM (JML<sup>\*</sup>), on bénéficie d’une amélioration notable de l’estimation des trois paramètres, comme en témoignent les erreurs quadratiques moyennes du Tab. IV.5. Ceci indique que l’échantillonnage stochastique permet une meilleure exploration de l’espace des paramètres et qu’il s’affranchit du problème des minima locaux. Toutefois, cette combinaison n’apporte pas d’amélioration par rapport à la déconvolution par JPM seule. Pour cette raison, nous n’avons pas illustré les résultats de (JML<sup>\*</sup>) dans la Fig.IV.10.

### Déconvolution par espérance *a posteriori* jointe

Avec la méthode JPM, les cartes de flux obtenues sont assez satisfaisantes, car elles permettent notamment de distinguer les deux régions de flux de la galaxie, qui sont fusionnées par la convolution spatiale. Nous pouvons observer à la fois visuellement et par les EQM l’influence de la valeur de la variance  $r_a$  de la loi *a priori* du flux  $a$  sur la carte de flux estimée. Avec la valeur la plus faible  $r_a = 100$  (JPM<sub>1</sub>), il subsiste une légère sous-estimation du flux, en particulier si le bruit est important. Avec la valeur plus élevée  $r_a = 1000$  (JPM<sub>2</sub>), certains pixels semblent au contraire sur-estimés et l’aspect arrondi de la galaxie est légèrement détérioré, notamment lorsque le niveau de bruit est faible. Il apparaît donc que la valeur de  $r_a$  doit être réglée suivant le niveau de bruit, avec une valeur d’autant plus grande que le niveau de bruit est élevé.

Concernant l’estimation de cartes déconvoluées de position et de largeur, les performances en terme d’erreur quadratique moyenne sont identiques pour les deux valeurs de  $r_a$  et comparables à celles obtenues sans déconvolution spatiale par la méthode PM. Cela pourrait donc laisser penser que l’amélioration apportée par la méthode JPM vis-à-vis de PM se limite à la déconvolution spatiale du flux. Mais on voit nettement sur les Fig. IV.10 et Fig. IV.11 que la résolution des cartes de position estimées par la méthode JPM est bien meilleure que celle des cartes de position estimées par PM sur la Fig. IV.9. En particulier, le passage des zones rouges aux zones bleues est plus rapide. Aussi, même si cela ne se traduit pas directement par une diminution de l’erreur quadratique moyenne, la déconvolution est efficace en terme d’amélioration de la résolution.

### Barres d’erreur

En calculant l’écart-type des échantillons générés par la méthode JPM à partir des données à RSB élevé, on obtient les cartes d’erreur illustrées par la Fig. IV.12. L’écart-type associé à la carte de flux est particulièrement élevé. Il n’est donc pas réaliste et ne fournit pas d’information.

Au contraire, les écarts-types de position et de largeur sont relativement faibles, variant autour de 2 dans la zone de fort flux pour une position située entre 6 et 10, et variant entre 0.3 et 0.6 pour une largeur située entre 1 et 3. Ils sont donc informatifs et peuvent être utiles pour l’étude de la cinématique de la galaxie.

### Statistiques des échantillons de la chaîne de Markov

La Fig. IV.13 affiche l’histogramme des  $P = 50\,000$  derniers échantillons des trois paramètres d’une chaîne obtenue par la méthode d’espérance *a posteriori* pour  $r_a = 100$ , pour deux pixels. Le premier a un flux élevé proche de 60, l’autre a un flux plus faible proche de 8. Il s’agit donc d’estimations de

leur loi *a posteriori* conditionnelle. La troisième ligne affiche également le tracé de ces échantillons pour le second pixel.

Nous pouvons observer que les histogrammes du pixel de fort flux ont une forme gaussienne. Nous pouvons donc dire que les échantillons oscillent autour d'une valeur moyenne qui est l'estimateur JPM. Les estimations du flux et de la position sont très proches de la valeur théorique. Dans le cas de la largeur, l'estimation est légèrement plus éloignée. Néanmoins, si l'on considère la largeur de la distribution, on peut estimer que la valeur théorique est proche de la valeur estimée.

Concernant l'estimation du flux et de la largeur du pixel de faible flux, on constate que la distribution des échantillons n'a pas de mode principal et qu'elle est étalée sur un large intervalle (sur l'intervalle de la loi *a priori* pour  $w$ ). D'après la 3<sup>ème</sup> ligne de la Fig. IV.13, cette large distribution n'est pas due à l'existence de plusieurs modes, mais à un continuum de variations. L'estimation par la moyenne des échantillons ne donne donc pas la valeur la plus probable de ces distributions, mais elle est finalement proche de la valeur théorique.

Concernant le paramètre de largeur, on remarque également que de nombreux échantillons de valeur proche de zéro ont été générés. Aussi, il pourrait être intéressant de se tourner vers des lois *a priori* pénalisant les valeurs proches de zéro telles que la loi inverse gamma.

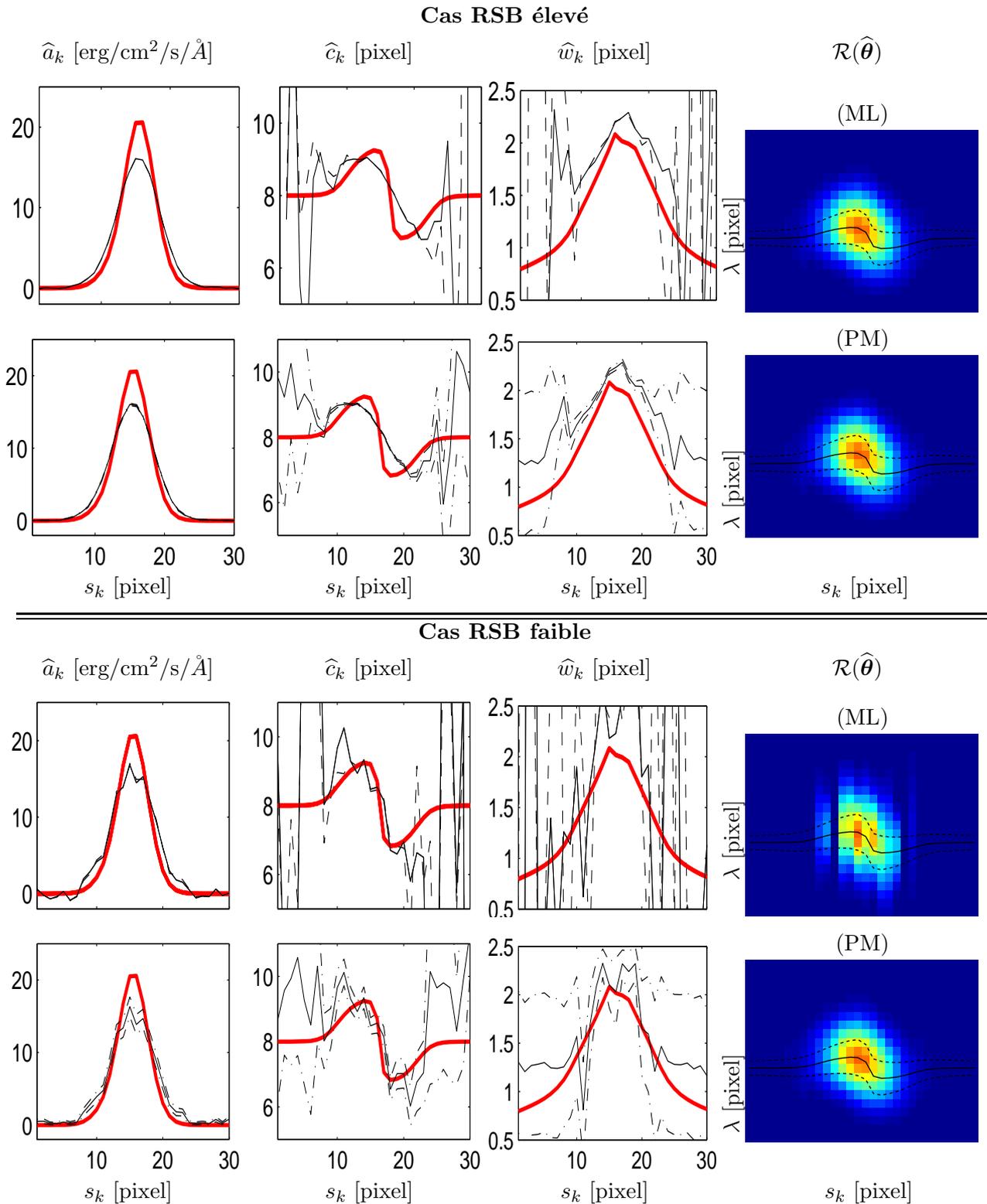
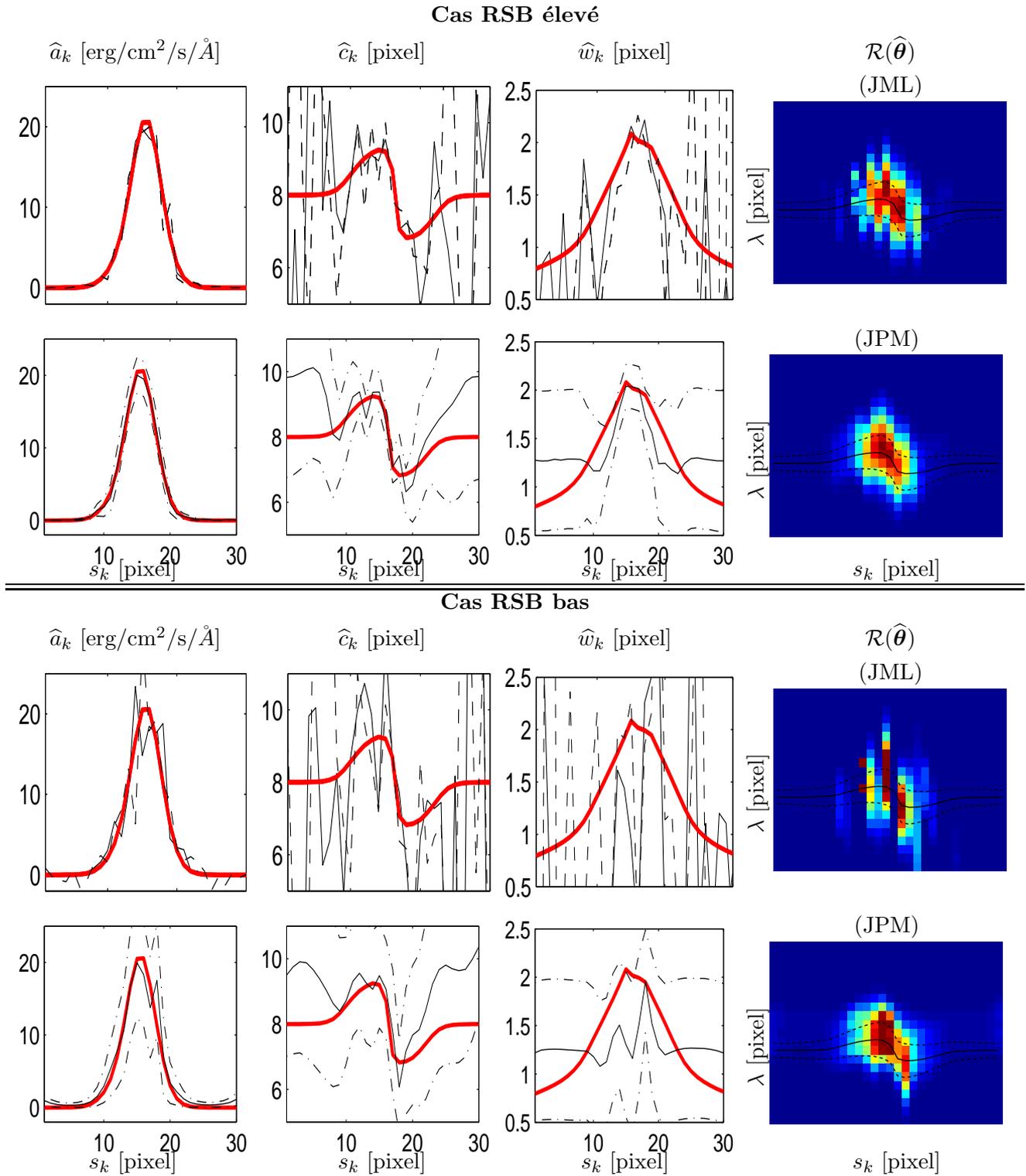


FIG. IV.6 – **Résultats d'estimation** : valeurs théoriques (ligne rouge épaisse) et estimées du flux, de la position et de la largeur en fonction de la position spatiale  $s_k$ . 1<sup>ère</sup> ligne : paramètres estimés par MM (- -) et par ML (-). 2<sup>nd</sup>e ligne : paramètres estimés par PM (-) +/- l'écart-type *a posteriori* (-·). Image de droite : modèle de raies 3D reconstitué pour les valeurs des courbes de cinématique estimées, auquel sont superposées la courbe de position théorique [ligne]  $\pm$  la largeur théorique [tirets].



