Spatio spectral reconstruction from low resolution multispectral data : application to the Mid-Infrared instrument of the James Webb Space Telescope

Mohamed Elamine Hadj-Youcef

université PARIS-SACLAY

NNT : 2018SACLS326

Thèse de doctorat

Spatio spectral reconstruction from low resolution multispectral data. Application to the Mid-Infrared instrument of the James Webb Space Telescope

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud au sein du
Laboratoire des Signaux et Systèmes (L2S) et de l'Institut d'Astrophysique Spatial (IAS)

Ecole doctorale n°580
Sciences et Technologies de l'Information et de la Communication (ED STIC)
Spécialité de doctorat : Traitement du signal et des images

Thèse présentée et soutenue à Gif-sur-Yvette, le 27/09/2018, par

MOHAMED ELAMINE HADJ-YOUCEF

Composition du Jury :

Thomas Rodet Professeur, École Normale Supérieure de Cachan (SATIE)	Président
André Ferrari Professeur, Université de Nice Sophia Antipolis (Laboratoire Lagrange)	Rapporteur
Sylvie Roques Directrice de Recherche, Université de Toulouse (IRAP)	Rapporteur
Éric Thièbaut Astronome, Observatoire de Lyon (CRAL)	Examinateur
Said Moussaoui Professeur, Ecole Centrale de Nantes (LS2N)	Examinateur
Aurélia Fraysse Maître de conférence, Université Paris-Sud (L2S)	Directrice de thèse
François Orieux Maître de conférence, Université Paris-Sud (L2S)	Encadrant
Alain Abergel Professeur, Université Paris-Sud (IAS)	Co-encadrant

Thesis conducted at

Laboratoire des Signaux et Systèmes (L2S)
CentraleSupélec, bât. Bréguet, 3, rue Joliot Curie
91190 Gif-sur-Yvette, France
http://www.l2s.centralesupelec.fr/

Institut d'Astrophysique Spatiale (IAS)
Université Paris-Sud, bât. 121, Rue Jean-Dominique Cassini
91405 Orsay, France
https://www.ias.u-psud.fr/

Funded by

Thesis duration 36 months

Reconstruction spatio-spectrale à partir de données multispectrales basse résolution. Application à l'instrument infrarouge moyen du Télescope spatial James Webb.

Mohamed Elamine HADJ-YOUCEF.

Résumé

Cette thèse s'inscrit dans le contexte de résolution d'un problème inverse pour l'astrophysique. L'objectif est de reconstruire un objet tridimensionnel, ayant une distribution spatiale et spectrale, à partir d'un ensemble d'images multi-spectrales de basse résolution fournies par l'imageur MIRI (Mid-InfraRed Instrument) à bord du prochain télescope spatial James Webb Space Telescope (JWST). Le JWST est le télescope spatial de la prochaine décennie. Il est développé en collaboration internationale entre la NASA, l'agence spatiale européenne (ESA) et l'agence spatiale canadienne (CSA). Le télescope est doté d'un grand miroir primaire (6,5 mètres) et il couvre une large bande spectrale (0,6 à 30 microns). Le lancement du JWST est prévu en mars 2021.

Les images multi-spectrales observées souffrent de plusieurs dégradations. Premièrement, une dégradation spatiale due au produit de convolution de l'objet d'intérêt avec la réponse optique dépendante de la longueur d'onde, appelée fonction d'étalement (où PSF-Point Spread Function). La résolution spatiale des images est donc limitée sous forme d'un flou spatial qui dépend de la longueur d'onde. Cet effet est dû à la diffraction de flux de photons (ou objet d'intérêt) sur le plan focal du télescope. Deuxièmement les images multispectrales portent très peu d'informations spectrales sur l'objet observé en raison d'une dégradation spectrale due au filtrage et l'intégration du flux de photons (objet d'intérêt) sûr de larges bandes spectrales par le détecteur infrarouge de l'imageur MIRI.

La reconstruction de l'objet d'intérêt original est un problème mal posé en raison du manque important d'informations spectrales dans l'ensemble de données multi-spectrales. La difficulté se pose alors dans le choix d'une représentation adapté de l'objet permettant la reconstruction de l'information spatiale et spectrale. Un modèle classique utilisé dans l'état de l'art repose sur l'utilisation d'une PSF moyennée sur la largeur spectrale de chaque bande. En effet, ce choix de PSF néglige la variation spectrale au sein d'une bande spectrale, et donc introduisant des erreurs dans le modèle de l'instrument spécialement pour de larges bandes. Cependant, ce modèle simpliste est convenqble que dans le cas d'un imageur à une bande spectrale très étroite, ce qui n'est pas le cas pour l'imageur de MIRI.

L'approche de reconstruction que nous avons proposée consiste à développer une méthode pour l'inversion qui se résume en quatre étapes: (1) Concevoir un modèle de l'instrument reproduisant les données multi-spectrales observées, prenant en compte la variation spectrale de la PSF et l'intégration spectrale sur de larges bandes. (2) Utiliser un modèle de l'objet adapté à la reconstruction. En effet, deux modèles ont été utilisés, un modèle multi-longueur d'onde et un modèle de mélange linéaire. (3) Exploiter conjointement l'ensemble des images multi-spectrales pour reconstruire la totalité de l'information spectrale et prendre en compte l'inter-corrélation entre les images, et enfin (4) développer des méthodes de reconstruction, efficaces et rapides, basées sur la minimisation des critères convexes en introduisant des priori à la solution. En effet, trois algorithmes de

reconstruction sont proposées.

Les résultats de reconstruction de l'objet spatio-spectral sont obtenus en utilisant neuf images multi-spectrales, simulées avec le modèle proposé de l'imageur de MIRI. Une augmentation significative des résolutions de l'objet, spatiale et spectrale, a été obtenu en utilisant le modèle de mélange linéaire. Où l'objet reconstruit montre l'effet de débruitage et de déconvolution sur la totalité des longueur d'ondes. De plus, nous avons obtenu une erreur relative n'excédant pas 5% à 30 dB et un temps d'exécution de 1 seconde pour l'algorithme de norm-l_2 et 20 secondes (avec 50 itérations) pour l'algorithme norm-l_2/l_1. C'est qui est 10 fois plus rapide que la solution classique calculée par l'algorithme de gradient conjugué.

Acknowledgments

First and foremost I would like to express my deep gratitude to my supervisors *François Orieux*, *Aurélia Fraysse*, and *Alain Abergel*. Thank you for your help, ideas, and discussions. I have learned a lot from your experience which helped me strengthening my knowledge and skills in a multidisciplinary environment.

I would like to express my deep gratitude for the reviewers, *Sylvie Roques* and *André Ferrari*, for the time they invested in reading my manuscript and provide me fruitful remarks. I also thank the examiners *Thomas Rodet*, *Éric Thiébaut* and *Said Moussaoui*, for evaluating my work and providing me constructive critics.

I would like to thank the *Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation* for having funded and supported my thesis during three years.

During my Ph.D. I met wonderful people with whom I spent valuable time. I would like to thank all member of the inverse problems group (L2S-GPI): *Mircea Dumitru*, *Thi Thanh Nguyen*, *Camille Chapdelaine*, *Li Wang*, *Mickaël Seznec*, *Alina Meresescu*, *Fangchen Feng*, *Guillaume Revillon*, *Maxime Martelli*, *Nicolas Gac*, *Charles Soussen*. In addition, I would like to thank my colleagues from L2S, *Patrice Brault* and *Messaoud Benidir*. A special thank goes to *Ali Mohammad-Djafari* for all the nourishing conversations. I would like also to thank my friends from the IAS, *Sahar Hassanzadeh*, *Gururaj Wagle* and *Jean-baptist Durrive*.

I devote this thesis to my lovely parents. They taught well and support me from the beginning. Words are not enough to thank you.

Although it is known that the distance separate people, but this remains false when it comes to my dear and best uncles, *Djillali Hadj-Youcef* and *Amar Arbaoui*. Thank you for your helps an supports from the bottom of my hearts.

Best Success.
M.Amine H-Y.

Contents

Chapter 1

Introduction

Contents

1.1 Context

The James Webb Space Telescope (JWST)[1] is the next flagship space telescope of NASA (National Aeronautics and Space Administration) with the collaboration of ESA (European Space Agency) and CSA (Canadian Space Agency) to be launched in 2021. The JWST mission will offer a unique combination of imagers and spectrometers in a broad range of wavelengths. After its launch, the JWST will be the largest telescope in space with a primary mirror diameter of 6.5 meters compared to 2.4 meters for Hubble telescope[2]. In this work, we are interested in the Mid-InfraRed Instrument (MIRI). It contains an imager and an Integral Field Unit (IFU) spectrometer both working between 5 and 28 microns. More specifically, in this thesis, we are interested in a problems related to the Imager of MIRI [Bouchet *et al.* 2015].

During the acquisition process of the 2D+λ spatio-spectral object of interest the telescope focuses the beam of photon flux at the focal plane. Unfortunately, the formed image is affected by diffraction. The result is modeled by a 2D spatial convolution with the optical response or PSF (Point Spread Function), which degrades the spatial resolution of the object in the form of a blur varying according to the wavelength. This degradation is well-known in diffraction theory [Goodman 2005]. The width of the PSF linearly increases with the wavelength. Then the blurred object is filtered spectrally over a broad band, and then integrated and sampled by the 2D detector of the MIRI imager. The main issue arising here is a severe degradation of the spectral resolution because of the integration (filter and detector response) adding to the limitation of the number of bands, e.g. nine bands for the imager of MIRI.

[1] https://www.jwst.nasa.gov/
[2] http://hubblesite.org/

The multispectral data observed by the imager of MIRI has the advantage of being rich in spatial information of the object due to the relatively wide field of view (FOV = 74" × 113" where " stands for arcseconds) compared to the FOV of a spectrometer. But they are poor in spectral resolution because of the spectral integration over broad band ($\lambda/\triangle\lambda \approx 5$). On the other hand, hyperspectral data observed by the Medium Resolution Spectrometer (MRS) of MIRI are rich in spectral information ($\lambda/\triangle\lambda \approx 1550 - 3250$) but are limited in spatial information because of the small FOV ($\sim$ 3.5" × 3.5"). In Figure 1.1 we illustrate the lack of spectral information in the multispectral data by showing the system input and output for an observation using the MIRI imager. The exploration and understanding of astrophysical objects are essentially based on the processing and analysis of the observed data. Therefore, it is important to consider the degradation of the measuring instrument and develop sophisticated methods for reconstruction of the best possible high-resolution spatio-spectral object without the instrument degradation.

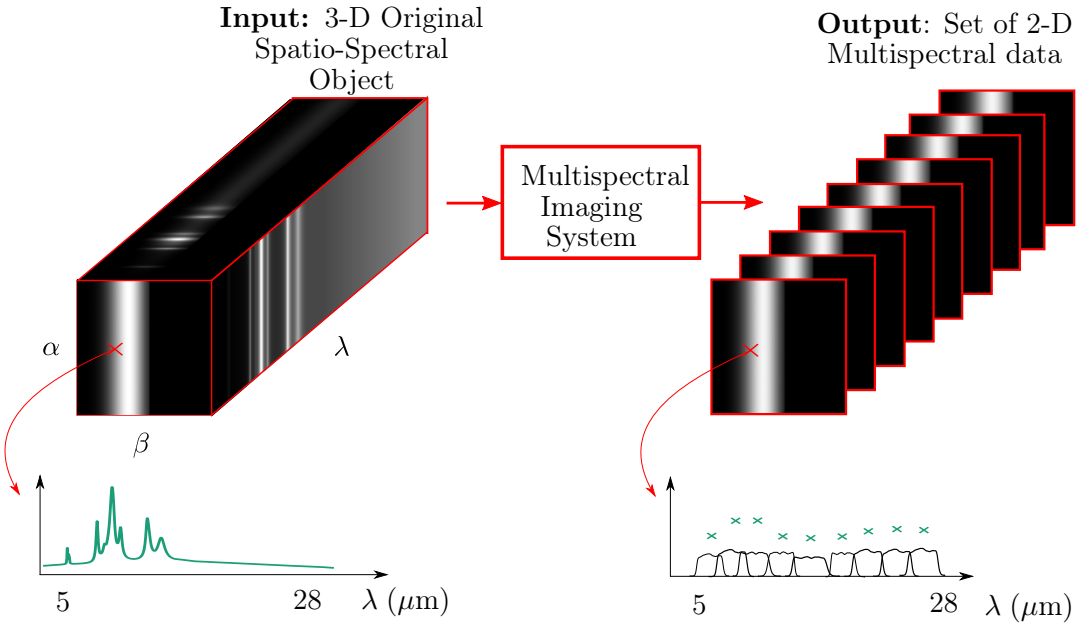


Figure 1.1: Illustration of the input and output for a multispectral imaging system such as the imager MIRI of the JWST [Bouchet *et al.* 2015]. The input is a 2D+λ continuous object having spatial and spectral distributions and the output a set of 2D low-resolution multispectral data degraded by the instrument response.

For the sake of clarity here is the meaning of words mostly used throughout the thesis. The *spatio-spectral object* is the object of interest we want to reconstruct. It has two spatial dimensions and one spectral dimension or 2D+λ, where λ symbolizes the wavelengths. *Multispectral data* is an observed image acquired by the observing instrument or multispectral imaging system. It results from the spectral integration of the object on a specific wavelength range. We use the word *channel* to mention a slice from the 2D+λ cube of the object at a particular wavelength. The word *band* is used to refer to the spectral response of the instrument. It is associated with the spectral transmission of the

filter and the spectral response of a detector. Bold-face capital letters refer to matrices (e.g. $\boldsymbol{H}$) and bold-face lower-case letters refer to vectors (e.g. $\boldsymbol{x}, \boldsymbol{y}$).

1.2 Objective

This thesis deals with an inverse problem in astronomy. The objective is the development of methods for the reconstruction of a 2D+λ astrophysical object with high spatial and spectral resolution from a set of 2D multispectral data with limited resolution obtained by the observation instrument.

The data processing is challenging because of the complex response of the segmented mirror of the JWST which varies according to the wavelength due to the diffraction, and the detector behavior which performs the spectral integration over broad bands and the spatial sampling. A combination of knowledge of the object and the observation system is necessary. In addition, a joint processing of multispectral data at multiple bands is required to reconstruct the object over the whole wavelengths range.

1.3 Methodology

In this thesis we propose to follow the methodology summarized in Figure 1.2. The reconstruction of a discrete version $\widehat{\phi}$ of the spatio-spectral object ϕ is achieved by building four major stages. The first stage is for the instrument model; this is a mathematical model able to reproduce the observed multispectral data $\boldsymbol{y}$ by accounting for the instrument response such as the optical response, the spectral filtering, the detector sampling and integration. Once the instrument model is developed we are able to simulate multispectral data for any object of interest and use it later for the reconstruction. The second stage is for the object model; it is the key model that restricts the object reconstruction to specific spatial and spectral distributions. In this thesis two models are used - the multiwavelength model and the mixture model. At the third stage comes the forward model; this model results from the instrument and object models, and it represents the joint processing of the whole multispectral dataset $\boldsymbol{y}$ in terms of the unknown parameter $\boldsymbol{x}$ (that describes the object ϕ) through the observation matrix $\boldsymbol{H}$. Next, we deal with the reconstruction in the fourth stage; it consists of the reconstruction of the spatio-spectral object using the description of the forward model and some additional priors information about the object, which are necessary in order to stabilize the solution.

1.4 Contributions

This section highlights my main contributions during the thesis. I first developed a mathematical model for a diffraction-limited instrument such as the MIRI imager on board the JWST. This model is used to develop the next stages of the inverse problem framework such as the forward model and the reconstruction. It is mainly adapted for the inversion algorithm, i.e. it allows the extraction of the observation matrix and its transpose. It

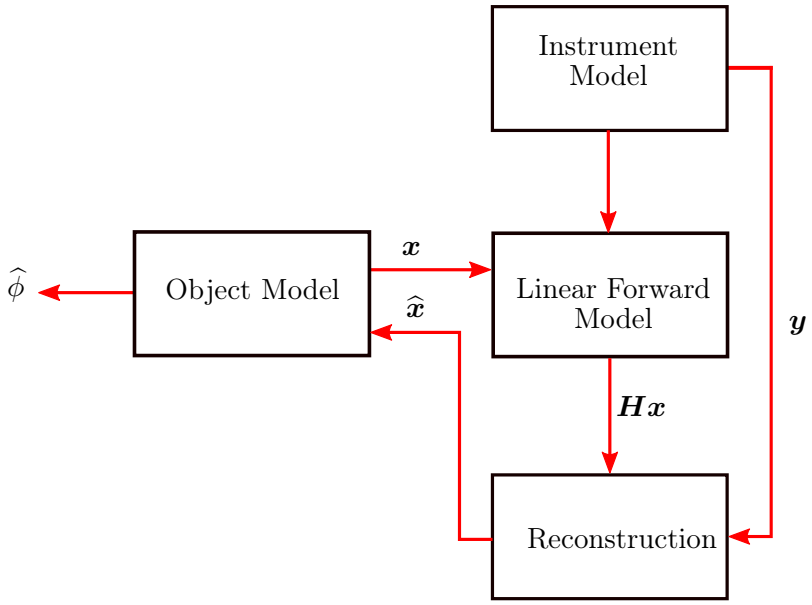


Figure 1.2: The proposed inverse problem paradigm for the reconstruction of a discrete 2D+λ object $\widehat{\phi}$ from a set of multispectral dataset $\boldsymbol{y}$. Here $\boldsymbol{x}$ represents an unknown parameter of the object model and $\boldsymbol{H}$ is the observation matrix.

accounts for the spectral shift-variant non-stationary optical response, the spectral integration over a set of P broad bands (or spectral responses) by the detector, the spatial sampling of the object over the detector matrix, and an additive term that accounts for modeling errors and noise generated by the detector.

Next, I represent the object of interest with two different linear models using firstly a piecewise linear function and secondly a mixing model. Both models preserve the spatial and spectral distribution of the object. Then I proposed two linear forward models where the discrete multichannel unknown object $\boldsymbol{x}$ is related to the discrete multispectral dataset $\boldsymbol{y}$ through an observation matrix $\boldsymbol{H}$. The contents and the meaning of $\boldsymbol{x}$ and $\boldsymbol{H}$ are specific for each object model. We end up with two linear forward models which are formulated by a sum of 2D convolutions between the elements of $\boldsymbol{x}$ and $\boldsymbol{H}$.

Several reconstruction methods are proposed based on the regularized least-squares method, one for the first object model and two for the second one. The solution is obtained as the minimizer of a convex cost function. The reconstruction problem is ill-posed and this is mainly due to the ill-conditioning of the Hessian matrix. This ill-posedness is corrected by adding different prior information of type l_2-norm and l_2/l_1-norm to enforce spatial smoothness or high spatial gradient of the solution respectively. For test and validation of algorithms, I simulated multispectral data using the developed instrument model of the JWST/MIRI imager from three spatio-spectral objects, one is a simplified model of an astrophysical object, and the others are synthetics objects having different spatial and spectral distributions. In addition, the multispectral dataset is corrupted with an additive Gaussian white noise of different levels. The overall reconstruction results

obtained on simulated data show a significant increase of spatial and spectral resolutions of the reconstructed object compared to the classical method.

1.5 Outline of the thesis

The following chapters of this thesis are divided as follows. Chapter 2 presents the inverse problem framework and a general overview of the most common linear forward models used in the literature. In addition, we highlight different methods used for reconstruction and present a mean of computation for each method. Chapter 3 describes the multispectral imaging system and its components and presents the development of an instrument model such as the MIRI imager of the JWST. Chapters 4 and 5 deal with the reconstruction of a high-resolution discrete spatio-spectral object from low-resolution multispectral data, using the object models - piecewise linear model and linear mixing model, respectively. Both chapters rely on the results reported in Chapter 3. Finally, a general conclusion is given in Chapter 6.

Chapter 2

General Overview

Contents

2.1 Inverse Problems Framework

In this thesis we are interested in a linear inverse problem application. In general an inverse problem is the process of restoration, reconstruction or estimation of an input vector $\boldsymbol{x}$ (the object of interest) from a degraded output $\boldsymbol{y}$ (observations or data), which is observed directly or indirectly. We denote $\boldsymbol{H}$ the observation matrix, a response of the linear system that relates the input of the observation system to the output. Figure 2.1 illustrates a general linear inverse problem and all terminology that are used. In fact, in order to deal with an inverse problem application we distinguish two main processes:

1. Forward or direct model which describes the formation of the observations.

2. Inverse problem or reconstruction is the reverse process that allows recovering of the input object from the observed data and knowledge about the system response.

Several inverse problems applications are defined based on (a) the type of the forward model, linear [Banham & Katsaggelos 1997] or non-linear [Chappell *et al.* 2009]. (b) The observation system, e.g. response of an imaging system such as telescope, camera, microscope. (c) The structure and the content of the observation matrix, e.g. the application is called deconvolution if $\boldsymbol{H}$ is a convolution matrix [Banham & Katsaggelos 1997], and inpainting if $\boldsymbol{H}$ is a mask [Bertalmio *et al.* 2000], a super-resolution problem if $\boldsymbol{H}$ is a downsampling operator [Park *et al.* 2003], denoising if $\boldsymbol{H}$ is an identity matrix [Donoho 1995, Buades *et al.* 2005]. (d) Single or multiple observations. e.g. 1D signal, 2D image, 3D

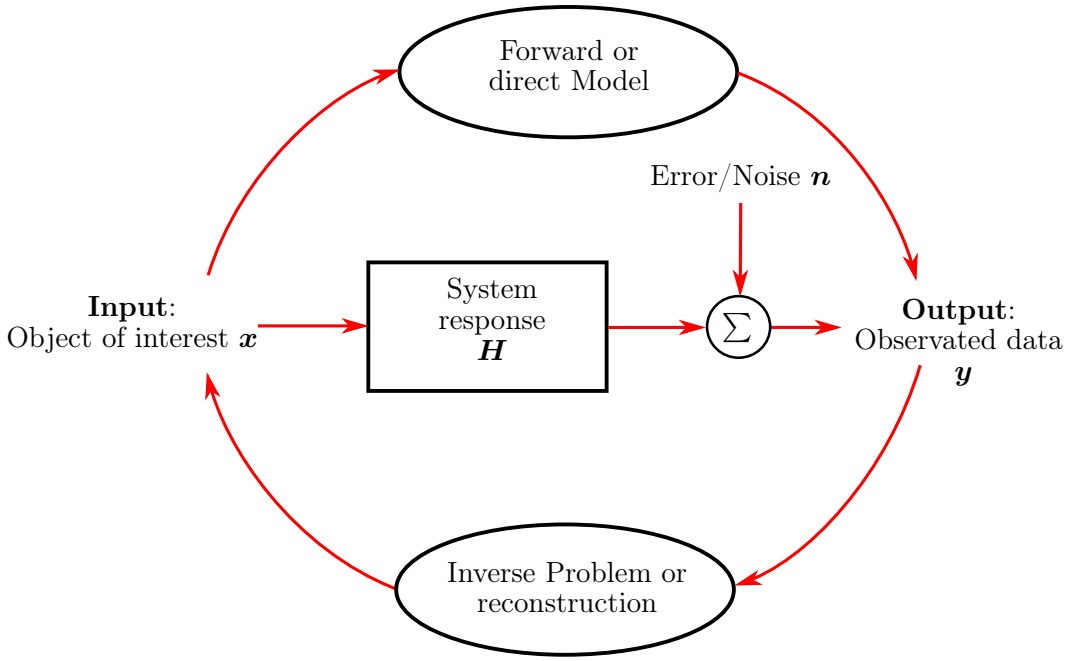


Figure 2.1: Illustration of the linear inverse problems framework.

datacube, 2D+λ (spatio-spectral object), 2D+t (spatio-temporel object). (e) The availability of information. In general $\boldsymbol{y}$ and $\boldsymbol{H}$ are known, otherwise the problem is called blind if $\boldsymbol{H}$ is unknown [Chan & Wong 1998, Molina *et al.* 2006], and it is called semi-blind or myopic if $\boldsymbol{H}$ is known partially [Mugnier *et al.* 2004, Orieux *et al.* 2010], or a system identification problem in case we only dispose of $\boldsymbol{x}$ and $\boldsymbol{y}$ [Ljung 1998].

If the next sections we provide a general overview for the forward model and the reconstruction methods.

2.2 Linear Forward Models

2.2.1 Mono-channel Linear Forward Model

The single-channel linear forward model is the most common model used in the literature [Bertalmio *et al.* 2000, Combettes & Pesquet 2004, Molina *et al.* 2006, Chambolle & Pock 2011]. Its classical formulation is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{2.1}$$

where $\boldsymbol{y} \in \mathbb{R}^{N_i N_j}$ represents the observed data in a vector form, here we provide details for an image type object of size $N_i \times N_j$ pixels. $\boldsymbol{n} \in \mathbb{R}^{N_i N_j}$ is an additive term added to the data in order to take into accounts noise and modeling errors. The vector $\boldsymbol{x} \in \mathbb{R}^{N_k N_l}$ is the original object of interest and $\boldsymbol{H} \in \mathbb{R}^{N i N_j \times N_k N_l}$ is the observation or degradation matrix.

The particularity of the single-channel forward model is that only one observation is considered at a time. Therefore, it is not suitable to for multiple observation such as multichannel or multispectral restoration.

2.2.2 Multichannel Linear Forward Model with a Block-Diagonal Matrix

In order to reconstruct an object which is observed at multiple times [Hunt & Kubler 1984] proposed a multichannel forward model where the object is observed with the same instrument, thus the instrument response is the same for all observations. Hence, the observation matrix $\boldsymbol{H}$ is a square block-diagonal matrix with the same block in the diagonal. Few years later [Galatsanos & Chin 1989] proposed an update of the previous model in order to take into account responses from different instruments, thus, the observation matrix is a block-diagonal observation matrix with $\boldsymbol{H} = \text{diag}\left(\boldsymbol{H}^1, \boldsymbol{H}^2, \ldots, \boldsymbol{H}^P\right)$. Moreover, the block matrices were all approximated with convolution matrices for faster computation in the Fourier domain. Consequently, $\boldsymbol{H}$ is block-circulant and has a block-Toeplitz form. The multichannel forward model is formulated as

$$\underbrace{\begin{pmatrix} \boldsymbol{y}^{(1)} \\ \boldsymbol{y}^{(2)} \\ \vdots \\ \boldsymbol{y}^{(P)} \end{pmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{pmatrix} \boldsymbol{H}^1 & & & \\ & \boldsymbol{H}^2 & & \\ & & \ddots & \\ & & & \boldsymbol{H}^P \end{pmatrix}}_{\boldsymbol{H}} \underbrace{\begin{pmatrix} \boldsymbol{x}^{(1)} \\ \boldsymbol{x}^{(2)} \\ \vdots \\ \boldsymbol{x}^{(P)} \end{pmatrix}}_{\boldsymbol{x}} + \underbrace{\begin{pmatrix} \boldsymbol{n}^{(1)} \\ \boldsymbol{n}^{(2)} \\ \vdots \\ \boldsymbol{n}^{(P)} \end{pmatrix}}_{\boldsymbol{n}} \tag{2.2}$$

where $\boldsymbol{x} \in \mathbb{R}^{PN_iN_j}$ represent the object of interest which is a stack of P channels in a vector form $\boldsymbol{x}^{(p)}, p = 1, 2, \ldots, P$. Each channel contains $N_k \times N_l$ pixels where $\boldsymbol{x}_{i,j}^{(p)}$ denotes the $(i,j)^{th}$ spatial position in the $(p)^{th}$ wavelength. $\boldsymbol{y} \in \mathbb{R}^{PN_iN_j}$ represents the stack of all multispectral data acquired via P spectral bands of the imaging system. The additive noise associated to the observation data set is represented by $\boldsymbol{n} \in \mathbb{R}^{PNiN_j}$. The block-diagonal matrix $\boldsymbol{H} \in \mathbb{R}^{PNiN_j \times PNkN_l}$ represents the full system response, where it maps every channel $\boldsymbol{x}^{(p)}$ to an observation $\boldsymbol{y}^{(p)}$ through its blocks or sub-matrices $\boldsymbol{H}^p \in \mathbb{R}^{NiNj \times N_kN_l}, p = 1, 2, \ldots, P$.

The block-diagonal forward model in Equation (2.2) has been widely used for multichannel deconvolution, e.g. [Benazza-Benyahia & Pesquet 2006, Bongard *et al.* 2011, Henrot *et al.* 2013, Song *et al.* 2016] address multichannel 2-D deconvolution problem for hyperspectral image deconvolution. It was also used for multichannel reconstruction of spectral computed tomography images [Rigie & La Rivière 2015].

However, this above forward model takes into account within channel degradation (or auto-degradation) but does not accounts for the between channels degradation (or cross-degradation) which can occur from the spectral correlation between channels. Therefore, this model is not suitable for multispectral imaging especially when there is a strong correlation between channels. In addition, the number of observations is not necessarily equal to the number of channels.

2.2.3 Multichannel Linear Forward Model in General

In [Galatsanos *et al.* 1991] a modified forward model is proposed to overcome limitations of the previous model for multichannel image restoration. It is given by

$$\underbrace{\begin{pmatrix} \boldsymbol{y}^{(1)} \\ \boldsymbol{y}^{(2)} \\ \vdots \\ \boldsymbol{y}^{(P)} \end{pmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{pmatrix} \boldsymbol{H}^{1,1} & \boldsymbol{H}^{1,2} & \dots & \boldsymbol{H}^{1,P} \\ \boldsymbol{H}^{2,1} & \boldsymbol{H}^{2,2} & \dots & \boldsymbol{H}^{2,P} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}^{P,1} & \boldsymbol{H}^{P,2} & \dots & \boldsymbol{H}^{P,P} \end{pmatrix}}_{\boldsymbol{H}} \underbrace{\begin{pmatrix} \boldsymbol{x}^{(1)} \\ \boldsymbol{x}^{(2)} \\ \vdots \\ \boldsymbol{x}^{(P)} \end{pmatrix}}_{\boldsymbol{x}} + \underbrace{\begin{pmatrix} \boldsymbol{n}^{(1)} \\ \boldsymbol{n}^{(2)} \\ \vdots \\ \boldsymbol{n}^{(P)} \end{pmatrix}}_{\boldsymbol{n}} \tag{2.3}$$

where the full system response $\boldsymbol{H}$ is represented by a block matrix corresponding to within and between channels degradation $\boldsymbol{H}^{m,p}, m, p = 1 \dots, P$. Here P is the total number of observations. The block $\boldsymbol{H}^{p,t}, t = m$ represents the direct (or auto) degradation, whereas the block $\boldsymbol{H}^{p,t}, t \neq m$ accounts for between or cross degradation occurring between channels.

Most of works used this model for color image restoration such as [Schultz & Stevenson 1995, Galatsanos *et al.* 2000, Molina *et al.* 2003, Wen *et al.* 2008, Yang *et al.* 2009a]. In this case $P = 3$ and $y^{(p)}$ corresponds to a degraded channel of the color image. Another use of the multichannel forward model in Equation (2.3) is for multichannel image reconstruction in spectral X-Ray computed tomography [Sawatzky *et al.* 2014]. In [Yang *et al.* 2009b] presented a method for deblurring multichannel images corrupted with impulsive noise is presented.

The above model is mostly used for applications where the number of channels and observations are the same, e.g. restoration of a colored image with three channels. In this thesis we will use this forward model but with different dimensions of $\boldsymbol{y}$ and $\boldsymbol{x}$, i.e. we will consider M channels in $\boldsymbol{x}$ and P multispectral data in $\boldsymbol{y}$ with $P \neq M$. In this case the observation block-matrix is $\boldsymbol{H} \in \mathbb{R}^{PN_iN_j \times MN_kN_l}, p = 1 \dots, P, m = 1 \dots, M$.

2.3 Reconstruction Methods

In this section we review different reconstruction method based on regularized inversion or Bayesian approach.

2.3.1 Regularization Method

A classical approach to reconstruct the object $\boldsymbol{x}$ from the observed data $\boldsymbol{y}$ giving a forward model is the minimization of an objective (or a cost) function $\mathcal{J}(\boldsymbol{x})$ having the solution $\widehat{\boldsymbol{x}}$ as a minimizer

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \ \{\mathcal{J}(\boldsymbol{x})\} \tag{2.4}$$

A naive choice for the objective function is to measure the fidelity or consistency of the solution to the data $\mathcal{J}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{y}, \boldsymbol{x})$ by measuring a distance between the data $\boldsymbol{y}$ and

the forward model $\boldsymbol{H}\boldsymbol{x}$. Several functions for $\mathcal{Q}(.)$ exist in the literature depending on the statistical distribution of the noise $\boldsymbol{n}$, e.g. l_2-norm in case of an additive Gaussian white noise [Rudin *et al.* 1992], or l_1-norm in case of Laplace noise [Alliney 1992]. l_∞-norm for a uniform noise. Here we consider the most common noise distribution which is a Gaussian white noise. A particular usefulness of this choice is that objective function is strictly convex and differentiable. The reconstruction problem is then given by

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \left\{ \mathcal{J}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 \right\} \tag{2.5}$$

which is known as the Least-Squares method [Demoment 1989]. In this case the solution of the problem is

$$\widehat{\boldsymbol{x}} = \left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{y} \tag{2.6}$$

where ^{T} symbolizes the transpose of a matrix.

Computing the solution in Equation (2.6) requires the inversion of the Hessian matrix $\boldsymbol{H}^T\boldsymbol{H}$ and its multiplication by $\boldsymbol{H}^T\boldsymbol{y}$. However, this solution is unstable for most inverse problems applications because of the ill-conditioning of the matrix $\boldsymbol{H}^T\boldsymbol{H}$, i.e. a small error in $\boldsymbol{H}^T\boldsymbol{y}$ leads to a high error in the solution, hence, the problem is called ill-posed. In order to overcome the ill-posedness of the problem we proceed by correcting the ill-conditioning of $\boldsymbol{H}^T\boldsymbol{H}$ by adding one or multiple regularization terms $\mathcal{R}(\boldsymbol{x})$ in order to stabilize the solution. Therefore Equation (2.5) becomes

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \left\{ \mathcal{J}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{x}, \boldsymbol{y}) + \mu\, \mathcal{R}(\boldsymbol{x}) \right\} \tag{2.7}$$

where μ is a regularization parameter which adjust the compromise between the data fidelity term and the regularization term.

2.3.1.1 Regularization Models

The regularization term can take different forms depending on the nature of the object but must enforce prior information about the object. Several choices can be found in the literature such as the Tikhonov regularization [Tikhonov & Arsenin 1977] to enforces smoothness of the solution by penalizing the difference between pixels in case of an image. This regularization has been used widely in the literature for its properties such as linearity, continuous differentiability, and speed of calculation of the solution.

$$\mathcal{R}_{l_2}(\boldsymbol{x}) = \sum_{k=1}^{N_k}\sum_{l=1}^{N_l} \left(\boldsymbol{x}_{k+1,l} - \boldsymbol{x}_{k,l}\right)^2 + \left(\boldsymbol{x}_{k,l+1} - \boldsymbol{x}_{k,l}\right)^2 = \|\boldsymbol{D}\boldsymbol{x}\|_2^2 \tag{2.8}$$

where $\boldsymbol{D}$ can be a 2D finite difference in case of a correlated (or smooth) object of an identity matrix in case of an uncorrelated image. Here N_k and N_l are number of pixels in row and column of the image $\boldsymbol{x}$. Often the operator $\boldsymbol{D}$ is implemented in the Fourier domain for faster computation through circular convolution, in this case circularity conditions are assumed $x_{N_k+1,l} = x_{1,l}$ and $x_{k,N_l+1} = x_{k,1}$.

The problem solution using Tikhonov regularization is given by

$$\widehat{\boldsymbol{x}}_{l_2} = \left(\boldsymbol{H}^T\boldsymbol{H} + \mu\boldsymbol{D}^T\boldsymbol{D}\right)^{-1}\left(\boldsymbol{H}^T\boldsymbol{y}\right). \tag{2.9}$$

However, the quadratic regularization creates ringing artifacts to the solution at a high value of the gradient such as the contour for an image. Therefore, a non-quadratic regularization helps in overcoming these artifacts such as the Total Variation (TV), isotropic total variation [Rudin *et al.* 1992] or anisotropic total variation [Yang *et al.* 2009b]. The TV regularization is used for an object with a piecewise constant distribution and it is defined by

$$\mathcal{R}_{TV}(\boldsymbol{x}) = \sum_{k=1}^{N_k}\sum_{l=1}^{N_l} |(\nabla x)_{k,l}| = \|\nabla \boldsymbol{x}\|_1 \tag{2.10}$$

where a ∇ is a 2D finite difference operator same as $\boldsymbol{D}$ in Equation 2.8. But is usually represented by ∇ in the total variation regularizer. It has two components, horizontal (h) and vertical (v), according to the dimension of the image $(\nabla x)_{k,l} = \left((\nabla x)^h_{k,l}, (\nabla x)^v_{k,l}\right)$, and the first order finite differences with Neumann boundary conditions [Chambolle 2004] are defined by

$$(\nabla x)^h_{k,l} = \begin{cases} x_{k+1,l} - x_{k,l} & \text{if } k < N_k \\ 0 & \text{if } k = N_k \end{cases}$$

and

$$(\nabla x)^v_{k,l} = \begin{cases} x_{k,l+1} - x_{k,l} & \text{if } l < N_l \\ 0 & \text{if } l = N_l \end{cases}$$

.

.

The isotropic TV is defined by

$$|(\nabla x)_{k,l}| = \sqrt{\left((\nabla x)^h_{k,l}\right)^2 + \left((\nabla x)^v_{k,l}\right)^2} \tag{2.11}$$

and the anisotropic TV

$$|(\nabla x)_{k,l}| = \left|(\nabla x)^h_{k,l}\right| + \left|(\nabla x)^v_{k,l}\right| \tag{2.12}$$

The problem solution using TV regularization cannot be directly computed because of the non-differentiability of the l_1-norm. Thus iterative algorithms are used instead such as the splitting Bregman method [Goldstein & Osher 2009] or more recently the primal-dual algorithm [Chambolle & Pock 2011].

Another nonlinear regularization is the half-quadratic regularization (e.g. l_2/l_1-norm). It promotes high gradient values to the solution and it is defined by a nonlinear function $\varphi(.)$

$$\mathcal{R}_{l_2/l_1}(\boldsymbol{x}) = \sum_{k=1}^{N_k}\sum_{l=1}^{N_l} \varphi\left((\boldsymbol{D}\boldsymbol{x})_{k,l}\right) \tag{2.13}$$

where several l_2/l_1 function for $\varphi(.)$ can be found in [Idier 2001], e.g. Huber function

$$\varphi(\delta) = \begin{cases} \delta^2 & \text{if } |\delta| < s \\ 2s|\delta| - s^2 & \text{otherwise} \end{cases}. \tag{2.14}$$

The half-quadratic regularization was first proposed in [Geman & Reynolds 1992] (GR) witch proposed a convex conjugate construction for the non linear as a minimum of two convex functions in which the quadratic function is multiplied by the auxiliary variable b. Few years later [Geman & Yang 1995] (GY) proposed another construction where the auxiliary variable is added inside the quadratic function.

$$\varphi_{GR}(\delta) = \min_{b} \left\{ b\delta^2 + \xi(b) \right\}, \tag{2.15}$$

$$\varphi_{GY}(\delta) = \min_{b} \left\{ (\delta - b)^2 + \xi(b) \right\}, \tag{2.16}$$

where $\xi(b)$ is the auxiliary function.

The solution of this problem is computed iteratively by performing an alternate minimization of the augmented objective function $\mathcal{J}^*(\boldsymbol{x}, \boldsymbol{b})$ with respect to $\boldsymbol{x}$ and $\boldsymbol{b}$

$$\begin{cases} \widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}}\ \mathcal{J}^*(\boldsymbol{x}, \boldsymbol{b}) & (2.17) \\ \widehat{\boldsymbol{b}} = \underset{\boldsymbol{b}}{\operatorname{argmin}}\ \mathcal{J}^*(\widehat{\boldsymbol{x}}, \boldsymbol{b}) & (2.18) \end{cases}$$

with

$$\mathcal{J}(\boldsymbol{x}) = \min_{\boldsymbol{b}} \mathcal{J}^*(\boldsymbol{x}, \boldsymbol{b}). \tag{2.19}$$

Equation (2.17) gives the following solution for the two half-quadratic constructions

$$\widehat{\boldsymbol{x}}_{GR} = \left(\boldsymbol{H}^T \boldsymbol{H} + \mu \boldsymbol{b} \boldsymbol{D}^T \boldsymbol{D} \right)^{-1} \left(\boldsymbol{H}^T \boldsymbol{y} \right). \tag{2.20}$$

$$\widehat{\boldsymbol{x}}_{GY} = \left(\boldsymbol{H}^T \boldsymbol{H} + \mu \boldsymbol{D}^T \boldsymbol{D} \right)^{-1} \left(\boldsymbol{H}^T \boldsymbol{y} + \mu \boldsymbol{D}^T \boldsymbol{b} \right). \tag{2.21}$$

Note that both solutions are linear and require the inversion of the Hessian matrix. In addition, it is worth noting that the solution $\widehat{\boldsymbol{x}}_{GY}$ can be computed directly and efficiently by inverting once the Hessian matrix in the Fourier domain for circular matrices $\boldsymbol{H}$ and $\boldsymbol{D}$ [Hunt 1971].

Other types of regularization exist among which the l_1-norm of the object to enforce sparsity of the solution [Mairal *et al.* 2014], i.e. a solution with less nonzero coefficients. Since the l_1-norm is not differentiable, problem solution is computed through minimization of the convex objective function iteratively until convergence using optimization algorithms, e.g. algorithms based on proximal operator [Beck & Teboulle 2009, Combettes & Pesquet 2011].

The regularization method is very sensitive to the noise and there is a difficulty to chose the amount of information introduced by the prior, which regularize the problem and compensate the lack of information in the data. In addition, the regularization parameter is set manually.

In the next section, we see another framework based on statistical distributions where the regularization parameter, also called hyper-parameter, can be tuned automatically.

2.3.2 Bayesian Interpretation

The reconstruction using the regularization method can be interpreted in the Bayesian framework using statistical distributions [Jaynes 2003]. The main difference between the two approaches is that the object to reconstruct is deterministic in the regularization method, whereas it is probabilistic in the Bayesian approach. Here the solution is obtained by maximizing the posterior distribution of the object, which is computed using the Baye's theorem

$$f(\boldsymbol{x}|\boldsymbol{y}) = \frac{f(\boldsymbol{y}|\boldsymbol{x})f(\boldsymbol{x})}{f(\boldsymbol{y})} \tag{2.22}$$

where $f(\boldsymbol{y}|\boldsymbol{x})$ is the likelihood distribution, $f(\boldsymbol{x})$ represents the prior distribution for the object, $f(\boldsymbol{x}|\boldsymbol{y})$ is the a posterior law, and $f(\boldsymbol{y})$ is the marginal law for the data.