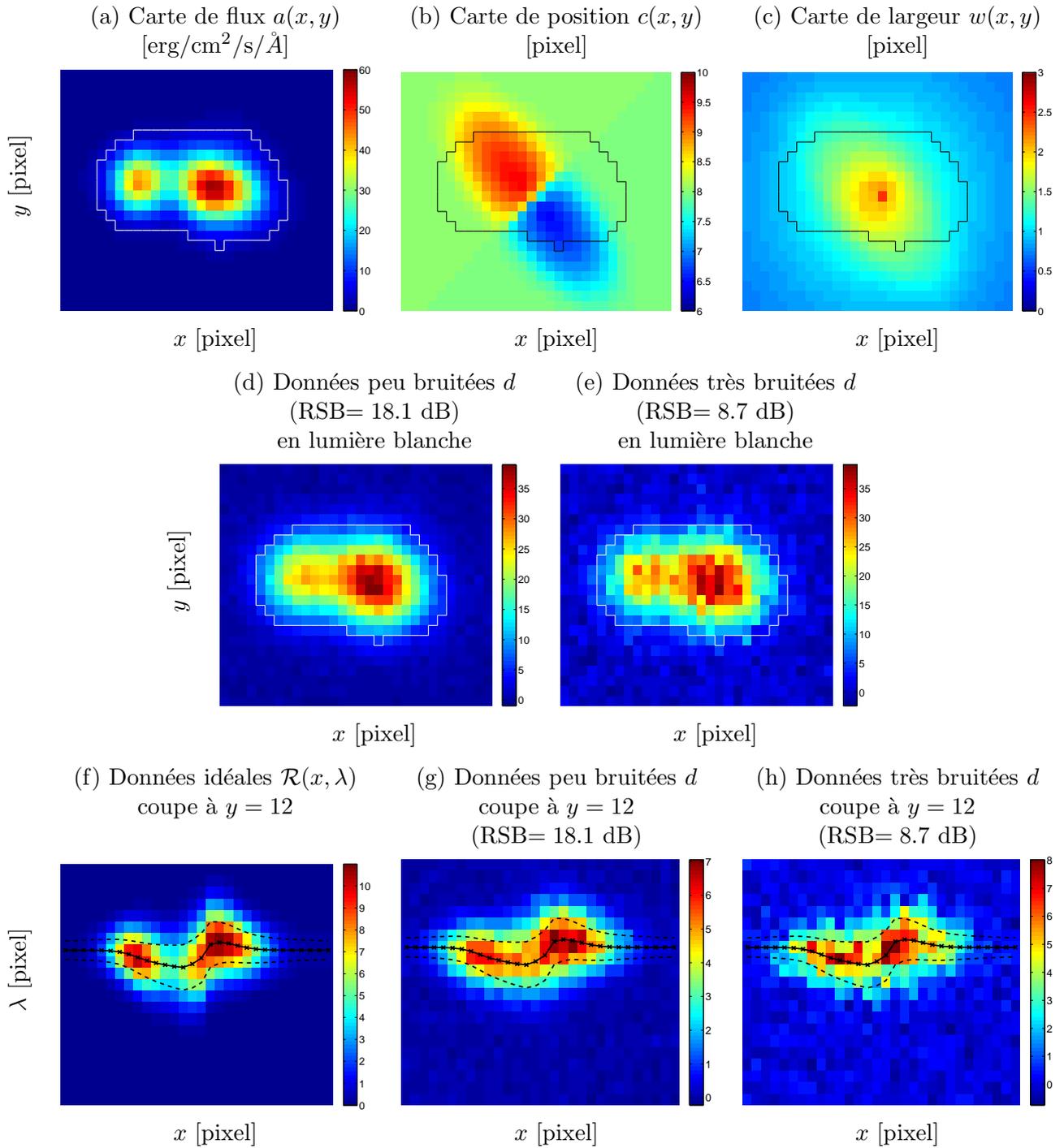


FIG. IV.8 – Données hyperspectrales 3D simulées pour l'estimation et la déconvolution 3D. Sur (a) (b) (c) (d) (e), la courbe noire délimite la zone contenant 95% du flux de la galaxie avant convolution. Sur (f) (g) (h), la courbe continue représente la position des raies, celles en tirets la position  $\pm$  la largeur.

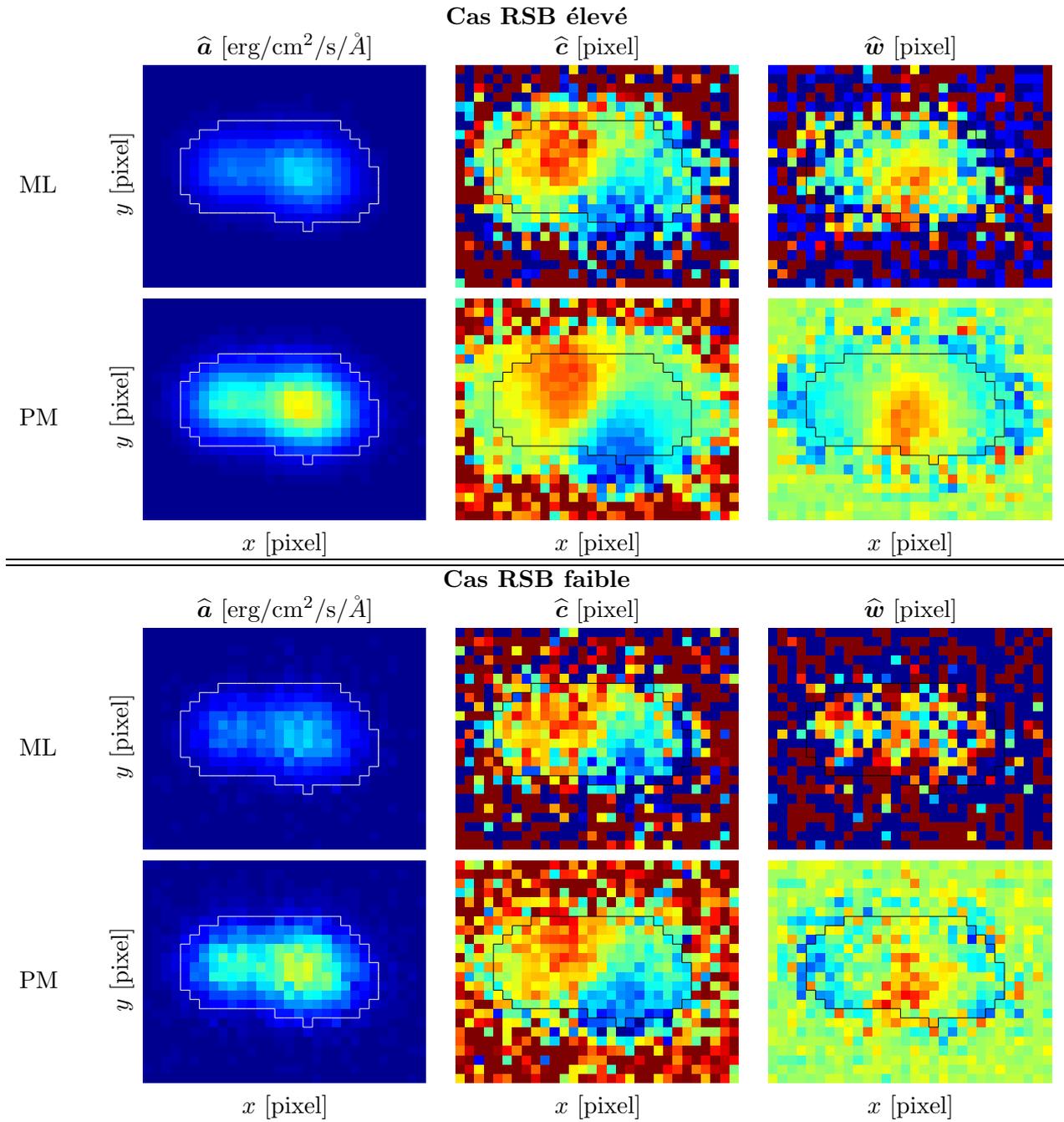


FIG. IV.9 – **Résultats d'estimation 3D** : Estimation de cartes sur des données 3D, à deux niveaux de bruit, pour les deux méthodes ML et PM. Les échelles de couleur utilisées sont celles des figures (a) (b) et (c) de la Fig. IV.8 ; les pixels en rouge (resp. en bleu foncé) ont une valeur égale ou supérieure (resp. inférieure) à la limite correspondante de l'échelle de couleur.

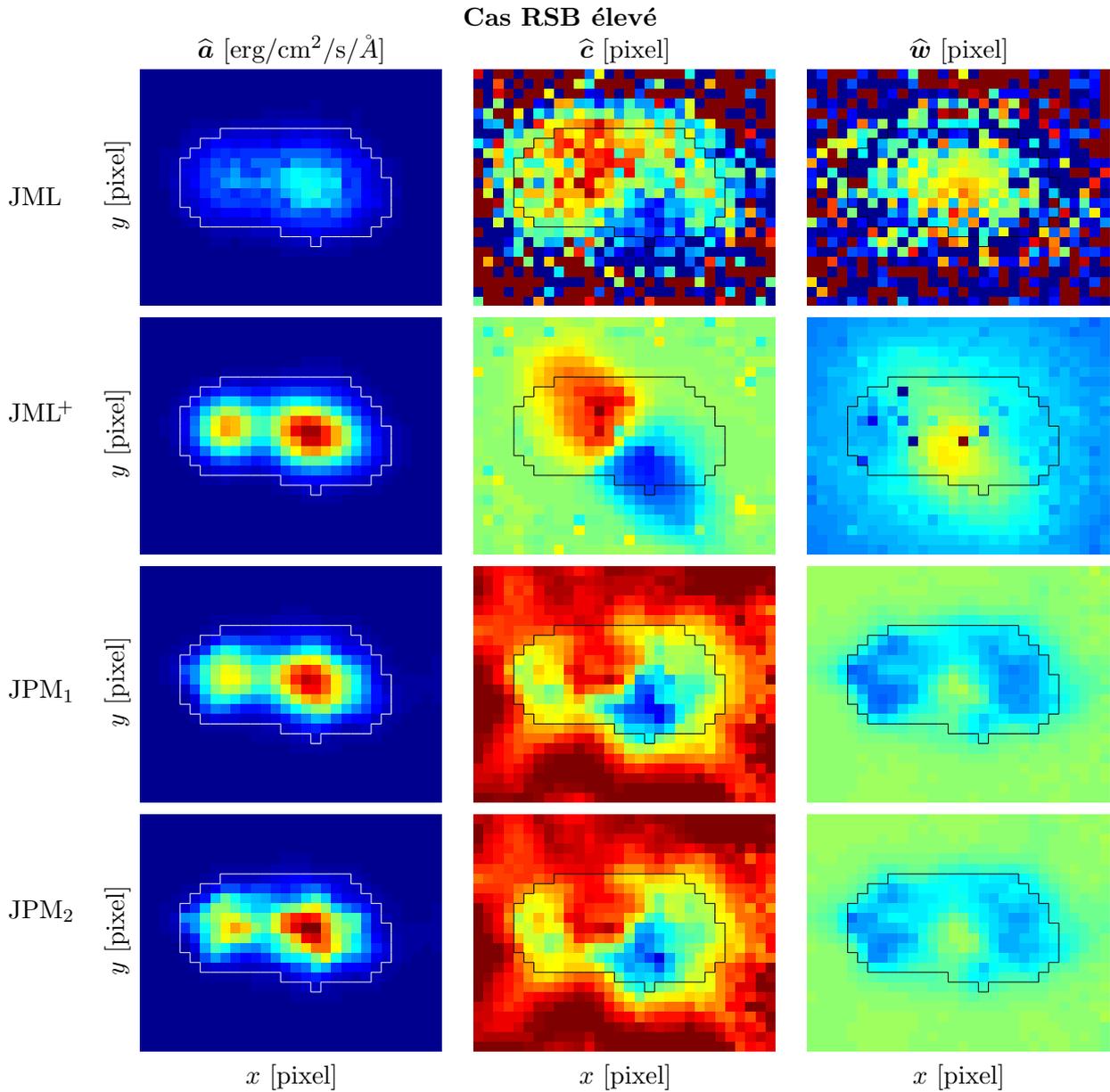


FIG. IV.10 – **Résultats de déconvolution 3D** : Déconvolution de cartes sur des données 3D, à un niveau de RSB élevé, pour les deux méthodes JML (initialisée par les résultats obtenus par la méthode ML et par les vraies valeurs pour JML<sup>+</sup>) et JPM (avec une valeur de la variance de la loi *a priori* du flux de  $r_a = 100$  pour JPM<sub>1</sub> et  $r_a = 1000$  pour JPM<sub>2</sub>).

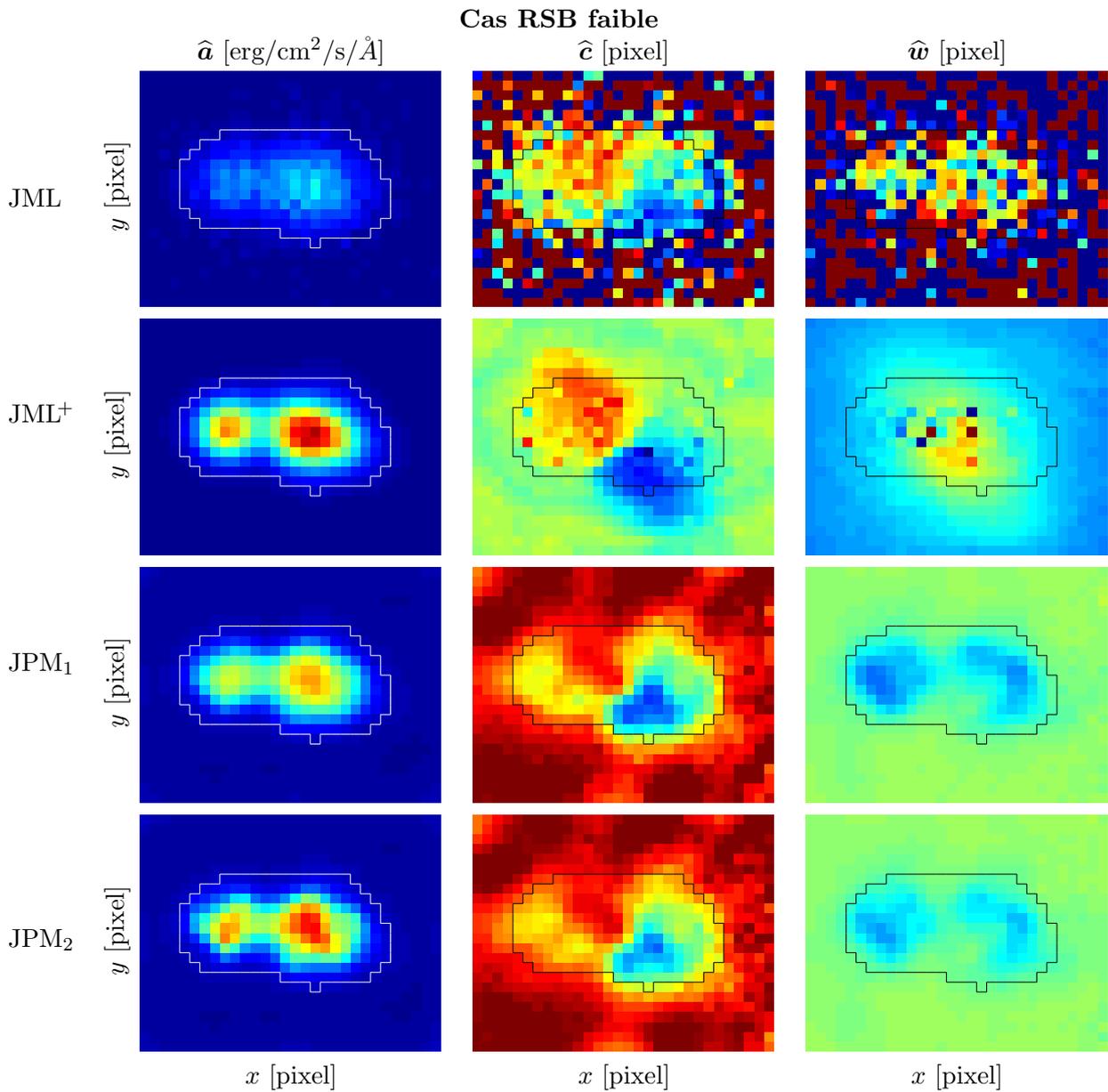


FIG. IV.11 – **Résultats de déconvolution 3D** : Déconvolution de cartes sur des données 3D, à un niveau de RSB faible, pour les deux méthodes JML et JPM (mêmes initialisations et paramètres que la Fig. IV.10).

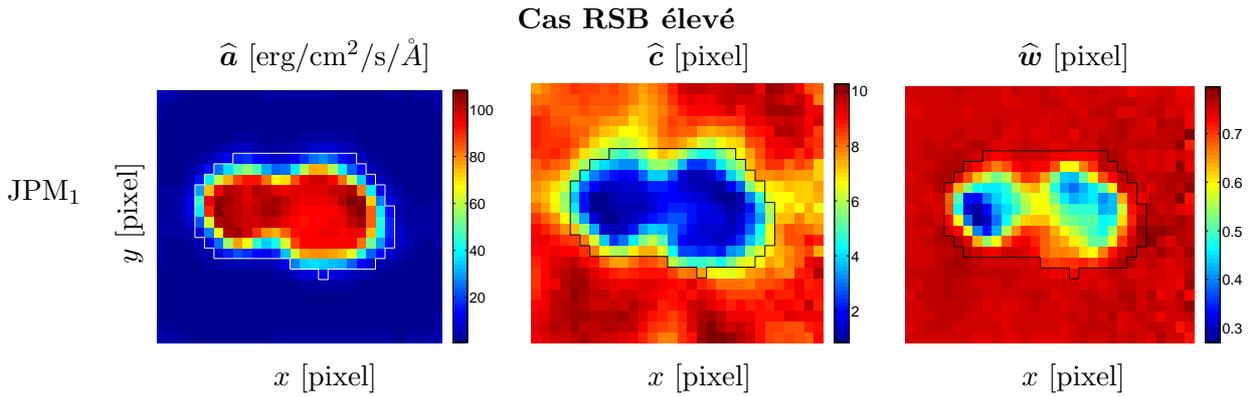


FIG. IV.12 – **Résultats de déconvolution 3D** : Écart-types obtenus par la méthode JPM dans le cas de RSB élevé.

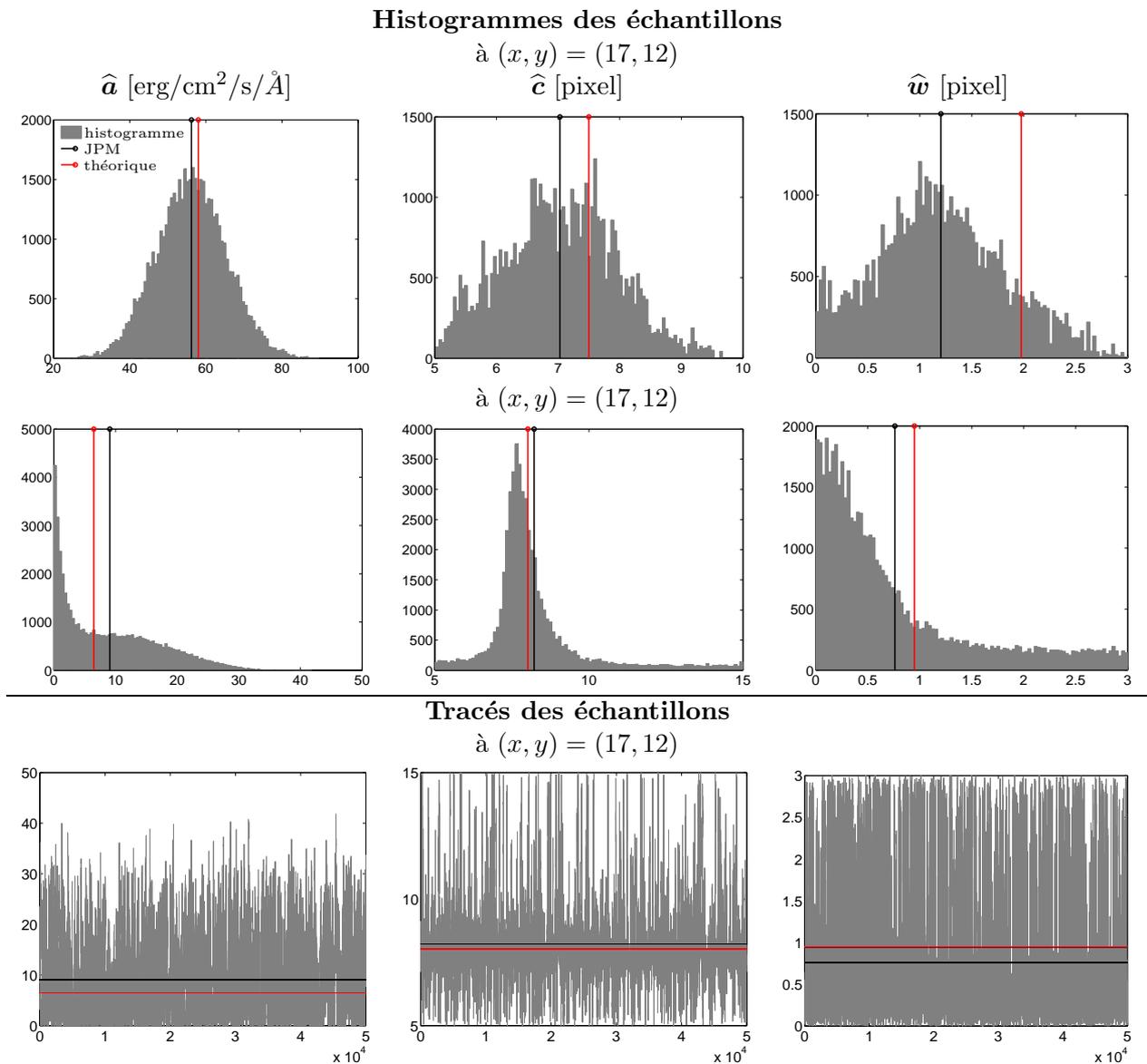


FIG. IV.13 – **Résultats de déconvolution 3D** : Histogrammes des échantillons d'une chaîne générée à partir des données à haut RSB et tracé des échantillons d'une chaîne. Barres rouges : vraies valeurs. Barres noires : moyenne des échantillons de la chaîne.

## Conclusions

Notre objectif était de développer de nouvelles méthodes d'estimation de cartes de flux, de vitesse et de dispersion de vitesse d'une galaxie en rotation, nécessaires à l'étude de sa cinématique.

Nous avons formalisé les données par un ensemble de raies gaussiennes convoluées spatialement et spectralement par la PSF du système, supposée connue. Le problème se traduit alors par l'estimation de cartes de flux, de position et de largeur de ces raies.

Notons que l'expression et la variabilité de la PSF n'ont pas d'influence sur les méthodes que nous proposons ici (ML, PM, JML, JPM), du moment que nous sommes capables de calculer la convolution ou composition des données par cette PSF. Elle peut donc être connue de façon analytique ou bien numérique suite à une calibration. Ainsi, dans le cadre de l'exploitation des données MUSE, la PSF pourra être connue analytiquement après avoir estimé les paramètres par la méthode présentée au Chapitre III, et la LSF sera disponible suite à une calibration instrumentale.

De même, la forme gaussienne des raies que nous avons estimées n'est pas nécessaire à l'application de ces méthodes. Nous pouvons utiliser n'importe quel modèle paramétrique formé d'une ou plusieurs raies. En effet, nous avons simulé et estimé les caractéristiques de doublets de raies à la section IV.5 et de raies simples à la section IV.6 sans distinction.

Les méthodes classiquement utilisées par la communauté astrophysique, sont la méthode des moments (MM) et l'estimateur du maximum de vraisemblance (ML), présentés dans la section IV.2, qui ne prennent pas en compte l'effet de convolution spatiale par la FSF.

Nous avons proposé d'estimer les cartes de flux, position et largeur par leur moyenne *a posteriori* (PM), en associant à chaque paramètre des lois *a priori* relativement peu informatives. Pour échantillonner la loi *a posteriori* du vecteur  $\theta_k$ , nous avons choisi l'échantillonneur de Gibbs, adapté au cas des vecteurs. Celui-ci permet de tirer profit de l'expression de la loi *a posteriori* conditionnelle de  $a$  que nous échantillonnons directement. Dans le cas des paramètres  $c$  et  $w$ , nous générons des candidats par une marche aléatoire, que nous sélectionnons par une étape de Metropolis-Hastings. L'estimation des cartes par échantillonnage stochastique de la loi *a posteriori* a l'avantage de fournir en outre des barres d'erreur, en calculant l'écart-type des échantillons.

Ces méthodes MM, ML et PM sont appliquées indépendamment sur le spectre de chaque pixel spatial et négligent la convolution spatiale par la FSF. La LSF peut néanmoins être prise en compte par l'estimation PM et ML et uniquement dans le cas particulier d'une LSF gaussienne pour la méthode des moments. À l'issue des expérimentations, nous pouvons conclure que la méthode des moments est relativement sensible au bruit, en particulier pour l'estimation de la position et de la largeur. En outre, l'estimation par maximum de vraisemblance est sensible à l'initialisation à cause de l'existence de minima locaux du critère. Enfin, la méthode d'estimation par espérance *a posteriori* a l'avantage d'être robuste au bruit, théoriquement indépendante de l'initialisation et de fournir des barres d'erreur.

Afin de prendre en compte l'effet de la convolution spatiale, nous avons proposé d'estimer conjointement l'ensemble des paramètres formant le vecteur  $\theta$  pour tous les pixels spatiaux. Nous avons donc étendu les méthodes précédentes par l'estimateur du maximum de vraisemblance conjointe (JML) et la moyenne *a posteriori* conjointe (JPM).

Cette méthode que nous avons proposée apporte indéniablement des améliorations par rapport aux méthodes classiques. Les cartes estimées par la méthode JML initialisée par ML font apparaître un phénomène d'alternance qui les rend physiquement non acceptables. L'initialisation de JML par les vraies valeurs montre cependant que cet estimateur peut fournir de très bons résultats mais que des minima locaux du critère rendent le problème d'optimisation très délicat. Les cartes de flux et de position estimées par l'espérance *a posteriori* conjointe sont relativement satisfaisantes, la carte de largeur reste cependant légèrement sous-estimée. Nous avons pu constater une amélioration de la résolution spatiale sur les cartes de position et de largeur estimées, ce qui atteste de l'efficacité de l'opération de déconvolution.



# Conclusion générale

## Principales contributions

Notre objectif était d'aborder la problématique de la déconvolution de données hyperspectrales, pour l'analyse des données de l'instrument MUSE. Pour cela, il nous a été nécessaire d'étudier au préalable la modélisation et l'estimation de la fonction d'étalement (PSF) de la chaîne d'acquisition.

Pour commencer au Chapitre II, nous nous sommes concentrés sur l'échelle, la forme et la variabilité de la PSF spatiale (FSF) en fonction de la position dans le champ et de la longueur d'onde. Pour le mode de fonctionnement grand champ sans correction par optique adaptative (OA), nous avons proposé plusieurs modèles de la FSF, principalement due à la turbulence atmosphérique et au télescope. Ceux-ci s'appuient sur la fonction de Moffat circulaire, dont nous modélisons la variation des paramètres en fonction de la longueur d'onde, mais également en fonction de paramètres atmosphériques.

Dans le cadre du mode grand champ corrigé par OA, la PSF spatiale varie également en fonction de la position de la source dans le champ. Nous avons approximé cette FSF par une fonction de Moffat elliptique, dont les paramètres varient à la fois en longueur d'onde et en fonction de coordonnées polaires dans le champ. Ainsi, nous avons montré que la FSF peut être décrite en tout point d'un cube de données hyperspectrales à l'aide d'un faible nombre de paramètres, avec des performances satisfaisantes.

Par la suite au Chapitre III, nous avons abordé le problème de l'estimation de la PSF spatiale. En effet, la connaissance de la FSF à toutes les longueurs d'onde est un enjeu crucial, non seulement pour évaluer la résolution des données, mais également pour appliquer plusieurs types de traitements, tels que la fusion de cubes de données, la séparation de sources et la déconvolution.

La FSF est la réponse impulsionnelle spatiale du système, elle peut donc être estimée à chaque longueur d'onde à partir de l'observation d'une source ponctuelle, telle une étoile isolée. Une des difficultés réside dans le fait que l'estimation de la FSF par exemple par maximum de vraisemblance nécessite celle de paramètres de nuisance, à savoir le spectre de l'étoile et la variance du bruit. L'estimation de la FSF se résume donc à celle des paramètres de nuisance et des paramètres du modèle analytique que nous avons développé, à partir des données.

Pour tirer profit de la nature hyperspectrale des données, nous avons proposé d'estimer les paramètres du modèle de FSF développé auparavant, par maximum de vraisemblance sur l'intégralité d'un cube d'observation bruité, en ayant estimé au préalable les paramètres du nuisance. La redondance d'information sur l'ensemble des longueurs d'onde assure des performances d'estimation satisfaisantes même pour des étoiles faiblement brillantes.

Ensuite, nous avons proposé une amélioration de l'estimation des paramètres de nuisance, s'appuyant sur le moyennage spectral des données. Dans ce cas, l'estimation de la FSF utilise des paramètres de nuisance moyennés dont la variance d'estimation est réduite. Nous avons pu constater que le moyennage spectral réduit significativement l'erreur d'estimation de la FSF, notamment pour les données très bruitées, et n'altère pas particulièrement les résultats dans les autres cas.

Nous avons montré que le spectre de l'étoile peut être efficacement corrigé en calculant sa valeur à chaque longueur d'onde par estimation conjointe du spectre et de la FSF, par un processus itératif. Cette correction améliore les performances de l'estimation de la FSF à partir d'étoiles peu brillantes. Nous avons également proposé d'estimer les paramètres du modèle de FSF par espérance *a posteriori* dans le cadre d'un modèle bayésien des données. Pour cela, nous avons généré des échantillons de la loi *a posteriori*, par une étape de de Metropolis-Hastings. Cette méthode a pour principal intérêt de surmonter le problème de l'existence de minima locaux, contrairement à l'estimation par optimisation d'un critère de maximum de vraisemblance. En outre, une telle approche fournit des barres d'erreur sur les valeurs estimées.