As explained in Section 2.3.1, the data fidelity term is chosen based on the distribution of the noise, e.g. the distribution of a zero-mean, white, Gaussian noise of variance σ_n or precision parameter $\gamma_n = \sigma_n^{-1}$, denoted by $f(\boldsymbol{n}|\gamma_n) = \mathcal{N}\left(\boldsymbol{0}, \gamma_n^{-1}\boldsymbol{I}_N\right)$, is given by

$$f(\boldsymbol{n}|\gamma_n) = (2\pi)^{-N/2}\ \gamma_n^{N/2}\ \exp\left[-\frac{1}{2}\gamma_n \left\|\boldsymbol{n}\right\|_2^2\right], \tag{2.23}$$

where N refers to the total number of element of $\boldsymbol{n}$. The likelihood of the object $\boldsymbol{x}$ associated to the data $\boldsymbol{y}$ (through the forward model $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$) is

$$f(\boldsymbol{y}|\boldsymbol{x}, \gamma_n) = \mathcal{N}(\boldsymbol{H}\boldsymbol{x}, \gamma_n^{-1}\boldsymbol{I}_N) = (2\pi)^{-N/2}\ \gamma_n^{N/2}\ \exp\left[-\frac{1}{2}\gamma_n \left\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\right\|_2^2\right]. \tag{2.24}$$

The data fidelity term appears in the exponential function of the likelihood and it is obtain through the co-logarithm by

$$\mathcal{Q}(\boldsymbol{y}, \boldsymbol{x}) = \left\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\right\|_2^2 = -k \log\left[f(\boldsymbol{y}|\boldsymbol{x}, \gamma_n)\right] + C \tag{2.25}$$

where $k = \frac{2}{\gamma_n}$ and $C = -\frac{N}{\gamma_n}\left(\log\gamma_n - \log 2\pi\right)$ are a multiplicative and an additive constants. Concerning the regularization term, it is introduced through a prior distribution of the object $f(\boldsymbol{x})$. For instance, Tikhonov regularization is enforced through a Gaussian distribution, $f(\boldsymbol{x}|\gamma_x) = \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{R} = \left(\gamma_x \boldsymbol{D}^T\boldsymbol{D}\right)^{-1}\right)$, which is given by

$$f(\boldsymbol{x}|\gamma_x) = (2\pi)^{-N/2}\ \gamma_x^{N/2}\det\left[\boldsymbol{D}^T\boldsymbol{D}\right]^{1/2}\ \exp\left[-\frac{1}{2}\gamma_x \boldsymbol{x}^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{x}\right] \tag{2.26}$$

where N is the size of the vector $\boldsymbol{x}$ same as for the data to simplify the notations. and γ_x is a precision parameter, inverse of the variance. The regularization term appears in the exponential function of the prior distribution and is obtained through the co-logarithm by

$$\mathcal{R}_{l_2}(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{x} = \left\|\boldsymbol{D}\boldsymbol{x}\right\|_2^2 = -k_x\ \log\left[f(\boldsymbol{x}|\gamma_x)\right] + C_x \tag{2.27}$$

where $k_x = \frac{2}{\gamma_x}$ and $C_x = -\frac{N}{\gamma_x}\left(\log\gamma_x - \log 2\pi\right) + \log[\det[\boldsymbol{D}^T\boldsymbol{D}]]$ are a multiplicative and an additive constants. The higher the regularity the higher the probability $f(\boldsymbol{x}|\gamma_x)$ and

the lower the regularization $\|\boldsymbol{D}\boldsymbol{x}\|_2^2$. Lastly the posterior distribution $f(\boldsymbol{x}|\boldsymbol{y},\gamma_x,\gamma_n)$ allow the estimate the object $\boldsymbol{x}$. From the Baye's theorem in Equation (2.22) we have

$$\begin{aligned} f(\boldsymbol{x}|\boldsymbol{y},\gamma_x,\gamma_n) &= \frac{f(\boldsymbol{y}|\boldsymbol{x},\gamma_n)f(\boldsymbol{x}|\gamma_x)}{f(\boldsymbol{y}|\gamma_x,\gamma_n)} \propto f(\boldsymbol{y}|\boldsymbol{x},\gamma_n)f(\boldsymbol{x}|\gamma_x) \\ &\approx \exp\left[\underbrace{-\frac{1}{2}\left(\gamma_n\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|_2^2-\gamma_x\|\boldsymbol{D}\boldsymbol{x}\|_2^2\right)}_{-\frac{\gamma_n}{2}\mathcal{J}(\boldsymbol{x})}\right] \end{aligned}$$

Here the objective function appears in the exponential function and is obtained through the co-logarithm of the posterior distribution. The regularization parameter $\mu = \frac{\gamma_x}{\gamma_n} = \frac{\sigma_n}{\sigma_x}$ is written as a signal to noise ratio.

Finally, the reconstruction solution can be computed using several estimators such the mean, the median or the maximum of the posterior distribution. These three estimators are the same in case of a Gaussian distribution. For instance, the Maximum A Posteriori (MAP) estimator minimizes the objective function. It is equal to the quadratic solution of the regularized least square through

$$\begin{aligned} \widehat{\boldsymbol{x}}_{\text{MAP}} &= \underset{\boldsymbol{x}}{\text{argmax}}\, f(\boldsymbol{x}|\boldsymbol{y},\gamma_x,\gamma_n) \propto \underset{\boldsymbol{x}}{\text{argmax}}\, f(\boldsymbol{y}|\boldsymbol{x},\gamma_n)f(\boldsymbol{x}|\gamma_x) \\ &= \underset{\boldsymbol{x}}{\text{argmin}}\, \mathcal{J}(\boldsymbol{x}) = \widehat{\boldsymbol{x}}_{l_2} \end{aligned} \tag{2.28}$$

Moreover, the estimation of the regularization parameter is possible in the Bayesian approach. The precision parameters of the object and the noise needs to be probabilized. For instance, a Gamma distribution or Jeffreys priors [Kass & Wasserman 1996].

2.4 Related Works

In this section we present key works in the literature that are related to our topic. In general, multispectral imaging has been used for many applications. For instance, [Nuzillard & Bijaoui 2000] present a set of blind source separation methods with an application for the analysis of multispectral astronomical images. [Levenson & Mansfield 2006] presents the importance and advantages of using multispectral images provided by a microscope for the improvement of the analysis of a biological object. Another application is developed in [Hedjam & Cheriet 2013] which uses multispectral imaging in order to develop an image enhancement algorithm for the preservation of the historical documents through digitization of ancient manuscripts. The idea is to scan the degraded documents by an infrared camera to detect only the artifact of the images.

As anticipated in Section 2.2.2, most of multispectral reconstruction in the literature deal with applications to remote sensing and color images [Galatsanos & Chin 1989, Katsaggelos & Paik 1988, Galatsanos *et al.* 1991, Schultz & Stevenson 1995] and in [Galatsanos *et al.* 2000, Molina *et al.* 2003, Wen *et al.* 2008, Yang *et al.* 2009a]. However, these methods cannot be easily applied to astronomical data where the dimensions of the

object are high and the number of observations is not necessarily equal to the number of channels of the spatio-spectral object. In addition the dynamic range of astronomical data can vary a lot as well as the structure of the images, e.g. sharp and sparse object.

Spatio-Spectral Reconstruction:

In recent years, researchers have started working on the reconstruction of spatio-spectral object in general, multispectral and hyperspectral data. In [Rodet *et al.* 2009] a method is proposed for the reconstruction of hyperspectral objects observed by an infrared slit spectrograph on board the Spitzer Space Telescope. The instrument model and the inversion method are carried out in the continuous spatio-spectral coordinates. The spatial distribution of the object is decomposed over a family of Gaussian functions, and the hyperspectral data are collected from different dithering.

Authors in [Bongard *et al.* 2011] present a general method for 3D deconvolution of hyperspectral astronomical data based on l_2-norm, where the reconstruction method is based on a multichannel quadratic regularization (see [Schultz & Stevenson 1995]). The prior information of the solution is enforced by using two types of regularization terms such as the spatial and the spectral smoothness. However, this reconstruction method is not suitable for our application because the observation matrix is block-diagonal along the spectral dimension. Thus, it does not take into account the cross correlation occurring between the channels.

An interesting work in [Soulez *et al.* 2013] deals with the development of a hyperspectral image restoration for integral field spectrographs (IFS) data where the spectral-variant PSF is approximated by a linear combination of few monochromatic PSFs for faster computation while preserving the PSF properties such as normalization, positivity, and symmetry. Moreover, the inversion method is based on regularization methods where separable regularization terms, spatial and spectral, are introduced to enforce prior information to the solution. In addition, a comparison between quadratic regularization and spatial sparsity regularization is provided and the reconstruction results are illustrated on simulations coming from the Multi Unit Spectroscopic Explorer (MUSE) instrument. In [Henrot *et al.* 2013] the authors deal with deconvolution of hyperspectral images. The output of the forward model is a 2D convolution of the channel l and the l-th channel PSF. The deconvolution method is based on regularization methods where two smoothness priors are enforced, spatial and spectral, and the positivity constraint. An iterative algorithm is derived for updating the solution and its constraint. The reconstruction results on simulated microscopy data, having a narrow wavelengths range of $0.3 - 0.6$ μm, are very good where a very low root mean square error is obtained. However, the developed forward model does not take into accounts the spectral variation of the PSF nor the cross-correlation between channels. Hence, this reconstruction method cannot be applied for multispectral data of the imager MIRI of the JWST.

Spectral unmixing:

Spectral unmixing is a technique used for multispectral/hyperspectral data reconstruc-

tion. It consists of decomposing a pixel spectrum into a collection of distinct spectra (end-members or spectral components) and estimating the corresponding abundances (weights or mixture coefficients or proportion) [Keshava & Mustard 2002, Plaza *et al.* 2004]. Authors in [Adams *et al.* 1986] proposed a linear mixing model in order to analyze a multispectral image. The spectral components in the image correspond to different materials (e.g. soil, rock, and shade), whereas the mixture coefficients represent the proportion of each spectral components. The linear mixing model has been used on a wide range of hyperspectral applications [Settle & Drake 1993, Haertel & Shimabukuro 2004, Berne *et al.* 2007]. Several works consider the spatial and spectral correlations between neighbors pixels in order to analyze the hyperspectral data. [Tarabalka *et al.* 2009] proposed a spectral–spatial classification scheme for hyperspectral images. The authors of [Guo *et al.* 2009] presented an hyperspectral image enhancement method based on the total variation regularizer to produce a higher visual quality hyperspectral image. Authors in [Dobigeon *et al.* 2009] proposed a Bayesian method for the extraction of end-members and the estimation of abundances from hyperspectral images. An overview of hyperspectral unmixing models and algorithms is presented in [Bioucas-Dias *et al.* 2012]. In a recent work [Loncan *et al.* 2015] proposed a comparison of several pansharpening techniques basically used for multispectral and adapted for hyperspectral images. It consists of improving the spatial resolution of hyperspectral data by fusing it with data characterized by sharper spatial information. Spectral unmixing have been very successfully used for the reconstruction and the analysis of hyperspectral data. In this thesis, we aim to use this technique in order to reconstruct a high spatio-spectral resolution object from a set of multispectral data which are severely limited in spectral resolution.

PSF Modeling:

Other works focused on with the variability of the PSF especially for image deconvolution such as [Denis *et al.* 2011], where the shift-variant PSF is approximated using a linear PSF interpolation. A few years later [Denis *et al.* 2015] presented an overview of fast PSF approximation and derived an improved PSF approximation based on PSF interpolation by proposing optimal interpolation weights and PSF samples. The obtained results are illustrated on the deconvolution of an image blurred with a spatial shift-variant PSF and it is shown that a low approximation error is obtained using the proposed approximation.

Yet a recent work in [Thiébaut *et al.* 2016] provides a consistent understanding of the shift-variant PSF models, PSF interpolation and modal PSF approximation. Authors shown that the PSF properties such as non-negativity, normalization, symmetries and invariances are well preserved with the approximation by the PSF interpolation using a set of calibrated PSF, whereas it not the case for the modal PSF approximation, which consists of decomposing the shift-variant PSF into a sum of multiplication of the (left and right) modes computed using the truncated singular value decomposition. This is because no constraints are imposed on the modes even though the model offers the best approximation in term of least-squares error.

PSF Homogenization:
When it comes to comparison between multispectral data issued from different instruments or spectral bands, the notion of PSF homogenization is usually used in practice, especially in astronomy [Aniano *et al.* 2011] or recently in [Boucaud *et al.* 2016], where each multispectral data has a broadband (or effective) PSF associated to it. PSF homogenization consists of convolving all observed multispectral data with specific kernels in order to unify the resolution of all data, practically the lowest resolution. However this approach degrades severally the spatial resolution of data especially the one with high resolution, thus, a considerable loss of information might occurs especially if there is a high range between the lowest resolution and highest resolution.

Additionally, more related works are presented in the *Problem Statement* section of Chapters 4 and 5.

The particularity of the proposed work lies in the observed data. A small number of low-resolution multispectral data degraded by a spatial blur and noise, in addition to severe lack of spectral distribution after integrations over broad bands by the detector. This limit the direct use of the methods applied on hyperspectral data, where a datacube with high spectral resolution is observed.

2.5 Conclusion

In the first section of this chapter, we presented the general paradigm of a linear inverse problem, where we seek the reconstruction of an object of interest, degraded by the response of a measuring system and corrupted with an additive noise. Then we presented several linear forward models that are used to model the degraded data with a particular interest for multichannel objects. We pointed out the difference between several linear forward models provided details for the forward model we will use later in Chapters 4 and 5. In fact, the response of the system is a block-matrix containing a direct degradation of the object channels in the observed data, and cross degradation occurring between channels. In addition, the dimension of the observation matrix depends on the number of observation and number of channels of the object which are not equals.

Next, we presented two frameworks for the reconstruction, regularization method and Bayesian estimation. Both frameworks present technical advantages and drawbacks and the choice between the two approaches is based on the application. (a) Regularization method offers the possibility to introduce a prior knowledge of the object using linear or nonlinear functions without any complications. However, this is not always the case in the Bayesian approach since the prior of the object corresponds to a defined statistical distribution. This restricts the choice of priors of the object and/or the hyper-parameter. (b) Regularization method is a supervised method because it requires to set manually a regularization parameter. Whereas the Bayesian approach can be unsupervised because the regularization parameter can be estimated through estimation of the precision parameters of the object and noise. (c) Although the Bayesian method offers an uncertainty of

the reconstruction and a possibility for an unsupervised algorithm, this approach remains limited in case of data with high dimensions. Our choice in this thesis is based on the regularization method since it allows a better flexibility of choosing the prior. This is very important in our application since we aim to reconstruct a high-resolution object with multiple components from multiple observations.

Finally, we presented in the last section related works that are useful for the rest of the thesis. From reconstruction of multispectral and hyperspectral data, spectral unmixing, PSF modeling, and PSF Homogenization.

Chapter 3

Multispectral Imaging System Model

Contents

3.1 Introduction

In this chapter we provide the multispectral imaging system response by developing an instrument model. We first present the considered multispectral imaging system which is the Mid-InfraRed Instrument (MIRI) Imager on board the James Webb Space Telescope (JWST)[1] in order to develop applications related to astrophysical data. We first highlight the optical issues related to this instrument. We then detail the response of each instrument components and propose a mathematical model of the multispectral imaging system. Lastly, to carry well the pre-processing of data clarify the physical units of the astrophysical object and the electronic units of the multispectral data and some useful conversions.

The multispectral imaging system is a diffraction-limited instrument. It consists of an optical system and a detector (or sensor). The non-stationarity of the optical response and the integration of the filtered object over broad bands are the main issues discussed here.

[1] https://www.jwst.nasa.gov/

In the following we present the JWST and provide details about the optical response, then we present the MIRI imager. Moreover, we develop an instrument model that will be used in Chapters 4 and 5.

3.2 JWST/MIRI Imager

In this section we first present the optical system response of the JWST. Then we address the multispectral imaging system of MIRI. It is worth mentioning that we only provide a description of the instrument components that are pertinent in this work. More specific details about the conception and characteristics of the instrument can be found in [Rieke *et al.* 2015, Bouchet *et al.* 2015, Glasse *et al.* 2015].

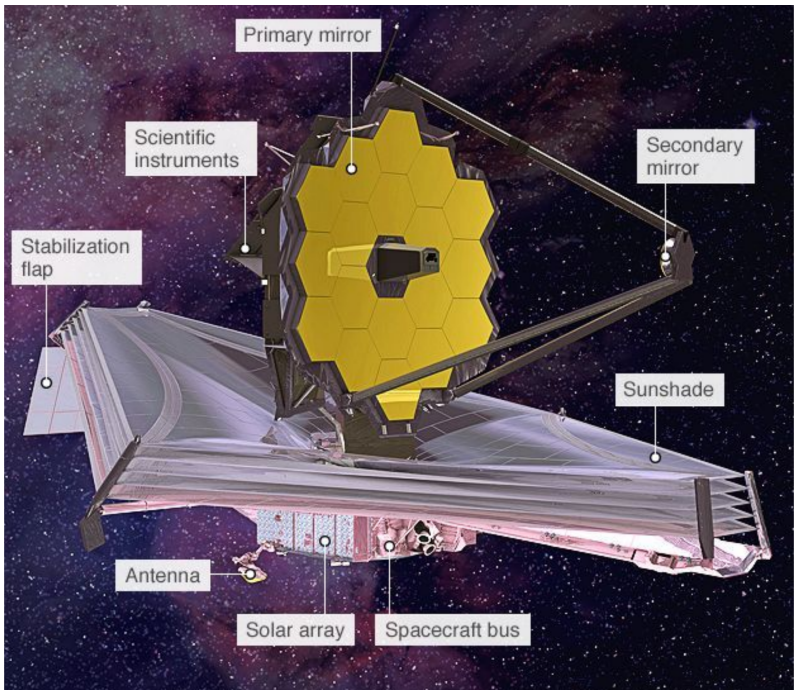


Figure 3.1: The James Webb Space Telescope (JWST) (image from NASA).

The JWST will be launched by an Ariane rocket in 2021. The JWST project is an international collaboration between NASA (National Aeronautics and Space Administration), the European Space Agency (ESA) and the Canadian Space Agency (CSA). With a primary mirror diameter of 6.5 meters, the JWST will be the largest telescope in space. The telescope has been fully optimized to observe in the InfraRed (IR) wavelength range from 2 to 10 microns. It will offer the possibility to observe down to 0.6 μm and up to 28 μm, with unprecedented performances.

In the JWST IR wavelength range, the previous space telescope was the Spitzer telescope, with a 0.85 m primary mirror. Going from a 0.85 m to a 6.5 m mirror improves the performances in terms of sensitivity and angular resolution. The angular resolution will be almost an order of magnitude (factor 7) better than that of Spitzer, the IR satellite

launched in 2003. The sensitivity will also be increased by a factor about $10-100$.

A set of four instruments will be located at the focal plane of the JWST and will allow various instrument observation modes, such as imaging, spectroscopy, and coronagraphy. The four instruments are : (1) A Near-IR Imager and Slitless Spectrograph ($0.6-5$ μm), NIRISS, provided by the Canadian Space Agency. (2) A Near-IR Camera ($0.6-5$ μm), NIRCam, provided by the University of Arizona (US). (3) A Near-IR Spectrometer ($1-5$ μm), NIRSpec, provided by ESA with components provided by NASA. (4) A Mid-IR Instrument ($5-28$ μm), MIRI, provided by a European Consortium of laboratories under the auspices of ESA and by the NASA Jet Propulsion Laboratory (JPL).

3.2.1 Effect of the Optical System of JWST

The photon flux emitted by the spatio-spectral astrophysical object (or source) in the sky is first reflected on a set of mirrors (primary and secondary) that composes the optical system then oriented to the focal plane as shown in Figure 3.2.

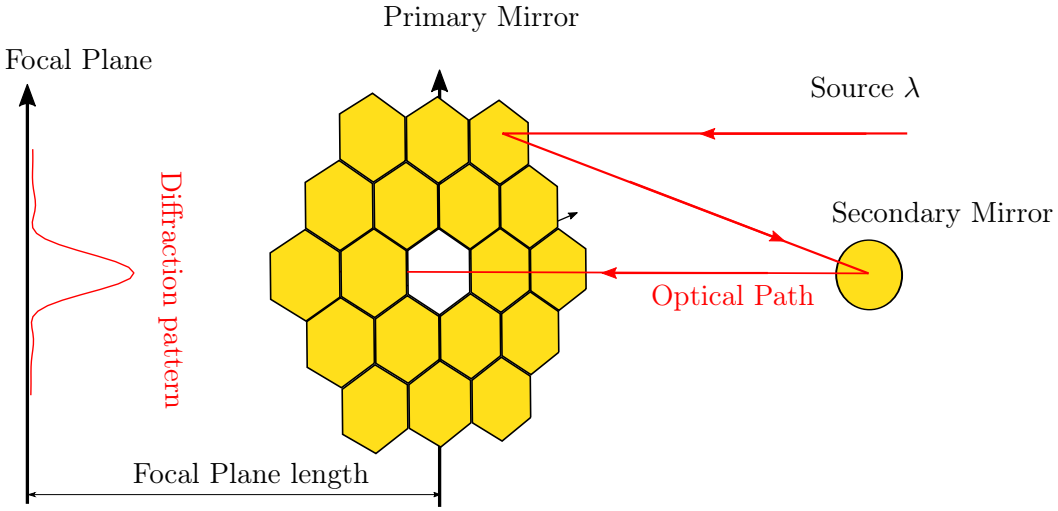


Figure 3.2: Illustration of the optical path from a source at wavelength λ to the focal plane of the JWST. Due to the diffraction theory a diffraction pattern is formed at the focal plane. The instrument MIRI is hosted on the back of the mirror and is not represented in the figure.

The optical system of the JWST is mainly composed of a 6.5 m primary mirror made by a set of 18 hexagonal-mirror segments aligned together to form one big mirror. The optical system of the telescope is limited by the diffraction as for all optical imaging system. The response of the optical system or Point Spread Function (PSF) depends on many parameters such as the focal plane length and the aperture transmittance function of the mirror, including its shape and size, e.g. for a circular aperture, the response is the Airy function (see Figure 3.3.a). In [Makidon *et al.* 2007] several methods of PSF calculation for JWST have been conducted with a comparison with PSF of different space telescopes such as Hubble and Spitzer. The calculation of the diffraction pattern for a far-field object, such as the astrophysical object, is based on the Fraunhofer diffraction

theory [Goodman 2005], where the intensity of the diffraction pattern intensity at the focal plane or the PSF is given by the square modulus $|\psi|^2$ of the 2D Fourier transform of the aperture transmittance function $A(x, y)$:

$$\psi(u,v) = \mathcal{F}\{A(x,y)\}(u,v) = \iint A(x,y) \exp\left\{-\frac{2\pi i(ux+vy)}{\lambda}\right\} dxdy \tag{3.1}$$

Therefore an analytic formula of the JWST's PSF can be obtained by performing the sum of 18 Fourier transforms of 18 off-centered hexagonal segments

$$\psi(u,v) = \psi_{\text{hex}}(u,v) \left[\sum_{i=1}^{18} e^{-\frac{i2\pi(x_i u + y_i v)}{\lambda}}\right] \tag{3.2}$$

where x_i and y_j are the coordinates of the *i-th* segment center for all 18 segments of the primary mirror of the JWST. ψ_{hex} is the amplitude for a single segment which is obtained by calculated the Fourier transform of an hexagonal aperture. The calculations details for this calculation are provided in [Mast *et al.* 1982, Sabatke *et al.* 2005] which we report here

$$\psi_{\text{hex}}(u,v) = \frac{\sqrt{3}a^2}{4\lambda}\left[\frac{\sin(3\beta'+\alpha')\sin(\beta'-\alpha')}{\beta'(\beta'-\alpha')} + \frac{\sin(3\beta'-\alpha')\sin(\beta'+\alpha')}{\beta'(\beta'+\alpha')}\right] \tag{3.3}$$

where $\alpha' = \frac{\sqrt{3}kau}{4}$ and $\beta' = \frac{kav}{4}$. $k = 2\pi/\lambda$ and a is the width of the segment. The PSF is then the square $|\psi|^2$.

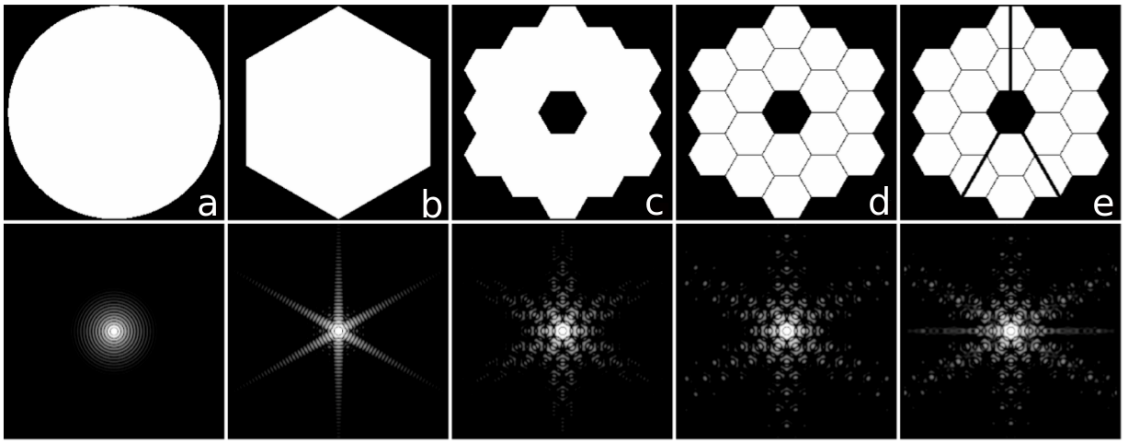


Figure 3.3: Illustration of the PSF in a logarithmic scale (2nd row) for different aperture transmittance function (1st row). The column (a) corresponds to a circular aperture and the classical Airy disk. Whereas, columns (d) and (e) illustrate the complexity of the PSF structure for the segmented JWST aperture, with and without secondary mirror supports, respectively [Makidon *et al.* 2007].

However the analytic expression of the PSF does not accounts for the misalignment of

the mirrors and the optical path difference (OPD). We use instead *WebbPSF*[2] [Perrin *et al.* 2012, Perrin *et al.* 2014] the official PSF simulation tool for the JWST mission in order to simulate a realistic PSF image because it considers the OPD maps precomputed by a detailed optical simulations of JWST. In addition *WebbPSF* takes into accounts the most recent JWST pupil and OPD models, thus it simulates PSFs for JWST based on up-to-date models for telescope wavefronts maps. It is developed by the Space Telescope Science Institute (STScI) to simulate monochromatic PSF and offers the possibility to compute broadband PSF weighted by a predefined source spectrum and allows tuning the spatial sampling as well as the size of the PSF image. It is available in both interfaces, graphical and scripting, which makes it easy to use in the programming language such as Python. Finally, it offers a suite of tools for quantifying PSF properties such as the Full Width at Half Maximum (FWHM), the *Strehl* ratio, the encircled energy.

We display in Figure 3.4 monochromatic PSFs at wavelengths from 5.6 to 25.5 μm as mentioned on top of each image. All PSF are normalized to 1 and displayed on a logarithmic scale. We clearly observe that the PSF structure is complex because of the hexagonal-shape segmented primary mirror. Moreover, we observe that the shape of the PSF depends on the wavelength, the larger the wavelength the wider the PSF, the shorter the wavelength the narrower the PSF ; we say that the PSF is spectral-variant or spectrally non-stationary. The PSF of JWST linearly depends on the wavelength as shown in Figure 3.5, where we display the Full-Width at Half-Maximum (FWHM) of the PSFs presented in Figure 3.4. We observe that the FWHM increases by a factor 5, as expected from the diffraction theory [Goodman 2005].

3.2.2 MIRI Imager

The Mid-InfraRed Instrument (MIRI) Imager [Bouchet *et al.* 2015] is one of the science instrument onboard the JWST (Figure 3.6). It covers a broad mid-infrared spectral range from 5 to 28 μm with a wavelength ratio around 5 between the shorter and the longer wavelength. MIRI imager will provide wide-field, broadband imaging that will continue the breathtaking astrophotography that has made Hubble [3] so universally admired.

The filter wheel contains nine broadband filters dedicated for imaging of MIRI to filter spectrally the photon flux, as shown in Figure 3.7. Their spectral characteristics are reported in Table 3.1 where p indicates the index (or position) of the filter. λ_c is the central wavelength and $\triangle\lambda$ is the bandwidth of the filter. It is between 2μm and 4μm except for the 1-*st* and 4-*th* filter. $R = \lambda/\triangle\lambda$ is the spectral resolving power. It varies from 3.5 to 16, which is very low compared to the Medium Resolution Spectrometer of MIRI [Wells *et al.* 2015], with R = 1500 to 3500. This indicates that the multispectral

[2] https://jwst.stsci.edu/science-planning/proposal-planning-toolbox/psf-simulation-tool-webbpsf

[3] https://www.nasa.gov/mission_pages/hubble/story/index.html

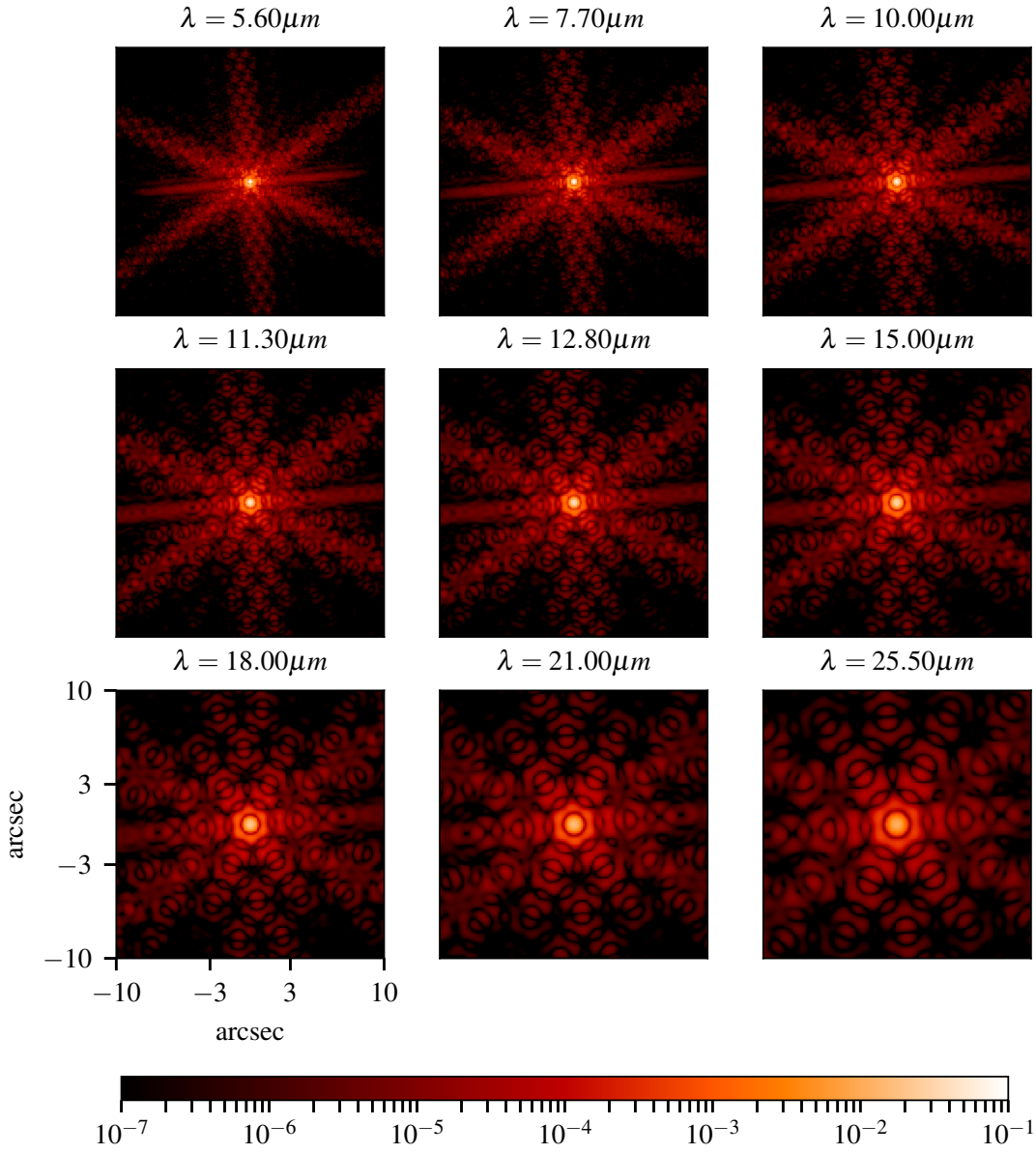


Figure 3.4: Monochromatic PSF of the JWST/MIRI imager simulated using *WebbPSF* [Perrin *et al.* 2012] and displayed in the same log scale. We clearly observe the dependency of the PSF to the wavelength.

data provided by MIRI imager are poor in spectral resolution.

The transmission profiles of the filters $\tau_p(\lambda)$ for $p = 1, \ldots, 9$ of MIRI imager filters are provided in [Bouchet *et al.* 2015] and reported in Figure 3.8. It is interesting to note that several filter profiles overlap which causes spectral cross-correlation between the multispectral data.

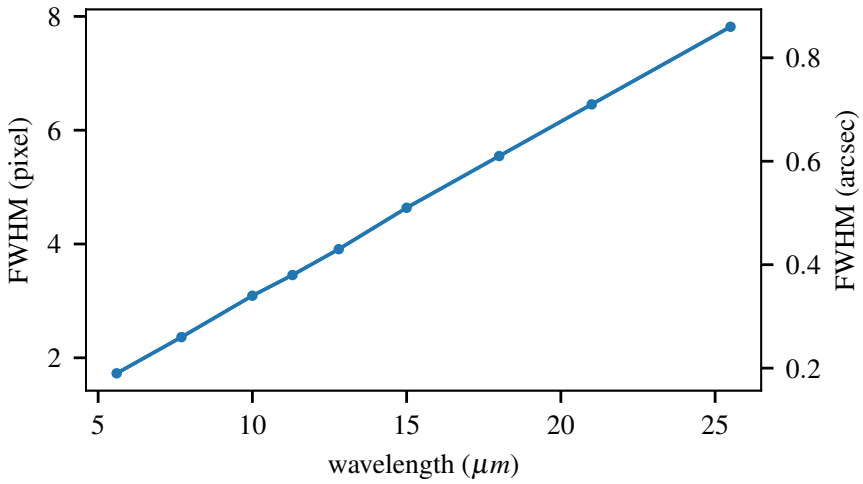


Figure 3.5: Illustration of the wavelength dependence of the PSF by displaying the Full-Width at Half-Maximum (FWHM) associated to the PSF in Figure 3.4. We notice a clear linear dependency of the PSF's FWHM to the wavelength, as expected from the diffraction theory.

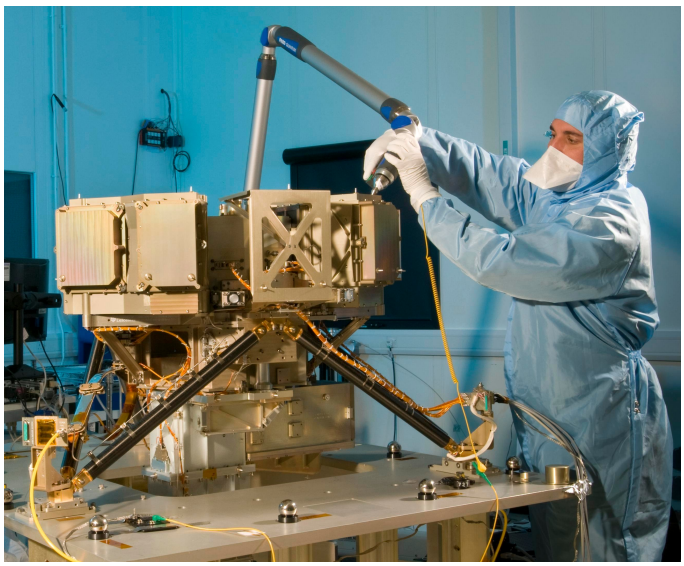


Figure 3.6: The Mid-infrared instrument Imager (`http://irfu.cea.fr`).

The detector is a device sensitive to infrared light. It converts the detected photon focused on the focal plane into a photocurrent in a detector pixel. It contains 1024 by 1024 pixels and a field of view of 113 arcsecond per 113 arcsecond (or "), whose 73.5" by 112.6" are dedicated to MIRI imager as shown in Figure 3.11. The pixel scale is 0.11 arcsecond which means that each pixel covers an area $\Omega_{pix} = 0.11^2$ arcsec^2. Additional technical details of MIRI imager detector can be in [Love *et al.* 2005, Rieke *et al.* 2015].

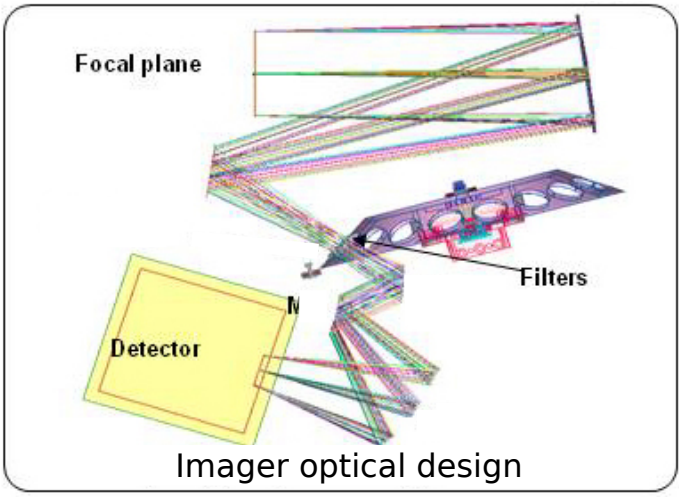


Figure 3.7: MIRI optical design shows the path of the photon flux inside the imager from the focal plane to the detector [Bouchet *et al.* 2015].

Table 3.1: Nine filters for MIRI Imager with their names, central wavelength and band-wavelength.

Filter position	$\lambda_c\,(\mu\text{m})$	$\triangle\lambda(\mu\text{m})$	$R = \lambda/\triangle\lambda$
$p = 1$	5.6	1.2	5.0
$p = 2$	7.7	2.2	3.5
$p = 3$	10.0	2.2	5.0
$p = 4$	11.3	0.7	16.0
$p = 5$	12.8	2.4	5.0
$p = 6$	15.5	3.0	5.0
$p = 7$	18.0	3.0	6.0
$p = 8$	21.0	5.0	4.0
$p = 9$	25.5	4.0	6.0

A non-ideal detector is characterized by its quantum efficiency (QE), which is measured in detected electrons per incident photon (e^-.photon^{-1}). Its spectral response $\eta(\lambda)$ depends on the wavelength and can be increased by applying an anti-reflection coating on the detector, see Figure 3.9.

The common spectral response used by the JWST community is the PCE (Photon Conversion Efficiency) [Glasse *et al.* 2015]. It is the product of filter transmission $\tau_p(\lambda)$ and detector quantum efficiency $\eta(\lambda)$. Thus, we define spectral bands response $\omega_p(\lambda)$ by

$$\omega_p(\lambda) = \tau_p(\lambda)\eta(\lambda), \quad \forall\lambda \in [\lambda_{\min}, \lambda_{\max}], \tag{3.4}$$

where a discrete values for the nine PCE profiles from 1 to 30 μm are publicly available on the website of the University of Arizona [4], see Figure 3.10.

[4] http://ircamera.as.arizona.edu/MIRI/ImPCE_TN-00072-ATC-Iss2.xlsx

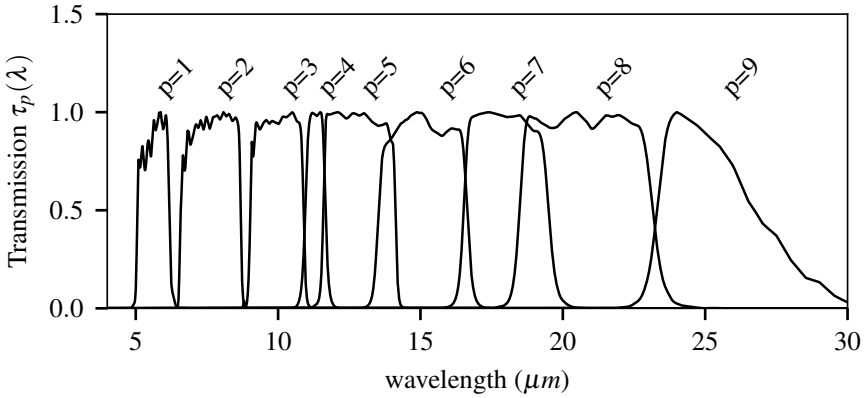


Figure 3.8: MIRI imager transmission profiles [Bouchet *et al.* 2015] covering the spectral range of 5 to 28 μm.

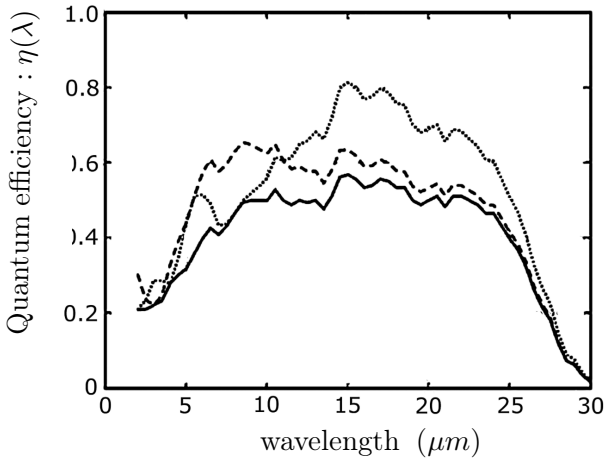


Figure 3.9: Measured quantum efficiency of bare detector material (solid line). The dashed line is a computed result assuming the array has an antireflection coating applied optimized for 6 μm, and the dotted line is for an AR coating optimized for 16 μm [Rieke *et al.* 2015].

3.3 Instrument Model

In this section we develop the multispectral imaging system response by establishing a relation between the input object to the output data. The multispectral imaging system we are considering is presented in Figure 3.12 as a block diagram, where its components are summarized in three blocks, namely: optical system, spectral bands and detector.

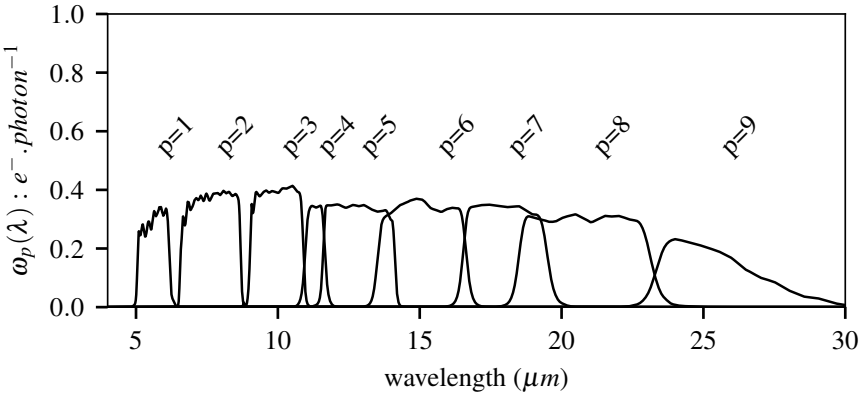


Figure 3.10: Nine spectral bands of the JWST/MIRI Imager [Glasse *et al.* 2015] also called PCE (Photon Conversion Efficiency).

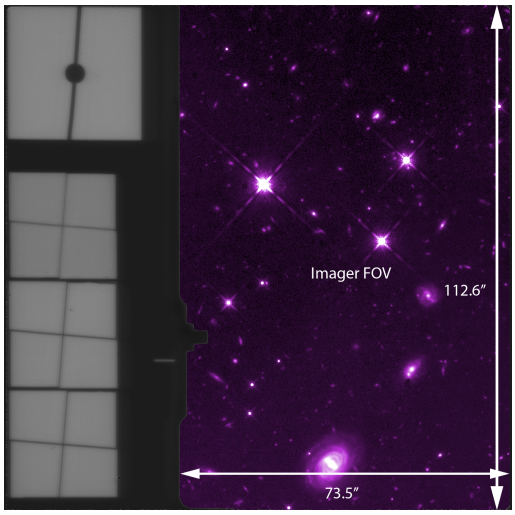


Figure 3.11: Illustration of sectioning of MIRI detector. A field of view (FOV) of $73.6'' \times 112.6''$ is dedicated to the imager while the rest is dedicted to the chronography and spectroscopy.

The parameters corresponding to each block are defined in the following sections. We first define the spatio-spectral object of interest as a function of three variables

$$\phi(\alpha, \beta, \lambda) : \mathbb{R}^3 \to \mathbb{R}$$

where $(\alpha, \beta) \in \mathbb{R}^2$ represent the spatial dimension and $\lambda \in \mathbb{R}_+$ represents the spectral one.

3.3.1 Optical System Response

Due to the light diffraction on the focal plane of the telescope, the optical system response is carried out by a spatial-convolution between the object of interest ϕ and a spectral-

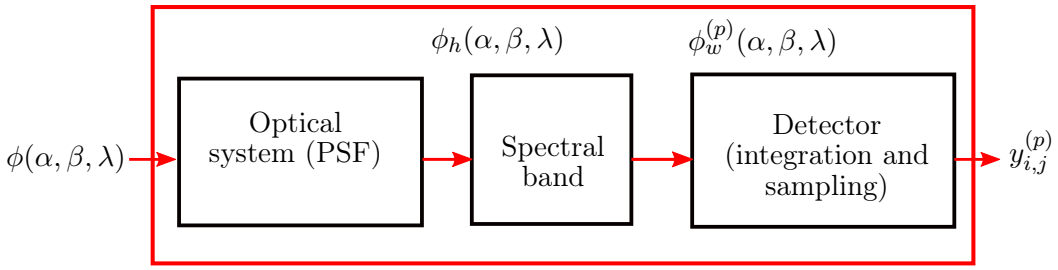


Figure 3.12: Illustration of the model of a multispectral imaging system with a block diagram.

variant PSF h [Goodman 2005] as follows

$$\begin{aligned}\phi_h(\alpha,\beta,\lambda) &= \phi \underset{\alpha,\beta}{*} h \\ &= \iint_{\mathbb{R}^2} \phi(\alpha',\beta',\lambda) h(\alpha-\alpha',\beta-\beta',\lambda) d\alpha' d\beta', \end{aligned} \tag{3.5}$$

where $\underset{\alpha,\beta}{*}$ stands for 2D spatial convolution.