Enfin au Chapitre IV, notre attention s'est portée sur un problème spécifique de déconvolution de données hyperspectrales, pour l'étude de la cinématique des galaxies. Il s'agit en effet d'un des thèmes de recherche astrophysique qui a motivé la construction du spectro-imageur MUSE. Ce thème s'appuie exclusivement sur des données hyperspectrales, car il requiert à la fois la dimension spatiale et la dimension spectrale pour analyser le spectre en différentes positions de la galaxie. Notre objectif était de développer de nouvelles méthodes d'estimation de cartes de flux, de vitesse et de dispersion de vitesse d'une galaxie en rotation, afin d'en améliorer la résolution. Nous avons formalisé les données par un ensemble de raies gaussiennes convoluées spatialement et spectralement par la PSF du système, supposée connue. Le problème s'est alors traduit par l'estimation de cartes de flux, de position et de largeur de ces raies, à partir d'un sous-cube de données extrait autour d'une raie d'émission de la galaxie.

Les méthodes classiquement utilisées par la communauté astrophysique estiment ces paramètres indépendamment à chaque position spatiale. Nous avons proposé d'estimer les cartes de flux, position et largeur conjointement à toutes les positions spatiales. Pour cela, nous avons présenté deux approches : la première par l'estimateur de maximum de vraisemblance, la seconde par l'espérance *a posteriori*, en associant à chaque paramètre des lois *a priori* relativement peu informatives. Pour échantillonner la loi *a posteriori* du vecteur de paramètres, nous avons eu recours à l'échantillonneur de Gibbs. Celui-ci nous a permis de tirer profit de l'expression de la loi *a posteriori* conditionnelle du flux que nous avons échantillonné directement. Dans le cas des paramètres de position et de largeur, nous avons généré des candidats par une étape de Metropolis-Hastings à marche aléatoire. L'estimation des cartes par échantillonnage stochastique de la loi *a posteriori* a l'avantage de fournir en outre des barres d'erreur, en calculant l'écart-type *a posteriori* des échantillons.

D'une part, nous avons comparé les méthodes d'estimation indépendantes à chaque position spatiale sans prendre en compte la FSF, telles que la méthode des moments, l'estimation par maximum de vraisemblance et celle par espérance *a posteriori*. D'autre part, nous avons confronté les méthodes de déconvolution réalisées sur l'ensemble des pixels en prenant en compte la FSF, telles que la méthode d'estimation par maximum de vraisemblance conjointe et celle par espérance *a posteriori* conjointe. Ces différentes comparaisons ont montré que l'estimateur de l'espérance *a posteriori* est relativement robuste au bruit. Il fournit des résultats plus réalistes que l'estimateur du maximum de vraisemblance, sensible à l'initialisation de l'algorithme d'optimisation.

## Perspectives

Nous sommes partis de l'hypothèse que la PSF peut s'écrire sous forme séparable comme la composition d'une PSF spectrale notée  $L$  par une PSF spatiale  $F$ . Cela se justifie par certaines approximations spécifiques aux problèmes que nous avons traités. Ainsi, pour la déconvolution de cartes, nous avons supposé que la LSF est localement constante en longueur d'onde ce qui permet d'écrire la PSF sous forme séparable :  $h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) = F_{\mathbf{z},\mu}(\mathbf{s} - \mathbf{z})L_{\mathbf{z},\mu}(\lambda)$ . Or, il a été récemment montré [CAR 12] que l'écriture de la PSF devrait être  $h_{\mathbf{z},\mu}(\mathbf{s}, \lambda) = F_{\mathbf{z},\mu}(\mathbf{s})L_{\mathbf{s},\mu}(\lambda)$  (indice  $\mathbf{s}$  et non  $\mathbf{z}$  pour la LSF), ce qui implique que la PSF n'est pas exactement séparable. Cette différence n'a pas d'impact sur le

développement théorique de nos méthodes mais devrait être mise en œuvre dans les algorithmes développés. Notons que l'approximation séparable est nécessaire pour certains problèmes d'analyse de données tels que la séparation de source [MEG 12] et que l'influence d'une telle approximation devrait être évaluée.

Concernant l'estimation de la FSF, il n'est pas toujours possible d'isoler une étoile brillante dans un cube d'observation. Il faut donc être capable d'estimer la PSF spatiale, soit à partir d'une étoile brillante non isolée, soit à partir d'une étoile de faible intensité. Aussi, il pourrait être intéressant de développer davantage les stratégies que nous avons évoquées sous forme de masques pour augmenter la robustesse de l'estimation de la FSF à partir d'une étoile non isolée. Ainsi, nous pourrions envisager de combiner les méthodes développées pour l'estimation de la FSF aux méthodes de séparation de sources [MEG 12]. De même, la modélisation du bruit d'acquisition par une loi poissonnienne plutôt qu'une approximation par une loi normale, pourrait améliorer l'estimation à partir d'étoile de très faible intensité. Néanmoins, cette formulation ne permettrait pas d'exprimer analytiquement l'estimateur du spectre pour l'estimation conjointe.

Nous avons évoqué dans la section III.4 la possibilité d'estimer conjointement le spectre de l'étoile et les paramètres du modèle de FSF, en intégrant à l'expression de l'anti-log-vraisemblance l'estimateur du spectre par moindres carrés pondérés. De même, comme nous l'avons suggéré dans la section III.5.3, nous pourrions échantillonner la loi *a posteriori* marginalisée par rapport au spectre de l'étoile, après lui avoir associé une loi *a priori*. Il serait intéressant de comparer les performances d'estimation de ces deux méthodes, par rapport à celles obtenues dans le cadre de l'estimation séparée de ces variables, tant par maximum de vraisemblance que par espérance *a posteriori*. Concernant l'estimation du spectre de l'étoile, nous avons constaté qu'elle est particulièrement sensible au bruit et que les méthodes d'estimation présentées dans le Chapitre III ont une efficacité limitée. Même si l'estimation du spectre n'était pas l'objectif premier de cette thèse, il serait profitable d'apporter une régularisation de douceur spectrale sur le spectre convolué par la LSF, ce qui pourrait également améliorer l'estimation de la FSF. Pour cela, une pénalisation quadratique sur la différence du spectre aux pixels voisins (modèle de Gauss-Markov) paraît intéressante car elle conserverait une expression analytique du minimum de la vraisemblance conjointe pour le spectre, ainsi que la possibilité de marginaliser la loi *a posteriori* conjointe pour l'estimation de l'espérance *a posteriori*.

Dans le cadre de la déconvolution de cartes de cinématique, nous avons été confrontés au problème de l'existence de minima locaux du critère de vraisemblance, auquel l'estimateur du maximum de vraisemblance conjointe est particulièrement sensible. Celui-ci pourrait fournir de meilleurs résultats, sous réserve de lui fournir davantage d'information. Ainsi, une régularisation spatiale appliquée au minimum sur la position et éventuellement sur les autres paramètres permettrait d'éviter le phénomène d'alternance que nous avons observé. Toutefois, il ne faudrait pas que cette régularisation détériore la résolution spatiale des données que l'on souhaite améliorer par l'opération de déconvolution.

Nous pourrions également opter pour l'estimateur du maximum *a posteriori* plutôt que celui de l'espérance et le calculer avec les méthodes stochastiques que nous avons utilisées ou bien des méthodes similaires. Par exemple, la méthode du recuit simulé qui fait appel à un paramètre de température sur la loi *a posteriori*, apporterait sa robustesse face à la présence éventuelle de minima locaux. Néanmoins, cette méthode est réputée pour avoir un coût de calcul élevé.

Par ailleurs, la carte de flux pourrait être calculée *a priori* à partir de l'observation de la galaxie par le télescope Hubble, observation qui est non convoluée par la FSF de l'atmosphère mais qui doit être déconvoluée de la PSF de l'instrument et ramenée à la résolution souhaitée. Cette carte fournirait un complément d'information selon deux stratégies possibles. La première serait de fixer la carte du flux à la valeur mesurée par Hubble et d'estimer, par n'importe quelle méthode, les cartes de position et de largeur. La seconde possibilité consisterait à définir la loi *a priori* du paramètre de flux par une gaussienne centrée autour de la valeur mesurée par Hubble, avec une variance correspondant à la variance d'observation du télescope spatial. Cette option contraindrait fortement l'estimation du flux

mais laisserait néanmoins une certaine marge de liberté.

Nous avons pu constater, tant au niveau de l'estimation de la FSF que de la déconvolution de cartes de cinématique, la difficulté d'utilisation des méthodes de MCMC. Le premier point critique est le choix des lois *a priori*. Nous avons choisi des lois relativement simples de type loi uniforme, Inverse Gamma et normale. Le choix d'une loi *a priori* plus informative, tout en restant cohérente avec la physique du problème, permettrait sans doute d'améliorer les résultats.

La deuxième difficulté a été le réglage des échelles des lois de proposition des étapes de Metropolis-Hastings par marche aléatoire. Le réglage adaptatif de ces paramètres au cours de l'échantillonnage permettrait sans doute de réduire le coût de calcul et peut-être de trouver de nouveaux modes dans la distribution. Par exemple, il serait profitable d'améliorer la loi de proposition, aussi bien des paramètres du modèle de FSF que des cartes de cinématique, par la méthode de Laplace, tel que cela a été suggéré par [BUI 12]. Nous remplacerions alors la marche aléatoire par une loi normale dont la moyenne et la variance dépendraient des données. Avec une procédure de réglage automatisée, nous pourrions envisager de faire varier les échelles des lois de proposition en fonction du flux de la raie. Pour automatiser le choix de la loi de proposition, nous pourrions estimer les paramètres en échantillonnant une population de Monte-Carlo (*Population Monte-Carlo sampling* [DJU 11]). Cette méthode adapte de manière itérative la loi de proposition en tenant compte des échantillons et de leur pondération en fonction de leur probabilité. Une telle méthode a l'avantage de tirer profit de l'ensemble des échantillons générés et de ne pas nécessiter de période de *burn-in*. Enfin, concernant l'exploration de l'espace des paramètres, nous pourrions générer plusieurs chaînes en parallèle (*parallel tempering*), chacune ayant une température différente [ATC 11]. Alors, l'échange d'échantillons entre des chaînes voisines permettrait d'échapper aux minima locaux et de converger plus rapidement.

Concernant l'application des méthodes développées durant cette thèse, les membres du consortium MUSE mettent en place une base de données dans l'environnement AstroWise, pour gérer les observations effectuées par MUSE pour le consortium. Cet environnement fournit à la fois l'accès aux données mais également un ensemble de services d'analyse et de traitement de ces données. En particulier, le portail *MUSE Quality Assessment* a été créé pour évaluer la qualité des données de la chaîne instrumentale et du DRS. Aussi, nous avons décidé d'y intégrer notre méthode d'estimation de la FSF. Pour cela, Camille Parisel, ingénieur à l'IRAP, a adapté en Python l'algorithme d'estimation conjointe de la FSF et du spectre de l'étoile par maximum de vraisemblance que nous avons proposé, en y intégrant l'estimation du résidu du fond du ciel par moindres carrés (paramètre de nuisance supplémentaire). Au moment de la rédaction de la thèse, cette estimation est réalisée à partir d'un cube de données simulé issu du DRS. Dans l'avenir, nous espérons effectuer cette opération directement à partir des images du CCD, ce qui permettrait d'éviter les étapes d'interpolation du logiciel de réduction des données (DRS). Cette estimation fournirait la PSF nécessaire à l'opération de fusion de cubes de données. À terme, cette collaboration permettra également aux utilisateurs de connaître la FSF en plusieurs points du cube de données et donc sa variation spatiale et spectrale. Celle-ci donnera un indice de qualité à la fois des données et des traitements effectués par le DRS.

Par ailleurs, nous nous sommes concentrés durant cette thèse sur certains modèles en particulier, tel que le modèle circulaire constant dans le champ, avec  $\alpha$  linéaire en longueur d'onde et  $\beta$  fixé. L'estimation pourrait être légèrement améliorée avec un modèle de variation décrit par des fonctions puissances, exponentielles, etc. et également en considérant la variation d'autres paramètres tels que la position de l'étoile en fonction de la longueur d'onde. Il serait intéressant que l'utilisateur puisse choisir la complexité de la variation spectrale de la FSF, en fonction de l'utilisation qu'il souhaite en faire. La méthodologie proposée reste valable à partir du moment où l'on dispose d'un modèle pour la PSF et pour l'étoile (position et spectre) sur l'intégralité du cube et dépendant d'un faible nombre de paramètres.

Concernant le stockage de la FSF dans la base de données, le modèle hyperspectral de FSF apporte

davantage de compressibilité, par rapport à celui de FSF numériques ou de modèles paramétriques indépendants à chaque longueur d'onde, car il n'est constitué que d'un faible nombre de paramètres.

Enfin, il existe deux méthodes alternatives pour l'estimation et la déconvolution de cartes de cinématique de galaxies. En effet, Benoît Epinat [EPI 08] a proposé de déconvoluer *a posteriori* les cartes de cinématique estimées à partir d'un sous-cube de données hyperspectrales extrait autour d'une raie. Une seconde approche [ALV 12] consiste à estimer les paramètres d'un modèle paramétrique pour la cinématique de la galaxie tout en prenant en compte la PSF de la chaîne instrumentale. Une telle approche peut là encore exploiter des techniques d'optimisation pour l'estimation par maximum de vraisemblance [BOO 06] ou des techniques d'échantillonnage stochastique comme cela a été effectué lors du stage de Valentin Alvarez à l'IRAP. Il serait profitable pour la communauté astrophysique de comparer les performances de ces différentes méthodes en terme de résolutions spatiale et spectrale recouvrées et en terme de variance sur les paramètres.



# Bibliographie

- [AIR 35] G.B. AIRY. On the diffraction of an object-glass with circular aperture. *Transactions of the Cambridge Philosophical Society*, 5 :283, 1835.
- [ALV 12] V. ALVAREZ. Cinématique 3D des galaxies. Rapport de stage de master 2, sous la direction de nicolas bouché et Hervé Carfantan, IRAP, Toulouse, septembre 2012.
- [ATC 11] YVES ATCHADÉ, GERSENDE FORT, ERIC MOULINES AND PIERRE PRIOURET. *Adaptive Markov Chain Monte Carlo : Theory and Methods*, pages 1–31. Cambridge University Press, 2011.
- [BAC 95] R. BACON, G. ADAM, A. BARANNE, G. COURTES, D. DUBET, J. P. DUBOIS, E. EMSELLEM, P. FERRUIT, Y. GEORGELIN, G. MONNET, E. PECONTAL, A. ROUSSET AND F. SAY. 3D spectrography at high spatial resolution. I. concept and realization of the integral field spectrograph TIGER. *Astronomy and Astrophysics Supplement Series*, 113 :347, octobre 1995.
- [BAC 01] R. BACON, Y. COPIN AND G MONNET. The SAURON project - I. the panoramic integral field spectrograph. *Monthly Notices of the Royal Astronomical Society*, 326(1) :23–35, septembre 2001.
- [BAC 09] R. BACON. Analysis of the spectral PSF. MUSE report, VLT-TRE-MUS-14670-0870, Centre de Recherche Astrophysique de Lyon, Lyon, 2009.
- [BOO 06] F. BOONE, D. MUDERS, P. SCHILKE, C. COMITO, S. LEURINI AND B. PARISE. Advanced science analysis package and the prototype DALIA. In *Astronomical Data Analysis Software and Systems XV ASP Conference Series*, San Lorenzo de El Escorial, Spain, octobre 2006.
- [BOU 10] S. BOURGUIGNON, D. MARY AND E. SLEZAK. Sparsity-based denoising of hyperspectral astrophysical data with colored noise : Application to the MUSE instrument. In *2nd IEEE Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Reykjavik, Iceland, juin 2010.
- [BOU 11a] S. BOURGUIGNON, H. CARFANTAN, E. SLEZAK AND D. MARY. Sparsity-based spatial-spectral restoration of MUSE astrophysical hyperspectral data cubes. In *3rd IEEE Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Lisbon, Portugal, juin 2011.
- [BOU 11b] S. BOURGUIGNON, D. MARY AND E. SLEZAK. Restoration of astrophysical spectra with sparsity constraints : Models and algorithms. *IEEE Journal of Selected Topics in Signal Processing*, 5(5) :1002–1013, septembre 2011.
- [BOU 12] S. BOURGUIGNON, D. MARY AND E. SLEZAK. Processing MUSE hyperspectral data : Denoising, deconvolution and detection of astrophysical sources. *Statistical Methodology*, 9(1-2) :32–43, janvier 2012.
- [BRO 11] S. BROOKS, A. GELMAN, G. JONES AND X.-L. MENG. *Handbook of Markov Chain Monte Carlo*. Handbooks of Modern Statistical Methods. Chapman and Hall/CRC, mai 2011.
- [BUI 12] P. BUI QUANG, C. MUSSO AND F. LE GLAND. Multidimensional Laplace formulas for nonlinear Bayesian estimation. In *Proceedings of the 20th European Signal Processing Conference*, Bucharest, Romania, août 2012.

- [CAR 09] H. CARFANTAN AND E. ANTERRIEU. ANR DAHLIA : Modélisation de la PSF spectrale intégrée de MUSE. Rapport technique ANR Dahlia, IRAP, Toulouse, France, septembre 2009.
- [CAR 12] H. CARFANTAN, A. JARNO, D. SERRE, V. MAZET, M. PETREMAND, S. BOURGUIGNON AND E. VILLENEUVE. ANR DAHLIA : Modélisation de la PSF de MUSE. Rapport technique ANR Dahlia, version 5.6, IRAP, Toulouse, France, juin 2012.
- [DJU 11] P.M. DJURIC, B. SHEN AND M.F. BUGALLO. Population Monte Carlo methodology a la Gibbs sampling. pages 669–673, août 2011.
- [DOU 01] A. DOUCET AND C. ANDRIEU. Iterative algorithms for state estimation of jump Markov linear systems. *IEEE Transactions on Signal Processing*, 49(6) :1216–1227, juin 2001.
- [DOU 05] A DOUCET AND X. WANG. Monte Carlo methods for signal processing : a review in the statistical signal processing context. *IEEE Signal Processing Magazine*, 22(6) :152–170, novembre 2005.
- [EIS 03] F. EISENHAEUER. SINFONI - integral field spectroscopy at 50 milli-arcsecond resolution with the ESO VLT. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4841, pages 1548–1561, mars 2003.
- [EPI 08] BENOÎT EPINAT. *Des galaxies proches aux galaxies lointaines : étude cinématique et dynamique*. Thèse de doctorat, Université de Provence, novembre 2008.
- [FEL 97] P. FELENBOK, J. GUÉRIN, A. FERNANDEZ, V. CAYATTE, C. BALKOWSKI AND R. C. KRAAN-KORTEWEG. The performance of MEFOS, the ESO multi-object fibre spectrograph. *Experimental Astronomy*, 7(2) :65–85, juin 1997.
- [FRI 66] D.D. L. FRIED. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *Journal of the Optical Society of America*, 56(10) :1372–1379, octobre 1966.
- [GEL 92] A. GELMAN AND D. B. RUBIN. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4) :457–472, novembre 1992.
- [GEL 96] A. GELMAN, G. O. ROBERTS AND W. R. GILKS. Efficient Metropolis jumping rules. *Bayesian statistics*, 5 :599–607, 1996.
- [GEY 92] C. J. GEYER. Practical Markov Chain Monte Carlo. *Statistical Science*, 7(4) :473–483, novembre 1992.
- [GIL 96] W. R. GILKS, S. RICHARDSON AND D. J. SPIEGELHALTER. *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC, London, UK, 1996.
- [IDI 01] JÉROME IDIER, Ed.. *Approche bayésienne pour les problèmes inverses*. Traité IC2, Série Traitement du signal et de l'image. Hermès Science Publications, Paris, France, novembre 2001.
- [JAL 11] A. JALOBEANU, M. PETREMAND AND C. COLLET. Bayesian fusion of hyperspectral astronomical images. In *30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, volume 1305, pages 295–302. AIP Conference Proceedings, 2011.
- [JAR 08] AURÉLIEN JARNO. *Développement d'un modèle numérique de l'instrument MUSE/VLT*. Thèse de doctorat, INSA Lyon, juillet 2008.
- [JOL 06] L. JOLISSAINT, J.-P. VÉRAN AND R. CONAN. Analytical modeling of adaptive optics : foundations of the phase spatial power spectrum approach. *Journal of the Optical Society of America (A)*, 23(2) :382–394, février 2006.
- [JOL 08] L. JOLISSAINT, H. CARFANTAN AND E. ANTERRIEU. Exploring the impact of PSF reconstruction errors on the reduction of astronomical adaptive optics based data. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7015, pages 70152K–70152K–12, Marseille, France, juin 2008.

- [JOL 09] LAURENT JOLISSAINT. Step-by-step procedure for computing numerically the seeing limited point spread function from optical turbulence phase structure function equation. Technical report, University of Leiden, The Netherlands, 2009.
- [KIR 61] G. KIRCHHOFF AND R. BUNSEN. XLII. Chemical analysis by spectrum-observations. second memoir. *Philosophical Magazine Series 4*, 22(148) :329–349, 1861.
- [KIT 95] C. ROBERT. KITCHIN. *Optical astronomical spectroscopy*. Institute of Physics Publishing, Bristol, United Kingdom, 1995.
- [KOL 91] A. N. KOLMOGOROV. Dissipation of energy in the locally isotropic turbulence. *Royal Society of London Proceedings Series A*, 434 :15–17, juillet 1991.
- [LAG 98] J. C. LAGARIAS, J. A. REEDS, M. H. WRIGHT AND P. E. WRIGHT. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1) :112–147, décembre 1998.
- [Le 03] O. LE FÈVRE, M. SAISSE, D. MANCINI, S. BRAU-NOGUE, O. CAPUTI, L. CASTINEL, S. D’ODORICO, B. GARILLI, M. KISSLER-PATIG, C. LUCUIX, G. MANCINI, A. PAUGET, G. SCIARRETTA, M. SCODEGGIO, L. TRESSE AND G. VETTOLANI. Commissioning and performances of the VLT-VIMOS instrument. In M. IYE AND A. F. M. MOORWOOD, Eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4841, pages 1670–1681, mars 2003.
- [LEN 08] P. LENA, D. ROUAN, F. LEBRUN, F. MIGNARD AND D. PELAT. *L’observation en astrophysique*. EDP Sciences, juin 2008.
- [LEN 09] P. LENA, H. DOLE, A.-M. LAGRANGE, D. ROUAN, P. BINÉTRUY AND F. GENOVA. *L’observation en astronomie*. Ellipses, 2009.
- [MAG 98] P. MAGAIN, F. COURBIN AND S. SOHY. Deconvolution with correct sampling. *The Astrophysical Journal*, 494(1) :472–477, février 1998.
- [MAI 11] LOUIS MAIRE. Identification de PSF spectrale pour calibration de l’instrument MUSE. Rapport de stage de master 2, sous la direction de Hervé Carfantan, IRAP, Toulouse, septembre 2011.
- [MAR 00] F. MARTIN, R. CONAN, A. TOKOVININ, A. ZIAD, H. TRINQUET, J. BORGNINO, A. AGABI AND M. SARAZIN. Optical parameters relevant for high angular resolution at paranal from GSM instrument and surface layer contribution. *Astronomy and Astrophysics Supplement Series*, 144(1) :39–44, mai 2000.
- [MAZ 05] VINCENT MAZET. *Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies*. Thèse de doctorat, Université Henri Poincaré, Nancy 1, décembre 2005.
- [MCL 08] I. S. MCLEAN. *Electronic imaging in astronomy. Detectors and instrumentation*. Springer-Praxis books in astronomy and planetary sciences. Springer and Praxis Publishing, second edition, 2008.
- [MEG 11] I. MEGANEM, Y. DEVILLE, S. HOSSEINI, H. CARFANTAN AND M. S. KAROUI. Extraction of stellar spectra from dense fields in hyperspectral MUSE data cubes using non-negative matrix factorization. In *3rd IEEE Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Lisbon, Portugal, juin 2011.
- [MEG 12] I. MEGANEM. *Méthodes de séparation aveugle de sources pour l’imagerie hyperspectrale. Application à la télédétection et à l’astrophysique*. Thèse de doctorat, Université Toulouse III, décembre 2012.
- [MOF 69] A. F. J. MOFFAT. A theoretical investigation of focal stellar images in the photographic emulsion and application to photographic photometry. *Astronomy and Astrophysics*, 3 :455–461, décembre 1969.

- [MUG 04] L. M. MUGNIER, T. FUSCO AND J.-M. CONAN. MISTRAL : a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images. *Journal of the Optical Society of America (A)*, 21(10) :1841–1854, octobre 2004.
- [MUL 91] P. MULLER. A generic approach to posterior integration and Bayesian sampling. Technical report # 91-09, Department of Statistics, Purdue University, février 1991.
- [ORI 10] F. ORIEUX, J.-F. GIOVANNELLI AND T. RODET. Bayesian estimation of regularization and point spread function parameters for wiener–hunt deconvolution. *Journal of the Optical Society of America (A)*, 27(7) :1593–1607, juillet 2010.
- [PAS 02] L. PASQUINI, G. AVILA, A. BLECHA, C. CACCIARI, V. CAYATTE, M. COLLESS, F. DAMIANI, R. DE PROPRIIS, H. DEKKER, P. DI MARCANTONIO, T. FARRELL, P. GILLINGHAM, I. GUINOARD, F. HAMMER, A. KAUFER, V. HILL, M. MARTEAUD, A. MODIGLIANI, G. MULAS, P. NORTH, D. POPOVIC, E. ROSSETTI, F. ROYER, P. SANTIN, R. SCHMUTZER, G. SIMOND, P. VOLA, L. WALLER AND M. ZOCCALI. Installation and commissioning of FLAMES, the VLT multifibre facility. *The Messenger*, 110 :1–9, décembre 2002.
- [PET 10] M. PETREMAND, L. MICHEL AND M. LOUYS. Visualization and logical binding of hyperspectral data using QuickViz and SAADA. In *Proceedings of Astronomical Data Analysis Software and Systems XX (ADASS) conference*, ASP Conference Proceedings, page 673, Boston, USA, novembre 2010.
- [PET 11] M. PETREMAND, C. COLLET AND A. JALOBÉANU. Fusion bayésienne d’images hyperspectrales astronomiques. In *Actes du 23<sup>e</sup> colloque GRETSI*, Bordeaux, France, septembre 2011.
- [PET 12] M. PETREMAND, A. JALOBÉANU AND C. COLLET. Optimal bayesian fusion of large hyperspectral astronomical observations. *Statistical Methodology, Special issue on Astrostatistics*, 9(1) :44–54, janvier 2012.
- [ROB 96] C. ROBERT. *Méthodes de Monte Carlo par chaînes de Markov*. Économica, Paris, France, 1996.
- [ROB 97] G. O. ROBERTS, A. GELMAN AND W. R. GILKS. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1) :110–120, 1997.
- [ROB 01] G. O. ROBERTS AND J. S. ROSENTHAL. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical Science*, 16(4) :351–367, novembre 2001.
- [ROD 81] F. RODDIER. The effects of atmospheric turbulence in optical astronomy. *Progress in optics*, 19 :281–376, 1981.
- [ROD 99] F. RODDIER. *Adaptive Optics in Astronomy*. University of Hawaii, Manoa, 1999.
- [SAR 90] M. SARAZIN AND F. RODDIER. The ESO differential image motion monitor. *Astronomy and Astrophysics*, 227(1) :294–300, janvier 1990.
- [SER 10] D. SERRE, E. VILLENEUVE, H. CARFANTAN, L. JOLISSAINT, V. MAZET, S. BOURGUIGNON AND A. JARNO. Modeling the spatial PSF at the VLT focal plane for MUSE WFM data analysis purpose. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7736, juillet 2010.
- [TOK 02] ANDREI TOKOVININ. From differential image motion to seeing. *Publications of the Astronomical Society of the Pacific*, 114(800) :1156–1166, octobre 2002.
- [TRU 01] I. TRUJILLO, A. AGUERRI, J. CEPÁ AND C. M. GUTIERREZ. The effects of seeing on Sersic profiles. II. the Moffat PSF. *Monthly Notices of the Royal Astronomical Society*, 328(3) :977–985, décembre 2001.
- [VIL 11a] E. VILLENEUVE, H. CARFANTAN, A. JARNO, D. SERRE, V. MAZET AND S. BOURGUIGNON. Modélisation et estimation de la PSF d’un instrument hyperspectral au sol pour l’astrophysique. In *Actes du 23<sup>e</sup> colloque GRETSI*, Bordeaux, France, septembre 2011.