The output of the optical system ϕ_h is a blurred version of ϕ, depending on the wavelength λ. This limits the spatial resolution of the observed object especially at long wavelengths, e.g. observations by the JWST/MIRI Imager.

3.3.2 Spectral Band Response

The diffracted object is spectrally filtered over P broad bands $\omega_p(\lambda)$, $p = 1,\ldots,P$. This operation is modeled by a multiplication with the spectral bands response as

$$\phi_w^{(p)}(\alpha,\beta,\lambda) = \omega_p(\lambda)\phi_h(\alpha,\beta,\lambda). \tag{3.6}$$

3.3.3 Detector Response

Finally, the object within a spectral band is integrated and sampled pixel-by-pixel on the detector matrix forming discrete multispectral data, $\mathcal{G}_{\text{samp}} = \{\boldsymbol{\theta}_{i,j}\}_{i,j=1}^{N_i,N_j}$, i and j indicates the pixel position, N_j and N_j are the number of pixel according to dimensions α and β, respectively. $\boldsymbol{\theta}_{i,j} = (\alpha_i,\beta_j)$ refers to the 2D angular position of the pixel (i,j). A basis function $b_{\text{samp}}^{i,j}(\boldsymbol{\theta})$ is defined over the pixel sensitive area Ω_{pix} to carry out spatial sampling of the pixel (i,j). Moreover, a detector noise and modeling errors term $n_{i,j}^{(p)}$ is added for each pixel (e.g. readout noise of the detector). The detector response for a single pixel is given by

$$y_{i,j}^{(p)} = \int_{\mathbb{R}_+} \left(\iint_{\Omega_{\text{pix}}} \phi_w^{(p)}(\boldsymbol{\theta},\lambda)\, b_{\text{samp}}^{i,j}(\boldsymbol{\theta}) d\boldsymbol{\theta} \right) d\lambda + n_{i,j}^{(p)}, \tag{3.7}$$

we discuss the choice of the basis function $b_{\text{samp}}^{i,j}$ later in Section 4.4.

3.3.4 Complete Model

Finally, the full equation of the instrument model is obtained by substituting Equations (3.5) and (3.6) in Equation (3.7). This yields

$$y_{i,j}^{(p)} = \iint_{\Omega_{\text{pix}}} \left(\int_{\mathbb{R}_+} \omega_p(\lambda) \left(\phi(\alpha,\beta,\lambda) \underset{\alpha,\beta}{*} h(\alpha,\beta,\lambda) \right) d\lambda \right) b_{\text{samp}}^{i,j}(\alpha,\beta)\, d\alpha d\beta + n_{i,j}^{(p)}. \quad (3.8)$$

The final instrument model in Equation (3.8) establishes a relation between the continuous spatio-spectral object $\phi(\alpha,\beta,\lambda)$ at the entrance of the multispectral imaging system to the discrete multispectral data $y_{i,j}^{(p)}$ through a complex instrument response, which includes spectral windowing and five integrations, two from spatial convolution, two for spatial sampling and one spectral integration. However, the above model does not include any non-ideal characteristic of the detector such as rejecting saturated data or bad data based on a predefined bad pixel mask, removing common noise components, correcting for anomalies in the initial frames in a detector integration caused by the reset [Gordon *et al.* 2015], which are assumed to be corrected upstream.

3.4 Unit Conversion : Physical Units and Electronic Units

The physical units of the object of interest at the input of the imaging system change throughout the observation and becomes electronic units by the output of the imaging system or multispectral data. It is interesting the highlight the transformation of the unit and determines the electronic units associated to the physical units that are used specifically for an astrophysical object.

3.4.1 Physical Units

The physical units of an astrophysical object or photon flux is *photon per seconds per pixel per microns* (photon.s^{-1}.pixel^{-1}.μm^{-1}) but a common units used is *Mega Jansky per steradian* (MJy.sr^{-1}) because flux densities are extremely small. The *Jansky* (*Jy*) is a unit usually used in radio and IR astronomy with 1 Jy = 10^{-26} J.s^{-1}.m^{-2}.Hz^{-1}.

In the following we provide details for the physical units conversion from MJy.sr^{-1} to photon.s^{-1}.pixel^{-1}.μm^{-1}. In fact we need basics from physical law of light, unit conversion and instrument parameters (telescope and MIRI imager). Let's first recall the energy of one photon in *Joule* (or *J*):

$$E_{\text{photon}} = \frac{hc}{\lambda_0} \text{ J}, \quad (3.9)$$

where $h = 6.62607 \times 10^{-34}$ J.s is the Planck's constant. $c = 2.998 \times 10^{8}$ m.s^{-1} is the speed of light in vacuum thus $hc = 1.98 \times 10^{-25}$ J.m. λ_0 is the wavelength of the light source in meter (m).

$$1 \text{ J} = \frac{\lambda_0}{hc} \text{ photons}. \quad (3.10)$$

The relation between wavelength λ and frequency ν of light is given by

$$\nu = \frac{c}{\lambda}. \quad (3.11)$$

Then we need the relation to convert the flux density per frequency F_ν to the flux density per wavelength F_λ:

$$F_\lambda d\lambda = F_\nu d\nu. \tag{3.12}$$

Also from Equation (3.11) we have

$$d\nu = \left| -\frac{c}{\lambda^2} \right| d\lambda \tag{3.13}$$

By replacing Equation (3.13) in (3.12) we can derive the photon flux per frequency

$$F_\nu = \frac{c}{\lambda^2} F_\lambda \tag{3.14}$$

We also need the aperture surface of the telescope $A_{\rm tel} = 25.03\, m^2$, and additional wavelength independent transmission terms such as $\tau_{\rm tel} = 0.88$, a transmission of the clean telescope optics at the start of the mission and $\tau_{\rm EOL} = 0.8$ is then used to account for the loss in transmission of all elements in the optical train up to the "end-of-life" of the nominal 5-year mission [Glasse *et al.* 2015].

The pixel scale of the MIRI imager detector is 0.11 arcsec.pixel^{-1} which means that each pixel covers an area $\Omega_{pix} = 0.11^2$ arcsec^2. Thus the solid angle per pixel in steradian (sr) is given by

$$\begin{aligned} \Omega_{\rm pix} &= 0.11^2 \times (2.35 \times 10^{-11}) \text{ sr.pixel}^{-1} \\ &= 2.84 \times 10^{-13} \text{ sr.pixel}^{-1} \end{aligned} \tag{3.15}$$

with

$$\begin{cases} 1 \text{ sr} & = 1 \text{ rad}^2 \\ 2.35 \times 10^{-11} \text{ sr} & = 1 \text{ arcsec}^2 \end{cases} \tag{3.16}$$

with $1^\circ = 60$ arcminutes $(') = 3600$ arcseconds$('')$.

Finally, unit conversion of a photon flux $[X]$ from MJy.sr^{-1} to photon.s^{-1}.pixel^{-1}.μm^{-1} at wavelength λ_0 is:

$$[X] \text{ MJy.sr}^{-1} \times C(\lambda_0) = [X] \text{ photon.s}^{-1}\text{.pixel}^{-1}\text{.}\mu\text{m}^{-1}$$

with

$$\begin{aligned} C(\lambda_0) &= 10^6.10^{-26}.\frac{\lambda_0}{hc}.\tau_{\rm tel} A_{\rm tel} \left(\frac{c}{\lambda_0^2} \right) \times 10^{-6}.\Omega_{\rm pix} \\ &= \left(\frac{\tau_{\rm tel} A_{\rm tel}}{\lambda_0} \right) 4.286 \times 10^{-6}. \end{aligned} \tag{3.17}$$

For instance let's consider a photon flux $F = 0.4$ MJy.sr^{-1} at $\lambda_1 = 5$ μm. The computation of the conversion coefficient using Equation (3.17) gives

$$\begin{aligned} C(\lambda_1) &= \left(\frac{0.88 \times 25.03}{5 \times 10^{-6}} \right) 4.286 \times 10^{-6} \\ &= 18.87 \end{aligned}$$

Thus, the converted photon flux is $F = 0.4 \times 18.87 = 7.55$ photon.s^{-1}.pixel^{-1}.μm^{-1}. We obtain the same values reported in [Glasse *et al.* 2015] and confirm the expression of the conversion coefficient of photon flux from MJy.sr^{-1} to photon.s^{-1}.pixel^{-1}.μm^{-1} in Equation (3.17). Throughout this rapport, the physical units of the spatio-spectral object is photon.s^{-1}.pixel^{-1}.μm^{-1} even if it is not mentioned explicitly.

3.4.2 Electronic Units

The electronic units of the multispectral data can be derived simply by using the instrument model in Equation (3.8) and the physical units of the object of interest, i.e.

$$\begin{cases} \tau_p(\lambda) & : \text{e}^-.\ \text{photon} \\ \phi(\alpha,\beta,\lambda) & : \text{photon.s}^{-1}.\text{pixel}^{-1}.\mu\text{m}^{-1} \\ h(\alpha,\beta,\lambda) & : \text{arcsec}^{-2} \text{ or } \text{sr}^{-1} \\ d\alpha d\beta & : \text{arcsec}^{2} \text{ or } \text{sr}^{1} \\ d\lambda & : \mu\text{m} \end{cases} \tag{3.18}$$

Therefore, the electronic units of the multispectral data $[X]$ for a single pixel are

$$y_{i,j}^{(p)} = [X]\ \text{e}^-.\text{s}^{-1}.\text{pixel}^{-1} \tag{3.19}$$

Same as above, the electronic units of the multispectral data is e$^-$.s^{-1}.pixel^{-1} even if it will not be mentioned explicitly.

3.5 Conclusion

This chapter deals with a model of a multispectral imaging system, particularly the MIRI imager on board the JWST. We first present the instrument and then highlight the issues related to its components, especially the limit of the spatial resolution of the object by the optical response (PSF) because of the diffraction of light on the focal plane of the telescope. Moreover, the spectral variance of the optical response of the JWST linearly depends on the wavelength, the larger the wavelength the wider the PSF, and the shorter the wavelength the narrower the PSF, as expected from the diffraction theory. In addition, the spectral resolution of the multispectral data is limited because of the spectral integration over broad bands, with an important wavelength ratio of ~ 5 and a very low spectral resolving power from 3.5 to 16, which implies that the multispectral data are poor in the spectral information of the original spatio-spectral object. Furthermore, an analytic expression of the PSF is computed thanks to the Fraunhofer theory, through a sum of 18 Fourier transform of 18 hexagonal-shape segments. In practice we use *WebbPSF*, a PSF simulator tool developed exclusively for the JWST mission due to many practical advantages.

We end up with an instrument model which rely on the input of the imaging system, a continuous original object ϕ, to the output a discrete multispectral data $y_{i,j}^{(p)}$. This

model is complex and takes into account the diffraction limitation of the optical system, which is modeled by a 2D spatial convolution of the object with a spectral-variant PSF, spectral integration over few broad bands, spatial sampling of the multispectral data over the detector matrix, and an additive term that account for modeling errors and noise.

In the next two chapters, we show how we incorporate the developed instrument model to develop spatio-spectral reconstruction methods.

Chapter 4

Multichannel Reconstruction of a Spatio-Spectral object from Low-Resolution Multispectral Data

Contents

4.1 Introduction

In this chapter we address the reconstruction of a discrete 2D+λ spatio-spectral object from a set of few 2D multispectral data, where the original object of interest is degraded by an instrument that suffers from diffraction, because of the limited size of its optical system, and from the spectral integration over broad bands. The instrument model considered is given by the Equation (3.8).

The linear forward model characterizes the relationship between the input (or unknown) of the multispectral imaging system and the multispectral data at the output. It accounts for degradation occurring within and between channels degradations or auto and cross-channel degradations, where the number of multispectral data is much lower than the number of spectral channels.

The reconstruction of the spatio-spectral object is based on regularization methods. We see that a naive least-squares method leads to an unstable solution because of the ill-conditioning of the observation matrix, thus the problem is ill-posed. A regularization term is added to correct the ill-posedness such as the spatial and spectral smoothness. In the last section we present reconstruction results with a comparison to multichannel 2D deconvolution for an application to the MIRI Imager on board the JWST.

4.2 Problem Statement

The accuracy of the instrument model developed in Chapter 3 depends on the knowledge of the instrument characteristics such as the monochromatic PSF. The real monochromatic PSF is the one measured during an observation by the MIRI imager, thus after the launch of the telescope. However, this task remains difficult to achieve because of the large wavelength number required in order to obtain a monochromatic PSF with high spectral resolution over a broad wavelength range of the instrument, i.e. 5 to 28 μm. In fact, conventional approaches generally neglect the spectral variation of the PSF in order to simplify the lack of knowledge of the real PSF [Guillard *et al.* 2010]. Several works consider 2D images and therefore need only a wavelength-independent 2D PSF. Another common approach, particularly in astrophysics, is to use a broadband (or effective) 2D PSF, e.g. [Geis & Lutz 2010, Aniano *et al.* 2011, R. Gastaud 2018], where the broadband PSFs are computed with an integral of the monochromatic PSFs, weighted by the spectral response of the instrument (filter and detector) and the object spectrum (or Source Energy Distribution) of a given astronomical object. Thus, its formulation is given by

$$h_{\text{broad}}(\alpha,\beta) = \frac{\int_{\mathbb{R}_+} \omega(\lambda) s(\lambda)\, h(\alpha,\beta,\lambda)\, d\lambda}{\int_{\mathbb{R}_+} \omega(\lambda) s(\lambda)\, d\lambda} \tag{4.1}$$

where the function $s(.)$ is the spectrum of the object, and $\omega(.)$ is the broad band spectral response of the instrument, e.g. Figure 3.10.

The drawbacks of this approach rely on three major points. (1) The spectral distribution (or spectrum) of the object: since the spectrum of an object being observed by the instrument is unknown, a predefined spectrum of another object (e.g. a black body at a given temperature or a flat spectrum) is used to compute the broadband PSF. An interesting question arises concern the choice of the predefined spectrum. in all cases, the broadband PSF is inaccurate unless it is computed from the original spectrum. (2) The spectral-independence of the broadband PSF: meaning that the same 2D PSF is used across the whole range of a band, whereas we know that the PSF of a diffraction-limited instrument is wavelength-dependent according to the diffraction theory. In fact, a unique PSF can only be used in case of narrow-band imaging, e.g. hyperspectral data, but not for a broad-band imaging, e.g. multispectral data. (3) The limitation of the spectral resolution: reconstruction of the object using the broadband 2D PSF is done band per band [Aniano *et al.* 2011, Boucaud *et al.* 2016]. Thus, the spectral resolution of multispectral data is limited by the number of instrument bands.

We illustrate these drawbacks on Figure 4.1. We clearly see the inaccuracy of spectral distribution of the reconstructed object (in red) using five broadband PSF. In addition, an ambiguity arises when the bands overlap where it is hard to restore the intensity of the object in the overlapping zone. This ambiguity becomes a serious limitation especially when the overlapping zone is large.

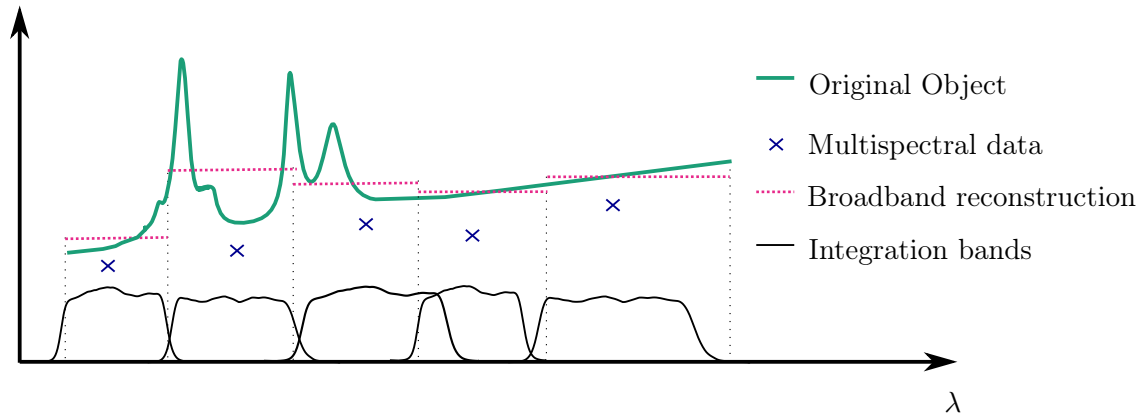


Figure 4.1: Inaccuracy of the broadband reconstruction method illustrated on five low-resolution multispectral data. The spectral distribution of the original object of interest is illustrated for a single spatial position over the wavelength range.

In [Soulez *et al.* 2013] authors proposed to use linear interpolation to model the spectral variation of the PSF, an idea that was used in [Denis *et al.* 2011] to model the spatial variation of the PSF. More recently, [Thiébaut *et al.* 2016] presented the existing models for spatial variant PSF based on speed and accuracy.

In the following sections, we propose to preserve the spectral variations of the PSF. Moreover, here instead of using a predefined spectrum, we propose to model the spectral distribution of the object by a piecewise linear function in order to approach the real one.

4.3 Object Model: Piecewise Linear Function

Having a set of discrete multispectral data, our goal is to reconstruct a discrete version of the object of interest from a set of low-resolution multispectral data. We model the continuous distribution of the object by a sum of discrete coefficients $\phi(\alpha_k, \beta_l, \lambda_m)$ weighted by a basis function $b(\alpha, \beta, \lambda)$, with k, l and m referring to the spatial and spectral positions of the sample, respectively. The basis function is defined as separable functions in order to distinguish the spatial and spectral reconstructions [Bongard *et al.* 2011], i.e.

$$b\,(\alpha, \beta, \lambda) = b_{\text{spat}}\,(\alpha, \beta)\, b_{\text{spec}}(\lambda). \tag{4.2}$$

where the subscript $_{\text{spat}}$ stands for spatial, and $_{\text{spec}}$ stands for spectral.

Moreover, as in Section 3.3.3, we define two discrete grids for the object $\mathcal{G}_{\text{spat}} = \{\alpha_k, \beta_l\}_{k,l=1}^{N_k,N_l}$ and $\mathcal{G}_{\text{spec}} = \{\lambda_m\}_{m=1}^{N_\lambda}$, N_k, N_l and N_λ refers to the number of pixels in

the grid according to dimensions α, β and λ, respectively. Therefore, the object model is given by

$$\phi(\alpha,\beta,\lambda) = \sum_{m=1}^{N_\lambda}\sum_{k=1}^{N_k}\sum_{l=1}^{N_l} \phi(\alpha_k,\beta_l,\lambda_m)\, b_{\text{spat}}^{k,l}(\alpha,\beta) b_{\text{spec}}^{m}(\lambda). \tag{4.3}$$

In order to handle the lack of spectral information in the multispectral data and to increase the spectral resolution of the object to reconstruct, we propose to model the spectral distribution of the object $b_{\text{spec}}(\lambda)$ by a uniform piecewise linear function, with N_λ channels [Hadj-Youcef *et al.* 2017a], see Figure 4.2. This choice allows us to obtain a simple model that preserves the spectral distribution of the object with less complexity, whereas the broadband model does not because the spectral variability of the PSF is neglected. The parameter N_λ controls the sharpness of the spectral sampling and compromises between this sharpness and the number of unknown channels. In fact, N_λ is to be set experimentally in order to satisfy a good reconstruction result.

We consider a uniform shift-basis function, i.e. $b_{\text{spec}}(\lambda) = b_{\text{spec}}(\lambda - \lambda_m)$ with $\lambda_m = m\triangle_\lambda$ and $\triangle_\lambda = \frac{\lambda_{N_\lambda}-\lambda_1}{N_\lambda - 1}$. This spectral uniform shift-basis function can be seen as the first-order uniform B-spline function [Thévenaz *et al.* 2000].

Hereafter we denote the discrete coefficients $\phi(\alpha_k,\beta_l,\lambda_m)$ by $x_{k,l}^{(m)}$ to denote the unknown parameter of the spatio-spectral object at (k,l)-*th* spatial position and m-*th* spectral channel in a finite dimensional space.

$$x_{k,l}^{(m)} = \phi(\alpha_k,\beta_l,\lambda_m). \tag{4.4}$$

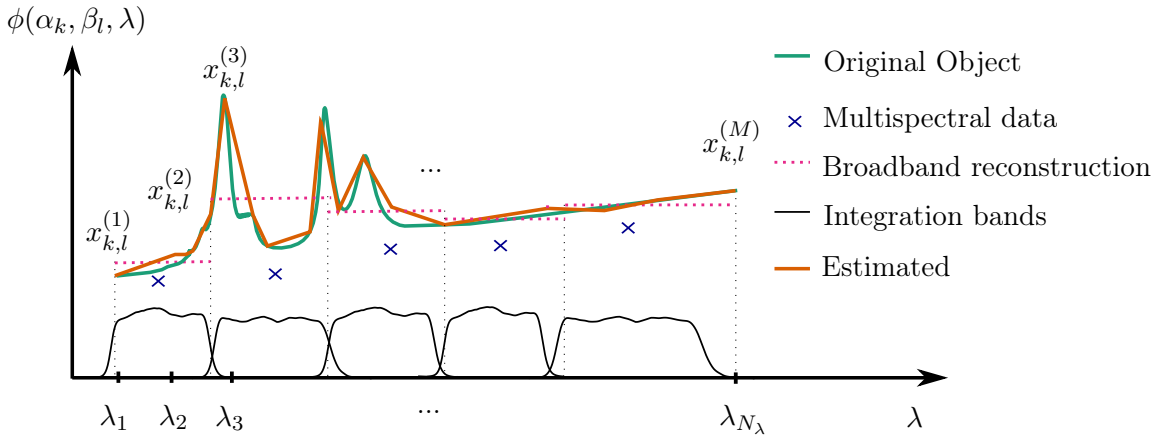


Figure 4.2: Illustration of the proposed modeling using a piecewise linear spectral model and comparison with the broadband reconstruction method on five low–resolution multispectral data. The spectral distribution of the original object is illustrated for a single spatial position (α_k,β_l) over the spectral band $[\lambda_1,\lambda_{N_\lambda}]$.

4.4 Forward Model: Definition of the Observation Matrix

In this section we develop a linear forward model characterizing the relationship between the discrete unknown parameter $\boldsymbol{x} \in \mathbb{R}^{N_\lambda N_k N_l}$ and the discrete multispectral data $\boldsymbol{y} \in \mathbb{R}^{PN_iN_j}$. It accounts for within channel and between channels degradations, respectively, where the number of multispectral data is much lower than the number of spectral channels, i.e. $N_\lambda \ll P$ [Hunt & Kubler 1984, Galatsanos *et al.* 1991, Schultz & Stevenson 1995]. The set of P low-resolution multispectral data is degraded by a spectral-variant PSF and integrated over broad spectral bands (or windows).

The linear forward model is obtained by substituting Equation (4.3) of the object model in the instrument model (3.8)

$$y_{i,j}^{(p)} = \iint_{\Omega_{\text{pix}}} \left(\int_{\mathbb{R}_+} \omega_p(\lambda) \left(\left(\sum_{m=1}^{N_\lambda} \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} x_{k,l}^{(m)} b_{\text{spat}}^{k,l}(\alpha,\beta) b_{\text{spec}}^{m}(\lambda) \right) \underset{\alpha,\beta}{*} h(\alpha,\beta,\lambda) \right) d\lambda \right) b_{\text{samp}}^{i,j}(\alpha,\beta) d\alpha d\beta + n_{i,j}^{(p)}.$$

where i, j indicates the pixel coordinates, and p refers to the band.

Since we are not interested in a super-resolution problem [Park *et al.* 2003] and we dispose multispectral data with high spatial resolution, we define $b_{\text{samp}}(.)$ and $b_{\text{spat}}(.)$ on a uniform or regular sampling grid with same sampling steps (or discretization intervals) $(\triangle\alpha)_{\text{samp}} = (\triangle\alpha)_{\text{spat}} = \triangle\alpha, (\triangle\beta)_{\text{samp}} = (\triangle\beta)_{\text{spat}} = \triangle\beta$ according to dimensions α and β, respectively. The sampling functions are called shift basis functions

$$b_{\text{samp}}^{i,j}(\alpha,\beta) = b_{\text{samp}}(\alpha - \alpha_i, \beta - \beta_j)$$

and

$$b_{\text{spat}}^{k,l}(\alpha,\beta) = b_{\text{spat}}(\alpha - \alpha_k, \beta - \beta_l)$$

with $\alpha_t = t\triangle\alpha$ and $\beta_t = t\triangle\beta$, for $t = \{i, j, k, l\}$.

By rearranging all wavelength-dependent terms in $h_{\text{int}}^{p,m}$ we obtain

$$y_{i,j}^{(p)} = \sum_{m=1}^{N_\lambda} \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} x_{k,l}^{(m)} \left(\iint_{\Omega_{\text{pix}}} \left(b_{\text{spat}} \underset{\alpha-\alpha_k,\beta-\beta_l}{*} h_{\text{int}}^{p,m} \right) b_{\text{samp}}(\alpha-\alpha_i, \beta-\beta_j) d\alpha d\beta \right) + n_{i,j}^{(p)}, \tag{4.5}$$

with

$$h_{\text{int}}^{p,m}(\alpha,\beta) = \int_{\mathbb{R}_+} \omega_p(\lambda) b_{\text{spec}}^{m}(\lambda) h(\alpha,\beta,\lambda)\, d\lambda. \tag{4.6}$$

$h_{\text{int}}^{p,m}$ is a wavelength independent PSF-like term. It is obtained by averaging monochromatic PSFs over the *p-th* spectral band, weighted by b_{spec}^{m} to account for spectral linear variations around the wavelength λ_m. Thus, the accurate the spectral variation in b_{spec}^{m}

the accurate the system response.

For computational purpose, we carry out the spectral integration in Equation (4.6) by defining a spectral grid with a uniform sampling $\mathcal{G}_\lambda = \{\lambda_i\}_{i=1}^{N_\lambda}$, N_λ is total number of wavelengths. $\lambda_i = i\triangle\lambda$ with $\triangle\lambda$ refers to the spectral sampling step. It is chosen to be sufficiently high in order to account properly the spectral variation of the PSF, e.g. we set N_λ in order to have $\triangle\lambda = 2.9\,10^{-2}\,\mu$m.

Furthermore Equation (4.5) can be simplified as

$$y_{i,j}^{(p)} = \sum_{m=1}^{N_\lambda}\sum_{k=1}^{N_k}\sum_{l=1}^{N_l} H_{i,j;k,l}^{p,m}\, x_{k,l}^{(m)} + n_{i,j}^{(p)}, \tag{4.7}$$

with

$$H_{i,j;k,l}^{p,m} = \iint_{\Omega_{\text{pix}}} \left(\left(\int_{\mathbb{R}_+} \omega_p(\lambda) h(\alpha,\beta,\lambda) b_{\text{spec}}^m(\lambda) d\lambda \right) \underset{\alpha,\beta}{*} b_{\text{spat}}(\alpha-\alpha_k, \beta-\beta_l) \right) b_{\text{samp}}(\alpha-\alpha_i, \beta-\beta_j) d\alpha d\beta \tag{4.8}$$

where $H_{i,j;k,l}^{p,m}$ represents an element of the observation matrix which models the detector spatial sampling and optical response through spatial convolution between the basis function of the object model and a PSF-like term $h_{\text{int}}^{p,m}$. This later is obtained after spectral integration of all wavelength-dependent terms, which includes the instrument parameters and the spectral model of the object, i.e. uniform piecewise linear function.

For the sampling function, we consider a separable rectangular impulse function over one pixel area $\Omega_{\text{pix}} = \triangle\alpha\triangle\beta$ [Yaroslavsky 2013]:

$$b_{\text{samp}}^{i,j}(\alpha,\beta) = \frac{1}{\triangle\alpha}\Pi\left(\frac{\alpha - i\triangle\alpha}{\triangle\alpha}\right) \times \frac{1}{\triangle\beta}\Pi\left(\frac{\beta - j\triangle\beta}{\triangle\beta}\right). \tag{4.9}$$

Other choices are possible such as the cardinal-sine [Bongard *et al.* 2011], Fourier transform of the rectangular function, or the 2D circular function [Yaroslavsky 2012].

Thus, Equation (4.8) becomes

$$\begin{aligned} H_{i,j;k,l}^{p,m} &= \frac{1}{\triangle\alpha}\frac{1}{\triangle\beta}\int_{i\triangle\alpha}^{(i+1)\triangle\alpha}\int_{j\triangle\beta}^{(j+1)\triangle\beta} \left(b_{\text{spat}} \underset{\alpha-\alpha_k,\beta-\beta_l}{*} h_{\text{int}}^{p,m} \right) d\alpha d\beta \\ &= \overline{\left[b_{\text{spat}} \underset{\alpha_i-\alpha_k,\beta_j-\beta_l}{*} h_{\text{int}}^{p,m} \right]} \\ &= \overline{\left[b_{\text{spat}} \underset{\alpha_{i-k},\beta_{j-l}}{*} h_{\text{int}}^{p,m} \right]} \\ &= H_{i-k,j-l}^{p,m} \end{aligned} \tag{4.10}$$

where $\overline{f(\alpha_i,\beta_j)}$ is the mean value of $f(\alpha,\beta)$ over the area of the pixel (i,j). Thus H naturally becomes a discrete convolution kernel.

Therefore, by rewriting Equation (4.7) with a matrix-vector notation, this yields

$$\boldsymbol{y}^{(p)} = \sum_{m=1}^{N_\lambda} \boldsymbol{H}^{p,m}\boldsymbol{x}^{(m)} + \boldsymbol{n}^{(p)}. \tag{4.11}$$

where the p-th multispectral image $\boldsymbol{y}^{(p)} \in \mathbb{R}^{N_i N_j}$ is a sum of N_λ discrete 2D convolutions of the spectral channels (unknown parameter) $\boldsymbol{x}^{(m)} \in \mathbb{R}^{N_k N_l}, m = 1, \ldots, N_\lambda$, with convolution matrices $\boldsymbol{H}^{p,m}, p = 1, \ldots, P$ and $m = 1, \ldots, N_\lambda$.

The forward model described by Equation (4.11) has many advantages. The observation matrices $\boldsymbol{H}^{p,m}$ are known and mainly depend on the accuracy of h_{int}, which is parametrized by the instruments parameters and the spectral model of the object. The model is linear, thus simple linear algebra can be applied to implement it. In addition, it allows us to obtain an explicit solution for the reconstruction (see next section). Computation of discrete convolution can be done efficiently through diagonalization in the Fourier domain [Hunt 1971] under the circular approximation assumption.

By combining all multispectral data set we obtain from Equation (4.11) the following multi-observation forward model

$$\underbrace{\begin{pmatrix} \boldsymbol{y}^{(1)} \\ \boldsymbol{y}^{(2)} \\ \vdots \\ \boldsymbol{y}^{(P)} \end{pmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{pmatrix} \boldsymbol{H}^{1,1} & \boldsymbol{H}^{1,2} & \cdots & \boldsymbol{H}^{1,N_\lambda} \\ \boldsymbol{H}^{2,1} & \boldsymbol{H}^{2,2} & \cdots & \boldsymbol{H}^{2,N_\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}^{P,1} & \boldsymbol{H}^{P,2} & \cdots & \boldsymbol{H}^{P,N_\lambda} \end{pmatrix}}_{\boldsymbol{H}} \underbrace{\begin{pmatrix} \boldsymbol{x}^{(1)} \\ \boldsymbol{x}^{(2)} \\ \vdots \\ \boldsymbol{x}^{(N_\lambda)} \end{pmatrix}}_{\boldsymbol{x}} + \underbrace{\begin{pmatrix} \boldsymbol{n}^{(1)} \\ \boldsymbol{n}^{(2)} \\ \vdots \\ \boldsymbol{n}^{(P)} \end{pmatrix}}_{\boldsymbol{n}}, \tag{4.12}$$

where $\boldsymbol{x} \in \mathbb{R}^{N_\lambda N_k N_l}$ is the stack of N_λ spectral channels represented in a vector form and each spectral channel $\boldsymbol{x}^{(m)}$ contains $N_k \times N_l$ pixels. Here $\boldsymbol{y} \in \mathbb{R}^{P N_i N_j}$ is the stack of all multispectral data observed with P broad-bands of the multispectral imaging system. $\boldsymbol{n} \in \mathbb{R}^{P N_i N_j}$ represents an additive unknown error associated to multispectral data in $\boldsymbol{y}$.

The block matrix $\boldsymbol{H} \in \mathbb{R}^{P N_i N_j \times N_\lambda N_k N_l}$ is defined by a set of $P \times N_\lambda$ Toeplitz and circulant sub-matrices $\boldsymbol{H}^{p,m} \in \mathbb{R}^{N_i N_j \times N_k N_l}, p = 1 \ldots, P, m = 1 \ldots, N_\lambda$. The block $\boldsymbol{H}^{p,t}, t = m$ represents the within (or auto) degradation, whereas the block $\boldsymbol{H}^{p,t}, t \neq m$ accounts for degradation occurring between channels. All sub-matrices are ill-conditioned, meaning that $\boldsymbol{H}$ is also ill-conditioned, which leads to an ill-posed problem. In addition, it is worth noting that the spectral distribution of the object is carried out by the spectral channels of $\boldsymbol{x}$ and are weighted by the block of the observation matrix.

We are interested in the particular case where we dispose a small number of low-resolution multispectral observed data compared to spectral channels, i.e. $P \ll N_\lambda$, which means there is a lack of spectral information in the data. For instance, for multispectral data observed by the JWST/MIRI imager we have $p = 9$ and $N_\lambda = 1000$. Hence the size of $\boldsymbol{H}$ is 9×256^2 by 1000×256^2, for a 256×256 pixel detector ($N_i = N_k$, $N_j = N_l$).

4.5 Reconstruction

4.5.1 Regularized Least-Squares

In this section we aim to reconstruct a multichannel spatio-spectral object through the reconstruction of its unknown parameter $\boldsymbol{x}$ using a set of multispectral data $\boldsymbol{y}$. A naive reconstruction consists of applying a Least-Squares method, where the solution $\widehat{\boldsymbol{x}}$ is obtained as the minimizer of a cost function

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \left\{ \sum_{p=1}^{P} \sum_{m=1}^{N_\lambda} \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \left(\boldsymbol{y}^{(p)} - \boldsymbol{H}^{p,m} \boldsymbol{x}^{(m)} \right)_{i,j}^2 \right\}. \tag{4.13}$$

Since the cost quadratic function is differentiable and $\boldsymbol{H}$ is linear, it gives a linear solution

$$\widehat{\boldsymbol{x}} = (\boldsymbol{H}^T \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{y}. \tag{4.14}$$

However, the matrix $\boldsymbol{H}^T \boldsymbol{H}$, composed of convolution block-matrices, is ill-conditioned. In fact it has a large condition number, defined by $\kappa(\boldsymbol{H}^T \boldsymbol{H}) = \left|\lambda_{\max}(\boldsymbol{H}^T \boldsymbol{H})\right| / \left|\lambda_{\min}(\boldsymbol{H}^T \boldsymbol{H})\right|$. If κ is large which is the case here, even a small error in $\boldsymbol{H}^T \boldsymbol{y}$ (e.g. due to the noise in $\boldsymbol{y}$) can cause a large error in $\boldsymbol{x}$. Therefore, the inversion problem in Equation (4.14) is ill-posed. The Least-squares solution is thus unstable because of the inversion of $\boldsymbol{H}^T \boldsymbol{H}$ [Neumaier 1998].

The simplest way to correct this ill-posedness is by adding some prior information about the solution, this leads to the regularized least-squares method [Demoment 1989]. It consists of adding a regularization term to the cost function in order to correct the ill-conditioning of $\boldsymbol{H}^T \boldsymbol{H}$ and stabilize the solution. In this chapter we are interested in a smooth spatio-spectral object. Hence two types of regularizations are of interest, spatial regularization $\mathcal{R}_{\text{spat}}(\boldsymbol{x})$ and spectral regularization $\mathcal{R}_{\text{spec}}(\boldsymbol{x})$. In this case the minimization problem becomes

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \left\{ \mathcal{J}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{x}, \boldsymbol{y}) + \mu_{\text{spat}}\, \mathcal{R}_{\text{spat}}(\boldsymbol{x}) + \mu_{\text{spec}}\, \mathcal{R}_{\text{spec}}(\boldsymbol{x}) \right\}, \tag{4.15}$$

where the first term measures the consistency of the sought object to the data,

$$\mathcal{Q}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2_{\boldsymbol{C}_n^{-1}} = (\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x})^t\, \boldsymbol{C}_n^{-1}\, (\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}), \tag{4.16}$$

with $\boldsymbol{C}_n$ refers to the covariance matrix of the multichannel noise vector $\boldsymbol{n}$. Since all multispectral data are observed with the same imaging system, we assume a special case of an identically independent distributed white Gaussian noise, i.e. $\boldsymbol{C}_n = \sigma_n \boldsymbol{I}_{PN_iN_j}$, with $\boldsymbol{I}_{PN_iN_j}$ refers to a, identity matrix of size PN_iN_j by PN_iN_j. Thus Equation (4.16) becomes

$$\mathcal{Q}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2. \tag{4.17}$$

The second term $\mathcal{R}_{\text{spat}}(\boldsymbol{x})$ is a spatial regularization. It enforces spatial smoothness between pixels of the sought object. It is defined by

$$\begin{aligned} \mathcal{R}_{\text{spat}}(\boldsymbol{x}) &= \sum_{m=1}^{N_\lambda} \left(\sum_{k=1}^{N_k} \sum_{l=1}^{N_l} \left(x_{k+1,l}^{(m)} - x_{k,l}^{(m)} \right)^2 + \left(x_{k,l+1}^{(m)} - x_{k,l}^{(m)} \right)^2 \right) \\ &= \|\boldsymbol{D}_{\text{spat}} \boldsymbol{x}\|_2^2 \end{aligned} \tag{4.18}$$

where $\boldsymbol{D}_{\text{spat}}$ is defined by

$$\boldsymbol{D}_{\text{spat}} = \begin{pmatrix} \boldsymbol{D} & & & \\ & \boldsymbol{D} & & \\ & & \ddots & \\ & & & \boldsymbol{D} \end{pmatrix} \in \mathbb{R}^{N_\lambda N_k N_l \times N_\lambda N_k N_l} \tag{4.19}$$

and $\boldsymbol{D}$ is the second-order finite difference operator along the spatial dimension of $\boldsymbol{x}$ under circularity conditions $x^{(m)}_{N_k+1,l} = x^{(m)}_{1,l}$ and $x^{(m)}_{k,N_l+1} = x^{(m)}_{k,1}$. The last term $\mathcal{R}_{\text{spec}}(\boldsymbol{x})$ refers to the spectral regularization. It enforces the similarity between intensity values of corresponding pixels in neighboring channels. It is defined by

$$\begin{aligned} \mathcal{R}_{\text{spec}}(\boldsymbol{x}) &= \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} \left(\sum_{m=1}^{N_\lambda} \left(x^{(m+1)}_{k,l} - x^{(m)}_{k,l} \right)^2 \right) \\ &= \|\boldsymbol{D}_{\text{spec}} \boldsymbol{x}\|_2^2 \end{aligned} \tag{4.20}$$

where $\boldsymbol{D}_{\text{spec}} \in \mathbb{R}^{N_\lambda N_k N_l \times N_\lambda N_k N_l}$ is the first-order finite difference operator along the spectral dimension of $\boldsymbol{x}$, defined in Section 4.5.3, under circularity condition $x^{(N_\lambda+1)} = x^{(1)}$. Regularization terms $\mu_{\text{spat}} \geq 0$ and $\mu_{\text{spec}} \geq 0$ are regularization parameters. They are set to adjust the trade-off between fidelity to data and spatial smoothness for μ_{spat}, and spectral smoothness between channels for μ_{spec}, respectively.

By rewriting the cost function $\mathcal{J}(\boldsymbol{x})$ in Equation (4.15), this yields

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\text{argmin}} \left\{ \mathcal{J}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \mu_{\text{spat}} \ \|\boldsymbol{D}_{\text{spat}} \boldsymbol{x}\|_2^2 + \mu_{\text{spec}} \ \|\boldsymbol{D}_{\text{spec}} \boldsymbol{x}\|_2^2 \right\}, \tag{4.21}$$

the resulting cost function $\mathcal{J}(\boldsymbol{x})$ is a sum of quadratic terms, therefore it is differentiable and strictly convex. Thus, the solution is obtained by solving :

$$\frac{\partial}{\partial \widehat{\boldsymbol{x}}} \mathcal{J}(\widehat{\boldsymbol{x}}) = \boldsymbol{0},$$

where

$$\begin{aligned} \frac{\partial}{\partial \widehat{\boldsymbol{x}}} \mathcal{J}(\widehat{\boldsymbol{x}}) &= \frac{\partial}{\partial \widehat{\boldsymbol{x}}} \Big(\boldsymbol{y}^T \boldsymbol{y} - \boldsymbol{y}^T \boldsymbol{H} \widehat{\boldsymbol{x}} - \widehat{\boldsymbol{x}}^T \boldsymbol{H}^T \boldsymbol{y} + \widehat{\boldsymbol{x}}^T \boldsymbol{H}^T \boldsymbol{H} \widehat{\boldsymbol{x}} + \mu_{\text{spat}} \, \widehat{\boldsymbol{x}}^T \boldsymbol{D}_{\text{spat}}^T \boldsymbol{D}_{\text{spat}} \widehat{\boldsymbol{x}} \\ &\quad + \mu_{\text{spec}} \, \widehat{\boldsymbol{x}}^T \boldsymbol{D}_{\text{spec}}^T \boldsymbol{D}_{\text{spec}} \widehat{\boldsymbol{x}} \Big) \\ &= -2\boldsymbol{H}^T \boldsymbol{y} + 2\boldsymbol{H}^T \boldsymbol{H} \boldsymbol{x} + 2\mu_{\text{spat}} \boldsymbol{D}_{\text{spat}}^T \boldsymbol{D}_{\text{spat}} \boldsymbol{x} + 2\mu_{\text{spec}} \boldsymbol{D}_{\text{spec}}^T \boldsymbol{D}_{\text{spec}} \boldsymbol{x} \\ &= -2 \left(\boldsymbol{H}^T \boldsymbol{y} - \left(\boldsymbol{H}^T \boldsymbol{H} + \mu_{\text{spat}} \boldsymbol{D}_{\text{spat}}^T \boldsymbol{D}_{\text{spat}} + \mu_{\text{spec}} \boldsymbol{D}_{\text{spec}}^T \boldsymbol{D}_{\text{spec}} \right) \boldsymbol{x} \right) \end{aligned} \tag{4.22}$$

Therefore, the global minimizer of $\mathcal{J}(\boldsymbol{x})$ is given by

$$\widehat{\boldsymbol{x}} = \underbrace{\left(\boldsymbol{H}^T \boldsymbol{H} + \mu_{\text{spat}} \, \boldsymbol{D}_{\text{spat}}^T \boldsymbol{D}_{\text{spat}} + \mu_{\text{spec}} \, \boldsymbol{D}_{\text{spec}}^T \boldsymbol{D}_{\text{spec}} \right)}_{\boldsymbol{Q}}^{-1} \boldsymbol{H}^T \boldsymbol{y}, \tag{4.23}$$

where $\boldsymbol{Q}$ is the Hessian matrix. It is a square block matrix of size $N_\lambda N_k N_l$ by $N_\lambda N_k N_l$ containing $N_\lambda \times N_\lambda$ Toeplitz matrices $\boldsymbol{Q}^{(p,m)}$. However, $\boldsymbol{Q}^{(p,m)} \neq \boldsymbol{Q}^{(p+t,m+t)} \quad \forall t \in \mathbb{N}_+$, thus, $\boldsymbol{Q}$ is block-Toeplitz but not Toeplitz itself. In addition, the matrix $\boldsymbol{Q}$ is very large and cannot be inverted in the spatio-spectral domain. In the next sections, we focus on the computation of the solution $\widehat{\boldsymbol{x}}$ iteratively without inverting $\boldsymbol{Q}$, and directly by inverting $\boldsymbol{Q}$ in the Fourier domain.

4.5.2 Optimization Algorithms

4.5.2.1 Gradient Descent Algorithm

The first algorithm is the gradient descent or steepest descent algorithm [Shewchuk 1994]. It computes the solution without inverting the matrix $\boldsymbol{Q}$ through an iterative scheme by updating the solution $\boldsymbol{x}_n$ for the iteration n and this is done by taking steps $\boldsymbol{r}_n$ in the opposite direction of the cost function gradient, i.e. $\boldsymbol{r}_n = -\partial\mathcal{J}(\boldsymbol{x}_n)/\partial\boldsymbol{x}_n$, which is equal to $\boldsymbol{H}^T\boldsymbol{y} - \boldsymbol{Q}\boldsymbol{x}_n$ in case of the cost function in Equation (4.21). All the steps are controlled by a convergence parameter a_n, where a small value of a_n will slow down the convergence to the global minimum whereas a high value will speed up the convergence but could lead to divergence. The convergence parameter can be computed automatically at every iteration (see [Shewchuk 1994]) or chosen manually. Moreover, an arbitrary initialization $\widehat{\boldsymbol{x}}_0$ is required in order to update the solution, see Algorithm 4.1.

Algorithm 4.1 Gradient Descent Algorithm

Input: $\boldsymbol{H}, \boldsymbol{y}, \boldsymbol{Q}, a$
Initialization : $\widehat{\boldsymbol{x}}_0 = \boldsymbol{0}$
for $n = 0 : N_{iter}$ **do**
  Directions of the steepest descent
  $\boldsymbol{r}_n = \boldsymbol{H}^T\boldsymbol{y} - \boldsymbol{Q}\widehat{\boldsymbol{x}}_n$
  Computation of the next iteration
  $\widehat{\boldsymbol{x}}_{n+1} = \widehat{\boldsymbol{x}}_n + a\,\boldsymbol{r}_n$
return $\widehat{\boldsymbol{x}}$

4.5.2.2 Conjugate Gradient Algorithm

The gradient descent algorithm needs a lot of iterations to reach convergence, especially for a multi-variable cost function. Another optimization algorithm called the Conjugate Gradient (CG) algorithm requires fewer iterations than the gradient descent to converge. It consists of minimizing the cost function in the opposite direction of n Q-orthogonal conjugated vectors $\left\{\boldsymbol{d}_{(0)}, \ldots, \boldsymbol{d}_{(n-1)}\right\}$ so the conjugate gradient algorithm does not make a step in the same direction as earlier steps. Moreover, same as the gradient descent algorithm the convergence parameter a_n is updated at every iteration, see Algorithm 4.2. The full description of the algorithm is provided in [Shewchuk 1994].

Subsequently, we keep the conjugate gradient algorithm for implementation. In addition, the heaviest calculation in the algorithm is the computation of high-dimensional

Algorithm 4.2 Conjugate Gradient Algorithm [Shewchuk 1994]

Input $\boldsymbol{H}, \boldsymbol{y}, \boldsymbol{Q}$
Initialization : $\widehat{\boldsymbol{x}}_0 = \boldsymbol{0}, \boldsymbol{r}_0 = \boldsymbol{d}_0 = \boldsymbol{H}^T\boldsymbol{y} - \boldsymbol{Q}\widehat{\boldsymbol{x}}_0$
for $n = 0 : N_{iter}$ **do**
  Convergence parameter
  $a_n \leftarrow \frac{\boldsymbol{r}_n^T\boldsymbol{r}_n}{\boldsymbol{d}_n^T\boldsymbol{Q}\boldsymbol{d}_n}$
  Computation of the next iteration
  $\widehat{\boldsymbol{x}}_{n+1} \leftarrow \widehat{\boldsymbol{x}}_n - a_n\,\boldsymbol{d}_n$
  Conjugate directions
  $\boldsymbol{r}_{n+1} \leftarrow \boldsymbol{r}_n + a_n\boldsymbol{Q}\boldsymbol{d}_n$
  $b_n \leftarrow \frac{\boldsymbol{r}_{n+1}^T\boldsymbol{r}_{n+1}}{\boldsymbol{r}_n^T\boldsymbol{r}_n}$
  $\boldsymbol{d}_{n+1} \leftarrow \boldsymbol{r}_{n+1} + b_{n+1}\boldsymbol{d}_n$
return $\widehat{\boldsymbol{x}}$

matrix-vector multiplication $\boldsymbol{Q}\boldsymbol{d}_n$ for every iteration. Thus, we consider the circular approximation of the blocks of $\boldsymbol{Q}$ in order to speed up the computation in the Fourier domain.

4.5.3 Diagonalization in the Fourier domain

In this section we propose a second method to compute explicitly the solution in Equation (4.23), i.e.

$$\widehat{\boldsymbol{x}} = \underbrace{\left(\boldsymbol{H}^T\boldsymbol{H} + \mu_{\text{spat}}\,\boldsymbol{D}_{\text{spat}}^T\boldsymbol{D}_{\text{spat}} + \mu_{\text{spec}}\,\boldsymbol{D}_{\text{spec}}^T\boldsymbol{D}_{\text{spec}}\right)}_{\boldsymbol{Q}}{}^{-1}\boldsymbol{H}^T\boldsymbol{y}, \tag{4.24}$$

with

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}^{(1)} \\ \boldsymbol{x}^{(2)} \\ \vdots \\ \boldsymbol{x}^{(N_\lambda)} \end{pmatrix} \text{ and } \boldsymbol{H}^T\boldsymbol{y} = \begin{pmatrix} \sum_p (\boldsymbol{H}^{p,1})^T\boldsymbol{y}^{(p)} \\ \sum_p (\boldsymbol{H}^{p,2})^T\boldsymbol{y}^{(p)} \\ \vdots \\ \sum_p (\boldsymbol{H}^{(p,N_\lambda)})^T\boldsymbol{y}^{(p)} \end{pmatrix} \tag{4.25}$$

This method consists of computing the solution by inverting the Hessian matrix $\boldsymbol{Q}$ explicitly through diagonalization of its blocks $\boldsymbol{Q}^{i,j}, \quad i,j = 1,\ldots,N_\lambda$ in the Fourier domain [Galatsanos *et al.* 1991]. The inverting procedure is detailed in the following and is summarized in four steps:

- Determination of the analytic expression of $\boldsymbol{Q}^{i,j}$.
- Diagonalization of $\boldsymbol{Q}^{i,j}$ in the Fourier domain: $\Rightarrow \boldsymbol{\Lambda}^{i,j}$.
- Determination of $\boldsymbol{\Lambda}_{\boldsymbol{Q}}$
- Inverting $\boldsymbol{\Lambda}_{\boldsymbol{Q}}$ by using a parallel computation

Determination of the analytic expression of $\boldsymbol{Q}^{i,j}$:

We first detail expressions of $\boldsymbol{Q}^{i,j}, i, j = 1, \ldots, N_\lambda$ and then proceed to the inversion. The matrix $\boldsymbol{Q}$ is a sum of three block-circulant matrices, $\boldsymbol{H}^T\boldsymbol{H}$, $\boldsymbol{D}_{\text{spat}}^T\boldsymbol{D}_{\text{spat}}$ and $\boldsymbol{D}_{\text{spec}}^T\boldsymbol{D}_{\text{spec}}$. Each matrix is defined in the following.

1. The first matrix $\boldsymbol{H}^T\boldsymbol{H}$ is given by

$$\boldsymbol{H}^T\boldsymbol{H} = \begin{pmatrix} \left(\boldsymbol{H}^{1,1}\right)^T & \left(\boldsymbol{H}^{2,1}\right)^T & \cdots & \left(\boldsymbol{H}^{P,1}\right)^T \\ \left(\boldsymbol{H}^{1,2}\right)^T & \left(\boldsymbol{H}^{2,2}\right)^T & \cdots & \left(\boldsymbol{H}^{P,2}\right)^T \\ \vdots & \vdots & \ddots & \vdots \\ \left(\boldsymbol{H}^{1,N_\lambda}\right)^T & \left(\boldsymbol{H}^{2,N_\lambda}\right)^T & \cdots & \left(\boldsymbol{H}^{P,N_\lambda}\right)^T \end{pmatrix} \begin{pmatrix} \boldsymbol{H}^{1,1} & \boldsymbol{H}^{1,2} & \cdots & \boldsymbol{H}^{1,N_\lambda} \\ \boldsymbol{H}^{2,1} & \boldsymbol{H}^{2,2} & \cdots & \boldsymbol{H}^{2,N_\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}^{P,1} & \boldsymbol{H}^{P,2} & \cdots & \boldsymbol{H}^{P,N_\lambda} \end{pmatrix}$$

Hence,

$$\boldsymbol{H}^T\boldsymbol{H} = \begin{pmatrix} \sum_{p=1}^{P} \left(\boldsymbol{H}^{p,1}\right)^T \boldsymbol{H}^{p,1} & \cdots & \sum_{p=1}^{P} \left(\boldsymbol{H}^{p,1}\right)^T \boldsymbol{H}^{p,N_\lambda} \\ \sum_{p=1}^{P} \left(\boldsymbol{H}^{p,2}\right)^T \boldsymbol{H}^{p,1} & \ddots & \sum_{p=1}^{P} \left(\boldsymbol{H}^{p,2}\right)^T \boldsymbol{H}^{p,N_\lambda} \\ \vdots & \ddots & \vdots \\ \sum_{p=1}^{P} \left(\boldsymbol{H}^{p,N_\lambda}\right)^T \boldsymbol{H}^{p,1} & \cdots & \sum_{p=1}^{P} \left(\boldsymbol{H}^{p,N_\lambda}\right)^T \boldsymbol{H}^{p,N_\lambda} \end{pmatrix} \tag{4.26}$$

2. By using the definition of $\boldsymbol{D}_{\text{spat}}$ in Equation (4.19), the second matrix $\boldsymbol{D}_{\text{spat}}^T\boldsymbol{D}_{\text{spat}}$ is

$$\begin{aligned} \boldsymbol{D}_{\text{spat}}^T\boldsymbol{D}_{\text{spat}} &= \begin{pmatrix} \boldsymbol{D}^T & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}^T & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{D}^T \end{pmatrix} \begin{pmatrix} \boldsymbol{D} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{D} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{D}^T\boldsymbol{D} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}^T\boldsymbol{D} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{D}^T\boldsymbol{D} \end{pmatrix} \end{aligned} \tag{4.27}$$

where $\boldsymbol{D} \in \mathbb{R}^{N_k N_l \times N_k N_l}$ is a second-order finite difference matrix.

3. Concerning the expression of the third matrix $\boldsymbol{D}_{\text{spec}}^T\boldsymbol{D}_{\text{spec}}$, we first define $\boldsymbol{D}_{\text{spec}}$ by a first-order finite difference matrix explicitly. For instance the first-order difference matrix $\boldsymbol{C}$ for a vector of size N_λ is given by

$$\boldsymbol{C} = \begin{pmatrix} -1 & 1 & 0 & \ldots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & -1 & 1 \end{pmatrix} \in \mathbb{R}^{(N_\lambda - 1) \times N_\lambda}.$$

However, the matrix $\boldsymbol{C}$ needs to be a circular matrix in order to carry out the diagonalization in the Fourier domain, thus we consider a circular approximation of $\boldsymbol{C}$ by adding one row :

$$\boldsymbol{C} \approx \left(\begin{array}{ccccc} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 & 1 \\ \hline 1 & 0 & \dots & 0 & -1 \end{array}\right) \in \mathbb{R}^{N_\lambda \times N_\lambda}, \tag{4.28}$$

It remains to perform a difference between intensities of the same pixel from different channels of $\boldsymbol{x}$ (a vectorized version of the 2D+λ cube). We do this by introducing a Kronecker product $\otimes$ between $\boldsymbol{C}$ and an identity matrix $\boldsymbol{I}_{N_k N_l}$ of size $N_k N_l$ by $N_k N_l$:

$$\boldsymbol{D}_{\text{spec}} = \boldsymbol{C} \otimes \boldsymbol{I}_{N_k N_l}. \tag{4.29}$$

Thus,

$$\begin{aligned} \boldsymbol{D}_{\text{spec}}^T \boldsymbol{D}_{\text{spec}} &= \left(\left(\begin{array}{ccccc} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & 0 & -1 & \ddots & 0 \\ 0 & & \ddots & \ddots & 1 \\ 1 & 0 & \dots & 0 & -1 \end{array}\right)^T \left(\begin{array}{ccccc} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & 0 & -1 & \ddots & 0 \\ 0 & & \ddots & \ddots & 1 \\ 1 & 0 & \dots & 0 & -1 \end{array}\right)\right) \otimes \boldsymbol{I}_{N_k N_l} \\ &= \left(\begin{array}{cccccc} 2 & -1 & 0 & \cdots & 0 & -1 \\ -1 & 2 & -1 & 0 & & 0 \\ 0 & -1 & 2 & -1 & \ddots & \vdots \\ \vdots & 0 & -1 & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & 2 & -1 \\ -1 & 0 & \cdots & 0 & -1 & 2 \end{array}\right) \otimes \boldsymbol{I}_{N_k N_l} \end{aligned}$$

Lastly, the third matrix $\boldsymbol{D}_{\text{spec}}^T \boldsymbol{D}_{\text{spec}}$ is then defined as

$$\boldsymbol{D}_{\text{spec}}^T \boldsymbol{D}_{\text{spec}} = \left(\begin{array}{cccccc} 2\boldsymbol{I}_{N_k N_l} & -\boldsymbol{I}_{N_k N_l} & \boldsymbol{0} & \cdots & \boldsymbol{0} & -\boldsymbol{I}_{N_k N_l} \\ -\boldsymbol{I}_{N_k N_l} & 2\boldsymbol{I}_{N_k N_l} & -\boldsymbol{I}_{N_k N_l} & \boldsymbol{0} & & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_{N_k N_l} & 2\boldsymbol{I}_{N_k N_l} & -\boldsymbol{I}_{N_k N_l} & \ddots & \vdots \\ \vdots & \boldsymbol{0} & -\boldsymbol{I}_{N_k N_l} & \ddots & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & & \ddots & \ddots & 2\boldsymbol{I}_{N_k N_l} & -\boldsymbol{I}_{N_k N_l} \\ -\boldsymbol{I}_{N_k N_l} & \boldsymbol{0} & \cdots & \boldsymbol{0} & -\boldsymbol{I}_{N_k N_l} & 2\boldsymbol{I}_{N_k N_l} \end{array}\right) \tag{4.30}$$

Therefore, the analytic expression of $\boldsymbol{Q}^{i,j}$ using Equations (4.26)-(4.27)-(4.30) are:

$$\begin{cases} \boldsymbol{Q}^{m,m} = \sum_p \left(\boldsymbol{H}^{p,m}\right)^T \boldsymbol{H}^{(p,m)} + \mu_{\text{spat}} \boldsymbol{D}^T \boldsymbol{D} - 2\mu_{\text{spec}} \boldsymbol{I}_{N_k N_l} & m = 1, \ldots, N_\lambda \\ \boldsymbol{Q}^{m+1,m} = \sum_p \left(\boldsymbol{H}^{p,m+1}\right)^T \boldsymbol{H}^{(p,m)} - \mu_{\text{spec}} \boldsymbol{I}_{N_k N_l} & m = 1, \ldots, N_\lambda - 1 \\ \boldsymbol{Q}^{m,m+1} = \sum_p \left(\boldsymbol{H}^{p,m}\right)^T \boldsymbol{H}^{(p,m+1)} - \mu_{\text{spec}} \boldsymbol{I}_{N_k N_l} & m = 1, \ldots, N_\lambda - 1 \\ \boldsymbol{Q}^{m,m+2} = \sum_p \left(\boldsymbol{H}^{p,m}\right)^T \boldsymbol{H}^{(p,m+2)} & m = 1 : N_\lambda - 2 \\ \boldsymbol{Q}^{m+2,m} = \sum_p \left(\boldsymbol{H}^{p,m+2}\right)^T \boldsymbol{H}^{(p,m)} & m = 1 : N_\lambda - 2 \\ \boldsymbol{Q}^{N_\lambda,1} = \sum_p \left(\boldsymbol{H}^{p,m}\right)^T \boldsymbol{H}^{(p,m)} - \mu_{\text{spec}} \boldsymbol{I}_{N_k N_l} & \\ \boldsymbol{Q}^{1,N_\lambda} = \sum_p \left(\boldsymbol{H}^{p,m}\right)^T \boldsymbol{H}^{(p,m)} - \mu_{\text{spec}} \boldsymbol{I}_{N_k N_l} & \end{cases} \tag{4.31}$$

Diagonalization of $\boldsymbol{Q}^{i,j}$ in the Fourier domain:

The diagonalization of the matrices $\boldsymbol{Q}^{i,j}, i, j = 1, \ldots, N_\lambda$ implies the diagonalization of the circulant matrices $\boldsymbol{H}^{p,m}, \boldsymbol{D}$ and $\boldsymbol{I}_{N_k N_l}$ in the Fourier domain, see [Hunt 1971]. This yields

$$\begin{cases} \boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m} & = \boldsymbol{F} \boldsymbol{H}^{p,m} \boldsymbol{F}^\dagger, \quad p = 1, \ldots, P \text{ and } m = 1, \ldots, N_\lambda \\ \boldsymbol{\Lambda}_{\boldsymbol{D}} & = \boldsymbol{F} \boldsymbol{D} \boldsymbol{F}^\dagger \\ \boldsymbol{\Lambda}_{\boldsymbol{I}} & = \boldsymbol{F} \boldsymbol{I}_{N_k N_l} \boldsymbol{F}^\dagger = \boldsymbol{I}_{N_k N_l} \end{cases} \tag{4.32}$$

where $\boldsymbol{F}$ and $\boldsymbol{F}^\dagger$ are the unitary discrete Fourier transform (DFT) matrix and its conjugate, respectively, such as $\boldsymbol{F}^\dagger \boldsymbol{F} = \boldsymbol{I}$. † symbolizes the conjugate transpose. Moreover, $\boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m}$ and $\boldsymbol{\Lambda}_{\boldsymbol{D}}$ are diagonal matrices whose elements are the eigenvalues of $\boldsymbol{H}^{p,m}$ and $\boldsymbol{D}$ [Hunt 1971], respectively. Therefore, we can write

$$\boldsymbol{\Lambda}^{i,j} = \boldsymbol{F} \boldsymbol{Q}^{i,j} \boldsymbol{F}^\dagger, \quad i, j = 1, \ldots, N_\lambda \tag{4.33}$$

with

$$\begin{cases} \boldsymbol{\Lambda}^{m,m} = \sum_p \left(\boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m}\right)^\dagger \boldsymbol{\Lambda}_{\boldsymbol{H}}{}^{(p,m)} + \mu_{\text{spat}} \boldsymbol{\Lambda}_{\boldsymbol{D}}^\dagger \boldsymbol{\Lambda}_{\boldsymbol{D}} - 2\mu_{\text{spec}} \boldsymbol{\Lambda}_{\boldsymbol{I}} & m = 1, \ldots, N_\lambda \\ \boldsymbol{\Lambda}^{m+1,m} = \sum_p \left(\boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m+1}\right)^\dagger \boldsymbol{\Lambda}_{\boldsymbol{H}}{}^{(p,m)} - \mu_{\text{spec}} \boldsymbol{\Lambda}_{\boldsymbol{I}} & m = 1, \ldots, N_\lambda - 1 \\ \boldsymbol{\Lambda}^{m,m+1} = \sum_p \left(\boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m}\right)^\dagger \boldsymbol{\Lambda}_{\boldsymbol{H}}{}^{(p,m+1)} - \mu_{\text{spec}} \boldsymbol{\Lambda}_{\boldsymbol{I}} & m = 1, \ldots, N_\lambda - 1 \\ \boldsymbol{\Lambda}^{m,m+2} = \sum_p \left(\boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m}\right)^\dagger \boldsymbol{\Lambda}_{\boldsymbol{H}}{}^{(p,m+2)} & m = 1 : N_\lambda - 2 \\ \boldsymbol{\Lambda}^{m+2,m} = \sum_p \left(\boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m+2}\right)^\dagger \boldsymbol{\Lambda}_{\boldsymbol{H}}{}^{(p,m)} & m = 1 : N_\lambda - 2 \\ \boldsymbol{\Lambda}^{N_\lambda,1} = \sum_p \left(\boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m}\right)^\dagger \boldsymbol{\Lambda}_{\boldsymbol{H}}{}^{(p,m)} - \mu_{\text{spec}} \boldsymbol{\Lambda}_{\boldsymbol{I}} & \\ \boldsymbol{\Lambda}^{1,N_\lambda} = \sum_p \left(\boldsymbol{\Lambda}_{\boldsymbol{H}}^{p,m}\right)^\dagger \boldsymbol{\Lambda}_{\boldsymbol{H}}{}^{(p,m)} - \mu_{\text{spec}} \boldsymbol{\Lambda}_{\boldsymbol{I}} & \end{cases} \tag{4.34}$$

Determination of $\boldsymbol{\Lambda_Q}$:

As a result, the Hessian matrix $\boldsymbol{Q}$ becomes

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{Q}^{1,1} & \boldsymbol{Q}^{1,2} & \dots & \boldsymbol{Q}^{1,N_\lambda} \\ \boldsymbol{Q}^{2,1} & \boldsymbol{Q}^{2,2} & \dots & \boldsymbol{Q}^{2,N_\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{Q}^{N_\lambda,1} & \boldsymbol{Q}^{N_\lambda,2} & \dots & \boldsymbol{Q}^{N_\lambda,N_\lambda} \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} \boldsymbol{F}^\dagger & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}^\dagger & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{F}^\dagger \end{pmatrix}}_{\overline{\boldsymbol{F}}^\dagger} \underbrace{\begin{pmatrix} \boldsymbol{\Lambda}^{1,1} & \boldsymbol{\Lambda}^{1,2} & \dots & \boldsymbol{\Lambda}^{1,N_\lambda} \\ \boldsymbol{\Lambda}^{2,1} & \boldsymbol{\Lambda}^{2,2} & \dots & \boldsymbol{\Lambda}^{2,N_\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Lambda}^{N_\lambda,1} & \boldsymbol{\Lambda}^{N_\lambda,2} & \dots & \boldsymbol{\Lambda}^{N_\lambda,N_\lambda} \end{pmatrix}}_{\boldsymbol{\Lambda_Q}} \underbrace{\begin{pmatrix} \boldsymbol{F} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{F} \end{pmatrix}}_{\overline{\boldsymbol{F}}} \tag{4.35}$$

where $\overline{\boldsymbol{F}} = diag\{\boldsymbol{F}, \boldsymbol{F}, \dots, \boldsymbol{F}\}$ and $\overline{\boldsymbol{F}}^\dagger$ are the multichannel unitary DFT matrices. The matrix $\boldsymbol{\Lambda_Q}$ is a non-diagonal block-diagonal (NDBD) matrix [Galatsanos *et al.* 1991].

Inverting $\boldsymbol{\Lambda_Q}$ through parallel computation:

The inversion of a NDBD matrix such as $\boldsymbol{\Lambda_Q}$ relies on extracting $N_k N_l$ matrices of size N_λ by N_λ, i.e. $\boldsymbol{R}^{(k)} \in \mathbb{R}^{N_\lambda \times N_\lambda}, \quad k = 1, \dots, N_k N_l$, and inverting them separately. These two steps are illustrated in the following:

1. In Figure 4.3 we present an illustration on extraction of the matrix $\boldsymbol{R}$. This extraction process is done in parallel so we have instantaneously all matrices $\boldsymbol{R}^{(k)}, \quad k = 1, \dots, N_k N_l$. This operation is formulated as

$$\boldsymbol{R}^k(i,j) = \boldsymbol{\Lambda}_{\boldsymbol{Q}}^{i,j}(k,k), \quad k = 1, 2, \dots, N_k N_l \text{ and } i, j = 1, 2, \dots, N_\lambda. \tag{4.36}$$

2. After extraction, we proceed to the inversion of the matrix under the condition that $\boldsymbol{R}^{(k)}$ is not a singular matrix, i.e. $\boldsymbol{T}^{(k)} = \left(\boldsymbol{R}^{(k)}\right)^{-1}, \quad k = 1, \dots, N_k N_l$. Therefore, ${\boldsymbol{\Lambda_Q}}^{-1}$ is computed as

$$\left(\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{\boldsymbol{inv}}\right)^{(i,j)}(k,k) = \boldsymbol{T}^{(k)}(i,j), \quad k = 1, 2, \dots, N_k N_l \text{ and } i, j = 1, 2, \dots, N_\lambda. \tag{4.37}$$

Finally, the solution in (4.23) is computed by performing inverse multichannel DFT as:

$$\begin{aligned} \widehat{\boldsymbol{x}} &= \left(\boldsymbol{H}^T\boldsymbol{H} + \mu_{\text{spat}}\boldsymbol{D}_{\text{spat}}^T\boldsymbol{D}_{\text{spat}} + \mu_{\text{spec}}\boldsymbol{D}_{\text{spec}}^T\boldsymbol{D}_{\text{spec}}\right)^{-1}\boldsymbol{H}^T\boldsymbol{y} \\ &= \boldsymbol{Q}^{-1}\boldsymbol{H}^T\boldsymbol{y} \\ &= \left(\overline{\boldsymbol{F}}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{\boldsymbol{inv}}\overline{\boldsymbol{F}}\right)\left(\overline{\boldsymbol{F}}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{H}}^\dagger\overline{\boldsymbol{F}}\right)\boldsymbol{y} \\ &= \overline{\boldsymbol{F}}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{\boldsymbol{inv}}\boldsymbol{\Lambda}_{\boldsymbol{H}}^\dagger\mathring{\boldsymbol{y}} \end{aligned} \tag{4.38}$$

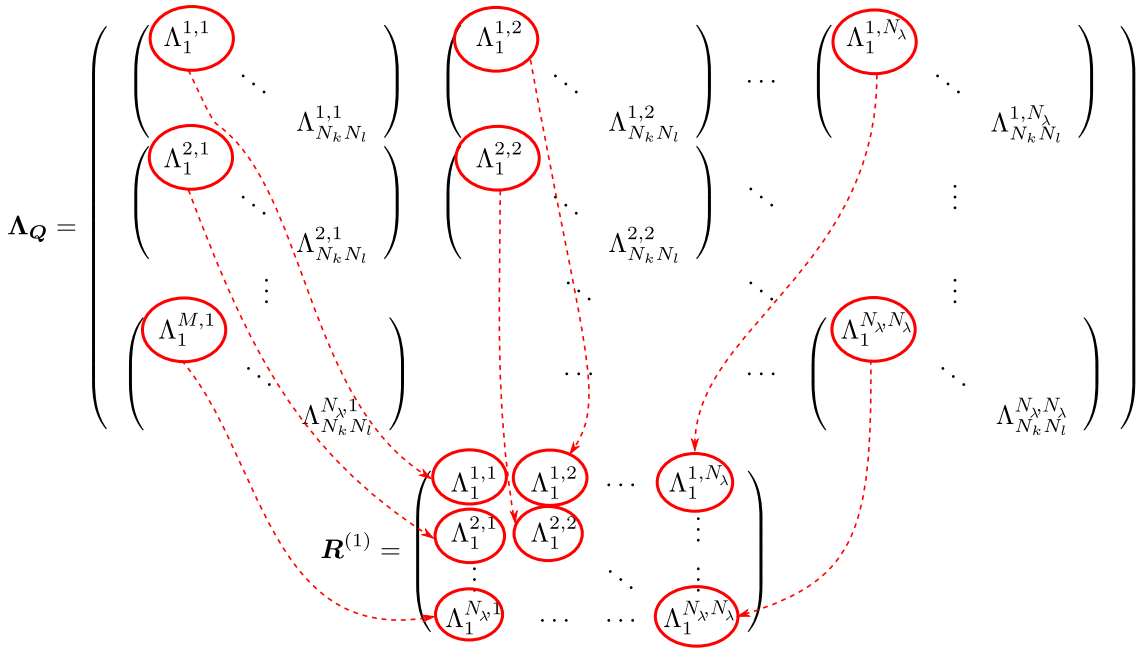


Figure 4.3: Illustration of the extraction procedure of $\boldsymbol{R}^{(1)}$ for the inversion of a non-diagonal bloc-diagonal matrix $\boldsymbol{\Lambda_Q}$.

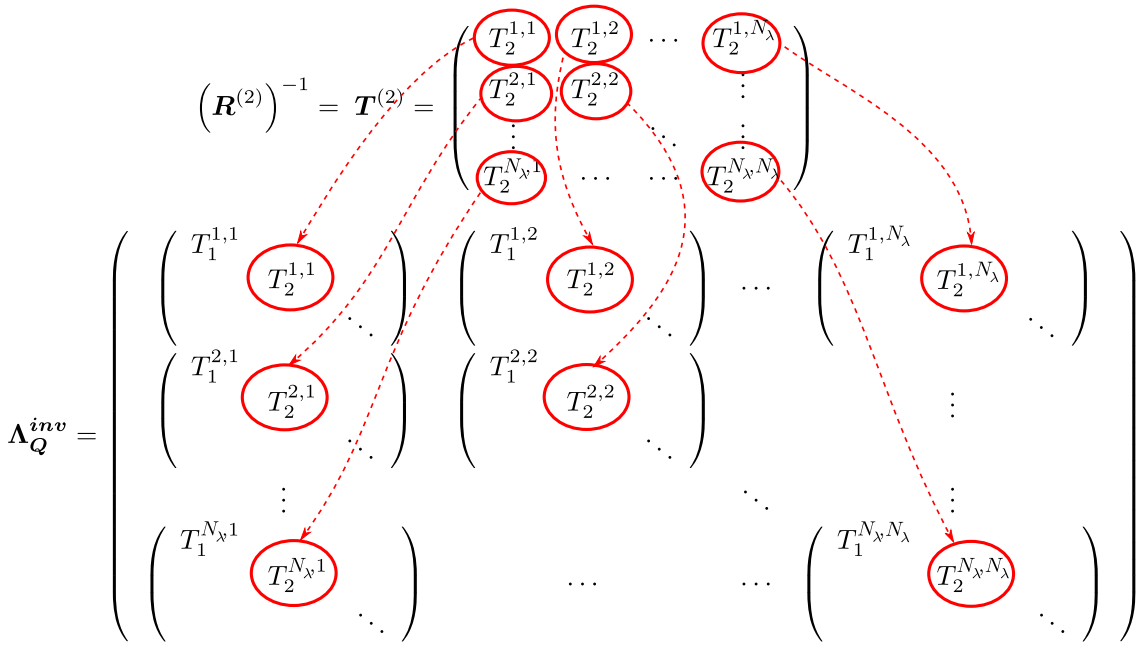


Figure 4.4: Illustration of the inversion of the matrix $\boldsymbol{\Lambda_Q}$ with an example for $k = 2$.

with $\mathring{\boldsymbol{y}} = \overline{\boldsymbol{F}}\boldsymbol{y}$ is the multichannel DFT of the multispectral data set. The presented method is called Multichannel Discrete Fourier transform (MDFT) and is summarized in

a pseudo-algorithm form in Algorithm 4.3.

The solution in (4.38) has several key advantages. It is explicit and does not need to be updated thanks to the diagonalization operation in the Fourier domain. The runtime to compute the solution directly is faster and more efficient compared to the iterative one (using CG algorithm) due to the parallel computation in Fourier. In addition, since all block matrices are reals valued, their Fourier transform are Hermitian, which allows us to reduce the cost and speed up the computation. Finally, the solution is linear and easy to implement directly on a calculating machine using simple linear algebra. Thus it is much efficient than the iterative solution such as the CG optimization.

Algorithm 4.3 Multichannel Discrete Fourier Transform (MDFT)

Input: $\boldsymbol{H}, \boldsymbol{D}, \boldsymbol{C}, \boldsymbol{y}, \mu_{\text{spat}}, \mu_{\text{spec}}$

Compute the Hessian matrix:
$\boldsymbol{D}_{\text{spat}} = \text{diag}\{\boldsymbol{D}, \boldsymbol{D}, \ldots, \boldsymbol{D}\}$ ▷ Equation (4.19)
$\boldsymbol{D}_{\text{spec}} = \boldsymbol{C} \otimes \boldsymbol{I}_{N_k N_l}$ ▷ Equations (4.28)-(4.29)
$\boldsymbol{Q} \leftarrow \boldsymbol{H}^T\boldsymbol{H} + \mu_{\text{spat}}\boldsymbol{D}_{\text{spat}}^T\boldsymbol{D}_{\text{spat}} + \mu_{\text{spec}}\boldsymbol{D}_{\text{spec}}^T\boldsymbol{D}_{\text{spec}}$ ▷ Equation (4.34)

Diagonalize $\boldsymbol{Q}$ (Non-Circulant Block Circulant) :
$\boldsymbol{\Lambda_H} \leftarrow \left\{\boldsymbol{\Lambda}_{\boldsymbol{H}}^{i,j} = \boldsymbol{F}\boldsymbol{H}^{i,j}\boldsymbol{F}^\dagger\right\}_{i,j=1}^{P,N_\lambda}$
$\boldsymbol{\Lambda_Q} \leftarrow \left\{\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{i,j} = \boldsymbol{F}\boldsymbol{Q}^{i,j}\boldsymbol{F}^\dagger\right\}_{i,j=1}^{N_\lambda,N_\lambda}$ ▷ Equation (4.35)
$N_\lambda, N_\lambda, \overline{N}_k, \overline{N}_l = \text{size}(\boldsymbol{\Lambda_Q})$

Invert $\boldsymbol{\Lambda_Q}$ (Non-Diagonal Block Diagonal) :
for $k = 0 : \overline{N}_k$ **do**
  for $l = 0 : \overline{N}_l$ **do**
    $\boldsymbol{R} = \boldsymbol{\Lambda}_{\boldsymbol{Q}}^{:,:}[k,l]$ ▷ Equation (4.36)
    $\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{\text{inv}\,:,:}[k,l] = \boldsymbol{R}^{-1}$ ▷ Equation (4.37)

Compute the solution:
$\widehat{\boldsymbol{x}} \leftarrow \overline{\boldsymbol{F}}^\dagger\,\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{\text{inv}}\boldsymbol{\Lambda}_{\boldsymbol{H}}^\dagger\overline{\boldsymbol{F}}\,\boldsymbol{y}$ ▷ Equation (4.38)
return $\widehat{\boldsymbol{x}}$

4.6 Simulation Results

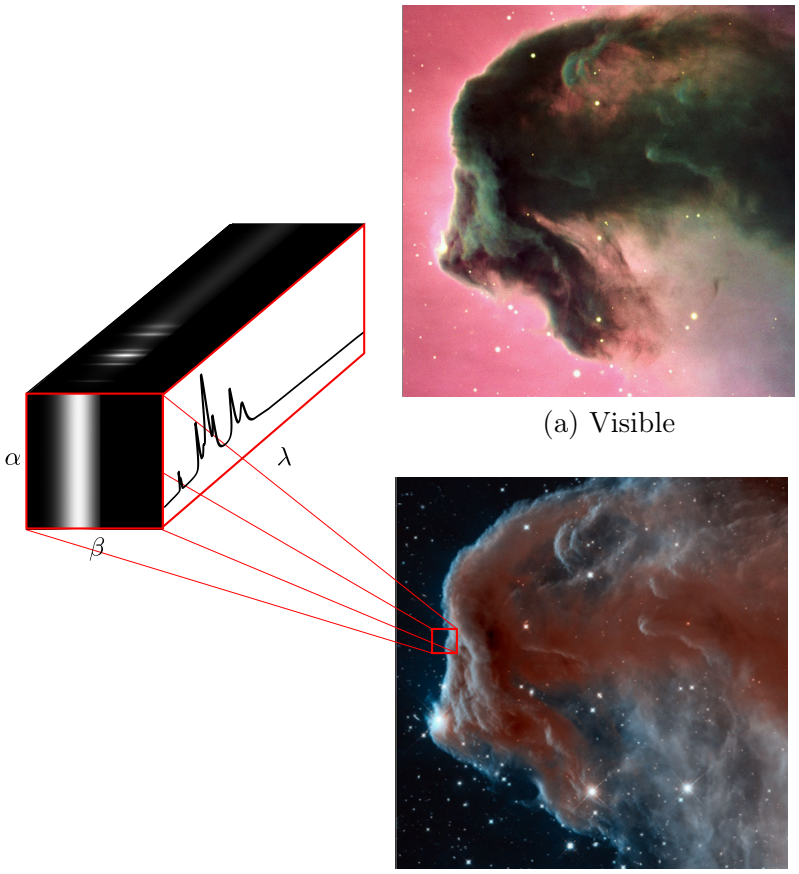
4.6.1 Description of Spatio-Spectral Object: *HorseHead nebula*

In this section we present an astrophysical object to test our reconstruction method. The object is a simplified spatio-spectral model of the *HorseHead nebula* [Abergel *et al.* 2003] modeling a cloud of matter (dust and gas) illuminated by a star. The datacube has been computed at IAS[1] using a state-of-the-art model of interstellar dust particles. It computes

[1] https://www.ias.u-psud.fr/

the emission spectrum in the infrared at different positions, for a given dust population and a given geometry of the object and of the illumination conditions.

Figure 4.5:(a)-(b) illustrates a comparison of the *HorseHead nebula* observed in the visible light using the Very Large Telescope (VLT)[2] and in the near-infrared spectrum using the space telescope Hubble[3]. It is clear that the near-infrared observation uncovers a lot of information about the object, such as the sharpness of structure and embedded or background stars.



(a) Visible

(b) Near-Infrared

Figure 4.5: Illustration of the spatial distribution of the *HorseHead nebula* in (a) The visible. (b) The near-infrared. The illuminating star is on the right (but outside the images). (c) Represents a small region from the sky taken for simulation [Abergel *et al.* 2003].

The object datacube is distributed on a spatial grid with a pixel scale of 0.5"/pixel. The intensity of each element of the cube has a unit of MJy/sr−Mega Jansky.Steradian^{-1} (see Section 3.4.1 for more details). In addition, the spectrum of the object is non-uniformly sampled. In practice we need to re-sample first the spatial and the spectral distributions of the object so they meet the JWST/MIRI imager requirements, i.e. wavelength range from 1 to 28μm and a pixel scale of 0.11"/pixel. Therefore, a spatial region of $N_k = 256 \times N_l =$

[2] http://www.eso.org/public/teles-instr/paranal-observatory/vlt/
[3] https://www.nasa.gov/mission_pages/hubble/story/index.html

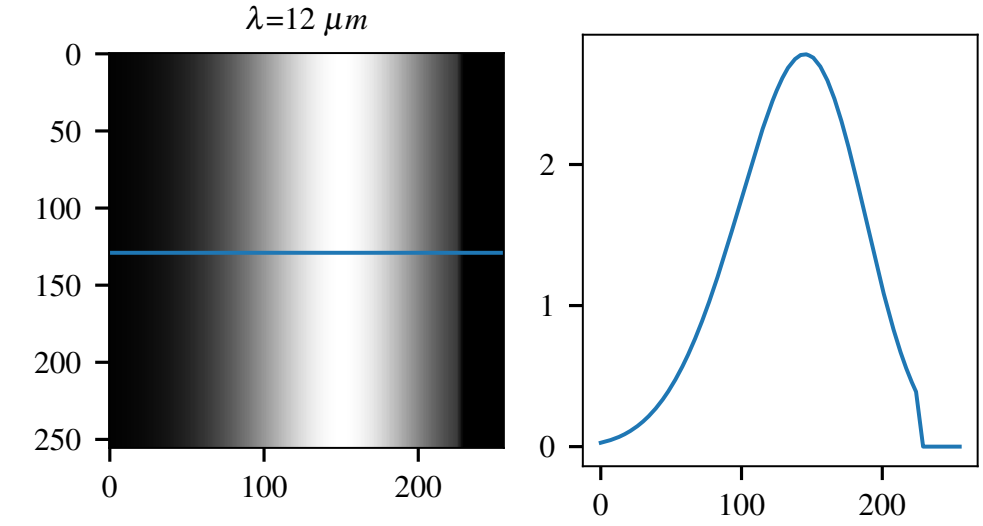
256 pixels is taken for simulations with $N_\lambda = 1000$ spectral samples uniformly distributed. Figure 4.6 shows the spatial and spectral distributions of the re-sampled object. Note that the horizontal spatial distribution of the object is smooth and the vertical spatial distribution is constant.

4.6.2 Setup of the Experiment

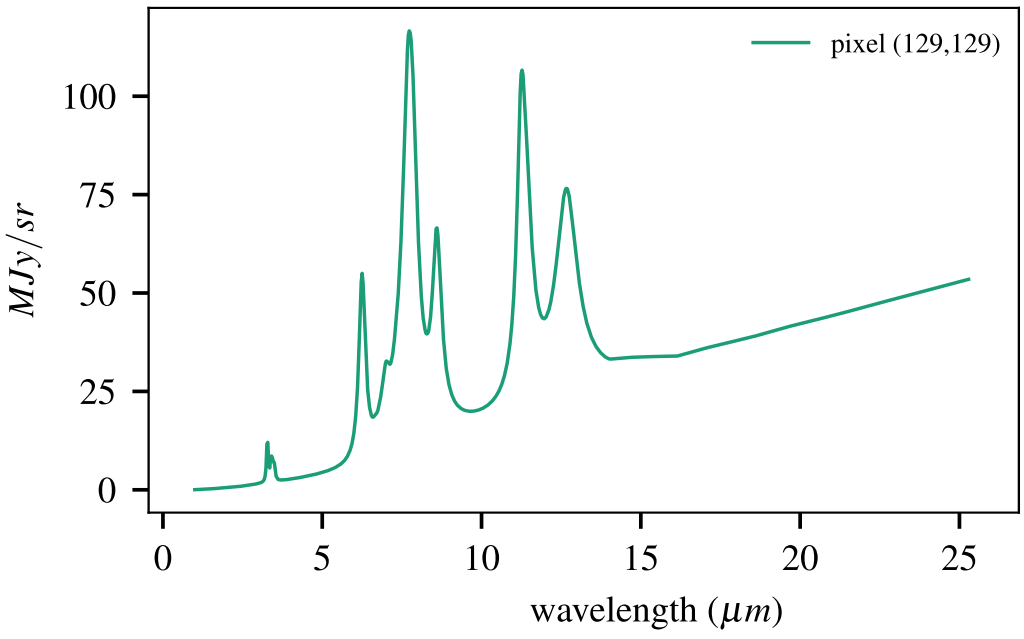
Nine multispectral data ($p = 9$) are simulated using Equation (3.8), with a zero-mean white Gaussian noise added in order to obtain a global Signal-to-Noise Ratio (SNR) of $30, 20, 10$ dB, defined by

$$SNR_{(dB)} = 10 \log_{10} \left(\frac{\frac{1}{PN_iN_j} \|\boldsymbol{y}\|_2^2}{\sigma_n^2} \right),$$

where σ_n is the standard deviation of the noise, P the number of multispectral data and N_iN_j the total number of pixels in the multispectral data.



(a)



(b)

Figure 4.6: Display of (a) the spatial distribution at 12μm and (b) the spectral distribution at pixel position $(129, 129)$ of the *HorseHead nebula*.

4.6.2.1 Influence of the Regularization Parameters: μ_{spat} and μ_{spec}

In order to tune the regularization parameters μ_{spat} and μ_{spec} we adopt a supervised strategy, running the code for different values in a range and keep the pair of parameters

that minimizes the convex cost function

$$\widehat{\mu}_{\text{spat}}, \widehat{\mu}_{\text{spec}} = \underset{\mu_{\text{spat}}, \mu_{\text{spec}}}{\operatorname{argmin}} \mathcal{J}\left(\widehat{\boldsymbol{x}}; \mu_{\text{spat}}, \mu_{\text{spec}}\right). \tag{4.39}$$

We are conscious that this approach is time-consuming for a wide range of parameters, but this is not an issue since the solution is computed in few seconds thanks to the implementation of the algorithm in the Fourier domain.

For a quantitative comparison between the original spectral channel $\boldsymbol{f}_{orig}$ and the reconstructed $\boldsymbol{f}_{rec}$ we compute the relative reconstruction error defined by

$$Error(\%) = 100 \times \frac{\|\boldsymbol{f}_{orig} - \boldsymbol{f}_{rec}\|_2}{\|\boldsymbol{f}_{orig}\|_2}. \tag{4.40}$$

It is also called NMSE (Normalized Mean Square Error).

We display the obtained results in Figure 4.7 after running the reconstruction algorithm multiple times for different values of regularization parameters μ_{spat} and μ_{spec}. We can clearly observe that the performance of the reconstruction algorithm on both regularization parameters and that there is an optimal pair that satisfies a minimum reconstruction error (mentioned by a red dot). The reconstruction results are obtained by using a 30 dB multispectral dataset and a number of spectral channels of 60.