- [VIL 11b] E. VILLENEUVE, H. CARFANTAN AND D. SERRE. PSF estimation of hyperspectral data acquisition system for ground-based astrophysical observations. In *3rd IEEE Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Lisbon, Portugal, juin 2011.
- [VIL 12] E. VILLENEUVE AND H. CARFANTAN. Hyperspectral data deconvolution for galaxy kinematics with MCMC. In *Proceedings of the 20th European Signal Processing Conference*, Bucharest, Romania, août 2012.
- [VIV 08] S. VIVES, E. PRIETO, Y. SALAUN AND P. GODEFROY. New technological developments in integral field spectroscopy. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7018, juin 2008.
- [WEE 32] M. E. WEEKS. The discovery of the elements. XIII. some spectroscopic discoveries. *Journal of Chemical Education*, 9(8) :1413, août 1932.
- [WEI 08] P. WEILBACHER. Data Reduction Library Design. MUSE report, VLT-TRE-MUS-14670-0689, Leibniz Institute for Astrophysics Potsdam, 2008.
- [WOO 82] N. J. WOOLF. High resolution imaging from the ground. *Annual Review of Astronomy and Astrophysics*, 20 :367–398, septembre 1982.
- [WYN 88] C. G. WYNNE AND S. P. WORSWICK. A low-dispersion survey spectrograph. *The Observatory*, 108 :161–163, octobre 1988.



# Publications et communications de l'auteur

## Conférences internationales avec comité de lecture et avec actes

1. D. SERRE, E. VILLENEUVE, H. CARFANTAN, L. JOLISSAINT, V. MAZET, S. BOURGUIGNON, ET A. JARNO. Modeling the spatial PSF at the VLT focal plane for MUSE WFM data analysis purpose. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7736, juillet 2010.
2. E. VILLENEUVE, H. CARFANTAN, ET D. SERRE. PSF estimation of hyperspectral data acquisition system for ground-based astrophysical observations. In *3rd IEEE Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Lisbon, Portugal, juin 2011.
3. E. VILLENEUVE ET H. CARFANTAN. Hyperspectral data deconvolution for galaxy kinematics with MCMC. In *Proceedings of the 20<sup>th</sup> European Signal Processing Conference*, Bucharest, Romania, août 2012.

## Conférence nationale avec comité de lecture et avec actes

4. E. VILLENEUVE, H. CARFANTAN, A. JARNO, D. SERRE, V. MAZET ET S. BOURGUIGNON. Modélisation et estimation de la PSF d'un instrument hyperspectral au sol pour l'astrophysique. In *Actes du 23<sup>e</sup> colloque GRETSI*, Bordeaux, France, septembre 2011.

## Séminaires

5. Modélisation et estimation de la PSF d'un instrument hyperspectral au sol pour l'astrophysique. Séminaire Calibration Astrale, organisé par le Centre de Compétence Technique Traitement du Signal et des Images (CCT-TSI) du CNES. Toulouse, 5 Décembre 2011.
6. Déconvolution de données hyperspectrales pour l'étude de la cinématique des galaxies. Imagerie multi/hyperspectrale : problématiques et nouveaux domaines d'applications. Réunion du GDR-ISIS, Paris, 4 Juillet 2012.



# Annexes



# Annexe A

## Compléments astrophysiques

### A.1 Spectres de galaxies

Le spectre d'une galaxie est constitué d'un continuum, de raies d'émission et de raies d'absorption. Chacun de ces composants est le résultats de nombreux phénomènes physiques [KIT 95][LEN 09]. Ainsi, le continuum est principalement dû à la radiation de type corps noir émise par les corps solides, liquides et par les gaz denses. Une autre part du continuum est dû au fond extragalactique qui est le rayonnement de fond provenant des anciennes générations de galaxies, dans le domaine de l'infrarouge. À celui-ci s'ajoute le fond cosmologique, c'est-à-dire le rayonnement diffus de corps noir de température  $2.275 \pm 0.002 \text{ K}$ , qui emplît l'univers de manière isotrope.

Concernant les raies du spectre, il faut savoir que les atomes et les ions absorbent des radiations à des fréquences précises, quantifiées. L'absorption ou l'émission de ces radiations sous forme de photons provoque le changement de niveau d'énergie d'un électron, c'est-à-dire l'excitation ou la désexcitation de l'atome. La lumière est donc émise ou absorbée par les atomes à des longueurs d'onde spécifiques, ce qui fait apparaître des raies dans le spectre.

En analysant le spectre d'une source lumineuse, on peut notamment [LEN 09] :

- calculer la vitesse d'un astre et sa dispersion en vitesse via l'effet Doppler-Fizeau,
- connaître sa composition atomique ou moléculaire,
- déterminer l'état des éléments qui la constitue, à savoir atomique, ionisé, moléculaire, solide ou gazeux,
- évaluer l'intensité et la direction de son champ magnétique,
- connaître la gravité locale,
- connaître la température de l'astre,
- sonder l'intérieur d'une étoile.

### A.2 Influence du mouvement des galaxies sur leur spectre

#### A.2.1 Effet Doppler-Fizeau

L'effet Doppler est le décalage entre la fréquence d'une onde acoustique ou électromagnétique mesurée à la source et la fréquence mesurée par un observateur, lorsque la source ou l'observateur est en mouvement. Cet effet est couramment observé, dans le domaine des ondes acoustiques. Par exemple, la sirène d'un véhicule en mouvement paraît plus aiguë lorsque celui-ci s'approche et plus grave lorsqu'il s'éloigne, par rapport au son que l'on entend lorsque le véhicule est immobile. Dans le domaine des ondes électromagnétiques, on parle d'effet Doppler-Fizeau.

Dans le cadre de cette thèse (Chapitre IV), les vitesses sont exprimées le long de la ligne de visée. Dans ce cas particulier, la fréquence de l'onde émise  $\nu_0$ , mesurée à la source, est reliée à la fréquence

mesurée  $\nu_m$  selon l'expression [EPI 08] :

$$\nu_0 = \nu_m \sqrt{\frac{1 - \beta}{1 + \beta}}$$

dans laquelle  $\beta = v/c$  où  $v$  est la vitesse relative de la source par rapport à l'observateur et  $c$  la célérité de la lumière.

Dans le cadre de l'analyse de données hyperspectrales, la vitesse est exprimée en fonction de la longueur d'onde de la source  $\lambda_0 = \frac{c}{\nu_0}$  et celle mesurée  $\lambda_m = \frac{c}{\nu_m}$  par la relation :

$$v = c \frac{(z + 1)^2 - 1}{(z + 1)^2 + 1}$$

dans laquelle  $z = \frac{\lambda_m - \lambda_0}{\lambda_0}$  est le décalage spectral. Lorsque ce décalage est faible ( $z < 1$ ), cette relation est approximée par  $v \sim c \cdot z = c \frac{\lambda_m - \lambda_0}{\lambda_0}$ .

### A.2.2 Décalage spectral

Il est nécessaire de distinguer les effets du mouvement global de la galaxie dû à l'expansion de l'univers, de ceux la rotation de la galaxie sur elle-même.

Ce que l'on appelle le red-shift est le décalage de longueur d'onde vers le rouge de raies spectrales, qui est interprété comme une conséquence de l'expansion de l'univers [LEN 09] et donc du mouvement global de la galaxie. Il correspond à une mesure de la distance de la source. Les objets les plus proches, telles que les étoiles de notre galaxie, ont un décalage spectral  $z = 0$ , alors que les galaxies lointaines ont un  $z$  pouvant aller jusqu'à 10. Le rayonnement cosmologique est associé à  $z = 1000$ .

La rotation de la galaxie provoque un décalage de la longueur d'onde des raies observées, qui s'ajoute au red-shift, vers le bleu ou vers le rouge selon que la zone de la galaxie s'éloigne ou s'approche de l'observateur. On travaille généralement avec les vitesses relatives exprimées dans le référentiel de la galaxie. La vitesse relative  $v_r$  est reliée au décalage spectral par l'équation [EPI 08] :

$$v_r \sim c \frac{\lambda_m - \lambda_e}{\lambda_e}$$

où  $\lambda_e$  est la longueur d'onde observée décalée par le seul fait de l'éloignement de la galaxie et  $\lambda_m$  la longueur d'onde mesurée.

## Annexe B

# Mesures de la qualité d'image pour l'astronomie

### B.1 Notions de résolution

Dans le cadre des systèmes de mesure de signaux et d'acquisition d'images, la résolution est la plus petite distance entre deux points pouvant être différenciés par le système. L'étude de données hyperspectrales nécessite de définir à la fois une résolution spatiale et spectrale.

#### B.1.1 Résolution spatiale

La résolution spatiale d'un système d'acquisition de données astronomiques au sol dépend à la fois de l'atmosphère et du télescope.

D'après le critère de résolution de Rayleigh, deux objets de même magnitude peuvent être distingués si le maximum de la tache de diffraction du premier coïncide avec le premier minimum du second. La résolution théorique d'un télescope idéal d'ouverture circulaire<sup>1</sup> est alors donnée par le plus petit angle qui annule la fonction d'Airy :

$$1.22 \frac{\lambda}{D}$$

exprimé en radians où  $\lambda$  est la longueur d'onde étudiée et  $D$  le diamètre du télescope tous deux exprimés dans la même unité. Dans [LEN 08][LEN 09], on considère que seuls 85% du rayon angulaire de la tache de diffraction suffisent pour séparer les objets, ce qui correspond à  $\frac{\lambda}{D}$ . Cet angle peut être encore réduit en fonction des techniques d'observation employées, telle que l'interférométrie.

Ainsi, la résolution spatiale du télescope s'améliore lorsque la longueur d'onde diminue : le télescope est donc naturellement plus performant dans le bleu que dans le rouge. À l'opposé, la résolution spatiale due à l'atmosphère est plus fine dans le rouge que dans le bleu.

Si le diamètre du télescope dépasse 20 à 40 cm et si l'étude concerne le domaine du visible, alors la résolution spatiale est bien moins bonne que la limite de diffraction, en raison de l'effet des turbulences de l'atmosphère [LEN 09] (cf. section I.2).

#### B.1.2 Résolution spectrale

Dans le domaine spectral, la résolution est l'intervalle  $\Delta\lambda$  nécessaire pour distinguer deux raies ponctuelles. Elle peut être quantifiée par le ratio  $R = \frac{\lambda}{\Delta\lambda}$  [LEN 09], sans dimension. Pour observer des spectres et analyser différentes raies, il est logiquement souhaitable d'avoir la plus fine résolution possible, correspondant à un  $R$  élevé. Néanmoins, lorsqu'on étudie des sources peu lumineuses, on préfère opter pour une résolution plus grande, car le nombre de photons collectés sur le détecteur est

---

<sup>1</sup>Dans le cas d'un télescope à ouverture carrée, le plus petit angle qui annule la fonction d'Airy est  $\frac{\lambda}{D}$ .

faible. Ainsi, avec une résolution plus grande, le nombre de photons reçus sur un intervalle spectral est plus élevé, ce qui augmente le rapport signal sur bruit.

## B.2 Les critères de qualité d'une observation

Différents critères de qualité sont utilisés en astrophysique pour quantifier la qualité d'une observation.

### B.2.1 Rapport de Strehl

Par caractériser la qualité d'un système d'acquisition au sol, il est souvent fait référence au rapport de Strehl correspondant à la hauteur maximale de la PSF réelle du système comprenant généralement le télescope et l'atmosphère éventuellement corrigée par optique adaptative, sur la hauteur maximale de la tache de diffraction qui est la PSF d'un télescope idéal [LEN 08] [LEN 09]. Ainsi, le rapport de Strehl d'un télescope idéal dans l'espace est 100%. Un bon système de correction par optique adaptative permet d'atteindre 50% à une longueur d'onde de  $2 \mu\text{m}$ , alors que sans correction, le rapport est d'environ 1% pour un grand télescope.

### B.2.2 Énergie encerclée

L'énergie encerclée est le diamètre du cercle dans lequel est concentré un pourcentage donné (50% ou 80%) de l'énergie issue d'une source ponctuelle [LEN 08]. À 50%, on l'appelle également la largeur à mi-puissance du faisceau (*Half Power Beam Width*). Ce critère est utilisé dans le cahier des charges de l'instrument MUSE pour caractériser la qualité optique des différents composants du système.

### B.2.3 Paramètre de Fried

Le paramètre de Fried d'une observation, introduit par D.L. Fried ([FRI 66]), est le diamètre équivalent d'un télescope comportant une pupille circulaire, dont l'unique limite à la qualité de l'image est la diffraction, qui donnerait une image de même résolution angulaire que celle de l'image dégradée par la traversée de l'atmosphère [LEN 08].

Le paramètre de Fried est chromatique, généralement donné à 500 nm et dépend de la fonction de structure  $C_n^2$  qui décrit les turbulences atmosphériques à différentes altitudes  $z$ . Il est donné par l'expression suivante, valable au zénith [ROD 81] :

$$r_0(\lambda) = 0.185 \cdot \lambda^{\frac{6}{5}} \left[ \int_0^\infty C_n^2(z) dz \right]^{\frac{-3}{5}}$$

Il peut être mesuré par le DIMM (*Differential Image Motion Monitor*) [SAR 90].

### B.2.4 Seeing

Le *seeing* est la largeur à mi-hauteur (FWHM) de la *Point Spread Function* (PSF) dans le cas d'une observation au sol. Il est relié au paramètre de Fried par la formule  $s = \frac{\lambda}{r_0}$ . Ce dernier étant proportionnel à  $\lambda^{\frac{6}{5}}$ , le *seeing* varie en fonction de  $\lambda^{-\frac{1}{5}}$ . Ses valeurs typiques varient entre 0.1" et 10".

### B.2.5 Rapport signal sur bruit

L'acquisition des données est non seulement perturbée par la traversée de l'atmosphère, du télescope et de l'instrument, mais également par le bruit de comptage des photons et le bruit de lecture du détecteur. La proportion de signal et de bruit est quantifiée par le rapport signal sur bruit (RSB).

Celui-ci peut être calculé de différentes manières à partir des données  $D(\mathbf{s}, \lambda)$  et de la variance du bruit  $\sigma^2(\mathbf{s}, \lambda)$  suivant ce que l'on souhaite évaluer :

– rapport signal sur bruit d'un pixel  $D(\mathbf{s}, \lambda)$  :  $\text{RSB}(\mathbf{s}, \lambda) = \frac{D^2(\mathbf{s}, \lambda)}{\sigma^2(\mathbf{s}, \lambda)}$  ;

– rapport signal sur bruit d'un spectre  $D_{\mathbf{s}}(\lambda)$ , pour une position spatiale  $\mathbf{s}$  donnée :

$$\text{RSB}(\mathbf{s}) = \frac{\sum_{\lambda} D_{\mathbf{s}}^2(\lambda)}{\sum_{\lambda} \sigma_{\mathbf{s}}^2(\lambda)}$$
 ;

– rapport signal sur bruit d'une image  $D_{\lambda}(\mathbf{s})$  calculé à la longueur d'onde  $\lambda$  :

$$\text{RSB}(\lambda) = \frac{\sum_{\mathbf{s}} D_{\lambda}^2(\mathbf{s})}{\sum_{\mathbf{s}} \sigma_{\lambda}^2(\mathbf{s})}$$
 ;

– rapport signal sur bruit d'un cube  $D(\mathbf{s}, \lambda)$  :  $\text{RSB} = \frac{\sum_{\mathbf{s}, \lambda} D^2(\mathbf{s}, \lambda)}{\sum_{\mathbf{s}, \lambda} \sigma^2(\mathbf{s}, \lambda)}$ .