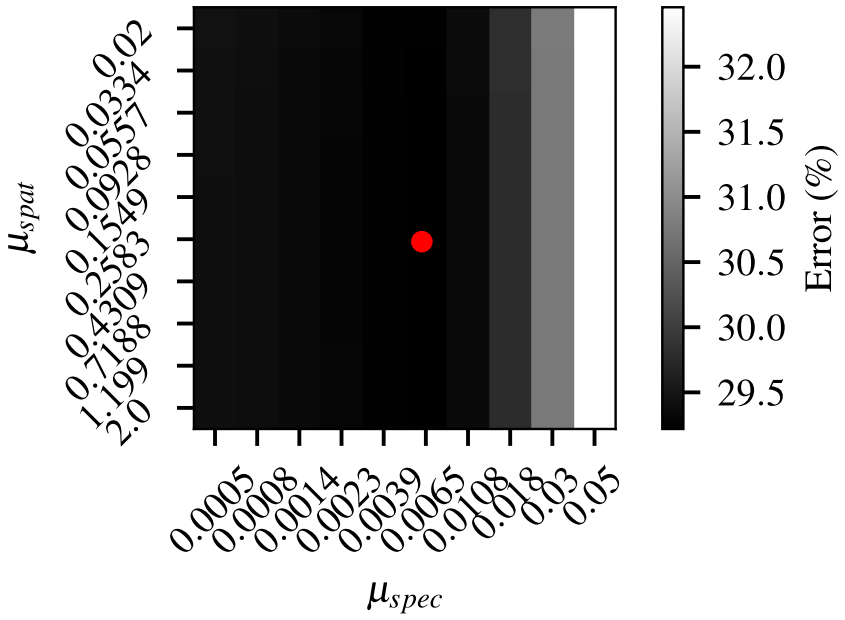


Figure 4.7: Influence of the regularization parameters on the reconstruction Algorithm 4.3 for the *HorseHead nebula* with SNR = 30 dB, $M = 60$ and $P = 9$. The red dot indicates the pair of parameters corresponding to the minimum of reconstruction error of the whole cube.

4.6.2.2 Simulated Multispectral Data

Figures 4.8 shows the simulated multispectral data of nine broad bands of the JWST/MIRI imager at 30 dB, indexed from $p = 1, \ldots, 9$ and associated to the objects *HorseHead nebula*. Figures 4.9 illustrates the spatial distribution of the multispectral data and the interaction between pixels, by displaying a cut from the central row of images of Figures 4.8. All images of the figure are represented with the same color-bar for better illustration and comparison.

The simulated multispectral data illustrate the complexity of the imaging system. In fact, we notice a presence of a blur that increases for larger wavelength, e.g. the 8-*th* and 9-*th* bands compared to the shorter wavelength, e.g. the 1-*st* and 2-*nd* bands. This is due to the spectral variability of the PSF. Moreover, we notice a difference in the intensities between images, where some images appear with high intensities and other with low intensities. This is due to the amount of information in the spectral components to be integrated within bands and the width of these bands. For instance the 1-*st* band is narrower then the 9-*th* band (see Figure 3.10). Furthermore, we notice domination of the noise in the multispectral data, especially those integrated from narrow bands, $p = 1, 3, 4$.

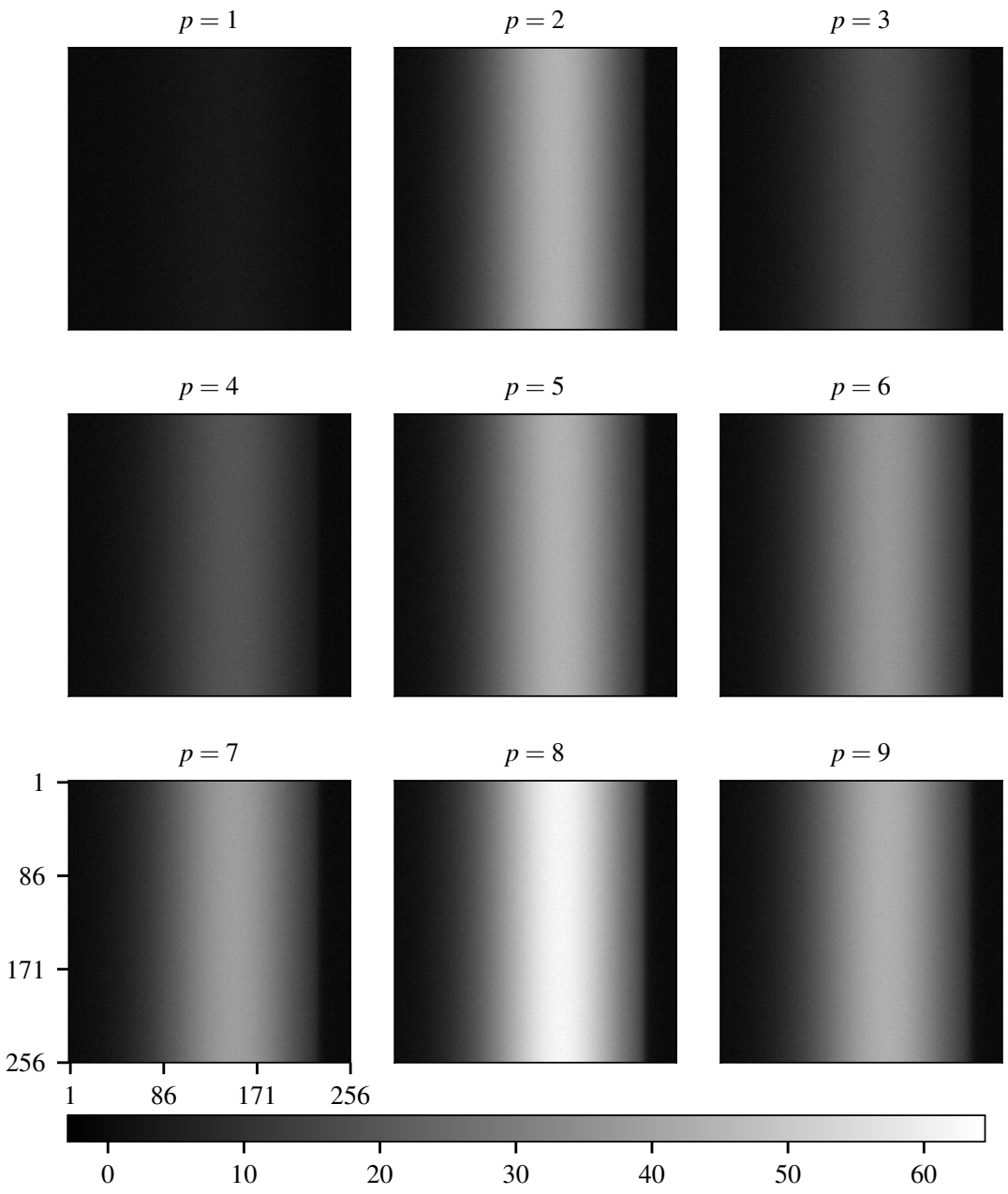


Figure 4.8: Simulation of nine multispectral data using Equation (3.8) for the *HorseHead nebula* object. All the multispectral data are corrupted with an additive white Gaussian noise so that SNR=30 dB.

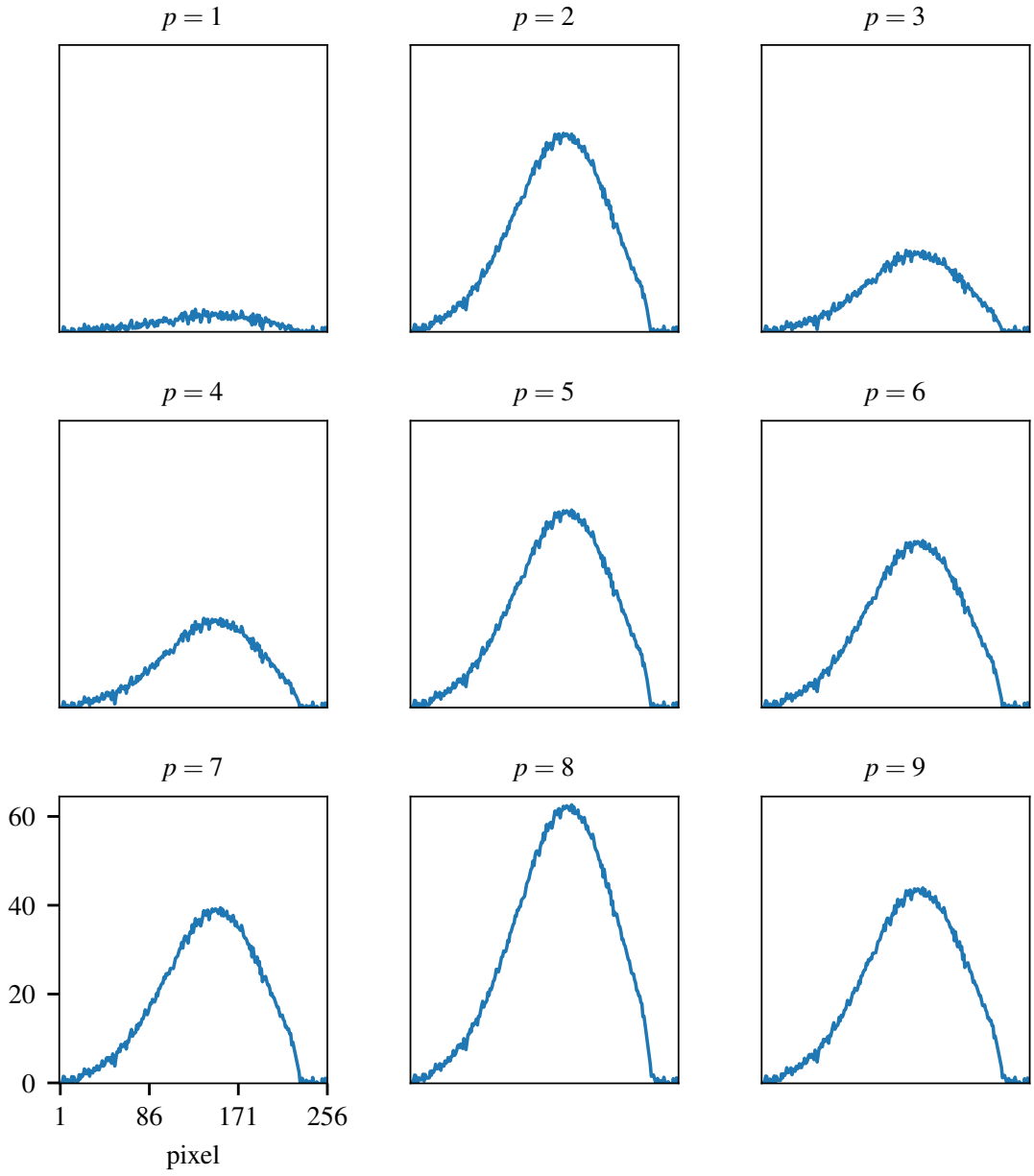


Figure 4.9: Illustration of the spatial distribution of multispectral data by displaying a slice from the central row of images in Figure 4.8.

4.6.3 Results and Discussion

The reconstruction results are summarized in Table 4.1 together with a comparison between the proposed method and the broadband reconstruction or multichannel 2D deconvolution method [Galatsanos & Chin 1989]. In Figure 4.10 we show the result of one single pixel $(127, 100)$, comparing the original spectrum ϕ_{orig}, the reconstructed spectrum using our method ϕ_{MDFT}, and the reconstructed spectrum using multichannel 2D deconvolution $\phi_{Broadband}$. The original spectrum is complex with spectral features at short wavelengths

$(4-13\,\mu\text{m})$ and continues. Therefore an accurate reconstruction using a few multispectral data is difficult, if not impossible, without using a strong prior knowledge of the spectrum of the object. In fact, the reconstructed spectrum computed with our method (using piecewise linear model) allows us to reconstruct an envelope like spectral distribution which significantly increases the spectral resolution compared to multichannel 2D deconvolution. Several values of $N_\lambda = \{20, 40, 60\}$ have been tested, and the reconstruction results for three wavelengths, 7.8, 16 and 21 μm, are reported in Table 5.1. Increasing N_λ improves the spectral resolution of the object model but increases the between-channel degradation and the number of unknowns. Moreover, we find that there is not much error improvement for $N_\lambda > 60$. In any case, our proposed reconstruction shows smaller reconstruction errors compared to the multichannel 2D deconvolution; this is due to our model accounting for within- and between-channels degradations.

The reconstruction results using the proposed method at different wavelengths are illustrated in Figure 4.11. As anticipated, a better reconstruction is obtained at $\lambda = 16\,\mu$m and $\lambda = 21\,\mu$m than at $\lambda = 7.8\,\mu$m (see the fourth row of the figure) since within the integration windows at long wavelengths the spectrum of the object does not contain any feature.

Table 4.1: Reconstruction results of the *HorseHead nebula* of size $1000 \times 256 \times 256$ using a multichannel quadratic regularization method.

SNR(dB)	$\lambda(\mu\text{m})$	Error (%)			
		MDFT			Broadband
		$N_\lambda = 20$	$N_\lambda = 40$	$N_\lambda = 60$	
30	7,8	49,44	42,37	41,42	52,85
	16,0	2,44	4,11	4,80	7,89
	21,0	1,87	3,82	4,26	11,92
20	7,8	49,50	43,07	41,46	52,84
	16,0	7,41	7,66	8,98	8,02
	21,0	4,42	5,40	5,77	11,97
10	7,8	50,71	43,71	42,38	52,84
	16,0	19,67	21,25	25,51	8,56
	21,0	10,85	11,31	13,38	12,13

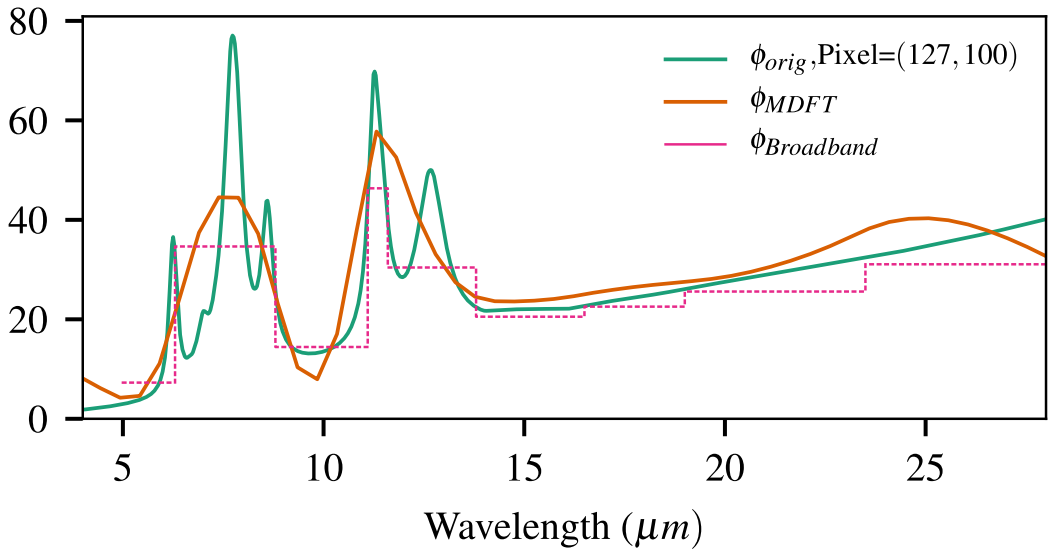


Figure 4.10: Comparison between one single pixel spectrum from the original object ϕ_{orig}, the proposed reconstruction ϕ_{MDFT} (with $N_\lambda = 60$) and the multichannel 2D deconvolution $\phi_{Broadband}$. The nine multispectral data ($p = 9$) were corrupted with zero-mean Gaussian noise of 30 dB.

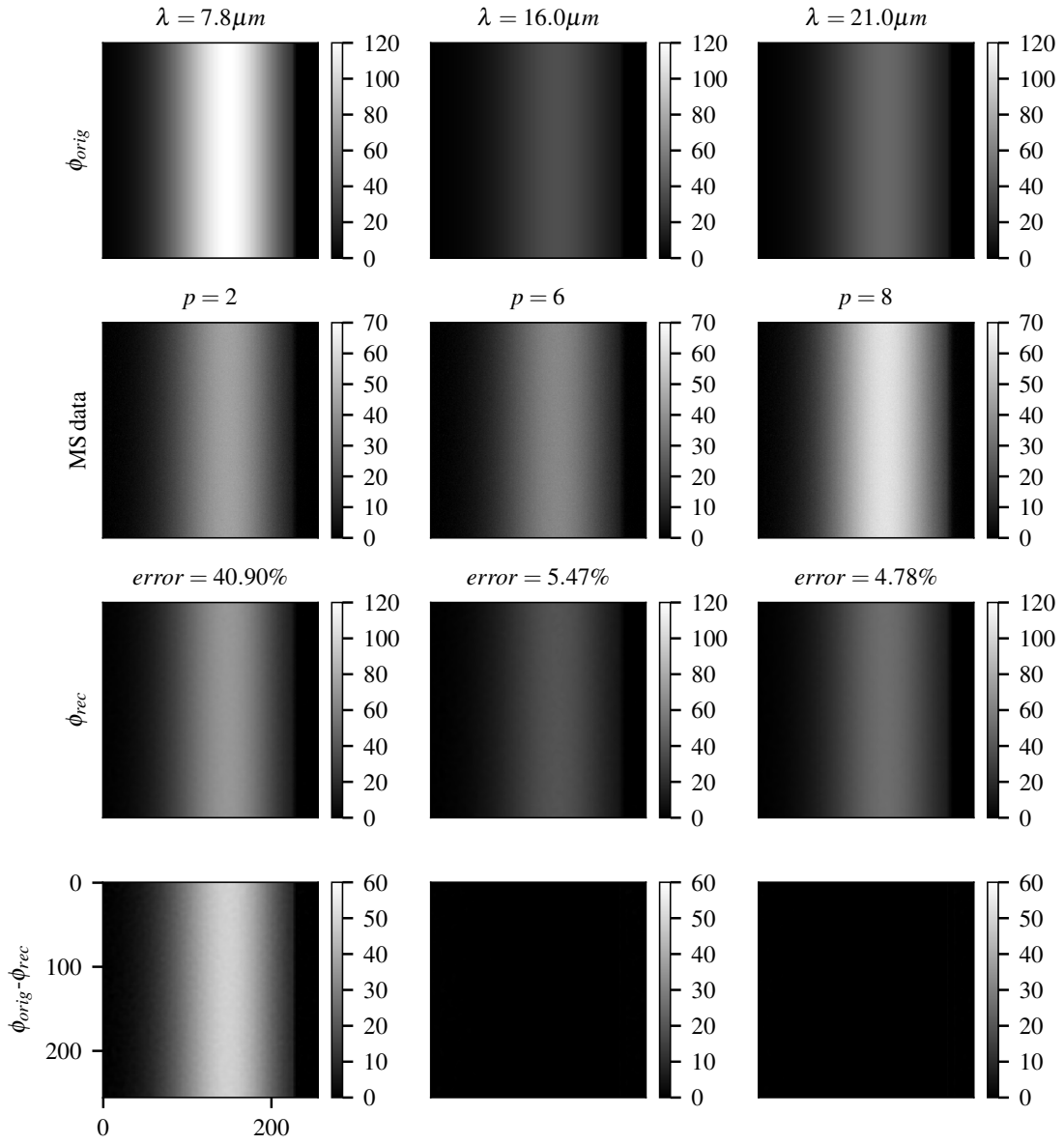


Figure 4.11: [1st row] Original channel of the *HorseHead nebula* at 7.8, 16 and 21 μm. [2nd row] Simulated multispectral data with 30 dB corresponding to the bands that includes wavelengths of the first row. [3rd row] Proposed reconstruction of the channel at 7.8, 16 and 21 μm. [4th row] Difference between the original and reconstructed spectral channels.

On the other hand, in figure 4.12 we display the influence of the number of channels N_λ on the runtime of the Algorithm 4.3. We clearly see that, the higher the value of N_λ the longer it takes to compute the solution and this is due to increase of the size of the Hessian matrix, hence, more block matrices are to inverse in the Fourier domain.

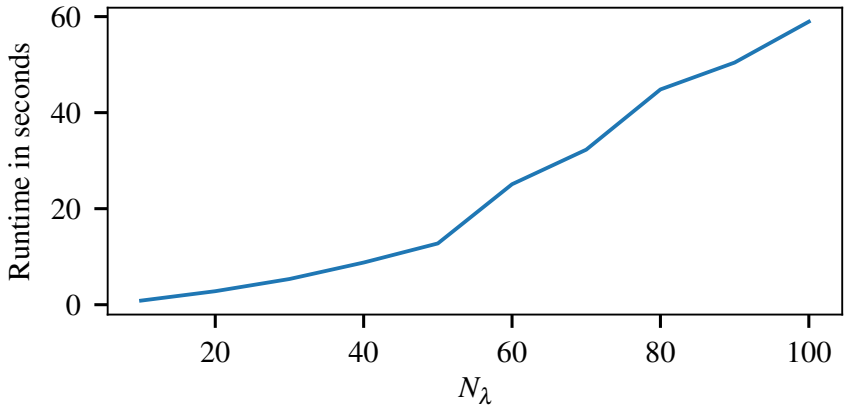


Figure 4.12: Display of the influence of number of channels N_λ on the runtime of the Algorithm 4.3.

4.7 Conclusion

This chapter presents a reconstruction method of a 2D+λ multichannel spatio-spectral object observed by a multispectral imaging system from a few low-resolution data. We first modeled the spectral distribution of the object with a piecewise linear function of one parameter in order to control the spectral sampling. Then we developed first a linear forward model where each multispectral data is a sum of discrete 2D convolution between the spectral channel, and the observation matrix that models the spectral degradation occurring between channels. Moreover, the reconstruction method is developed using regularization methods where we proposed to correct the ill-conditioning of the Hessian matrix by enforcing spatial and spectral smoothness to the solution. The computation of the multichannel quadratic solution consists mainly of inverting the Hessian matrix $\boldsymbol{Q}$. However the matrix $\boldsymbol{Q}$ is very large and cannot be inverted in the spatio-spectral domain. Hence we proposed to compute the solution with two approaches: (1) iteratively by using an optimization algorithm such as the conjugate gradient algorithm, and (2) directly by performing inverting the Hessian matrix in the Fourier domain, which is explicit and 10× faster than the iterative solution thanks to the diagonalization of circular matrices.

Finally, the obtained results on simulated data of the JWST/MIRI Imager highlights the complexity of the instrument response. In addition, an important increase of spatial and spectral resolution of the reconstructed spatio-spectral object is observed compared to multichannel 2D deconvolution method thanks to the modeling of the spectral distribution of the object and considering of the spectral variation of the PSF instead of using stationary broadband PSF per band.

However, the proposed reconstruction method presents some limitations because of the object model. Firstly, it is difficult to choose the number of channels to reconstruct, N_λ, and the higher it is the longer it takes for the algorithm to compute the solution. Secondly, the choice of a piecewise linear function is a good choice to model object with

linear spectral distribution within the integration bands, but integration bands in case of MIRI imager of the JWST are very large and the spectral distribution of the original object might vary a lot implying that the object model is not accurate anymore.

In the next chapter, we see how to overcome this limitation and improve the reconstruction results by using an accurate model for the object.

Chapter 5

Spatio-Spectral Reconstruction through Mixture Coefficient Estimation from Low-Resolution Multispectral Data

Contents

5.1 Introduction

The aim of this chapter is the reconstruction of an object with a high spatio-spectral resolution from a set of low-resolution multispectral data degraded by a multispectral imaging system, such as the JWST/MIRI imager described in Chapter 3. Rather than representing the spectral distribution of the object by a piecewise linear function, as in Chapter 4, we choose to use a linear mixing model (or dictionary representation), where each spectral distribution is represented by a linear combination of high-resolution spectral components. Therefore the reconstruction problem of a spatio-spectral object turns to an estimation of a set of mixture coefficients which are the weights of the spectral components.

As for the spatial distribution, two types of object are addressed which contains smooth or high gradients, respectively. Therefore two reconstruction methods are developed. A

comparison of simulation results is presented for a synthetic and a real-like astrophysical (*HorseHead nebula*) spatio-spectral object.

For the sake of consistency, the paradigm of this chapter is similar to the one used in Chapter 4, i.e. object model, linear forward model, reconstruction, simulation results.

5.2 Problem Statement

We have seen in Chapter 4 that the reconstruction of the spectral distribution of a spatio-spectral object from a small number of multispectral data, observed by an imager over broad bands, is not very accurate without enforcing a strong prior on the spectral distribution even though we obtained better reconstruction results in [Hadj-Youcef *et al.* 2017a] compared to the common reconstruction method [Galatsanos & Chin 1989].

In this chapter we use another linear model for the spatio-spectral object in order to improve the spatial and spectral resolution of the reconstructed object. We are interested in objects with spectral correlations where each spectral distribution is approximated by a linear combination of uncorrelated spectral components (also known as signatures or end-members) for all spatial coordinates. This model is called the linear mixing model (LMM). It was first proposed in [Adams *et al.* 1986] to analyze a multispectral image, where the spectral end-members in the image corresponds to materials such as soil, rock, and shade. In [Keshava & Mustard 2002] a spectral unmixing method is presented, i.e. decomposition of a mixed spatial position (or pixel) into a collection of distinct spectra, and a set of mixture coefficients. These coefficients represent the weight of each uncorrelated spectral components. The linear mixing model has been used on a wide range of hyperspectral applications, see [Bioucas-Dias *et al.* 2012] where an overview of hyperspectral unmixing methods is given. Moreover, several works consider the spatial and spectral correlation between neighbors pixels in order to analyze hyperspectral data, [Tarabalka *et al.* 2009] proposed a spectral–spatial classification scheme for hyperspectral images. The authors of [Guo *et al.* 2009] present an hyperspectral image enhancement method based on the total variation regularizer to produce a higher visual quality hyperspectral image.

The spectral components could be found in archives such as for stellar spectra libraries [Jacoby *et al.* 1984, Pickles 1998]. The spectral components can also be extracted using source separation techniques; for instance, Independent Component Analysis (ICA) [Hyvärinen *et al.* 2004] Principal Component Analysis (PCA) [Jolliffe 1986], or blind source separation techniques [Cichocki & Amari 2003, Comon & Jutten 2010]. Moreover, authors in [Dobigeon *et al.* 2009] proposed a Bayesian method for the extraction of end-members (or spectral components) and the estimation of abundances (or mixture coefficients) from hyperspectral images. Another approach consists of taking advantages of hyperspectral instruments such as a spectrometer, e.g. the spectrometer of JWST/MIRI [Wells *et al.* 2015].

We propose to represent the object by a small number of M uncorrelated spectral components (typically M is less than the number of spectral bands P). This allows us to make use of the linear mixing model to represent in an approximate manner the spatio-spectral object, where the M uncorrelated spectral components are weighted by a set of M mixture coefficients (or weights) $c^1, \ldots, c^M$ at each spatial position (α_k, β_l), as illustrated in Figure 5.1.

The particularity of the proposed work is that the reconstruction of the spatio-spectral object is done through the estimation of the mixture coefficients using a set of low-resolution multispectral data degraded by a spatial blur and noise, in addition to lack of spectral distribution after integrations over broad bands by the detector.

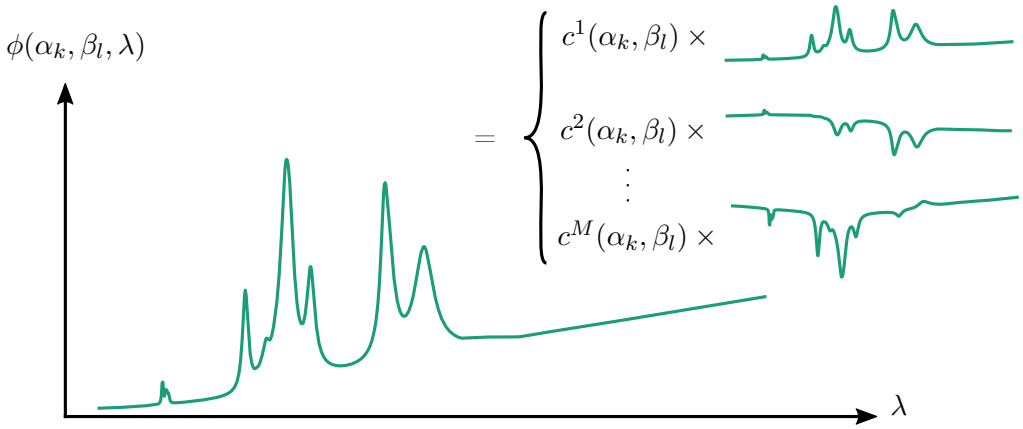


Figure 5.1: Illustration of the linear mixing model on a single spatial position (α_k, β_l) of the spatio-spectral object.

5.3 Object Model: Linear Mixing Model

In this section we are interested in representing the continuous spatio-spectral object by a set of discrete coefficients using the linear mixing model. The object is represented by a sum of M high-resolution spectral components assumed to be known, $s^m(\lambda), m = 1, \ldots, M$, weighted by mixture coefficients $c^m(\alpha, \beta)$ associated to each spatial position (α, β), as illustrated in Figure 5.2. This yields

$$\phi(\alpha, \beta, \lambda) = \sum_{m=1}^{M} c^m(\alpha, \beta) s^m(\lambda) \, . \tag{5.1}$$

Therefore in order to reconstruct the object, a set of M 2D mixture coefficients (or weights maps) for each spectral component must be estimated. A discrete representation of the mixture coefficients upon a shift basis function is given by the reconstruction formula [Yaroslavsky 2013]

$$c^m(\alpha, \beta) = \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} x^m_{k,l} \, b_{\text{rec}}(\alpha - \alpha_k, \beta - \beta_l), \tag{5.2}$$

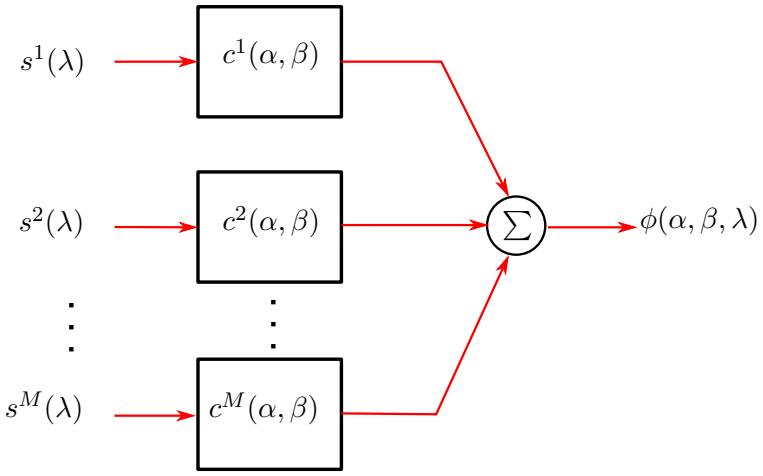


Figure 5.2: Representation of a spatio-spectral object using a linear mixing model.

where $x^m_{k,l}$ are the discrete coefficients of the (k,l)-*th* spatial position associated to the m-*th* spectral component. Moreover, we chose the shift basis function $b_{\text{rec}}(.)$ to be a rectangular impulse function defined on the grid $\mathcal{G}_{\text{rec}} = \{\alpha_k, \beta_l\}_{k,l=1}^{N_k,N_l}$. N_k, N_l refers to the total number of pixel according to dimensions α and β, respectively as in Chapter 4. $\alpha_k = k\triangle\alpha'$ and $\beta_l = l\triangle\beta'$, where $\triangle\alpha'$ and $\triangle\beta'$ are the sampling steps according to dimensions α and β, respectively. They are set to be equal to the sampling steps of the multispectral data $\triangle\alpha$ and $\triangle\beta$. A finer sampling steps could be defined in case of a super-resolution problem by setting the $\triangle\alpha'$ and $\triangle\beta'$ to be a fraction of $\triangle\alpha$ and $\triangle\beta$.

Finally, the model for the object in Equation (5.1) becomes

$$\phi(\alpha,\beta,\lambda) = \sum_{m=1}^{M} \left(\sum_{k=1}^{N_k} \sum_{l=1}^{N_l} x^m_{k,l}\, b_{\text{rec}}(\alpha-\alpha_k, \beta-\beta_l) \right) s^m(\lambda) \tag{5.3}$$

This model represents a continuous spatio-spectral object $\phi(\alpha,\beta,\lambda)$ by a sum of M spectral components weighted by discrete mixture coefficients (unknown parameter of the object) $x^k_{k,l}$. In the next section, we focus on the reconstruction of these coefficients using an inverse problem framework.

5.4 Linear Forward Model: Observation Matrix

The discrete forward model expressing the relationship between the discrete multispectral data and the discrete mixture coefficients is obtained by substituting Equation (5.3) of the object model in the instrument model (3.8). This yields

$$y^{(p)}_{i,j} = \iint_{\Omega_{\text{pix}}} \left(\int_{\mathbb{R}_+} \omega_p(\lambda) \left(\left(\sum_{m=1}^{M} \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} x^m_{k,l}\, b_{\text{rec}}(\alpha-\alpha_k, \beta-\beta_l) s^m(\lambda) \right) \underset{\alpha,\beta}{*} h(\alpha,\beta,\lambda) \right) d\lambda \right) b^{i,j}_{\text{samp}}(\alpha,\beta) d\alpha d\beta + n^{(p)}_{i,j}. \tag{5.4}$$

By rearranging terms of Equation (5.4), we obtain

$$y_{i,j}^{(p)} = \sum_{m=1}^{M} \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} H_{i,j;k,l}^{p,m} \, x_{k,l}^{m} + n_{i,j}^{(p)}, \tag{5.5}$$

with

$$H_{i,j;k,l}^{p,m} = \iint_{\Omega_{\text{pix}}} \left(\left(\int_{\mathbb{R}_+} \omega_p(\lambda) h(\alpha, \beta, \lambda) s^m(\lambda) d\lambda \right) \underset{\alpha,\beta}{*} b_{\text{rec}}(\alpha - \alpha_k, \beta - \beta_l) \right) b_{\text{samp}}(\alpha - \alpha_i, \beta - \beta_j) d\alpha d\beta \tag{5.6}$$

where the observation matrix $H_{i,j;k,l}^{p,m}$ only depends on the instrument and the object parameters. Hence, the accurate the knowledge of the instrument model and the modeling of the object, the more accurate is the observation matrix.

Moreover, by considering the same assumptions as in Section 4.4, i.e. setting $b_{\text{samp}}(.)$ to be an impulse rectangular function, the observation matrix becomes a convolution matrix $H_{i,j;k,l}^{p,m} = H_{i-k;j-l}^{p,m}$. Therefore, we end-up with the matrix-vector representation of the forward model

$$\boldsymbol{y}^{(p)} = \sum_{m=1}^{M} \boldsymbol{H}^{p,m} \boldsymbol{x}^m + \boldsymbol{n}^{(p)}, \quad p = 1, 2, \ldots, P. \tag{5.7}$$

where the p-th multispectral image $\boldsymbol{y}^{(p)}$ is a sum of M discrete 2D convolutions of discrete mixture coefficients $\boldsymbol{x}^m$ with circular convolution matrices $\boldsymbol{H}^{p,m}$.

In addition, by combining all multispectral dataset we obtain the following multi-observation forward model

$$\underbrace{\begin{pmatrix} \boldsymbol{y}^{(1)} \\ \boldsymbol{y}^{(2)} \\ \boldsymbol{y}^{(3)} \\ \vdots \\ \boldsymbol{y}^{(P)} \end{pmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{pmatrix} \boldsymbol{H}^{1,1} & \boldsymbol{H}^{1,2} & \cdots & \boldsymbol{H}^{1,M} \\ \boldsymbol{H}^{2,1} & \boldsymbol{H}^{2,2} & \cdots & \boldsymbol{H}^{2,M} \\ \boldsymbol{H}^{3,1} & \boldsymbol{H}^{3,2} & \cdots & \boldsymbol{H}^{3,M} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}^{P,1} & \boldsymbol{H}^{P,2} & \cdots & \boldsymbol{H}^{P,M} \end{pmatrix}}_{\boldsymbol{H}} \underbrace{\begin{pmatrix} \boldsymbol{x}^1 \\ \boldsymbol{x}^2 \\ \vdots \\ \boldsymbol{x}^M \end{pmatrix}}_{\boldsymbol{x}} + \underbrace{\begin{pmatrix} \boldsymbol{n}^{(1)} \\ \boldsymbol{n}^{(2)} \\ \boldsymbol{n}^{(3)} \\ \vdots \\ \boldsymbol{n}^{(P)} \end{pmatrix}}_{\boldsymbol{n}}, \tag{5.8}$$

where $\boldsymbol{x} \in \mathbb{R}^{MN_kN_l}$ is the stack of M mixture coefficients in a vector form. Each coefficient contains $N_k \times N_l$ pixels and $x_{k,l}^m$ denotes the $(k,l) - th$ spatial position of the m-th coefficient. Here $\boldsymbol{y} \in \mathbb{R}^{PN_iN_j}$ is the stack of all multispectral data observed with P broad bands of the imaging system. In addition, $\boldsymbol{n} \in \mathbb{R}^{PN_iN_j}$ represents an additive unknown error associated to multispectral data in $\boldsymbol{y}$.

It is worth noting that the linear system in Equation (5.8) is over-determined, whereas the one proposed in Equation (4.12) is under-determined. In addition, the spectral distribution is totally embedded in the 2D block matrices of $\boldsymbol{H}$ unlike in the previous chapter, where the spectral distribution of the object is carried out by the channels and is weighted by the block matrices of $\boldsymbol{H}$. Therefore, there is no way to enforce spectral prior here as we did in Chapter 4 because it is already embedded in $\boldsymbol{H}$.

5.5 Reconstruction by Mixture Coefficients Estimation

In this section we aim to reconstruct the spatio-spectral object through the estimation of a stack of mixture coefficients $\widehat{\boldsymbol{x}}$ by using a set of low-resolution multispectral data $\boldsymbol{y}$. A naive reconstruction such least-squares solution $\widehat{\boldsymbol{x}} = \left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{y}$ leads to an unstable solution because of the ill-conditioning of the matrix $\boldsymbol{H}^T\boldsymbol{H}$, similar to the problem encounter in Section 4.5. This ill-conditioning is corrected by adding a regularization term to the cost function. Therefore, $\widehat{\boldsymbol{x}}$ is obtained by solving

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \left\{\mathcal{J}\left(\boldsymbol{x}\right) = \mathcal{Q}(\boldsymbol{x},\boldsymbol{y}) + \mu\, \mathcal{R}(\boldsymbol{x})\right\}, \tag{5.9}$$

where $\mathcal{J}(\boldsymbol{x})$ is a convex cost function which is composed of a data fidelity term $\mathcal{Q}(\boldsymbol{x},\boldsymbol{y})$ in order to enforce agreement of the solution with the data. It is defined under the assumption of a stationary white Gaussian noise by

$$\begin{aligned}\mathcal{Q}(\boldsymbol{x},\boldsymbol{y}) &= \sum_{p=1}^{P}\sum_{m=1}^{M}\sum_{i=1}^{N_i}\sum_{j=1}^{N_j}\left(\boldsymbol{y}^{(p)} - \boldsymbol{H}^{p,m}\boldsymbol{x}^m\right)_{i,j}^2 \\ &= \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 .\end{aligned} \tag{5.10}$$

The second term of Equation (5.9) $\mathcal{R}(\boldsymbol{x})$ is a regularization term to enforce prior information of the sought object. Finally, $\mu \geq 0$ is a regularization parameter used to tune the trade-off between fidelity to the data and prior information. It is chosen depending on the spatial distribution of the unknown mixture coefficients we want to enforce, e.g. smoothness, sharp-edges, sparsity....

In this chapter, we consider two distributions: the first one is an object with spatial smoothness, and the second one is an object with strong spatial gradients or sharp-edges. In subsections 5.5.1 and 5.5.2 we detail two types of multichannel regularization and provide a reconstruction method in each case.

5.5.1 Multichannel Quadratic Regularization

The choice of a quadratic term enforce the spatial smoothness of the solution by penalizing the difference between the four (horizontal and vertical) neighboring pixels of $x_{k,l}^m$ for each mixture coefficient. The quadratic regularization $\mathcal{R}_{l_2}(\boldsymbol{x})$ is then defined by

$$\begin{aligned}\mathcal{R}_{l_2}(\boldsymbol{x}) &= \sum_{m=1}^{M}\sum_{k=1}^{N_k}\sum_{l=1}^{N_l}\left(\left(x_{k+1,l}^m - x_{k,l}^m\right)^2 + \left(x_{k,l+1}^m - x_{k,l}^m\right)^2\right) \\ &= \sum_{m=1}^{M}\|\boldsymbol{D}\boldsymbol{x}^m\|_2^2 = \left\|\overline{\boldsymbol{D}}\boldsymbol{x}\right\|_2^2,\end{aligned} \tag{5.11}$$

with

$$\overline{\boldsymbol{D}} = \begin{pmatrix}\boldsymbol{D} & & & \\ & \boldsymbol{D} & & \\ & & \ddots & \\ & & & \boldsymbol{D}\end{pmatrix}_{MN_kN_l \times MN_kN_l} \tag{5.12}$$

where $\overline{\boldsymbol{D}}$ is a block-diagonal matrix and $\boldsymbol{D} \in \mathbb{R}^{N_k N_l \times N_k N_l}$ is a 2D second-order finite difference operator, e.g. 2D *Laplacian* filter, with circularity conditions $x^m_{N_k+1,l} = x^m_{1,l}$ and $x^m_{k,N_l+1} = x^m_{k,1}$.

Therefore, the quadratic cost function $\mathcal{J}_{l_2}(\boldsymbol{x})$,

$$\mathcal{J}_{l_2}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \mu \left\|\overline{\boldsymbol{D}}\boldsymbol{x}\right\|_2^2, \tag{5.13}$$

is a sum of two convex functions, which ensures the existence of a global solution of the problem. In addition, $\mathcal{J}_{l_2}(\boldsymbol{x})$ is linear and differentiable, thus, the multichannel quadratic solution $\widehat{\boldsymbol{x}}_{l_2}$ is explicit and obtained by canceling the gradient of $\mathcal{J}_{l_2}(\boldsymbol{x})$. This yields

$$\widehat{\boldsymbol{x}}_{l_2} = \underbrace{\left(\boldsymbol{H}^T\boldsymbol{H} + \mu\overline{\boldsymbol{D}}^T\overline{\boldsymbol{D}}\right)}_{\boldsymbol{Q}_{l_2}}{}^{-1} \underbrace{\left(\boldsymbol{H}^T\boldsymbol{y}\right)}_{\boldsymbol{q}_{l_2}}, \tag{5.14}$$

where $\boldsymbol{Q}_{l_2} \in \mathbb{R}^{MN_kN_l \times MN_kN_l}$ is the Hessian of the cost function $\mathcal{J}_{l_2}(\boldsymbol{x})$ with a block structure, while $\boldsymbol{q}_{l_2}$ is the gradient of $\mathcal{J}_{l_2}$ at the origin ($\boldsymbol{x} = 0$). We compute the quadratic solution in Equation (5.14) by inverting the Hessian matrix. This is done by assuming circularity of the convolution matrices, $\boldsymbol{H}^{p,m} : p = 1, \ldots, P, m = 1, \ldots, M$, and diagonalizing them in the *Fourier* space. The Hessian matrix becomes a non-diagonal block-diagonal matrix [Galatsanos *et al.* 1991] where its inversion procedure is presented in Section 4.5.3. Algorithm 5.4 presents a pseudo-algorithm that summarizes the multichannel reconstruction with a quadratic regularization.

Algorithm 5.4 Multichannel Reconstruction with a Quadratic Regularization (l_2)

Input $\boldsymbol{H}, \boldsymbol{y}, \boldsymbol{D}, \mu, a = 0.5, N_{\text{iter}}$

$\overline{\boldsymbol{D}} = \text{diag}\{\boldsymbol{D}, \boldsymbol{D}, \ldots, \boldsymbol{D}\}$ ▷ Equation (5.28)
$\overline{\boldsymbol{F}} = \text{diag}\{\boldsymbol{F}, \boldsymbol{F}, \ldots, \boldsymbol{F}\}$

Compute the Hessian matrix :
$\boldsymbol{Q}_{l_2} \leftarrow \boldsymbol{H}^T\boldsymbol{H} + \mu\,\overline{\boldsymbol{D}}^T\overline{\boldsymbol{D}}$ ▷ Equation (5.29)

Diagonalize $\boldsymbol{Q}_{l_2}$ (Non-Circulant Block Circulant) :
$\boldsymbol{\Lambda_H} \leftarrow \left\{\boldsymbol{\Lambda}_{\boldsymbol{H}}^{i,j} = \boldsymbol{F}\boldsymbol{H}^{i,j}\boldsymbol{F}^{\dagger}\right\}_{i,j=1}^{P,M}$
$\boldsymbol{\Lambda_Q} \leftarrow \left\{\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{i,j} = \boldsymbol{F}\boldsymbol{Q}_{l_2}^{i,j}\boldsymbol{F}^{\dagger}\right\}_{i,j=1}^{M,M}$ ▷ Equation (4.35)
$M, M, \overline{N}_k, \overline{N}_l = \text{size}(\boldsymbol{\Lambda_Q})$

Invert $\boldsymbol{\Lambda_Q}$ (Non-Diagonal Block Diagonal) :
for $k = 0 : \overline{N}_k$ **do**
  for $l = 0 : \overline{N}_l$ **do**
    $\boldsymbol{R} = \boldsymbol{\Lambda}_{\boldsymbol{Q}}^{:,:}[k, l]$ ▷ Equation (4.36)
    $\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{:,:\ \text{inv}}[k, l] = \boldsymbol{R}^{-1}$ ▷ Equation (4.37)

Compute the solution:
$\widehat{\boldsymbol{x}}_{l_2} \leftarrow \overline{\boldsymbol{F}}^{\dagger}\,\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{\text{inv}}\boldsymbol{\Lambda}_{\boldsymbol{H}}^{\dagger}\overline{\boldsymbol{F}}\,\boldsymbol{y}$ ▷ Equation (5.29)
return $\widehat{\boldsymbol{x}}_{l_2}$

5.5.2 Multichannel Non-Quadratic Regularization : Half-Quadratic

Despite technical and mathematical advantages, quadratic regularization enforces smoothness of the solution but fails to restore strong spatial gradients (high frequencies of the gradient $\boldsymbol{D}\boldsymbol{x}$) and creates ringing artifacts. This limit the use of the quadratic regularization, especially for deconvolution problems where high frequencies are corrupted by the noise. A non-quadratic regularization is defined by a non-quadratic function in order to overcome these artifacts.

Several methods are found in the literature such as methods based on partial differential equation using a non-quadratic function [Perona & Malik 1990], Total Variation (l_1-norm of the gradient) [Rudin *et al.* 1992, Chambolle 2004] or Half-Quadratic regularization (l_2/l_1-norm) [Geman & Reynolds 1992, Geman & Yang 1995, Charbonnier *et al.* 1997]. We are particularly interested in the method proposed in [Geman & Yang 1995] for three practical reasons. Firstly, the minimization of the cost function is done through alternating two simple minimization problems, one is quadratic and the other is separable. Secondly, we want to exploit the advantages of the quadratic solution such as an explicit and a linear solution, so we can invert the Hessian matrix in the Fourier domain for faster and efficient computation. Thirdly, a variety of convex functions can be used to define the

non-quadratic l_2/l_1-norm (e.g. *Huber*, *Hyperbolic*, *LogCosh*,...).

The half-quadratic regularization proposed in [Geman & Yang 1995] consists of introducing a couple of $N_k \times N_l$ *auxiliary* variables $b = (b_h, b_v)$ horizontal and vertical, so that the non-quadratic regularization function φ is expressed as the minimum (with respect to b) of the sum of a quadratic function and an auxiliary function $\xi(b)$

$$\varphi(\delta) = \min_b \left\{ \frac{1}{2}(\delta - b)^2 + \xi(b) \right\}, \quad \forall \delta \in \mathbb{R}. \tag{5.15}$$

The horizontal and vertical auxiliary variables b_h, b_v are proportional to the gradient of x, i.e. the higher the gradient (sharp discontinuity), the higher the intensity of the auxiliary variables. They shift the quadratic function to a suitable position so the cost of the non-quadratic regularization at high gradient is lower compared to the cost of the quadratic regularization. In Figure 5.3 we display three regularization functions, quadratic (l_2-norm), absolute value (l_1-norm as for the total variation) and Huber function (l_2/l_1-norm) which is defined by

$$\varphi(\delta) = \begin{cases} \delta^2 & \text{if } |\delta| < s \\ 2s|\delta| - s^2 & \text{otherwise} \end{cases}, \tag{5.16}$$

where the threshold parameter s define the transition of the Huber function from quadratic function to absolute value function. We illustrate in the figure that a lower cost is obtained with Huber function (blue dot) compared to the quadratic function (red dot) for high value of δ, and this by shifting the quadratic function by a value of the auxiliary variable. This is exactly what we are looking for in order to decrease the cost of the regularization function at high gradient and avoid smoothing them. In a Bayesian framework, auxiliary variables are interpreted as the mean of a Gaussian distribution [Champagnat & Idier 2004, Giovannelli 2008].

In addition, the Huber functions can approximate the quadratic (l_2-norm) or the absolute function (l_1-norm), up to a certain limit of δ, by fixing a threshold parameter s to a high value or a low value receptively. This make the Huber function more reliable and flexible for smooth objects with or without high spatial gradients.

To address the non-quadratic reconstruction we define the non-quadratic cost function by

$$\mathcal{J}_{l_2/l_1}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \mu\, \mathcal{R}_{l_2/l_1}(\boldsymbol{x}) \tag{5.17}$$

with the multichannel non-quadratic regularization is given by

$$\begin{aligned} \mathcal{R}_{l_2/l_1}(\boldsymbol{x}) &= \frac{1}{a} \sum_{m=1}^{M} \left(\sum_{k=1}^{N_k} \sum_{l=1}^{N_l} \varphi_a \left(\boldsymbol{x}_{k,l+1}^m - \boldsymbol{x}_{k,l}^m \right) + \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} \varphi_a \left(\boldsymbol{x}_{k+1,l}^m - \boldsymbol{x}_{k,l}^m \right) \right) \\ &= \frac{1}{a} \sum_{m=1}^{M} \left(\sum_{k=1}^{N_k} \sum_{l=1}^{N_l} \varphi_a \left([\boldsymbol{D}_h \boldsymbol{x}^m]_{k,l} \right) + \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} \varphi_a \left([\boldsymbol{D}_v \boldsymbol{x}^m]_{k,l} \right) \right), \end{aligned} \tag{5.18}$$

where $\boldsymbol{D}_h, \boldsymbol{D}_v \in \mathbb{R}^{N_k N_l \times N_k N_l}$ are first-order finite difference operators between two pixels along the horizontal and vertical direction of $\boldsymbol{x}^m$, respectively, with circularity conditions

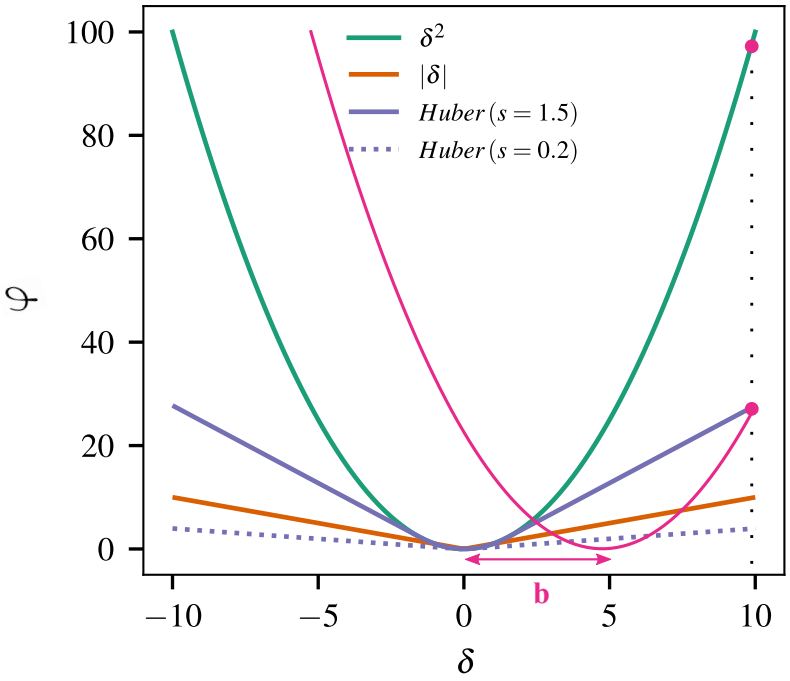


Figure 5.3: Comparison between three regularization functions: Quadratic function, Absolute value function and Huber function. The symbol b in blue indicates the auxiliary variable. The red and blue dots indicates the cost of the quadratic function at $\delta = 10$ and the Huber function respectively. This indicates the decrease of the cost function by adding an auxiliary variable to shift the quadratic function.

$x^m_{N_k+1,l} = x^m_{1,l}$ and $x^m_{k,N_l+1} = x^m_{k,1}$. φ is a non-quadratic regularization function such as Huber function. Same as [Idier 2001] we introduced a scale parameter $a > 0$ which controls the scale of the non-quadratic function, $\varphi_a = a\varphi$, and allows the use of a wide range of auxiliary functions, e.g. $a \in [0, \frac{1}{2}]$ for the Huber function and the auxiliary function is defined by

$$\xi_a(b) = a \begin{cases} \frac{1}{1-2a} b^2 & \text{if } |b| < (1-2a)s \\ 2s|b| - (1-2a)s^2 & \text{otherwise} \end{cases}. \tag{5.19}$$

Therefore, the non-quadratic cost function $\mathcal{J}_{l_2/l_1}(\boldsymbol{x})$ can be written as the minimum of an half-quadratic augmented criterion

$$\min_{\boldsymbol{b}_h, \boldsymbol{b}_v} \mathcal{J}^*_{l_2/l_1}(\boldsymbol{x}, \boldsymbol{b}_h, \boldsymbol{b}_v) = \mathcal{J}_{l_2/l_1}(\boldsymbol{x}), \tag{5.20}$$

where $\mathcal{J}^*_{l_2/l_1}(\boldsymbol{x}, \boldsymbol{b}_h, \boldsymbol{b}_v)$ is defined by replacing the non linear function (in Equation (5.18)) by Geman&Yang convex construction [Geman & Yang 1995] (in Equation (5.15)). This

yields

$$\mathcal{J}^*_{l_2/l_1}(\boldsymbol{x},\boldsymbol{b}_h,\boldsymbol{b}_v) = \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|_2^2 + \frac{\mu}{a}\sum_{m=1}^{M}\left(\sum_{k=1}^{N_k}\sum_{l=1}^{N_l}\left(\frac{1}{2}\left[\boldsymbol{D}_h\boldsymbol{x}^m - \boldsymbol{b}_h^m\right]_{k,l}^2 + \xi_a\left(\left[\boldsymbol{b}_h^m\right]_{k,l}\right)\right)\right.$$
$$\left. + \sum_{k=1}^{N_k}\sum_{l=1}^{N_l}\left(\frac{1}{2}\left[\boldsymbol{D}_v\boldsymbol{x}^m - \boldsymbol{b}_v^m\right]_{k,l}^2 + \xi_a\left(\left[\boldsymbol{b}_v^m\right]_{k,l}\right)\right)\right), \quad (5.21)$$

and

$$\boldsymbol{b}_h = \begin{pmatrix} \boldsymbol{b}_h^1 \\ \boldsymbol{b}_h^2 \\ \vdots \\ \boldsymbol{b}_h^M \end{pmatrix} \in \mathbb{R}^{MN_kN_l}, \text{ and } \boldsymbol{b}_v = \begin{pmatrix} \boldsymbol{b}_v^1 \\ \boldsymbol{b}_v^2 \\ \vdots \\ \boldsymbol{b}_v^M \end{pmatrix} \in \mathbb{R}^{MN_kN_l} \quad (5.22)$$

are vector representations of the stack of all auxiliary variables along the horizontal and vertical directions, respectively.

The augmented criteria in Equation (5.21) is composed of three terms. (1) A least-squares term measuring the fidelity to the data, (2) a quadratic term expressing the difference between neighbor pixels and depends on the auxiliary variables, and (3) an auxiliary function which depends only on auxiliary variables.

Therefore, the multichannel half-quadratic solution is obtained by minimizing the augmented cost function

$$\widehat{\boldsymbol{x}}_{l_2/l_1} = \underset{\boldsymbol{x},\boldsymbol{b}}{\operatorname{argmin}}\ \mathcal{J}^*_{l_2/l_1}(\boldsymbol{x},\boldsymbol{b}_h,\boldsymbol{b}_v), \quad (5.23)$$

by performing an alternate minimization with respect to $\boldsymbol{x}$ and $\boldsymbol{b}$. It results into two minimization problems

$$\begin{cases} \widehat{\boldsymbol{b}}_h, \widehat{\boldsymbol{b}}_v = \underset{\boldsymbol{b}_h,\boldsymbol{b}_v}{\operatorname{argmin}}\ \mathcal{J}^*_{l_2/l_1}(\boldsymbol{x},\boldsymbol{b}_h,\boldsymbol{b}_v) & (5.24) \\ \widehat{\boldsymbol{x}}_{l_2/l_1} = \underset{\boldsymbol{x}}{\operatorname{argmin}}\ \mathcal{J}^*_{l_2/l_1}(\boldsymbol{x},\widehat{\boldsymbol{b}}_h,\widehat{\boldsymbol{b}}_v) & (5.25) \end{cases}$$

The first problem in Equation (5.25) corresponds to the minimization of $\mathcal{J}^*_{l_2/l_1}(\boldsymbol{x},\boldsymbol{b}_h,\boldsymbol{b}_v)$ with respect to $\boldsymbol{x}$. This yields

$$\widehat{\boldsymbol{x}}_{l_2/l_1} = \underset{\boldsymbol{x}}{\operatorname{argmin}}\ \left\{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|_2^2 + \frac{\mu}{2a}\sum_{m=1}^{M}\left(\|\boldsymbol{D}_h\boldsymbol{x}^m - \boldsymbol{b}_h^m\|_2^2 + \|\boldsymbol{D}_v\boldsymbol{x}^m - \boldsymbol{b}_v^m\|_2^2\right)\right\} \quad (5.26)$$

or in a compact representation

$$\widehat{\boldsymbol{x}}_{l_2/l_1} = \underset{\boldsymbol{x}}{\operatorname{argmin}}\ \left\{\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|_2^2 + \frac{\mu}{2a}\left(\left\|\overline{\boldsymbol{D}}_h\boldsymbol{x}-\boldsymbol{b}_h\right\|_2^2 + \left\|\overline{\boldsymbol{D}}_v\boldsymbol{x}-\boldsymbol{b}_v\right\|_2^2\right)\right\} \quad (5.27)$$

with

$$\overline{\boldsymbol{D}}_h = \begin{pmatrix} \boldsymbol{D}_h & & & \\ & \boldsymbol{D}_h & & \\ & & \ddots & \\ & & & \boldsymbol{D}_h \end{pmatrix}, \overline{\boldsymbol{D}}_v = \begin{pmatrix} \boldsymbol{D}_v & & & \\ & \boldsymbol{D}_v & & \\ & & \ddots & \\ & & & \boldsymbol{D}_v \end{pmatrix} \in \mathbb{R}^{MN_kN_l \times MN_kN_l} \quad (5.28)$$

Therefore the minimizer of (5.25) is given by

$$\widehat{\boldsymbol{x}}_{l_2/l_1} = \underbrace{\left(\boldsymbol{H}^T\boldsymbol{H} + \frac{\mu}{2a}\left(\overline{\boldsymbol{D}}_h^T\overline{\boldsymbol{D}}_h + \overline{\boldsymbol{D}}_v^T\overline{\boldsymbol{D}}_v\right)\right)^{-1}}_{\boldsymbol{Q}_{l_2/l_1}} \underbrace{\left(\boldsymbol{H}^T\boldsymbol{y} + \frac{\mu}{2a}\left(\overline{\boldsymbol{D}}_h^T\boldsymbol{b}_h + \overline{\boldsymbol{D}}_v^T\boldsymbol{b}_v\right)\right)}_{\boldsymbol{q}_{l_2/l_1}}. \quad (5.29)$$

where $\boldsymbol{Q}_{l_2/l_1} \in \mathbb{R}^{MN_kN_l \times MN_kN_l}$ is a non-circulant block circulant matrix and $\boldsymbol{q}_{l_2/l_1} \in \mathbb{R}^{MN_kN_l}$ is a multichannel vector.

This solution is similar to the one obtained in Equation (5.14) with an additive term on the right to account for auxiliary variables. It is computed directly by inverting the Hessian matrix $\boldsymbol{Q}$, by considering circulant approximation of $\boldsymbol{H}^{p,m} : p = 1, \ldots, P, m = 1, \ldots, M$, $\boldsymbol{D}_h$ and $\boldsymbol{D}_v$, and diagonalizing them in the Fourier space. The inversion procedure is described in details in Section 4.5.3.

The second minimization problem in Equation (5.24) concerns the update of the auxiliary variables. This yields

$$\begin{aligned}\widehat{\boldsymbol{b}}_h^m, \widehat{\boldsymbol{b}}_v^m = \underset{\boldsymbol{b}_h^m, \boldsymbol{b}_v^m}{\operatorname{argmin}} \Bigg\{ &\sum_{k=1}^{N_k}\sum_{l=1}^{N_l} \underbrace{\left(\frac{1}{2}\left[\boldsymbol{D}_h\boldsymbol{x}^m - \boldsymbol{b}_h^m\right]_{k,l}^2 + \xi_\alpha\left(\left[\boldsymbol{b}_h^m\right]_{k,l}\right)\right)}_{\psi_h\left([\boldsymbol{b}_h]_{k,l}\right)} + \\ &\sum_{k=1}^{N_k}\sum_{l=1}^{N_l} \underbrace{\left(\frac{1}{2}\left[\boldsymbol{D}_v\boldsymbol{x}^m - \boldsymbol{b}_v^m\right]_{k,l}^2 + \xi_\alpha\left(\left[\boldsymbol{b}_v^m\right]_{k,l}\right)\right)}_{\psi_v\left([\boldsymbol{b}_v]_{k,l}\right)} \Bigg\} \\ = \sum_{k=1}^{N_k}\sum_{l=1}^{N_l} \underset{\boldsymbol{b}_h^m, \boldsymbol{b}_v^m}{\operatorname{argmin}} &\left\{\psi_h\left(\left[\boldsymbol{b}_h\right]_{k,l}\right) + \psi_v\left(\left[\boldsymbol{b}_v\right]_{k,l}\right)\right\} \end{aligned} \quad (5.30)$$

where ψ_h and ψ_v are two convex and differentiable functions.

The minimization problem in Equation (5.30) is separable for each pixel position (k, l) and mixture coefficient m, thus all elements of the auxiliary variables can be computed in parallel. A single element of the solution is then obtained by

$$\begin{cases} \left[\boldsymbol{b}_h^m\right]_{k,l} = \underset{\boldsymbol{b}_h^m}{\operatorname{argmin}} \left\{\psi_h\left(\left[\boldsymbol{b}_h^m\right]_{k,l}\right)\right\} & (5.31) \\ \left[\boldsymbol{b}_v^m\right]_{k,l} = \underset{\boldsymbol{b}_v^m}{\operatorname{argmin}} \left\{\psi_v\left(\left[\boldsymbol{b}_v^m\right]_{k,l}\right)\right\} & (5.32) \end{cases}$$

where the minimizers $\left[\boldsymbol{b}_h^m\right]_{k,l}$ and $\left[\boldsymbol{b}_v^m\right]_{k,l}$ verify $\psi_h'\left(\left[\widehat{\boldsymbol{b}}_h^m\right]_{k,l}\right) = 0$ and $\psi_v'\left(\left[\widehat{\boldsymbol{b}}_v^m\right]_{k,l}\right) = 0$, respectively, and this $\forall\, k \in [1, N_k]$ and $\forall\, l \in [1, N_l]$. Therefore

$$\psi_h'\left(\left[\widehat{\boldsymbol{b}}_h^m\right]_{k,l}\right) = \begin{cases} \left(\left[\widehat{\boldsymbol{b}}_h^m\right]_{k,l} - \left[\boldsymbol{D}_h\boldsymbol{x}^m\right]_{k,l}\right) + \frac{2a}{1-2a}\left[\widehat{\boldsymbol{b}}_h^m\right]_{k,l}, & \text{If } \left|\left[\widehat{\boldsymbol{b}}_h^m\right]_{k,l}\right| < (1-2a)s \\ \left(\left[\widehat{\boldsymbol{b}}_h^m\right]_{k,l} - \left[\boldsymbol{D}_h\boldsymbol{x}^m\right]_{k,l}\right) + a\,2s\,\operatorname{sign}\left(\left[\widehat{\boldsymbol{b}}_h^m\right]_{k,l}\right), & \text{Otherwise} \end{cases} \quad (5.33)$$

Thus, the obtained auxiliary variables in case of Huber function are

$$\left[\widehat{\boldsymbol{b}}_h^m\right]_{k,l} = \begin{cases} [\boldsymbol{D}_h\boldsymbol{x}^m]_{k,l} - 2a\,[\boldsymbol{D}_h\boldsymbol{x}^m]_{k,l}\,, & \text{If } \left|[\boldsymbol{D}_h\boldsymbol{x}^m]_{k,l}\right| < s \\ [\boldsymbol{D}_h\boldsymbol{x}^m]_{k,l} - a\,2s\,\text{sign}\left([\boldsymbol{D}_h\boldsymbol{x}^m]_{k,l}\right), & \text{Otherwise} \end{cases}$$
$$= [\boldsymbol{D}_h\boldsymbol{x}^m]_{k,l} - a\,\varphi'\left([\boldsymbol{D}_h\boldsymbol{x}^m]_{k,l}\right) \tag{5.34}$$

Finally,

$$\widehat{\boldsymbol{b}}_h^m = \boldsymbol{D}_h\boldsymbol{x}^m - a\,\varphi'\left(\boldsymbol{D}_h\boldsymbol{x}^m\right). \tag{5.35}$$

and similarly,

$$\widehat{\boldsymbol{b}}_v^m = \boldsymbol{D}_v\boldsymbol{x}^m - a\,\varphi'\left(\boldsymbol{D}_v\boldsymbol{x}^m\right). \tag{5.36}$$

where φ' is the first derivative of the Huber function φ and it is given in Appendix B.1.

We summarize the proposed multichannel reconstruction with an half-quadratic regularization (l_2/l_1) in a pseudo-algorithm form in Algorithm 5.5.

Algorithm 5.5 Multichannel Reconstruction with a Half-Quadratic Regularization (l_2/l_1)

Input $\left(\widehat{\boldsymbol{x}}_{l_2/l_1}\right)_0 = \mathbf{0}, \boldsymbol{H}, \boldsymbol{y}, \boldsymbol{D}_h, \boldsymbol{D}_v, \mu, a \in [0, \frac{1}{2}], N_{\text{iter}}$

$\overline{\boldsymbol{D}}_h = \text{diag}\left\{\boldsymbol{D}_h, \boldsymbol{D}_h, \ldots, \boldsymbol{D}_h\right\}$ ▷ Equation (5.28)
$\overline{\boldsymbol{D}}_v = \text{diag}\left\{\boldsymbol{D}_v, \boldsymbol{D}_v, \ldots, \boldsymbol{D}_v\right\}$ ▷ Equation (5.28)
$\overline{\boldsymbol{F}} = \text{diag}\left\{\boldsymbol{F}, \boldsymbol{F}, \ldots, \boldsymbol{F}\right\}$

Compute the Hessian matrix
$\boldsymbol{Q}_{l_2/l_1} \leftarrow \boldsymbol{H}^T\boldsymbol{H} + \frac{\mu}{2a}\left(\overline{\boldsymbol{D}}_h^T\overline{\boldsymbol{D}}_h + \overline{\boldsymbol{D}}_v^T\overline{\boldsymbol{D}}_v\right)$ ▷ Equation (5.29)

Diagonalize $\boldsymbol{Q}_{l_2/l_1}$ (Non-Circulant Block Circulant) :
$\boldsymbol{\Lambda}_{\boldsymbol{Q}} \leftarrow \left\{\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{i,j} = \boldsymbol{F}\boldsymbol{Q}_{l_2/l_1}^{i,j}\boldsymbol{F}^\dagger\right\}_{i,j=1}^{M,M}$
$M, M, \overline{N}_k, \overline{N}_l = \text{size}(\boldsymbol{\Lambda}_{\boldsymbol{Q}})$

Invert $\boldsymbol{\Lambda}_{\boldsymbol{Q}}$ (Non-Diagonal Block Diagonal) :
for $k = 0 : \overline{N}_k$ **do**
  for $l = 0 : \overline{N}_l$ **do**
    $\boldsymbol{R} = \boldsymbol{\Lambda}_{\boldsymbol{Q}}^{:,:}[k,l]$ ▷ Equation (4.36)
    $\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{:,:\,\text{inv}}[k,l] = \boldsymbol{R}^{-1}$ ▷ Equation (4.37)

for $n = 1 : N_{\text{iter}}$ **do**
  1) - Update the auxiliary variables
  $(\widehat{\boldsymbol{b}}_h)_n \leftarrow \overline{\boldsymbol{D}}_h\left(\widehat{\boldsymbol{x}}_{l_2/l_1}\right)_{n-1} - a\,\varphi'_H\left(\overline{\boldsymbol{D}}_h\left(\widehat{\boldsymbol{x}}_{l_2/l_1}\right)_{n-1}\right)$ ▷ Equation (5.35)
  $(\widehat{\boldsymbol{b}}_v)_n \leftarrow \overline{\boldsymbol{D}}_v\left(\widehat{\boldsymbol{x}}_{l_2/l_1}\right)_{n-1} - a\,\varphi'_H\left(\overline{\boldsymbol{D}}_v\left(\widehat{\boldsymbol{x}}_{l_2/l_1}\right)_{n-1}\right)$ ▷ Equation (5.36)
  $\left(\boldsymbol{q}_{l_2/l_1}\right)_n \leftarrow \boldsymbol{H}^T\boldsymbol{y} + \frac{\mu}{2a}\left(\overline{\boldsymbol{D}}_h^T\left(\widehat{\boldsymbol{b}}_h\right)_n + \overline{\boldsymbol{D}}_v^T\left(\widehat{\boldsymbol{b}}_v\right)_n\right)$ ▷ Equation (5.29)

  2) - Compute the solution:
  $\left(\widehat{\boldsymbol{x}}_{l_2/l_1}\right)_n \leftarrow \overline{\boldsymbol{F}}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{Q}}^{\text{inv}}\overline{\boldsymbol{F}}\,\left(\boldsymbol{q}_{l_2/l_1}\right)_n$ ▷ Equation (5.29)

return $\left(\widehat{\boldsymbol{x}}_{l_2/l_1}\right)_{N_{iter}}$

5.6 Simulation and Results

5.6.1 Description the Original Spatio-Spectral Objects

We apply the proposed reconstruction Algorithms 5.4 and 5.5 on simulated multispectral data observed by the Mid-InfraRed Instrument (MIRI) Imager [Bouchet *et al.* 2015] on-board the James Webb Space Telescope (JWST). For this experiment, we use the *HorseHead nebula* cube described in Section 4.6.1 and generate two other spatio-spectral objects with different spatial and spectral distributions in order to test the performance of the algorithms. We provide in the following the spatial and spectral distribution of these

three objects:

1. The first object is a simplified spatio-spectral model of the *HorseHead nebula*, see Section 4.6.1. It is a 256×256 pixels spatial region from the sky taken for simulation, with $N_\lambda = 1000$ spectral samples uniformly distributed within the mid-infrared wavelength range $4 - 28$ μm.

   In Figure 5.4 we display all spectral components of the *HorseHead nebula* cube. We notice a clear correlation between spectral components. Therefore, we perform a data reduction using the Principal Components Analysis method (PCA) [Jolliffe 1986] in order to extract a set of uncorrelated principal components which will be used as spectral components in the forward model (Equation (5.6)).

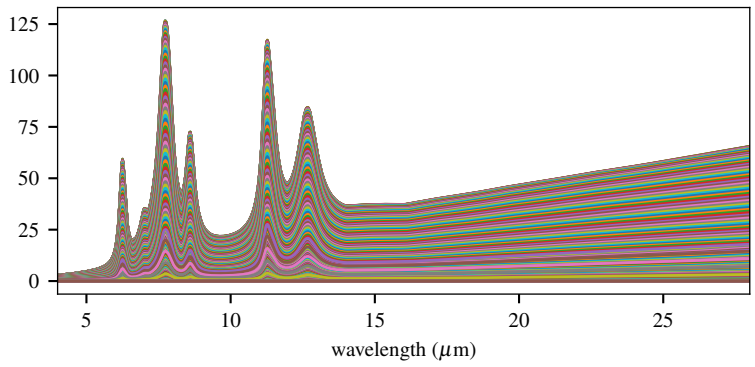


Figure 5.4: Display of all spectral components of the *HorseHead nebula* cube. We notice a clear correlation between spectral distributions of all pixels.

   After applying the PCA algorithm (in Appendix B.2) on the *HorseHead nebula* cube with 10 principal components, we display the explained variance in % for each principal component in Figure 5.5. It is clear that at most $M = 3$ principal components are sufficient to represent a linear combination of all spectral components of the cube. Figure 5.6 shows the three extracted principal components after projection of the data on the new reduced space.

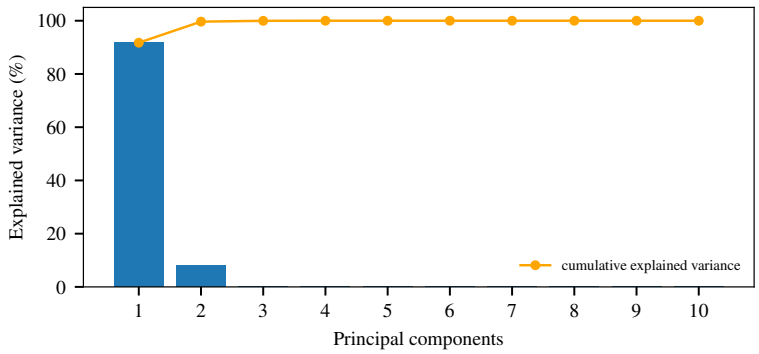


Figure 5.5: Explained variance (in %) by different principal component, extracted from the *HorseHead nebula* cube.

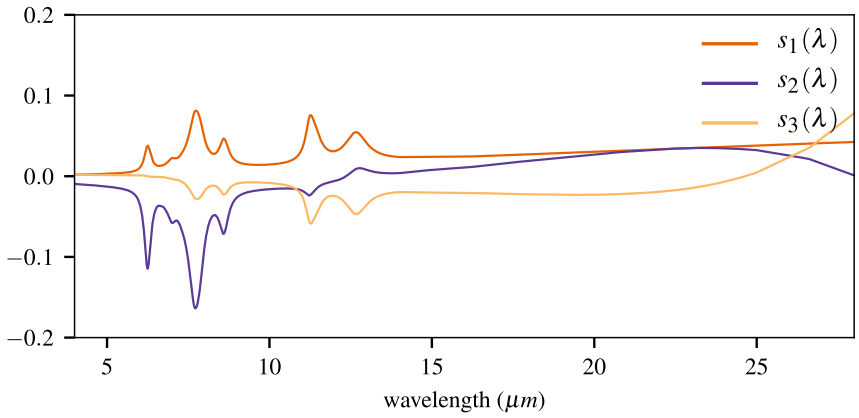


Figure 5.6: Extracted spectral components from the *HorseHead nebula* using the principal component analysis method [Jolliffe 1986].

2. The second object is Synthetic$_1$. It is simulated using three ($M = 3$) mixtures coefficients with strong spatial gradients patterns, see Figure 5.7. Then we associate three spectral components[1] to these mixture coefficients. They are extracted from Spitzer spectroscopy data by using a blind source separation method [Berne *et al.* 2007] and covers the mid-infrared wavelength range from 4 to 21 μm as shown in Figure 5.8.

[1] http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/469/575

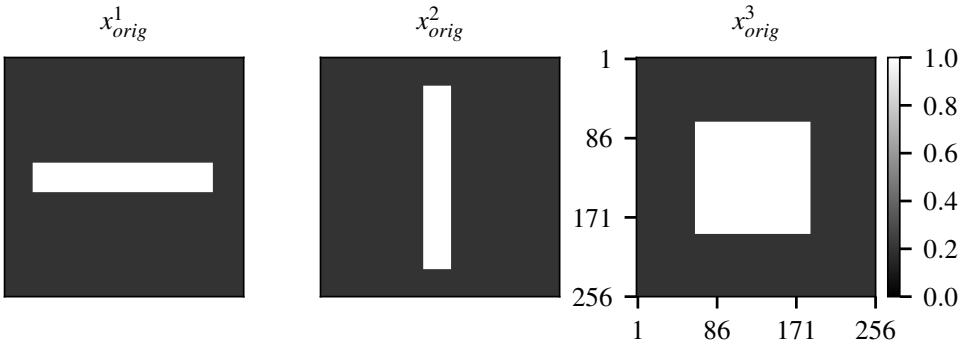


Figure 5.7: Original mixture coefficients used to construct the spatio-spectral object Synthetic_1.

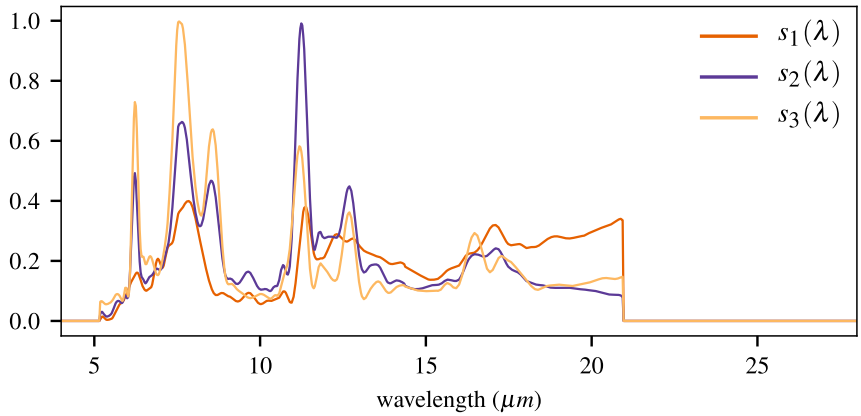


Figure 5.8: Original spectral components used to describe the original spatio-spectral object Synthetic_1. They are extracted from Spitzer spectro-imagery data by using a blind source separation [Berne *et al.* 2007].

3. The third spatio-spectral object Synthetic_2 is created by mixing two mixture coefficient ($M = 2$) having a smooth distribution, i.e. gradient image, and a sharp-edge distribution, e.g. cameraman, see Figure 5.9. The two spectral components associated to mixture coefficients are the ones presented in [Berne *et al.* 2007] and extracted from Spitzer spectroscopy data, see Figure 5.10.

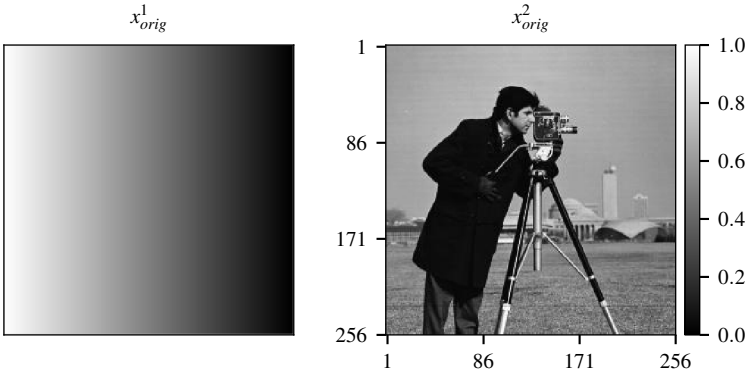


Figure 5.9: Original mixture coefficients used to construct the spatio-spectral object Synthetic_2.

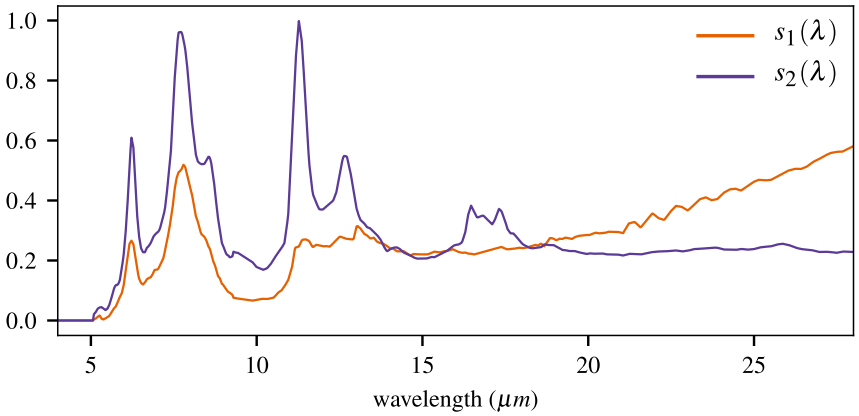


Figure 5.10: Original spectral components used to describe the original spatio-spectral object Synthetic_2. They are extracted from Spitzer spectro-imagery data by using a blind source separation [Berne *et al.* 2007].

The spectral components $s_m(\lambda), m = 1, \ldots, 3$ are re-sampled in order to fit the spectral grid $\mathcal{G}_\lambda$ that is defined in Section 5.4. Next, we simulate nine multispectral data for the three input objects using the developed instrument model in Equation (3.8). An additive zero-mean white Gaussian noise of different level, $\text{SNR} = 40, 30, 20, 10$ and 5 dB, is added to corrupt the multispectral data. The Signal-to-Noise Ratio (SNR) is defined by $\text{SNR}_{(dB)} = 10 \log_{10} \left(\frac{\frac{1}{PN_iN_j} \|\boldsymbol{y}\|_2^2}{\sigma_n^2} \right)$ where σ_n is the standard deviation of the noise. P is the number of multispectral data or bands, and N_iN_j is the total number of pixels in the multispectral data.

5.6.2 Choice of Algorithm Parameters: μ and s

Both quadratic and half-quadratic regularization require a regularization parameter μ. It must be chosen carefully since it compromises the solution between data fidelity and regularization terms. Our supervised approach consists of tuning the parameter by running the code for different values of μ, in a range $[10^{-3}, 10^{1}]$, and keep the value that minimizes the relative reconstruction error, which is defined by

$$Error(\%) = 100 \times \frac{\|\boldsymbol{f}_{orig} - \boldsymbol{f}_{rec}\|_2}{\|\boldsymbol{f}_{orig}\|_2}, \tag{5.37}$$

where $\boldsymbol{f}_{orig}$ is the original object and $\boldsymbol{f}_{rec}$ is the reconstructed one.

In addition, the same approach is used to select the threshold parameter s for the half-quadratic regularization. This approach is time-consuming for a wide range of parameters, but this is not an issue since the solution, quadratic or half-quadratic, is computed in just a few seconds (less than 20 seconds) thanks to the implementation of the algorithms in the Fourier domain.

Finally, all the experiments are coded in Python 2.7 on a laptop machine, with a memory of 16 GiB RAM and a processor Intel Core i7 CPU 2.50GHz ×8.

5.6.3 Results and Discussion

In this section we present the simulation results with a discussion. We first present the simulation results of the multispectral data then the estimated mixtures coefficients, followed by spatial and spectral distributions of the reconstruction object. In addition, we provide algorithm parameters and a comparison between algorithms.

5.6.3.1 Simulated Multispectral Data

Figures 5.11, 5.13 and 5.15 show the simulated multispectral data of nine broad bands of the JWST/MIRI imager at 30 dB, indexed from $p = 1, \ldots, 9$, and associated to the objects *HorseHead nebula*, Synthetic$_1$ and Synthetic$_2$, respectively. Figures 5.12, 5.14 and 5.16 illustrate the spatial distribution of the multispectral data by displaying a slice from the central row of images of Figures 5.11, 5.13 and 5.15. All Figure images are represented with the same color-bar for better illustration and comparison.

These figures illustrate the complexity of the imaging system. We notice the presence of a blur that increases for long wavelengths, e.g. in multispectral data with $p = 8, 9$, compared to short wavelengths, e.g. $p = 1, 2$, and this is due to the spectral variability of the PSF. This blur is also well visible in the multispectral data of the object Synthetic$_1$ in Figure 5.14, where all the strong gradient are smoothed. Moreover, we notice a difference in the intensities between images, where several images appear with a high intensity and others with low intensity. This is due to the relative amplitude of spectral components integrated within bands, and/or and the width of these bands. For instance the 1-*st* band is a narrower compared to the 9-*th* band (see Figure 3.10). Furthermore, we notice the

domination of the noise in the multispectral data for narrow bands, $p = 1, 3, 4$.

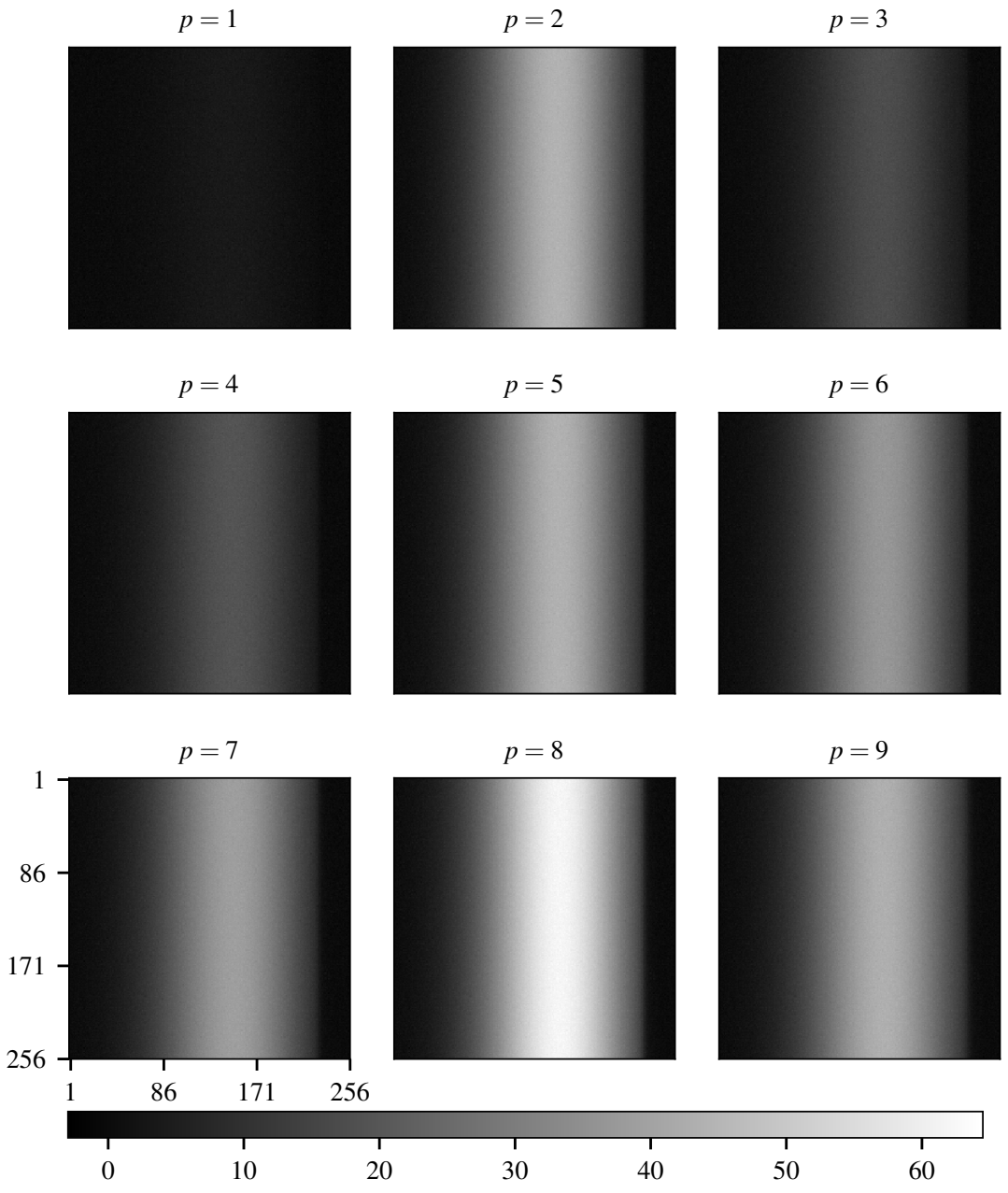


Figure 5.11: Simulation of nine multispectral data using the instrument model in Equation (3.8) for the *HorseHead nebula* object. All the multispectral data are corrupted with an additive white Gaussian noise so that SNR=30 dB.