Le rapport signal sur bruit peut être exprimé en décibels ; il est calculé à partir d'un rapport signal sur bruit standard par la relation :  $\text{RSB (dB)} = 10 \log_{10}(\text{RSB})$ .



# Annexe C

## Modélisation de la FSF

Avertissement : Tous les résultats conservés dans cet annexe ont été obtenus avec des simulations issues d'une version de PAOLA antérieure à la version 6.2 utilisée dans le chapitre II. Il est donc normal que l'on obtienne des résultats qui ne sont pas strictement identiques.

### C.1 Compléments pour la modélisation de la FSF en mode WFM avec optique adaptative

Cette section constitue un complément aux résultats présentés dans la section II.4 dans laquelle nous avons modélisé les variations des paramètres d'un modèle de Moffat elliptique variable dans le champ et par un modèle circulaire constant dans le champ. Nous avons voulu compléter cette modélisation de telle sorte que la variabilité de la FSF dans le champ soit décrite, sans tenir compte de l'ellipticité de la FSF. Pour cela, nous avons estimé la FSF par un modèle circulaire de FSF à chaque longueur d'onde  $\lambda$  et position dans le champ  $(\rho, \theta)$ . Puis nous avons modélisé les variations des paramètres de ce modèle en fonction de  $\lambda$  et  $\rho$ .

Le Tab. C.1 compare les performances en terme d'erreur quadratique  $\varepsilon_2$  des modèles circulaires dont les paramètres sont soit optimaux pour chaque  $\lambda$  et position  $(\rho, \theta)$ , soit modélisés en fonction de  $\lambda$  et  $\rho$ . Même si les données sont différentes de celles utilisées dans la section II.4, nous pouvons constater que la prise en compte de la variabilité de la FSF dans le champ permet de réduire de 4.65% à 3.28% l'erreur quadratique moyenne obtenue avec les paramètres optimaux. Avec les paramètres modélisés, l'amélioration est notable en terme d'erreur  $\varepsilon_2$  maximale (9.89% au lieu de 11.49%), ce qui signifie que l'erreur de modélisation de la FSF en bord de champ ( $\rho \geq 30''$ ) a été réduite. Néanmoins, les résultats pour ce modèle circulaire variable dans le champ sont moins satisfaisants que pour le modèle elliptique variable dans le champ avec  $\bar{\varepsilon}_2 = 2.69\%$  et  $\varepsilon_{2,max} = 4.71\%$ , soit une augmentation d'un facteur 2 environ.

Modèle de Moffat circulaire $M_{\alpha,\beta}(r) = \left[1 + \left(\frac{r}{\alpha}\right)^2\right]^{-\beta}$	$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$
Paramètres optimaux	3.28%	5.60%
Paramètres modélisés $\begin{cases} \alpha(\lambda, \rho, \theta) = 36.1182 - 0.0230 \lambda - 0.0263 \rho \\ \beta(\lambda, \rho, \theta) = 1.9502 - 4.6890 \lambda - 5.9320 \cdot 10^{-4} \rho \end{cases}$	5.07%	9.89%

TAB. C.1 – WFM avec optique adaptative : Performances du modèle circulaire dans le champ avec différentes valeurs des paramètres  $[\alpha, \beta, \gamma, \phi]$  : optimales ou modélisées. Valeur du *seeing* :  $s = 0.5''$ ,  $C_n^2$  médian. Moyennes et maxima des indices d'erreur sur  $\rho \in \{0; 20; 30; 40\}''$ ,  $\theta \in \{0; 15; 30; 45; 60; 75; 90\}$  degrés,  $\lambda \in \{465; 550; 650; 750; 850; 930\}$  nm.

## C.2 Influence du *seeing* sur la FSF en mode WFM avec optique adaptative

Dans la section II.3.4, nous avons étudié les variations des paramètres du modèle de Moffat pour l'approximation de la FSF en mode WFM sans optique adaptative. Dans cette section, nous voulons évaluer la pertinence du modèle de Moffat circulaire pour l'approximation de la FSF à différents *seeings* ainsi que l'influence du *seeing* sur la variation des paramètres du modèle.

La Fig. C.1 montre les variations des paramètres  $\alpha$  et  $\beta$  du modèle de Moffat circulaire constant dans le champ, en fonction de la longueur d'onde, pour différents *seeings*. En comparant cette figure à la Fig. II.10 p. 42, nous constatons que les paramètres varient davantage en fonction de la longueur d'onde lorsque le *seeing* augmente. En effet, si ces paramètres peuvent être raisonnablement approximatés par une fonction linéaire de  $\lambda$  pour  $s = 0.5''$ , ils devraient plutôt être approximatés par une fonction quadratique, pour les *seeings* plus élevés. Ce phénomène diffère du cas sans optique adaptative dans lequel le *seeing* agit essentiellement comme un facteur d'échelle. En effet, dans le cas présent, il faut tenir compte du fait que l'optique adaptative n'est pas aussi performante dans le bleu que dans le rouge.

Le Tab. C.2 rassemble les modèles de variation des paramètres  $\alpha$  et  $\beta$  en fonction de  $\lambda$  pour les différents *seeings* ainsi que les indices d'erreur entre la FSF et son approximation. Nous pouvons constater que les coefficients des termes linéaires et en particulier quadratiques augmentent avec le *seeing*. Ceci confirme l'observation précédente. L'indice d'erreur augmente légèrement avec le *seeing* mais change très peu lorsqu'on remplace les paramètres optimaux par leurs valeurs modélisées. Cela indique qu'il est possible d'approximer la FSF par un modèle dépendant d'un faible nombre de paramètres, avec une erreur limitée, quel que soit le *seeing*.

Notons que les résultats du Tab. C.2 concernant le *seeing*  $s = 0.5''$  ne peuvent pas être comparés à ceux du Tab. C.1. D'une part, les modélisations des paramètres  $\alpha$  et  $\beta$  sont différentes : celles du Tab. C.2 sont quadratiques en fonction de  $\lambda$  mais ne tiennent pas compte de la position  $\rho$ , alors que les précédentes sont bilinéaires en fonction de  $\lambda$  et  $\rho$ . D'autre part, puisque les données correspondent à des positions dans le champ polaires pour le Tab. C.1 et cartésiennes pour le Tab. C.2, les données sur lesquelles s'appuient les modélisations sont différentes.

Modèle MF circulaire	Statistiques sur $\{x_i, y_j, \lambda_k, \}_{i,j,k}$					
	$\bar{\varepsilon}_2$	$\varepsilon_{2,max}$	$\bar{\varepsilon}_1$	$\varepsilon_{1,max}$	$\bar{\varepsilon}_\infty$	$\varepsilon_{\infty,max}$
Paramètres optimaux pour $s = 0.5''$ $\begin{cases} \alpha(\lambda) = 46.81 - 5.40 \cdot 10^{-2} \lambda + 2.20 \cdot 10^{-5} \lambda^2 \\ \beta(\lambda) = 1.91 - 3.37 \cdot 10^{-4} \lambda - 7.67 \cdot 10^{-8} \lambda^2 \end{cases}$	3.30%	5.52%	5.31%	8.30%	3.62%	9.10%
	3.34%	5.70%	5.33%	8.34%	3.66%	10.10%
Paramètres optimaux pour $s = 0.7''$ $\begin{cases} \alpha(\lambda) = 88.87 - 1.18 \cdot 10^{-1} \lambda + 5.39 \cdot 10^{-5} \lambda^2 \\ \beta(\lambda) = 2.28 - 1.2 \cdot 10^{-3} \lambda + 4.54 \cdot 10^{-7} \lambda^2 \end{cases}$	3.10%	5.71%	4.39%	6.74%	3.03%	7.03%
	3.15%	5.89%	4.41%	6.73%	3.19%	8.09%
Paramètres optimaux pour $s = 1.0''$ $\begin{cases} \alpha(\lambda) = 181.24 - 3.25 \cdot 10^{-1} \lambda + 1.71 \cdot 10^{-4} \lambda^2 \\ \beta(\lambda) = 3.11 - 3.8 \cdot 10^{-3} \lambda + 2.14 \cdot 10^{-6} \lambda^2 \end{cases}$	4.61%	6.86%	9.37%	11.95%	5.53%	12.41%
	4.70%	7.20%	9.40%	12.63%	5.76%	13.57%
Paramètres optimaux pour $s = 1.3''$ $\begin{cases} \alpha(\lambda) = 260.11 - 4.19 \cdot 10^{-1} \lambda + 2.00 \cdot 10^{-4} \lambda^2 \\ \beta(\lambda) = 3.36 - 3.8 \cdot 10^{-3} \lambda + 1.87 \cdot 10^{-6} \lambda^2 \end{cases}$	4.43%	6.91%	7.71%	11.57%	5.54%	12.34%
	4.45%	6.83%	7.72%	11.59%	5.58%	12.67%

TAB. C.2 – Mode WFM avec optique adaptative : Performances du modèle circulaire constant dans le champ à différents *seeings*, pour un  $C_n^2$  médian - Statistiques des indices d'erreur sur  $(x, y) \in [0, 10, 20, 30]^2$  arcsec,  $\lambda \in \{465; 550; 650; 750; 850; 930\}$  nm - Comparaison aux performances du modèle de Moffat basé sur les paramètres modélisés par une fonction quadratique de  $\lambda$ .

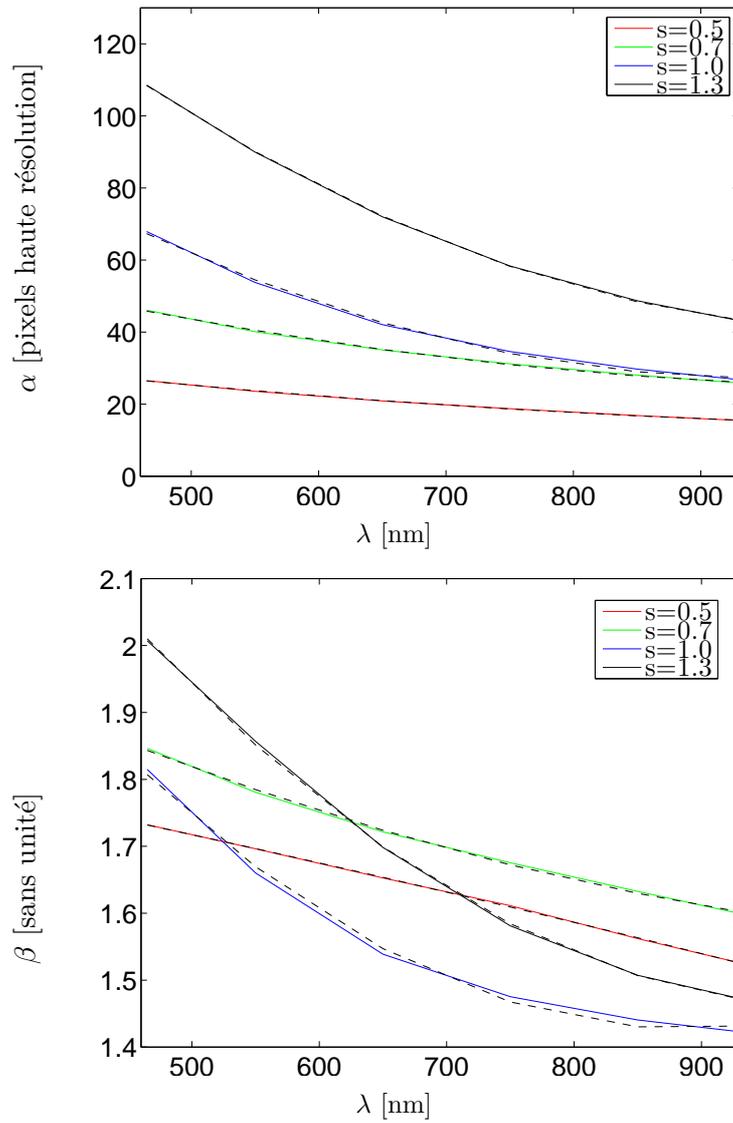


FIG. C.1 – WFM avec optique adaptative : Évolution des paramètres  $\alpha$  et  $\beta$  en fonction de la longueur d'onde  $\lambda$  pour différentes valeurs du *seeing*  $s$  dans le cadre d'une modélisation par une fonction de Moffat circulaire constante dans le champ - Superposition aux modèles quadratiques d'évolution de ces paramètres [tirets].



## Annexe D

# Compléments méthodologiques

### D.1 Méthode des moments : extension

L'objectif est de déterminer le flux, la position et la largeur notés respectivement  $[a, c, w]$  et  $[a_2 = \alpha \cdot a, c_2 = c + \Delta, w_2 = \beta \cdot w]$  de deux raies voisines non identiques, pour lesquelles les paramètres  $\alpha$ ,  $\Delta$  et  $\beta$  sont connus. Le modèle de doublet est donné par la somme des deux raies.

La première étape est d'évaluer le flux, la position et la largeur de ce doublet notés  $a_d$ ,  $c_d$  et  $w_d$ . On peut alors exprimer ces paramètres en fonction de ceux des raies individuelles :

**Flux** Le flux du doublet est la somme du flux de chaque raie :  $a_d = \sum_{\lambda} \mathcal{R}_d(\lambda) = a + \alpha \cdot a$ , alors  $a_d = (1 + \alpha)a$ .

On en déduit l'estimateur du flux en fonction de celui du flux du doublet :  $\hat{a} = \frac{\widehat{a}_d}{1+\alpha}$  et  $\hat{a} = \frac{\alpha \cdot \widehat{a}_d}{1+\alpha}$ .

**Position** Le moment d'ordre 1 donne le barycentre des centres des deux raies :

$$c_d = \frac{1}{a_d} \sum_{\lambda} \lambda \mathcal{R}_d(\lambda) = \frac{a}{a_d} c + \frac{\alpha \cdot a}{a_d} (c + \Delta). \text{ On obtient donc } c_d = c + \frac{\alpha}{1 + \alpha} \Delta.$$

On en déduit :  $\hat{c} = \widehat{c}_d - \frac{\alpha}{1 + \alpha} \Delta$  et  $\widehat{c}_2 = \widehat{c}_d + \frac{1}{1 + \alpha} \Delta$ .