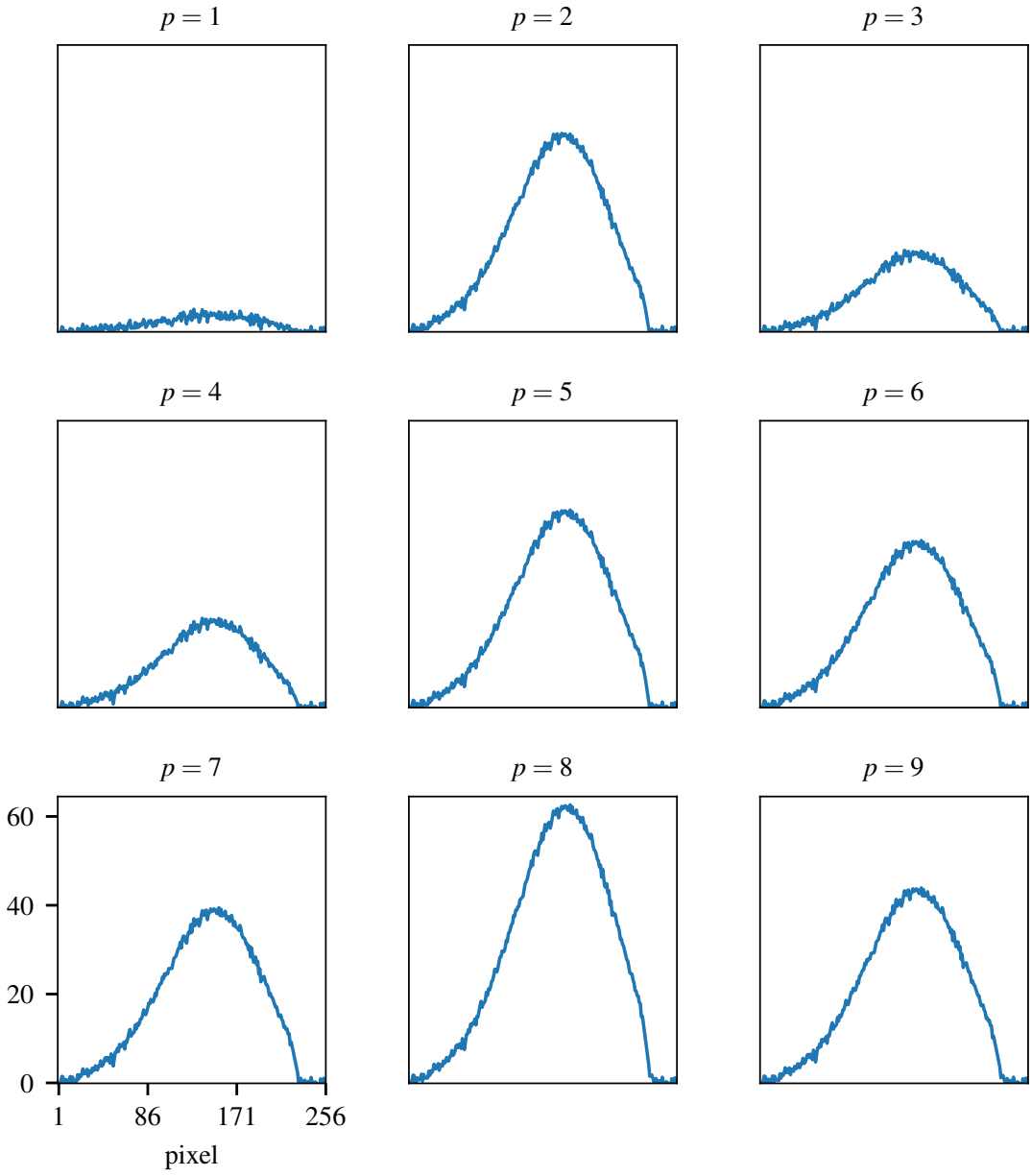


Figure 5.12: Display of slice from the central row of images in Figure 5.11.

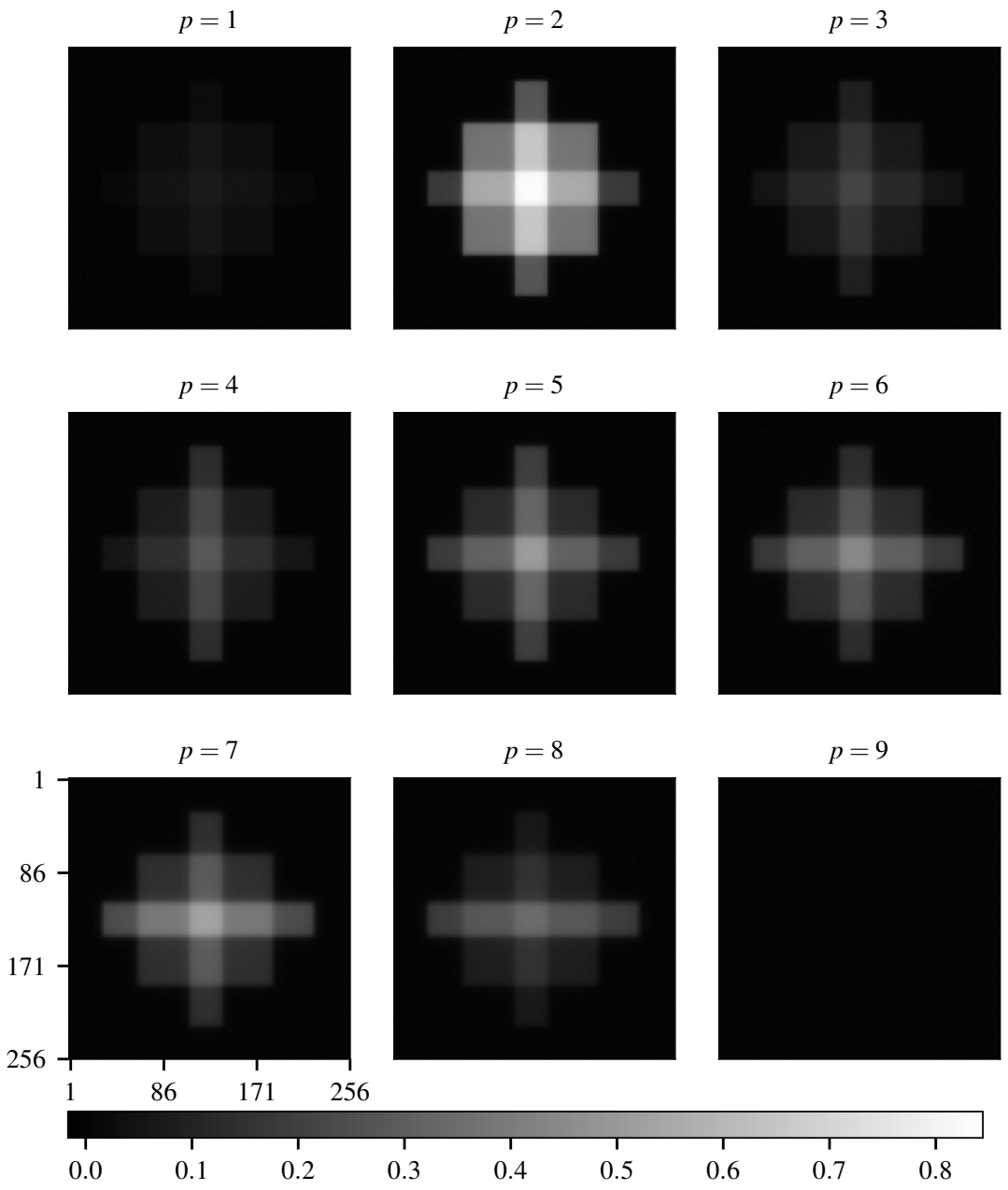


Figure 5.13: Simulation of nine multispectral data using the instrument model in Equation (3.8) for the object Synthetic_1. All the multispectral data are corrupted with an additive white Gaussian noise so that SNR=30 dB.

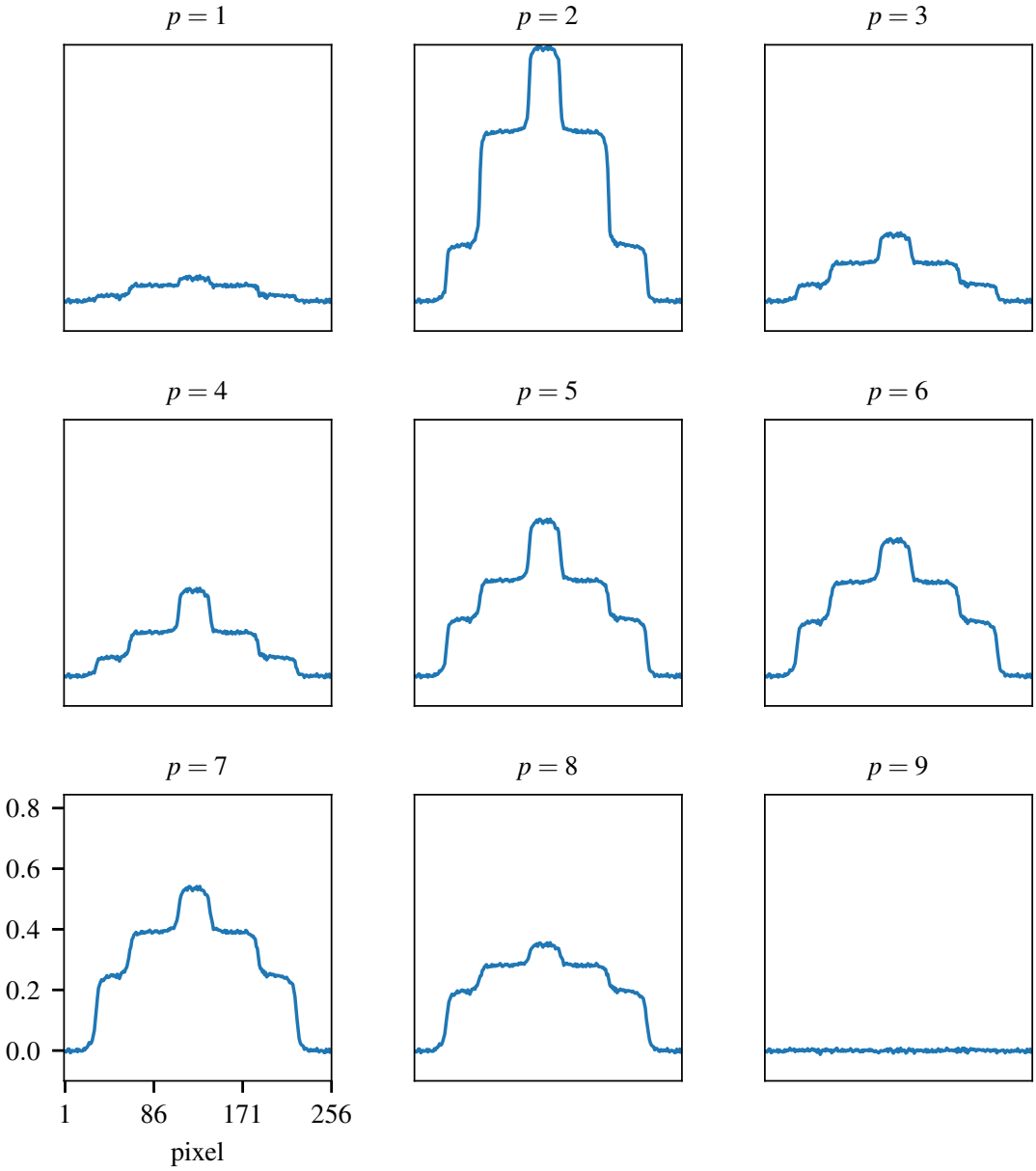


Figure 5.14: Display of slice from the central row of images in Figure 5.13.

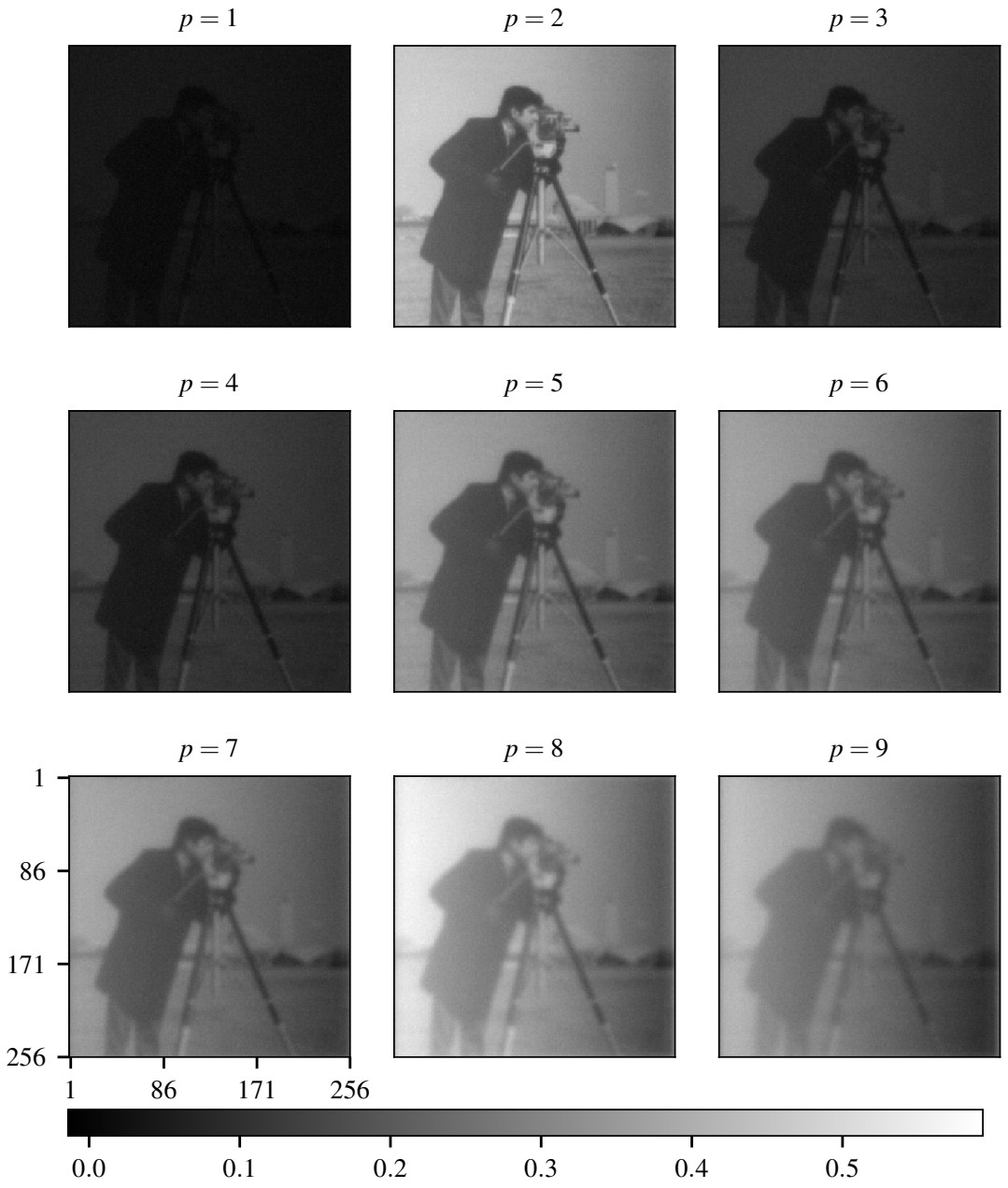


Figure 5.15: Simulation of nine multispectral data using the instrument model in Equation (3.8) for the object Synthetic_2. All the multispectral data are corrupted with an additive white Gaussian noise so that SNR=30 dB.

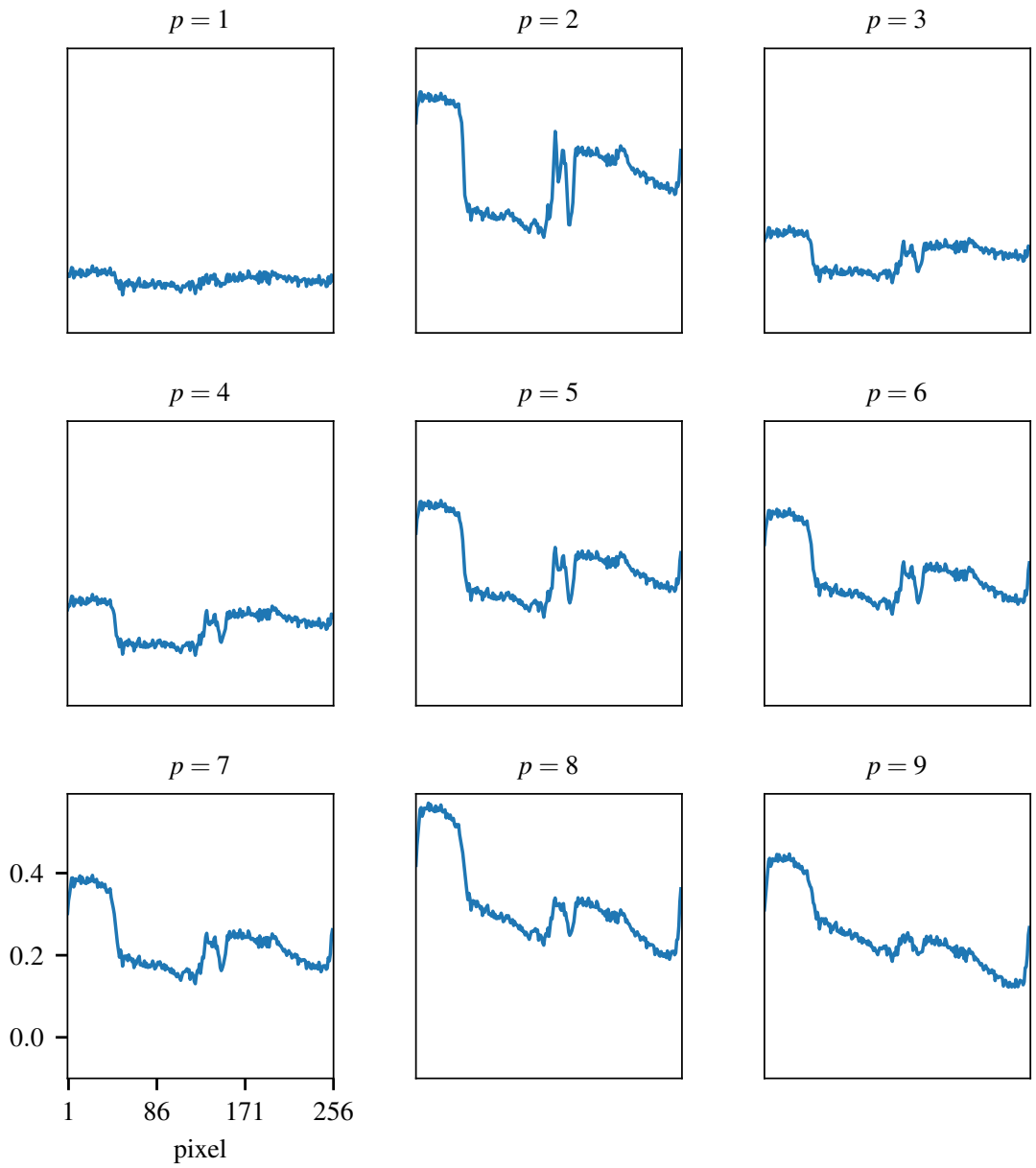


Figure 5.16: Display of slice from the central row of images in Figure 5.15.

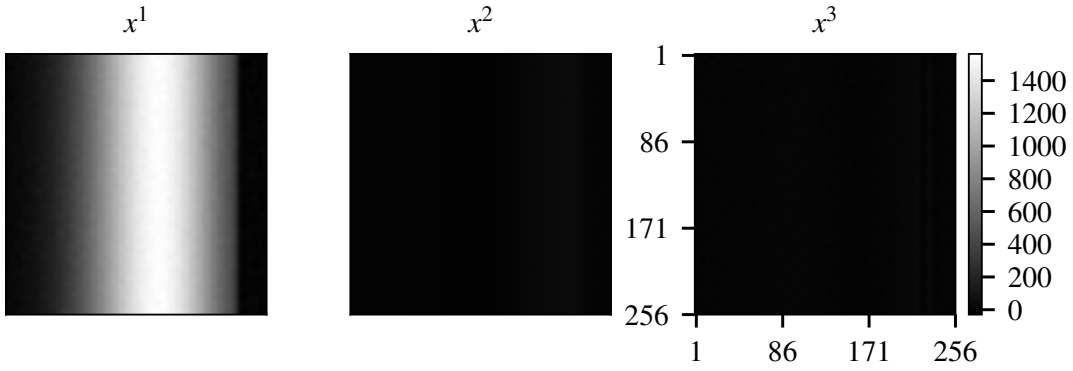
5.6.3.2 Estimated Mixture Coefficient

In Figures 5.17 and 5.18 we display estimation results of the mixture coefficients for the three spatio-spectral objects by using Algorithms 5.4 and 5.5, respectively.

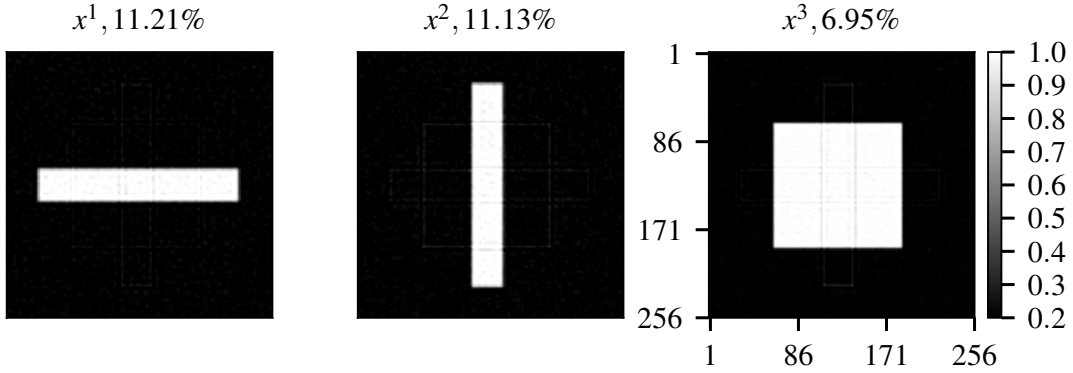
The estimated mixtures coefficients for the *HorseHead nebula* are shown in Figures 5.17.(a) and 5.18.(a) using l_2 and l_2/l_1 algorithms respectively. Both results looks similar where we estimated a high intensity of $\boldsymbol{x}^1$ and a very low intensities for $\boldsymbol{x}^2$ and $\boldsymbol{x}^3$. These results are expected due to the domination of the first spectral component $s_1(\lambda)$ (see Figure 5.6) in all object spectra (Figure 5.4). Unfortunately we can not judge which algorithm provide a better estimation because we do not dispose the original coefficients for the *HorseHead nebula* object, but we can make a comparison later once we presents the reconstructed objects.

Concerning estimated mixture coefficients for objects Synthetic_1 and Synthetic_2, they are shown in Figure 5.17.(b) and 5.17.(c), respectively. We can see clearly cross patterns in each estimated mixture coefficients by the l_2 algorithm. As expected since the quadratic regularization fails to reconstruct strong spatial gradients from the multispectral data. These cross patterns are removed or faded in Figure 5.18.(b) and 5.18.(c) using the half-quadratic regularization thanks to the auxiliary variables that prevent the algorithm from smoothing the strong gradient mixture coefficients. Improvement factors (the ratio error between $l_2/l_1/l_2$) of $2.63, 3.35$ and 3.99 are obtained using l_2/l_1 algorithm instead of the l_2 algorithm for the estimation of the mixture coefficients $\boldsymbol{x}^1, \boldsymbol{x}^2$ and $\boldsymbol{x}^3$ of the object Synthetic_1, and improvement factors of 1.83 and 1.17 are obtained using l_2/l_1 algorithm instead of the l_2 algorithm for the mixture coefficients $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ of the object Synthetic_2. The improvement factors for the object Synthetic_1 are higher than the object Synthetic_2. This is due to the spatial distribution of the original mixture coefficients (in Figure 5.7 and 5.9 respectively), i.e. the three coefficients of Synthetic_1 present higher spatial gradients than the two coefficients of Synthetic_2 (where only one coefficient has high spatial gradient and the other is smooth).

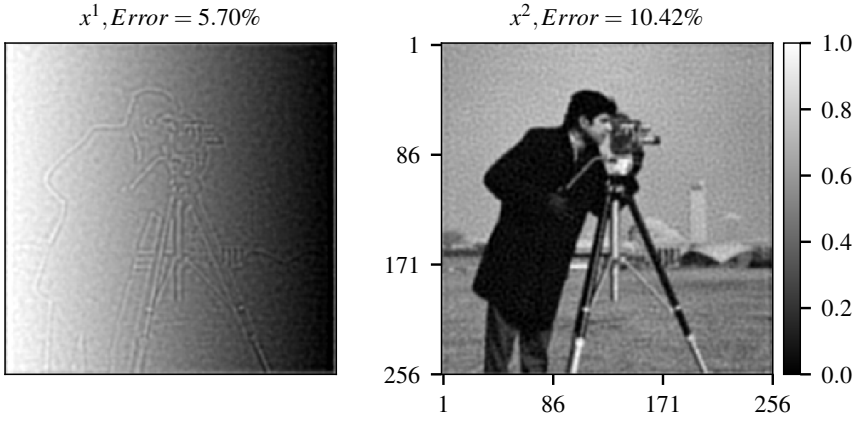
The corresponding auxiliary variables for the three objects are shown in Figures 5.19, 5.20 and 5.21. As expected from the half-quadratic construction in [Geman & Yang 1995], the estimated auxiliary variables mimic very well the gradient of the mixture coefficients, see the horizontal gradient b_h and the vertical gradient b_v of the mixture coefficients. As anticipated in Section 5.5.2, we notice that the threshold parameter s is very small (close to zero) for the objects Synthetic_1 and Synthetic_2 compared to the *HorseHead nebula*. This is due to high gradients of the mixture coefficients of the Synthetic objects. Moreover, the Huber function with a small threshold parameter s is equivalent to the l_1-norm function. This is equivalent to using the total variation regularization but with a lower cost thanks to the half-quadratic solution.



(a) *HorseHead nebula*

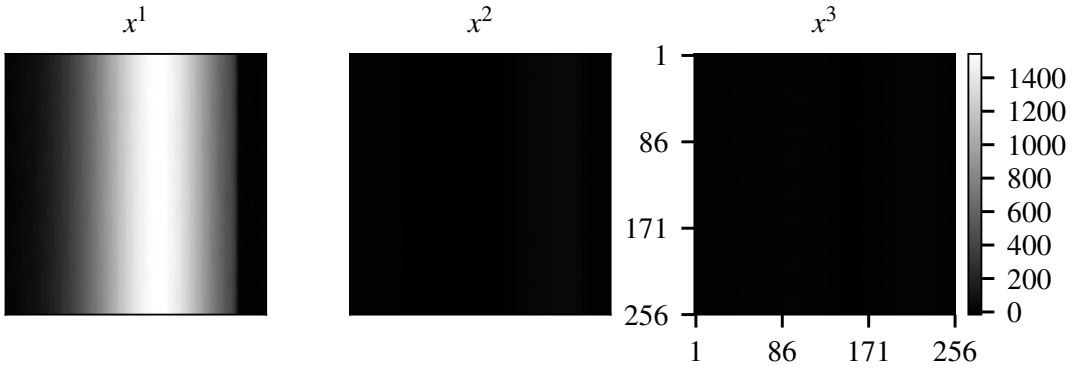


(b) Synthetic$_1$

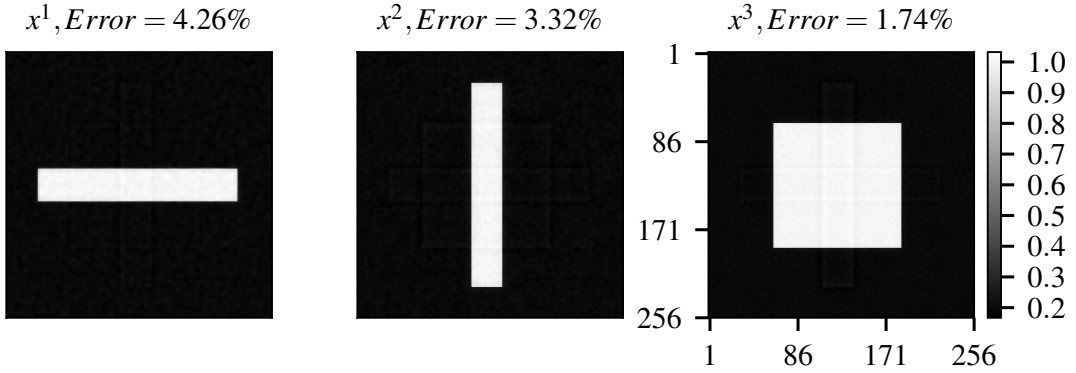


(c) Synthetic$_2$

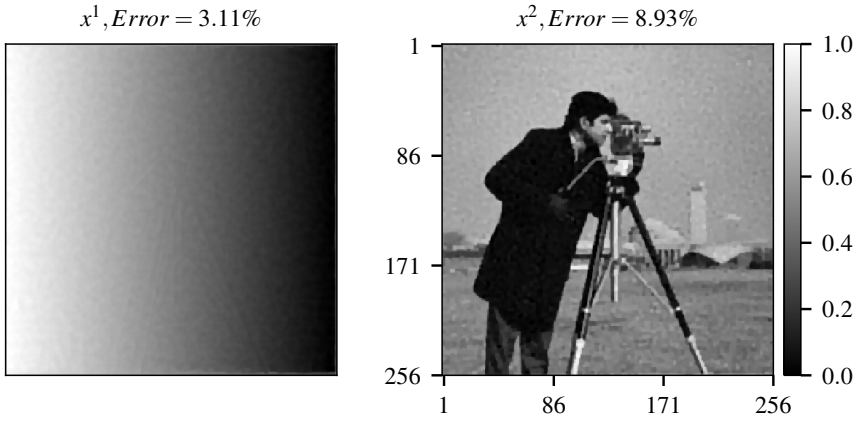
Figure 5.17: The estimation results of the mixture coefficients using the quadratic regularization (l_2-norm).



(a) *HorseHead nebula*



(b) Synthetic$_1$



(c) Synthetic$_2$

Figure 5.18: The estimation results of the mixture coefficients using the half-quadratic regularization (l_2/l_1-norm).

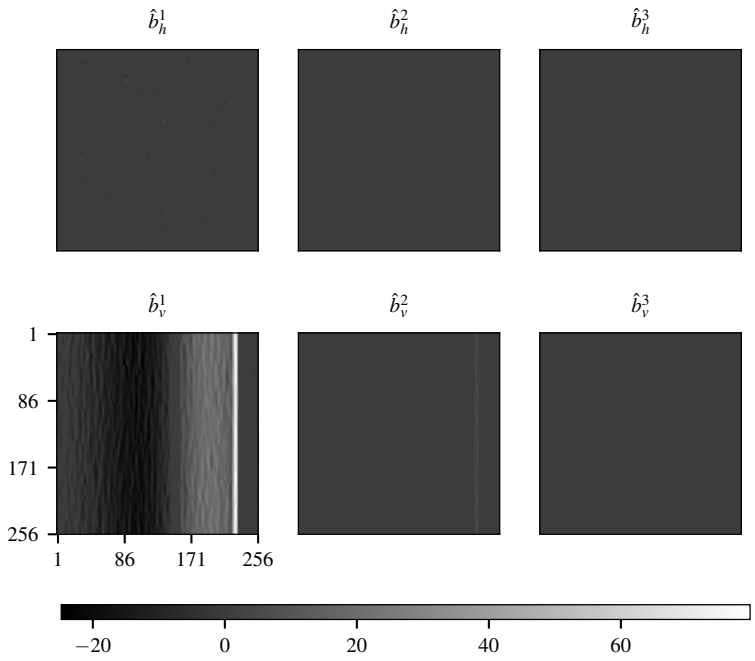


Figure 5.19: Estimated auxiliary variables for the *HorseHead nebula* object after 50 iterations with a threshold parameter $s = 1.17$ and a 30 dB multispectral data. Subscript h refers to horizontal and v for vertical.

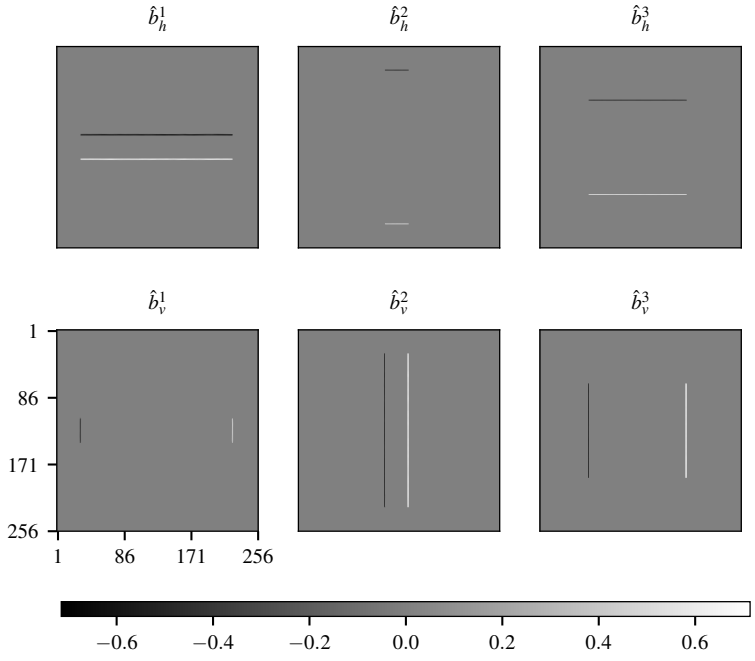


Figure 5.20: Estimated auxiliary variables for the object Synthetic_1 after 50 iterations with a threshold parameter $s = 0.03$ and a 30 dB multispectral data.

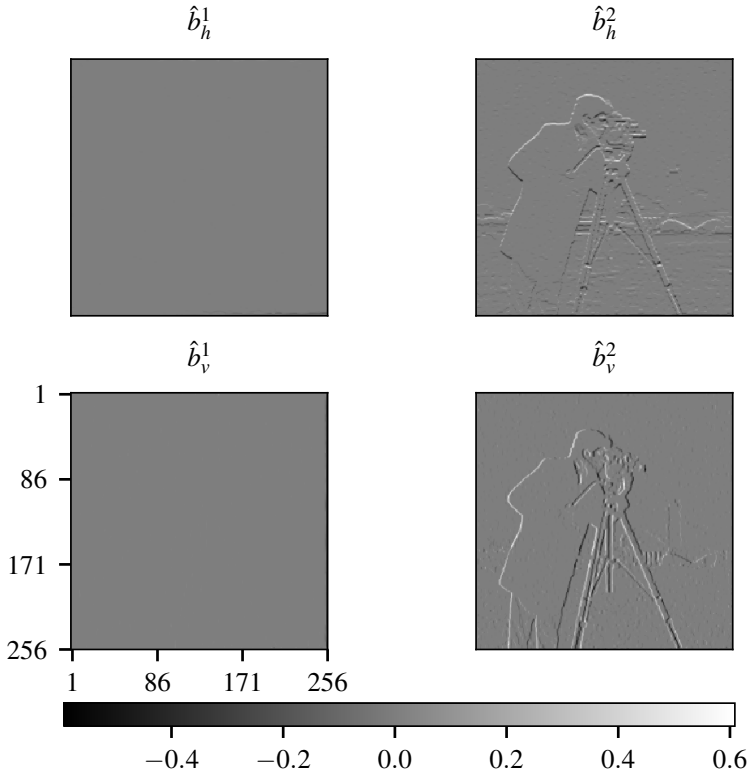


Figure 5.21: Estimated auxiliary variables for the object Synthetic$_2$ after 50 iterations with a threshold parameter $s = 0.01$ and a 30 dB multispectral data.

5.6.3.3 Reconstruction Results of the Spatio-Spectral Object

The discrete version of the spatio-spectral object is reconstructed using the linear mixing model in Equation (5.3). The reconstruction results of the three $1000 \times 256 \times 256$ spatio-spectral objects from a set of nine multispectral data, corrupted with a 30 dB zero-mean white Gaussian noise, are summarized in Table 5.1. A comparison between reconstruction results using l_2 and l_2/l_1 algorithms is provided. The reconstruction error in the third column of the table is computed using whole 2D+λ of the original spatio-spectral object and the reconstructed one.

Based on what we have seen on the estimated mixing coefficients in Section 5.6.3.2, a smaller reconstruction error is obtained using the l_2/l_1 algorithm compared the l_2 algorithm for the three objects at the cost of few more seconds. We obtained a gain of 0.48%, 3.38% and 1.04% for the objects *HorseHead nebula*, Synthetic$_1$ and Synthetic$_2$, respectively. Moreover, improvement factors of 1.73, 2.77 and 1.21 are obtained for the three objects by using l_2/l_1 algorithm compared to the l_2 one. These improvement factors confirm the increase of objects resolutions due to the good estimation of mixture coefficients using the l_2/l_1 algorithm.

The l_2 algorithm needs around 1 second to compute the final solution compared to

19 seconds of the l_2/l_1 algorithm (for 50 iterations). This is because the l_2/l_1 algorithm iterates both the solution and auxiliary variables. It is worth noting that both algorithms are 10 times faster than the iterative algorithm such as the conjugate gradients algorithm (see Algorithm 4.2) thanks to the matrix inversion in the Fourier domain.

Table 5.1: Reconstruction results of three $1000 \times 256 \times 256$ spatio-spectral objects from a set of nine ($P = 9$) multispectral data corrupted with a 30 dB zero-mean white Gaussian noise.

Spatio-Spectral Object ϕ	Regularization	Error (%)	Runtime (seconds)	N_{iter}	μ	s
HorseHead nebula	l_2	1.14	1.36	1	6.11×10^{-3}	
	l_2/l_1	0.66	20.33	50	5.99×10^{-2}	1.17
Synthetic_1	l_2	5.29	1.19	1	6.11×10^{-3}	
	l_2/l_1	1.91	19.98	50	4.64×10^{-2}	0.03
Synthetic_2	l_2	5.95	0.97	1	6.11×10^{-3}	
	l_2/l_1	4.91	18.50	50	5.99×10^{-3}	0.01

Spatial distribution: The comparison between original, observed and reconstructed objects (at $6, 12$ and 18 μm) are shown in Figures 5.22, 5.24 and 5.26 using the l_2 algorithm, and in Figures 5.23, 5.25 and 5.27 using the l_2/l_1 algorithm for objects *HorseHead nebula*, Synthetic_1 and Synthetic_2, respectively.

The reconstructed objects at $6, 12$ and 18 μm show a clear denoising and deconvolution of the multispectral data. The overall dynamic range and the spatial distribution are well reconstructed using both algorithms. However, we obtain better reconstruction results using the l_2/l_1 algorithm due to the good estimation of high gradients of the mixture coefficients. Improvement factors of $0.93, 0.2, 0.18$ are obtained for the *HorseHead nebula* at $6, 12$ and 18 μm, and $2.49, 2.57, 2.46$ are obtained for the object Synthetic_1, and $1.14, 1.15, 1.18$ are obtained for the object Synthetic_2.

Furthermore, for a better illustration and comparison of the spatial distribution, we display a slice (from the middle row) of the multispectral data, original and reconstructed object in Figure 5.28 (l_2 algorithm) and Figure 5.29 (l_2/l_1 algorithm). We see that ringing artifacts appear on strong gradients of the reconstructed object using the quadratic regularization ϕ_{l_2}. The amplitude of these artifacts is strongly reduced using the half-quadratic solution ϕ_{l_2/l_1}, e.g. the middle columns of Figures 5.28 and 5.29 illustrate the significant improvement of the reconstruction results using the l_2/l_1 algorithm over the l_2 one. The noise also appears reduced. Finally, the sharp edges of the object are well reconstructed. Therefore, the proposed method performs deconvolution and denoising of the data at all wavelengths.

Additionally, we depict in Figure 5.30 the difference between the original and the reconstructed object for both algorithms. It is clear that the l_2 algorithm fails to reconstruct

high gradient and that the l_2/l_1 algorithm helps in overcoming this limitation.

Spectral distribution: We display in Figure 5.31 the original spectral distribution ϕ_{orig}, the quadratic solution ϕ_{l_2} and the half-quadratic solution ϕ_{l_2/l_1}, for a single spatial position as indicated in the legend. The results in the first column fit the original distribution over the whole spectral range from 4 to 28 μm because the spatial position corresponds to a smooth region. Whereas, results in the second column show improvement in spectral reconstruction using the l_2/l_1 algorithm compared to the l_2 algorithm for a spatial position located at high gradients.

In Figure 5.32 we compare the reconstruction of the spectral distribution using different methods presented in this thesis. A same spatial position (127,100) of the *HorseHead nebula* is displayed for the original object ϕ_{orig}, the reconstructions proposed in this chapter ϕ_{l_2} and ϕ_{l_2/l_1}, the reconstruction using the proposed method in Chapter 4 ϕ_{MDFT}, and the broadband reconstruction presented also in Chapter 4 $\phi_{Broadband}$. The broadband method gives a bad reconstruction since it neglects of the spectral variations in the broadband PSF. Although ϕ_{MDFT} approximates ϕ_{orig} better than $\phi_{Broadband}$, the result remains insufficient and fails to reconstruct high spectral variations within a band, even if spatial and spectral priors have been enforced to the solution. Lastly, the two proposed solutions in this chapter ϕ_{l_2} and ϕ_{l_2/l_1} provide an accurate and satisfactory reconstruction that fit ϕ_{orig} over the whole range of wavelengths. This is mainly due to the joint processing of all multispectral data and the accuracy of the forward model, through the observation matrix $\boldsymbol{H}$ that describes the instrument and the object models.

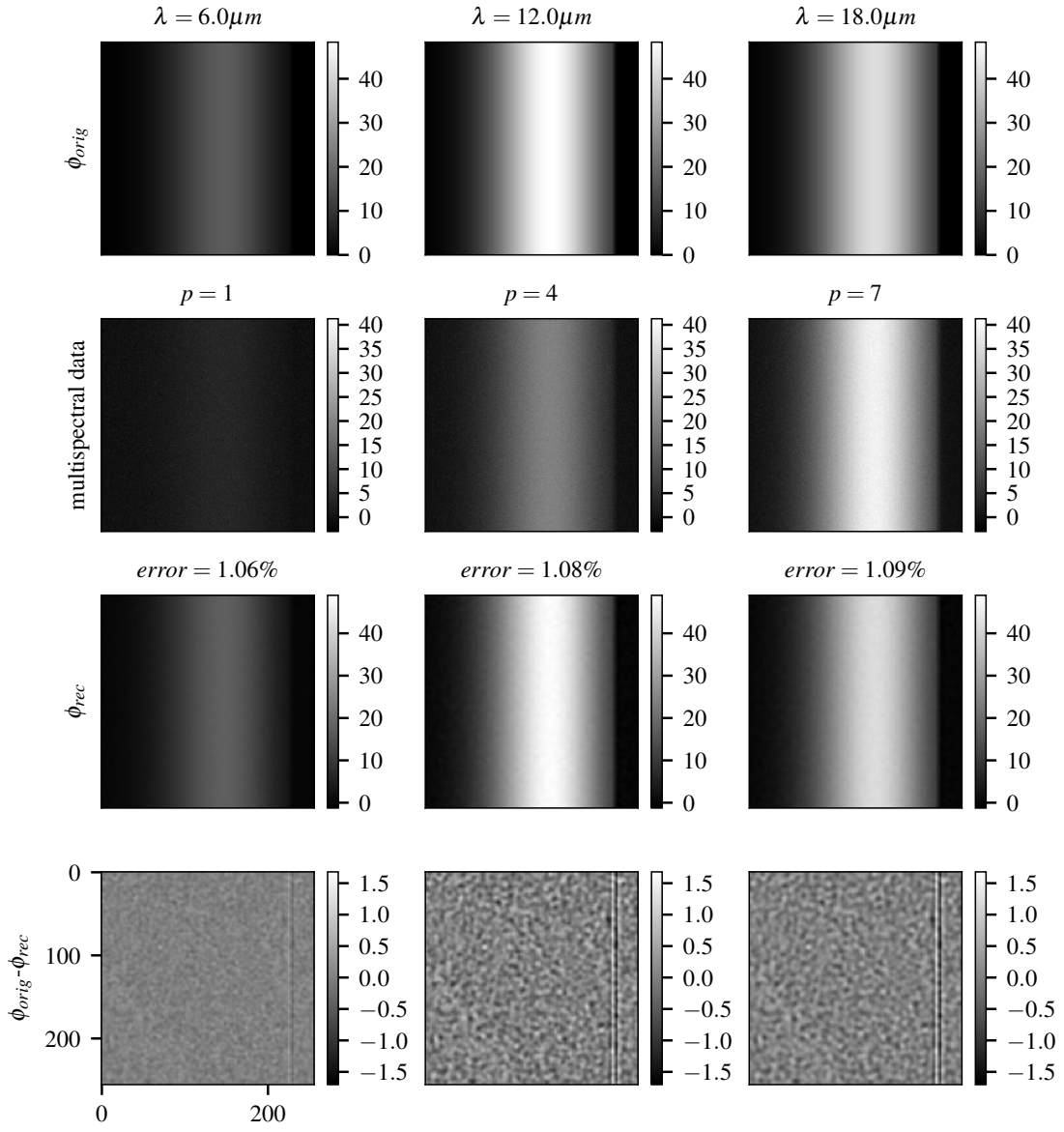


Figure 5.22: [1st row] Original *HorseHead nebula* object at 6, 12 and 18 μm. [2nd row] Simulated multispectral data with 30 dB associated to the bands that include wavelengths of the first row, i.e. $p = 1, 4, 7$. [3rd row] Reconstructed object at 6, 12 and 18 μm with the l_2 algorithm. [4th row] Difference between the original and reconstructed at 6, 12 and 18 μm..