**Largeur**

$$\begin{aligned} w_d^2 &= \frac{1}{a_d} \sum_{\lambda} \lambda^2 \mathcal{R}_d(\lambda) - c_d^2 \\ &= \frac{1}{(1 + \alpha)a} \sum_{\lambda} \lambda^2 (\mathcal{R}_1(\lambda) + \mathcal{R}_2(\lambda)) - \left( \frac{c + \alpha c_2}{1 + \alpha} \right)^2 \\ &= \frac{1}{1 + \alpha} \left( \frac{1}{a} \sum_{\lambda} \lambda^2 \mathcal{R}_1(\lambda) \right) + \frac{\alpha}{1 + \alpha} \left( \frac{1}{\alpha \cdot a} \sum_{\lambda} \lambda^2 \mathcal{R}_2(\lambda) \right) - \frac{c^2}{(1 + \alpha)^2} - \frac{\alpha^2 c_2^2}{(1 + \alpha)^2} - \frac{2\alpha c c_2}{(1 + \alpha)^2} \\ &= \frac{1}{1 + \alpha} \left( \frac{1}{a} \sum_{\lambda} \lambda^2 \mathcal{R}_1(\lambda) - c^2 \right) + \frac{\alpha}{1 + \alpha} \left( \frac{1}{\alpha \cdot a} \sum_{\lambda} \lambda^2 \mathcal{R}_2(\lambda) - c_2^2 \right) \\ &\quad + \frac{\alpha c^2}{(1 + \alpha)^2} + \frac{\alpha c_2^2}{(1 + \alpha)^2} - \frac{2\alpha c c_2}{(1 + \alpha)^2} \\ &= \frac{w^2 + \alpha w_2^2}{1 + \alpha} + \frac{\alpha c^2 + \alpha c_2^2 - 2\alpha c c_2}{(1 + \alpha)^2} \\ &= \frac{(1 + \alpha \beta^2)w^2}{1 + \alpha} + \frac{\alpha \Delta^2}{(1 + \alpha)^2} \end{aligned}$$

On en déduit l'expression des estimateurs de la largeur au carré de chaque raie :  $\widehat{w}^2 = \left( \widehat{w}_d^2 - \frac{\alpha \Delta^2}{(1+\alpha)^2} \right) \frac{1+\alpha}{1+\alpha\beta^2}$   
 et  $\widehat{w}_2^2 = \left( \widehat{w}_d^2 - \frac{\alpha \Delta^2}{(1+\alpha)^2} \right) \frac{(1+\alpha)\beta^2}{1+\alpha\beta^2}$ .

	Flux	Position	Largeur
Raie 1	$\widehat{a} = \frac{\widehat{a}_d}{1+\alpha}$	$\widehat{c} = \widehat{c}_d - \frac{\alpha}{1+\alpha} \Delta$	$\widehat{w} = \sqrt{\left( \widehat{w}_d^2 - \frac{\alpha \Delta^2}{(1+\alpha)^2} \right) \frac{1+\alpha}{1+\alpha\beta^2}}$
Raie 2	$\widehat{a}_2 = \frac{\alpha \cdot \widehat{a}_d}{1+\alpha}$	$\widehat{c}_2 = \widehat{c}_d + \frac{1}{1+\alpha} \Delta$	$\widehat{w}_2 = \beta \sqrt{\left( \widehat{w}_d^2 - \frac{\alpha \Delta^2}{(1+\alpha)^2} \right) \frac{1+\alpha}{1+\alpha\beta^2}}$

Si les deux raies ont la même largeur ( $\beta = 1$ ) et le même flux ( $\alpha = 1$ ), on retrouve  $\widehat{a} = \frac{\widehat{a}_d}{2}$ ,  $\widehat{c} = \widehat{c}_d \pm \frac{\Delta}{2}$   
 et  $\widehat{w} = \sqrt{\widehat{w}_d^2 - \left( \frac{\Delta}{2} \right)^2}$ .

## D.2 Intégration du spectre de l'étoile pour l'estimation de la FSF par espérance *a posteriori*

Dans la section III.5.3, nous proposons d'estimer conjointement le spectre  $s(\lambda)$  d'une étoile isolée et les paramètres  $\boldsymbol{\theta}$  du modèle de FSF à partir des données d'observation de cette étoile. Une autre solution consiste à intégrer le spectre de l'étoile hors du problème, comme cela a été fait dans [DOU 01]. Il s'agit alors de calculer la loi *a posteriori*  $p(\mathbf{D}|\boldsymbol{\theta})$  par marginalisation de  $p(\boldsymbol{\theta}, \{s(\lambda)\})_\lambda | \mathbf{D}$ .

Rappelons que la vraisemblance des données  $\mathbf{D}$  s'écrit :

$$p(\mathbf{D}|\boldsymbol{\theta}) \propto \exp \left( -\frac{1}{2} \|\mathbf{D} - s(\lambda) \cdot m(x, y, \lambda, \boldsymbol{\theta})\|_{\mathbf{\Gamma}^{-1}}^2 \right)$$

où  $m$  est le modèle de FSF de paramètres  $\boldsymbol{\theta}$  et  $\mathbf{\Gamma}$  la matrice diagonale contenant la variance du bruit  $\sigma^2(x, y, \lambda)$ .

La vraisemblance  $p(\mathbf{d}_\lambda|\boldsymbol{\theta})$  des données pour une longueur d'onde  $\lambda$  donnée est proportionnelle à

$$\exp \left( -\frac{1}{2} \|\mathbf{d}_\lambda - s \mathbf{m}_\lambda\|_{\mathbf{\Gamma}_\lambda^{-1}}^2 \right)$$

où  $\mathbf{d}_\lambda$  est un vecteur contenant les données  $D_\lambda(x, y)$ ,  $s$  est la valeur du spectre à la longueur d'onde fixée,  $\mathbf{m}_\lambda$  est un vecteur contenant le modèle de FSF de paramètres  $\boldsymbol{\theta}$  calculé à la longueur d'onde  $\lambda$  et  $\mathbf{\Gamma}_\lambda$  est une matrice diagonale contenant la variance du bruit correspondante. Le terme d'erreur quadratique aux données peut s'écrire de la manière suivante :

$$\|\mathbf{d}_\lambda - s \mathbf{m}_\lambda\|_{\mathbf{\Gamma}_\lambda^{-1}}^2 = (\mathbf{d}_\lambda - s \mathbf{m}_\lambda)^T \mathbf{\Gamma}_\lambda^{-1} (\mathbf{d}_\lambda - s \mathbf{m}_\lambda) = \mathbf{d}_\lambda^T \mathbf{\Gamma}_\lambda^{-1} \mathbf{d}_\lambda + (s - s_0) A (s - s_0) - s_0^2 A$$

avec  $A = \mathbf{m}_\lambda^T \mathbf{\Gamma}_\lambda^{-1} \mathbf{m}_\lambda$  et  $s_0 = A^{-1} \mathbf{m}_\lambda^T \mathbf{\Gamma}_\lambda^{-1} \mathbf{d}_\lambda$ .

Avec une loi *a priori* uniforme sur le spectre, la loi *a posteriori* conjointe du spectre à la longueur d'onde  $\lambda$  et des paramètres est :

$$p(s, \boldsymbol{\theta} | \mathbf{d}_\lambda) \propto p(\boldsymbol{\theta}) \exp \left( -\frac{1}{2} ((s - s_0) A (s - s_0) - s_0^2 A + \mathbf{d}_\lambda^T \mathbf{\Gamma}_\lambda^{-1} \mathbf{d}_\lambda) \right).$$

Notons qu'avec une loi *a priori* gaussienne  $\mathcal{N}(0, \sigma^2)$  sur le spectre, cette loi *a posteriori* conjointe a la même expression, en remplaçant  $A$  par  $A' = \mathbf{m}_\lambda^T \mathbf{\Gamma}_\lambda^{-1} \mathbf{m}_\lambda + \frac{1}{\sigma^2}$ .

En marginalisant l'expression de la loi *a posteriori* conjointe par rapport au spectre, on obtient :

$$\begin{aligned}
 p(\boldsymbol{\theta}|\mathbf{d}_\lambda) &= \int p(\boldsymbol{\theta}, s|\mathbf{d}_\lambda) ds \\
 &= p(\boldsymbol{\theta}) \exp\left(-\frac{1}{2}(-s_0^2 A + \mathbf{d}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{d}_\lambda)\right) \int \exp\left(-\frac{1}{2}(s - s_0)A(s - s_0)\right) ds \\
 &= p(\boldsymbol{\theta}) \exp\left(-\frac{1}{2}(-s_0^2 A + \mathbf{d}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{d}_\lambda)\right) \sqrt{2\pi A^{-1}}
 \end{aligned}$$

La loi *a posteriori* des paramètres conditionnellement à l'intégralité des données  $\mathbf{D}$  correspond au produit de la loi *a posteriori* marginale précédente, pour l'ensemble des longueurs d'onde :

$$\begin{aligned}
 p(\boldsymbol{\theta}|\mathbf{D}) &\propto p(\boldsymbol{\theta}) \prod_{\lambda} p(\mathbf{d}_\lambda|\boldsymbol{\theta}) \\
 &\propto p(\boldsymbol{\theta}) \frac{\sqrt{2\pi}^{N_\lambda}}{\sqrt{\prod_{\lambda} \mathbf{m}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{m}_\lambda}} \exp\left(-\frac{1}{2} \sum_{\lambda} \mathbf{d}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{d}_\lambda\right) \exp\left(\frac{1}{2} \sum_{\lambda} \frac{(\mathbf{m}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{d}_\lambda)^2}{\mathbf{m}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{m}_\lambda}\right) \\
 &\propto p(\boldsymbol{\theta}) \frac{1}{\sqrt{\prod_{\lambda} \mathbf{m}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{m}_\lambda}} \exp\left(\frac{1}{2} \sum_{\lambda} \frac{(\mathbf{m}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{d}_\lambda)^2}{\mathbf{m}_\lambda^T \boldsymbol{\Gamma}_\lambda^{-1} \mathbf{m}_\lambda}\right)
 \end{aligned}$$

L'échantillonnage du vecteur  $\boldsymbol{\theta}$  selon cette loi *a posteriori* peut être réalisé par la méthode de Metropolis-Hastings. Un estimateur de  $\boldsymbol{\theta}$  est l'espérance *a posteriori*, calculée par la moyenne des échantillons générés. Cette méthode a l'avantage de ne pas faire apparaître le spectre de l'étoile qui constitue dans les autres méthodes un paramètre de nuisance.



# Deconvolution of hyperspectral data for VLT instrument MUSE

Emma VILLENEUVE

---

## Abstract :

Hyperspectral instruments can acquire images simultaneously at a large number of wavelengths. Processing and analysing such data is a challenge for the signal-image community. For ground-based observations, one cornerstone is the Point Spread Function (PSF) and particularly its spatial component (FSF). During this thesis, we tackled a hyperspectral deconvolution problem, which requires to model and estimate the FSF.

First, we looked into the modelling of the FSF at every position of the hyperspectral data cube. We showed that this FSF can be approximated, at each wavelength, by a Moffat function. Without a correction by adaptive optics, modelling the scale and shape parameters of the Moffat function by a linear function with respect to the wavelength allows the FSF to be well approximated in the whole cube, with only 4 parameters.

Then, we studied the FSF estimation for all wavelengths, based on the former model, from the observation of an isolated star. We proposed several estimation strategies, using maximum likelihood and posterior mean estimators for the FSF model parameters, as well as for the nuisance parameters, such as the star's spectrum and position.

Finally, we focused on hyperspectral deconvolution for the study of galaxies kinematics. Thanks to the Doppler effect, this issue is reflected by the estimation of flux maps, position and width from an emission line in the galaxy spectrum, taking into account the 3D PSF. We proposed to estimate these parameters with posterior mean estimators, which are computed with Monte Carlo Markov Chains (MCMC) algorithms.

---

**Keywords :** signal processing, image processing, hyperspectral data, point spread function, estimation, deconvolution, Monte-Carlo Markov Chains, galaxy kinematics.

---

# Déconvolution de données hyperspectrales pour l'instrument MUSE du VLT

**Auteur :** Emma VILLENEUVE

**Directeurs de Thèse :** Hervé CARFANTAN, Eric ANTERRIEU

---

**Date et lieu de la soutenance :** le 10 Décembre 2012 à l'Observatoire Midi-Pyrénées.

---

## Résumé :

Le traitement et l'analyse de cubes de données hyperspectrales, tels que fournis par le spectro-imageur MUSE capable d'acquérir simultanément des images à un grand nombre de longueurs d'onde, est un challenge pour la communauté signal-image. Une difficulté particulière pour les observations depuis le sol est due à la réponse de la chaîne d'acquisition, plus précisément sa composante spatiale FSF. Cette thèse aborde un problème de déconvolution de cubes hyperspectraux, nécessitant au préalable la modélisation et l'estimation de la FSF.

Nous montrons que la FSF, due à l'instrument et aux turbulences atmosphériques éventuellement corrigées par l'optique adaptative (OA), peut être modélisée en tout point du cube par une fonction de Moffat. Dans le cas sans OA, une variation spectrale linéaire des paramètres de cette fonction permet une bonne approximation de la FSF dans tout le cube, avec seulement quatre paramètres.

Puis nous étudions l'estimation des paramètres de ce modèle de FSF pour l'ensemble des longueurs d'onde, à partir de l'observation d'une étoile isolée. Nous proposons des schémas d'estimation, au sens du maximum de vraisemblance et de la moyenne a posteriori, pour ces paramètres ainsi que pour des paramètres de nuisance tels que le spectre et la position de l'étoile.

Enfin, nous abordons un problème de déconvolution hyperspectrale pour l'étude de la cinématique de galaxies. Par effet Doppler, ce problème se traduit en l'estimation de cartes de flux, de position et de largeur d'une raie spectrale d'émission de la galaxie. Pour cela, nous proposons des estimateurs de type moyenne a posteriori calculés par des simulation de Monte Carlo par Chaînes de Markov (MCMC).

---

**Mots-clés :** Traitement du signal, traitement des images, données hyperspectrales, réponse instrumentale, estimation, déconvolution, Monte-Carlo Markov Chains, cinématique des galaxies.

---

**Discipline :** Traitement du signal et des images.

---

Institut de Recherche en Astrophysique et Planétologie - UMR 5277  
Observatoire Midi-Pyrénées  
14, Avenue Edouard Belin  
31400 Toulouse - FRANCE