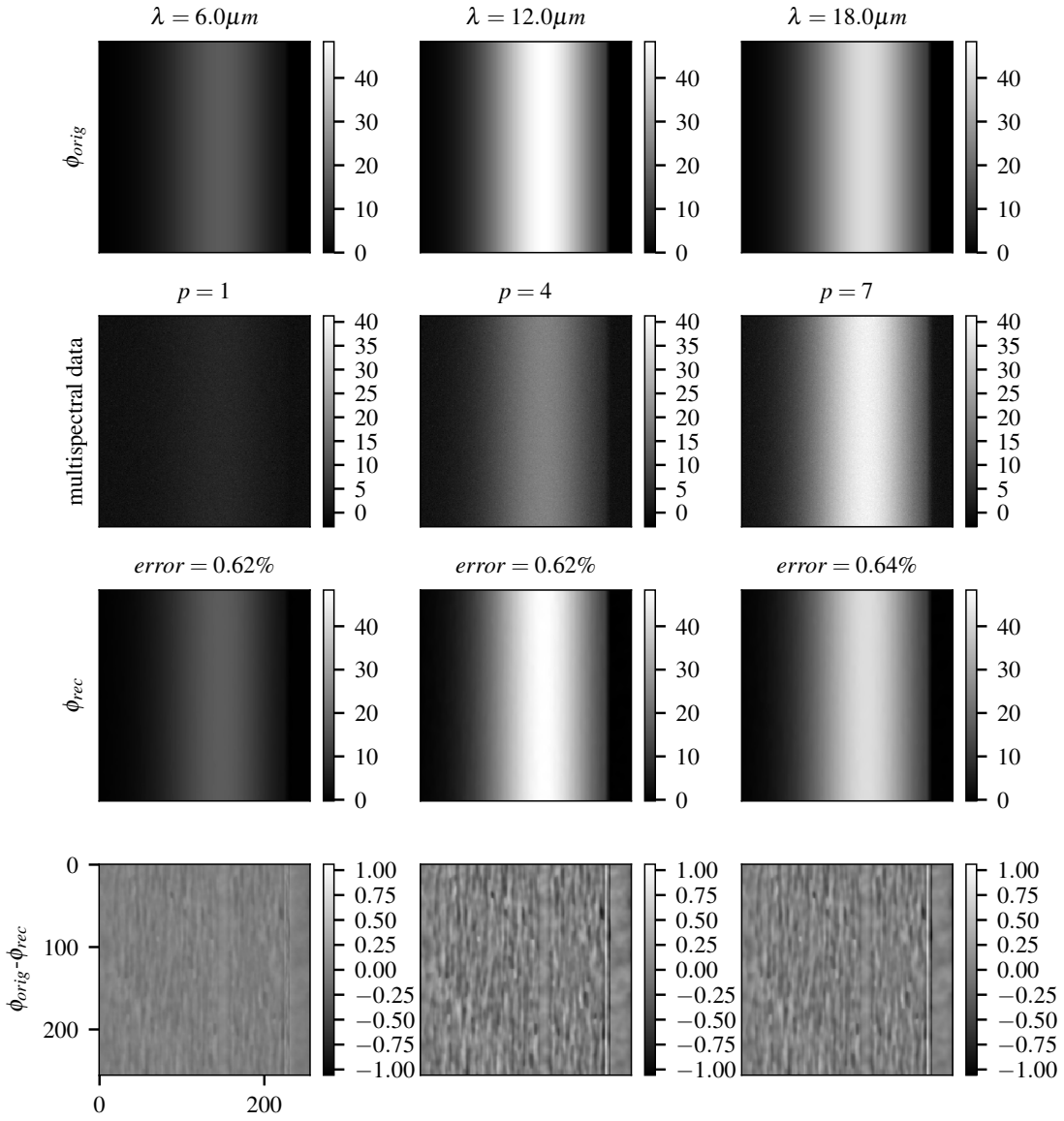


Figure 5.23: Same as in Figure 5.22 but here the reconstructed results of the *HorseHead nebula* are obtain using the l_2/l_1 algorithm.

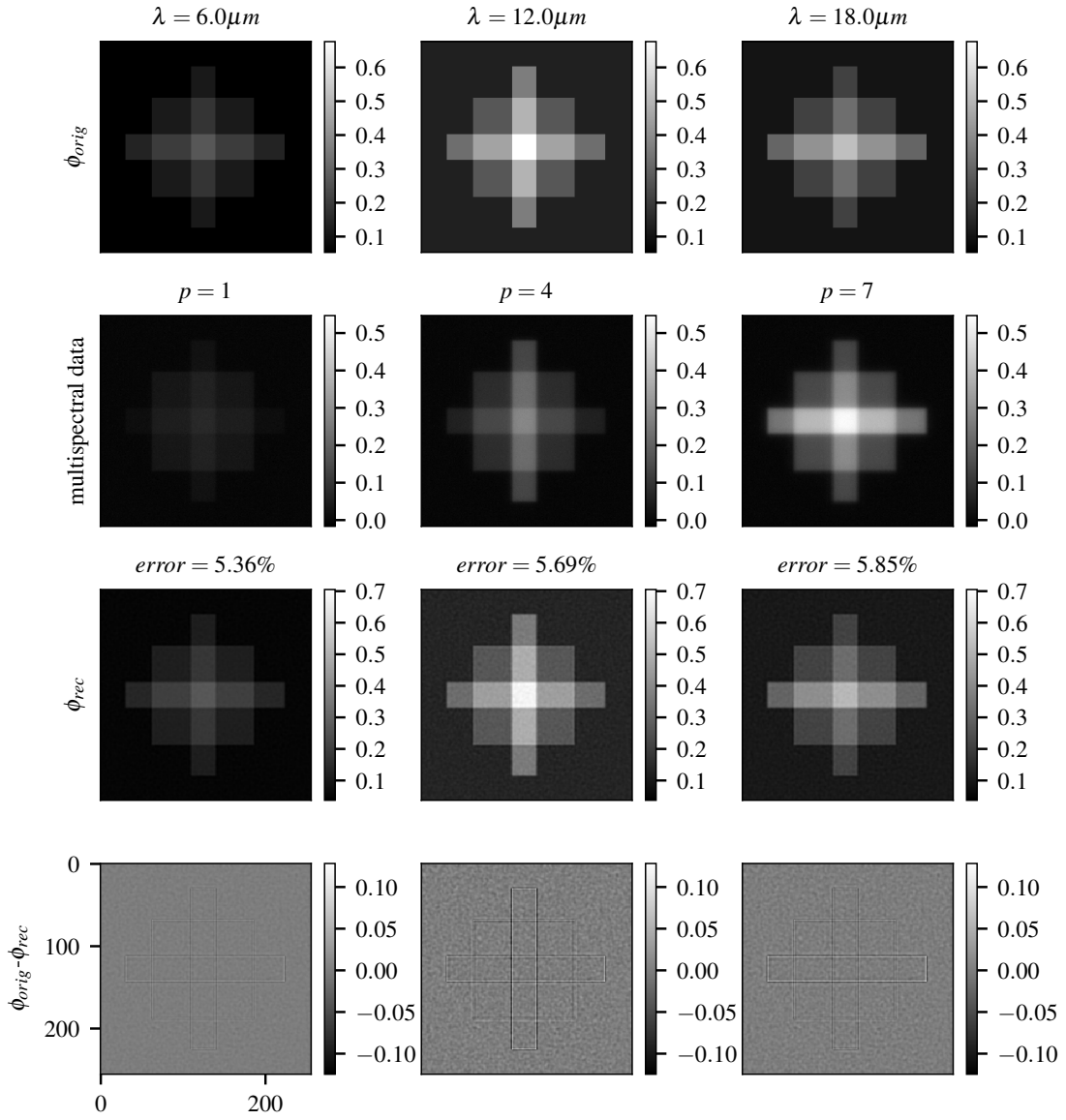


Figure 5.24: [1st row] Original Synthetic$_1$ object at 6, 12 and 18 μm. [2nd row] Simulated multispectral data with 30 dB associated to the bands that include wavelengths of the first row, i.e. $p = 1, 4, 7$. [3rd row] Reconstructed object at 6, 12 and 18 μm with the l_2 algorithm. [4th row] Difference between the original and reconstructed at 6, 12 and 18 μm..

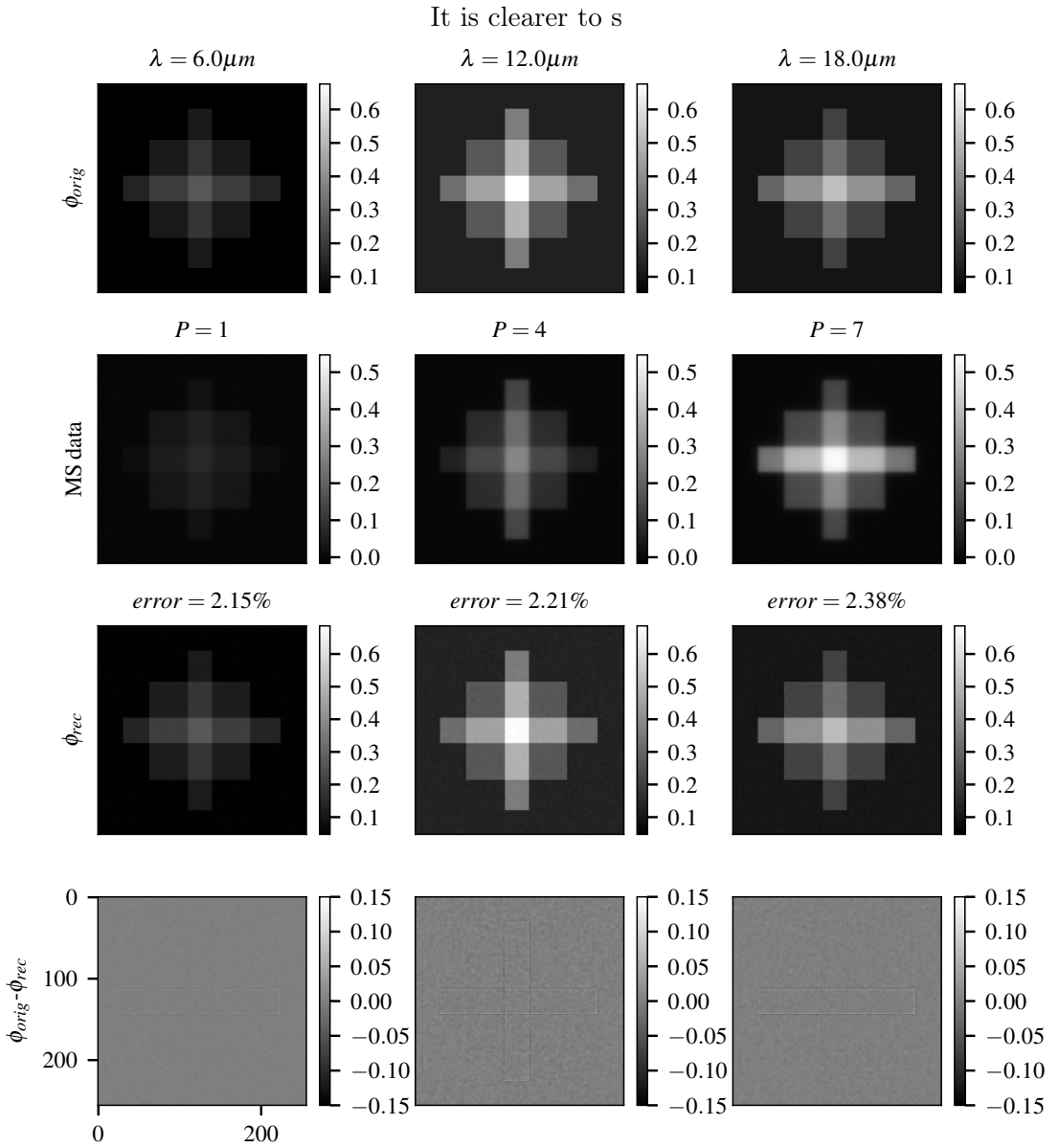


Figure 5.25: Same as in Figure 5.24 but here the reconstructed results of the Synthetic$_1$ are obtain using the l_2/l_1 algorithm.

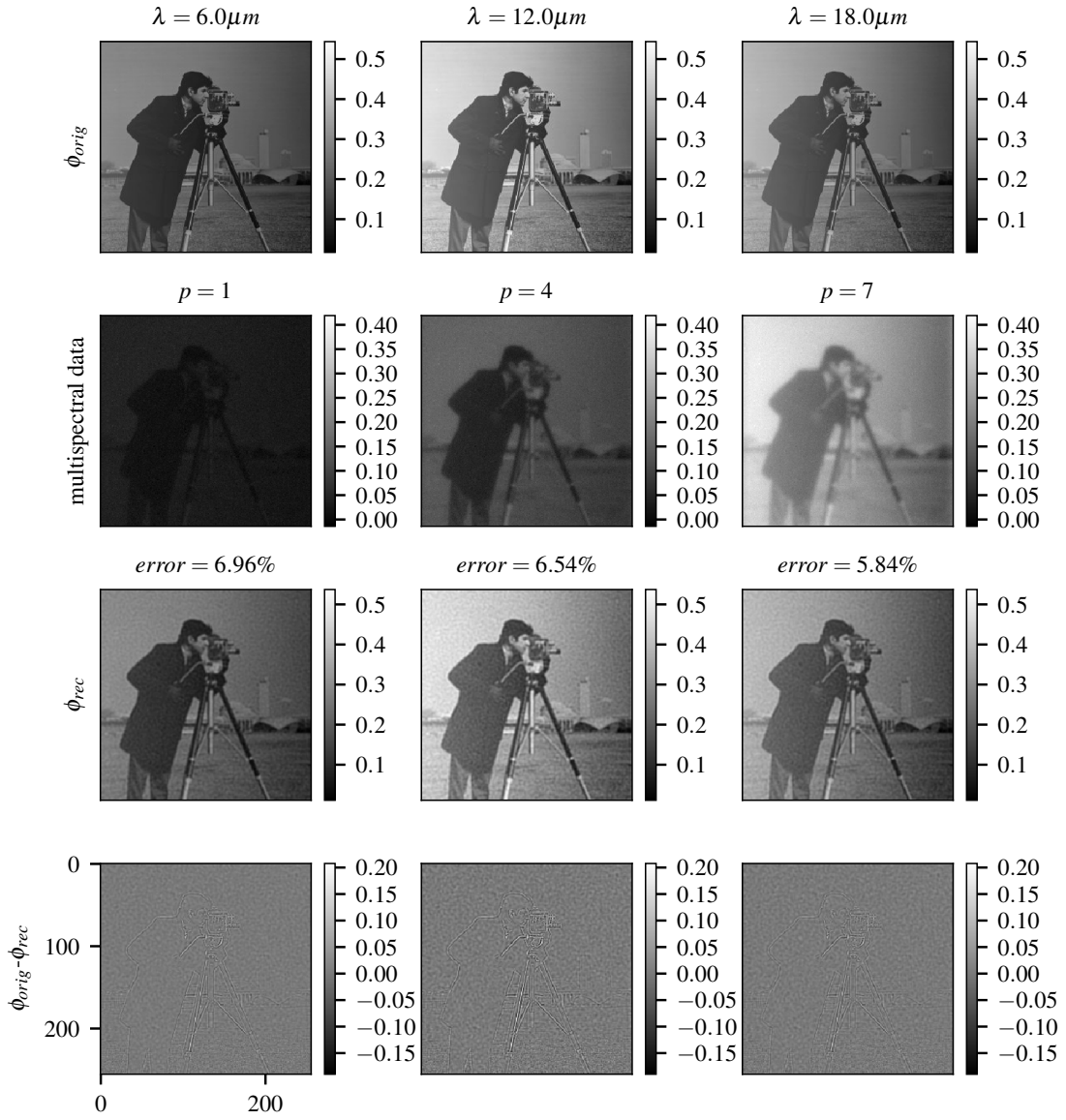


Figure 5.26: [1st row] Original Synthetic$_2$ object at 6, 12 and 18 μm. [2nd row] Simulated multispectral data with 30 dB associated to the bands that include wavelengths of the first row, i.e. $p = 1, 4, 7$. [3rd row] Reconstructed object at 6, 12 and 18 μm with the l_2 algorithm. [4th row] Difference between the original and reconstructed at 6, 12 and 18 μm.

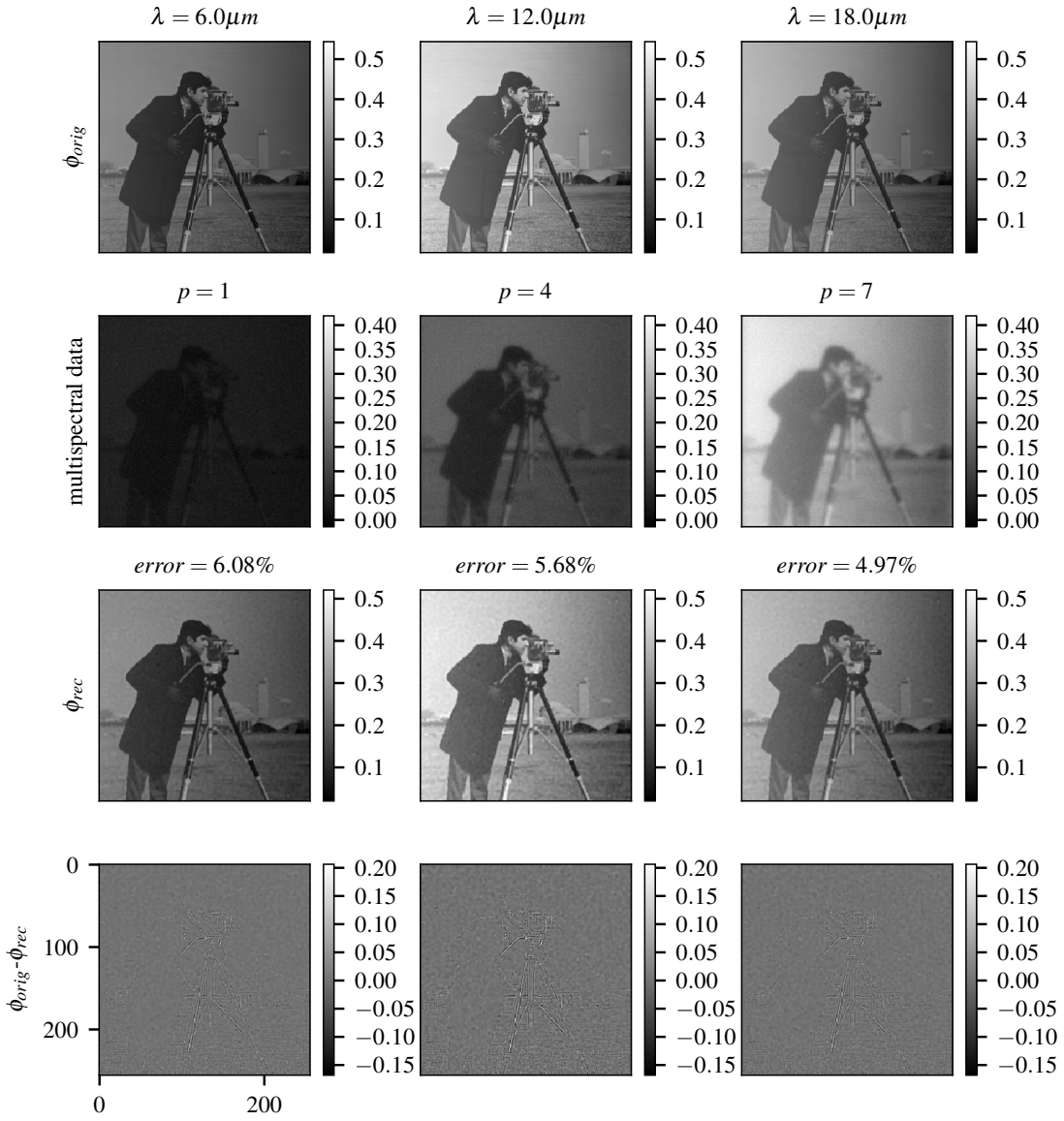


Figure 5.27: Same as in Figure 5.26 but here the reconstructed results of the Synthetic$_2$ are obtain using the l_2/l_1 algorithm.

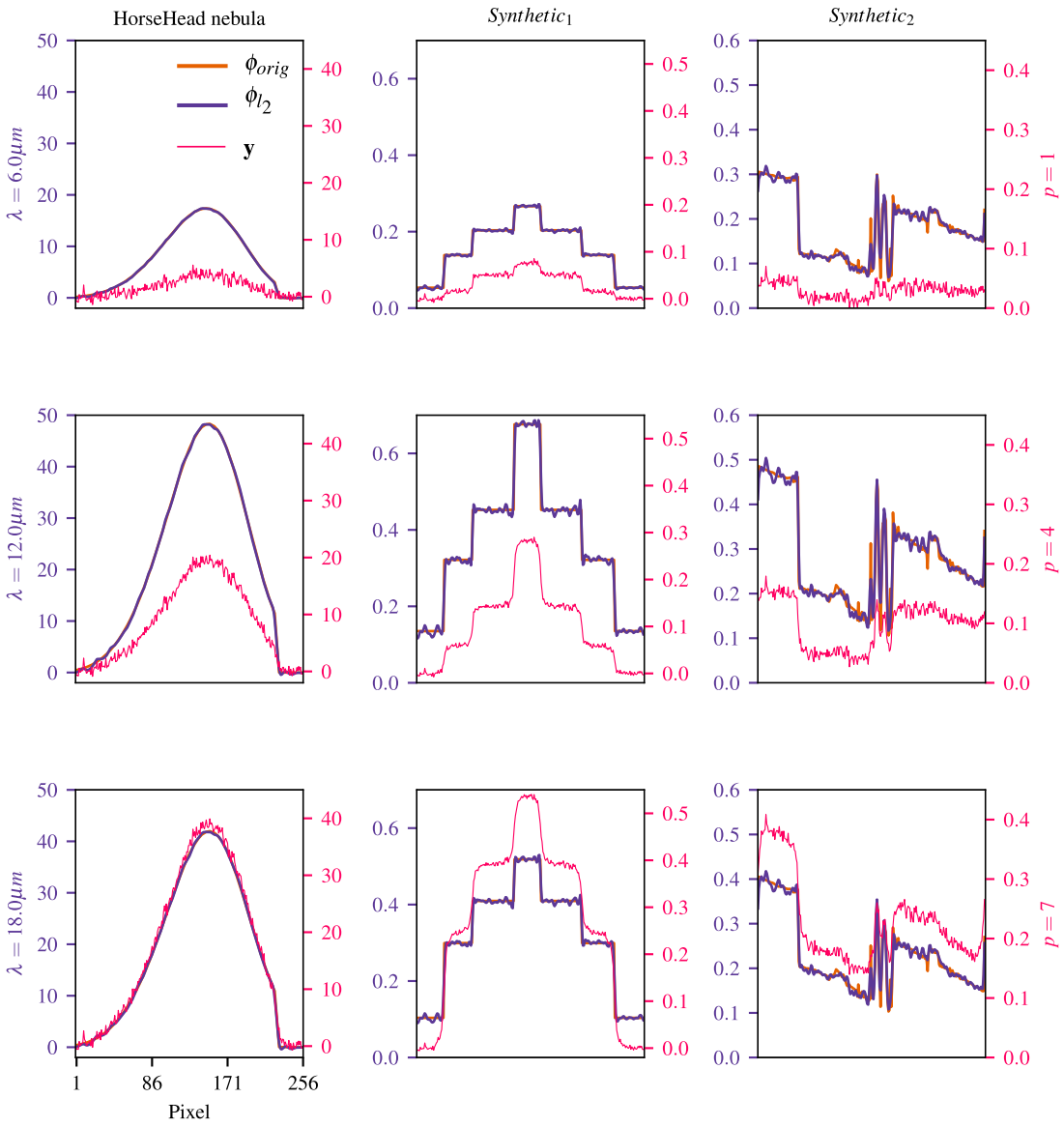


Figure 5.28: Illustration of the reconstruction results by displaying three spatial slices at 6, 12 and 18 μm selected from the middle row of the original object ϕ_{orig} and the reconstructed object using the quadratic regularization ϕ_{l_2}. [1st column] *HorseHead nebula* object. [2nd column] Synthetic$_1$ object. [3rd column] Synthetic$_2$ object. The multispectral data used for the three experiments were corrupted with a 30 dB white Gaussian noise. The magenta scale on the right of the plots corresponds to the electronic unit of the multispectral. The blue scale on the left of the plots corresponds to the physical unit of the spatio-spectral object (see Section 3.4).

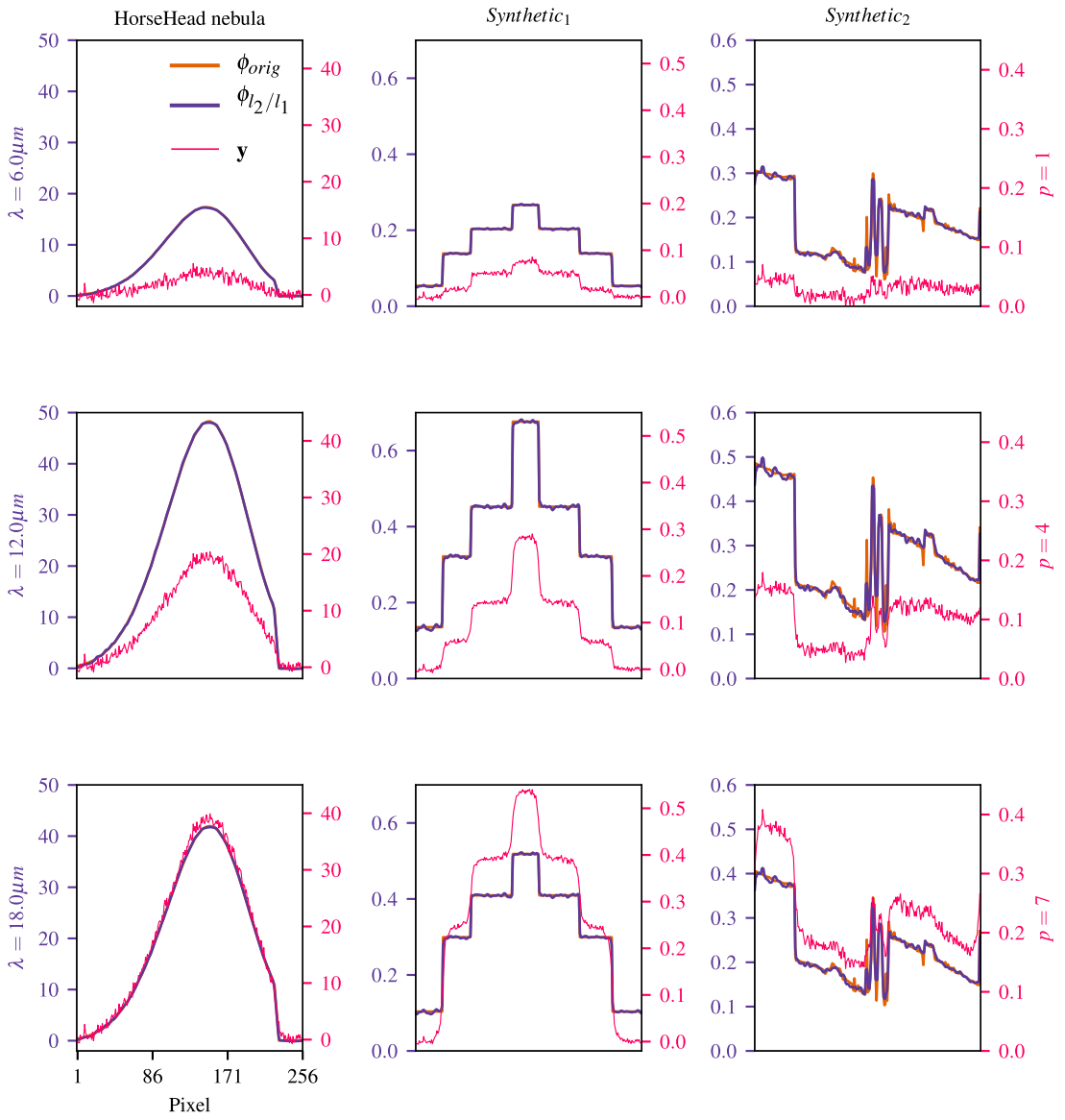


Figure 5.29: Same as Figure 5.28 but here the reconstrcted object is obtained using the l_2/l_1 algorithm.

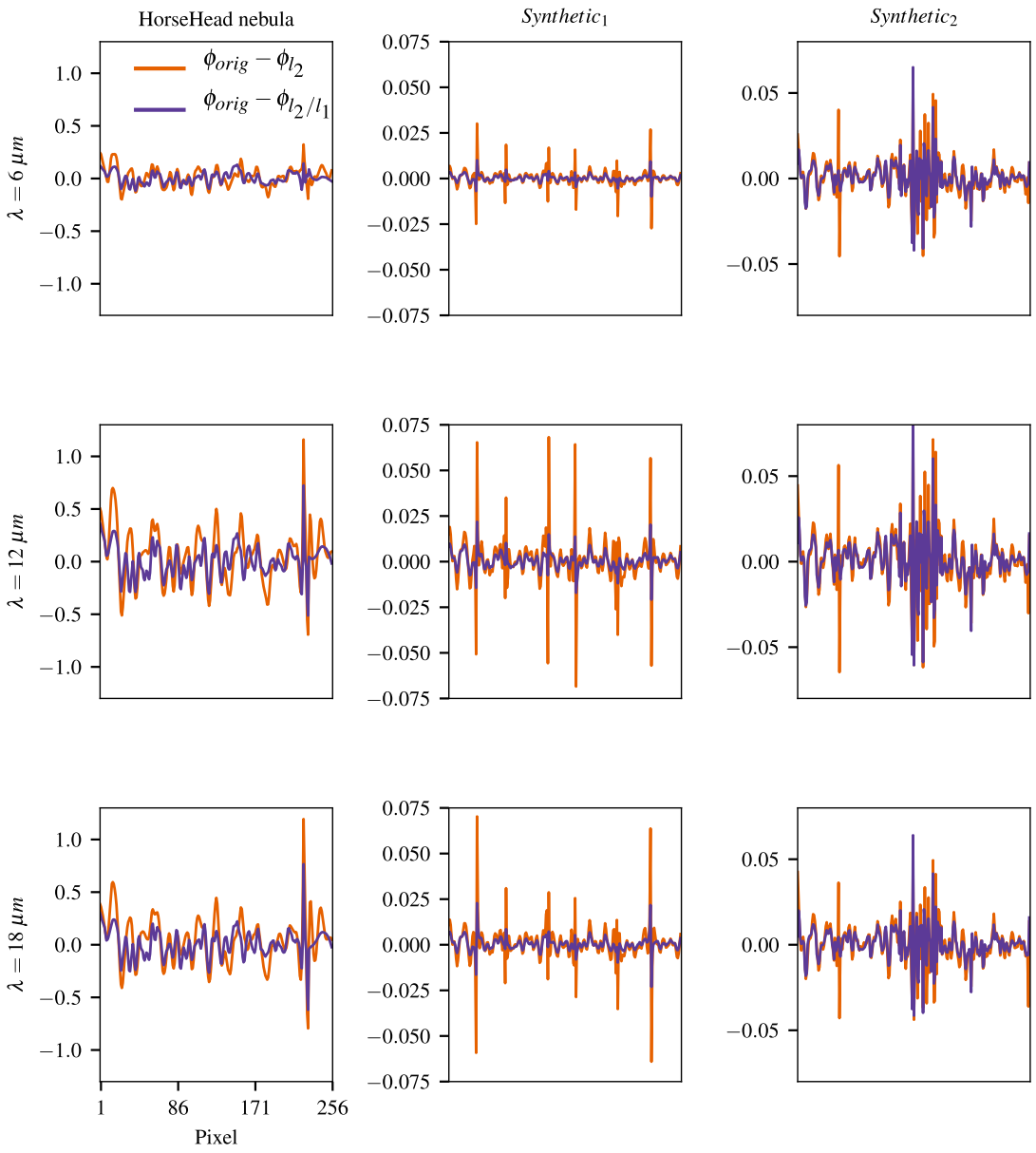
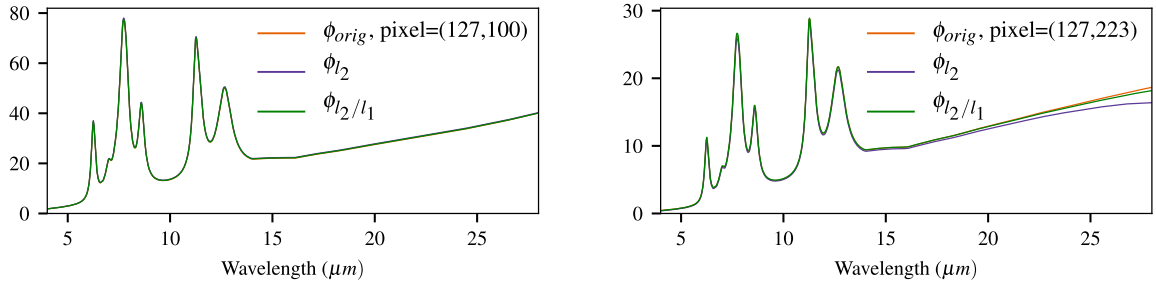
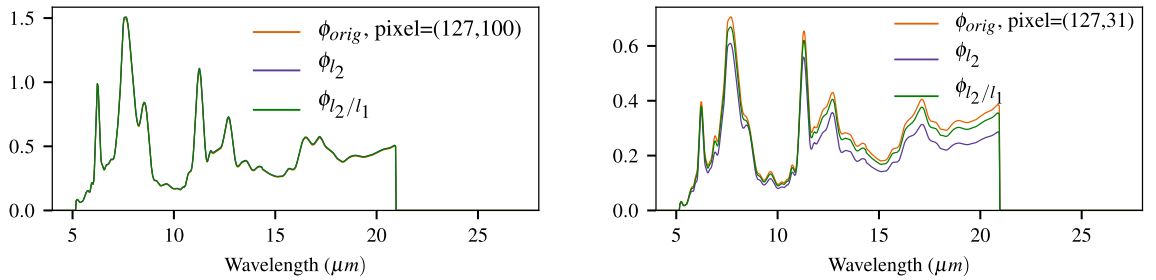


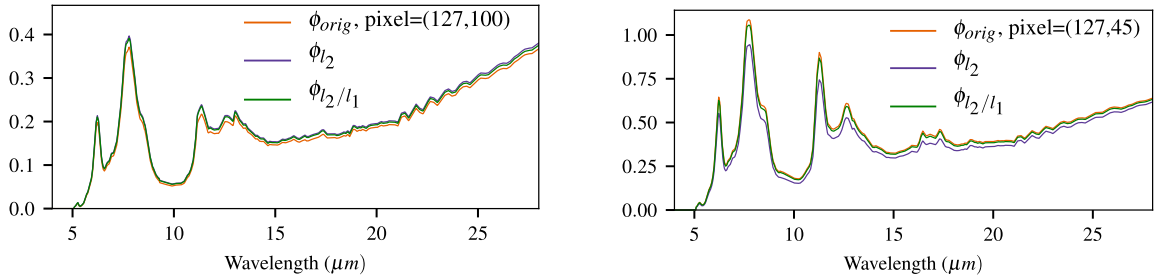
Figure 5.30: Display of the difference between slices of the original and the reconstructed object using the l_2 algorithm (in Figure 5.28) $\phi_{orig} - \phi_{l_2}$ (in blue), and using the l_2/l_1 algorithm (in Figure 5.29) $\phi_{orig} - \phi_{l_2/l_1}$ (in orange).



(a) *HorseHead nebula*



(b) Synthetic$_1$



(c) Synthetic$_2$

Figure 5.31: Comparison between spectral distributions of the original object ϕ_{orig}, and the reconstructed ones ϕ_{l_2} and ϕ_{l_2/l_1} for a single pixel. The first column corresponds spectral distributions of central spatial positions for the three objects. The second column corresponds to spectral distributions of spatial positions that are located on a high gradient as indicated on the legends.

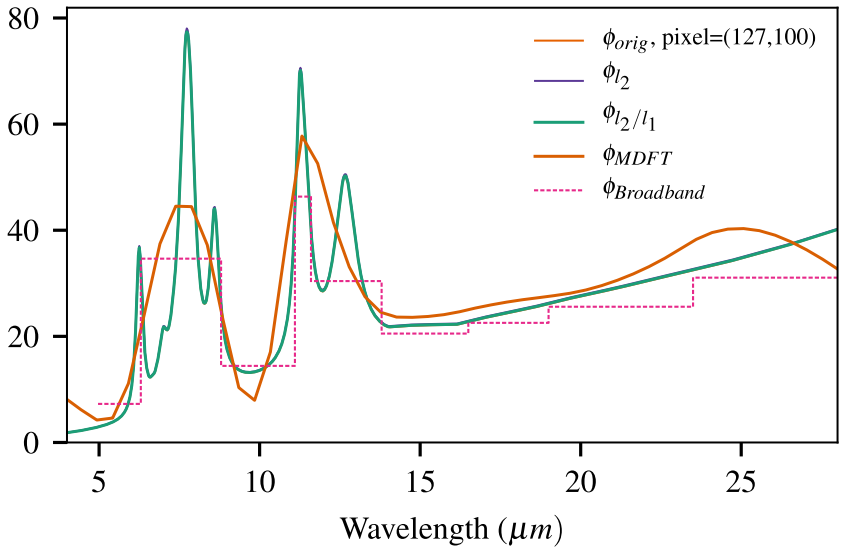
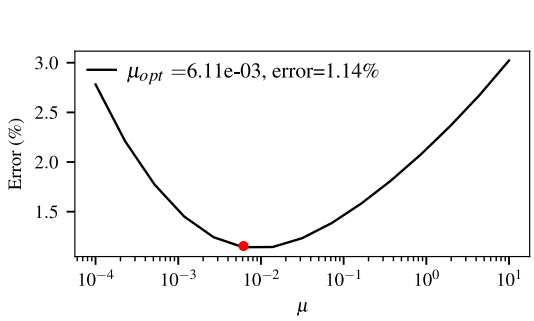


Figure 5.32: Comparison between different spectral distributions for a single pixel position $(127, 100)$ of the object *HorseHead nebula*. ϕ_{orig} is the original, ϕ_{l_2} and ϕ_{l_2/l_1} are the reconstruction proposed in this chapter using the l_2 and l_2/l_1 algorithm, ϕ_{MDFT} is the reconstruction result obtained the method proposed in Chapter 4, and $\phi_{Broadband}$ is the reconstruction result using 2D deconvolution band per band by considering broadband PSFs.

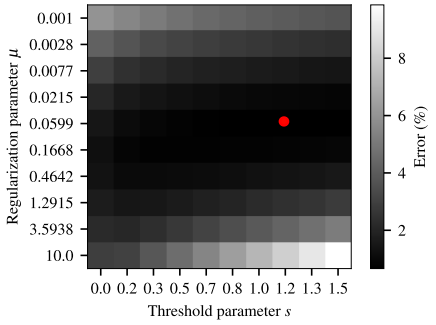
5.6.3.4 Influence of the Algorithm Parameters : μ and s

The influence of the parameters on the Algorithm 5.4 (l_2) and Algorithm 5.5 (l_2/l_1) are illustrated in Figure 5.33. As expected from the regularized least-squares method, we observe that both algorithms performances depend on the regularization parameter. Testing different values and evaluating the reconstruction error allows us to determine the optimal value that gives the best reconstruction quality of mixture coefficients.

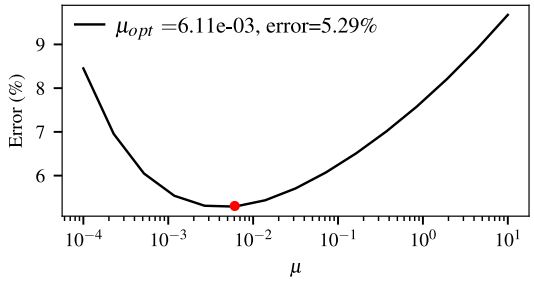
The optimal regularization parameters are depicted in red dots on the Figure 5.33. We see that it is enclosed in $\mu_{\text{opt}} \in]10^{-3}, 10^{-2}[$. In addition, we see as expected that l_2/l_1 the algorithm depends also on the threshold parameter s of Huber function, and $s_{\text{opt}} \in]0, 1.5[$ for the three experiments. For the *HorseHead nebula* object a value of $s = 1.17$ gives the best results. At the contrary, a very small value of s is necessary for objects Synthetic_1 and Synthetic_2. This is because of the spatial distributions of mixture coefficients, i.e the higher the gradient, the smaller s, and the lower the gradient the bigger s. As anticipated, the Huber function with s close to zero is equivalent to the l_1-norm function. It means the enforcement of an l_1-norm of the gradients, like the total variation, but with a lower computational cost thanks to the half-quadratic solution.



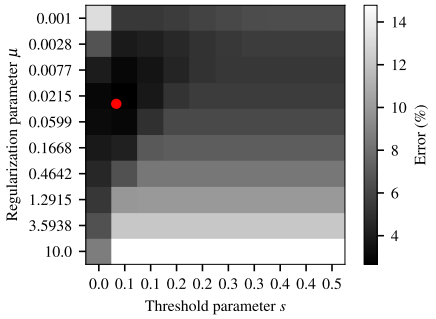
(a) l_2 algorithm



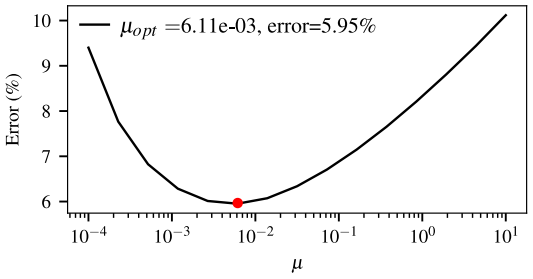
(b) l_2/l_1 algorithm



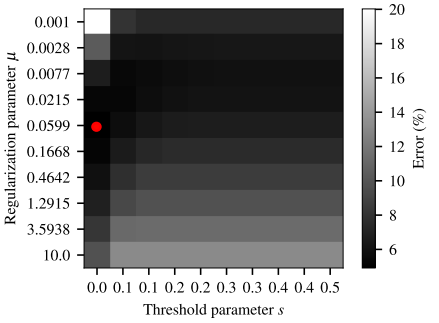
(c) l_2 algorithm



(d) l_2/l_1 algorithm



(e) l_2 algorithm



(f) l_2/l_1 algorithm

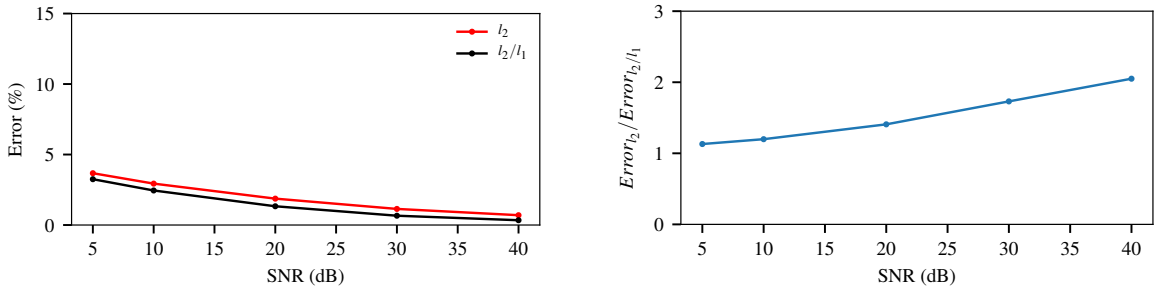
Figure 5.33: Influence of the parameters on [left column] the l_2 algorithm in 5.4 and [right column] the l_2/l_1 algorithm in 5.5, for three spatio-spectral objects: (a)-(b) *Horse-Head nebula*, (c)-(d) Synthetic$_1$ and (e)-(f) Synthetic$_2$. The red dot indicates the optimal parameters corresponding to the minimum reconstruction error.

5.6.3.5 Influence of the Noise Level

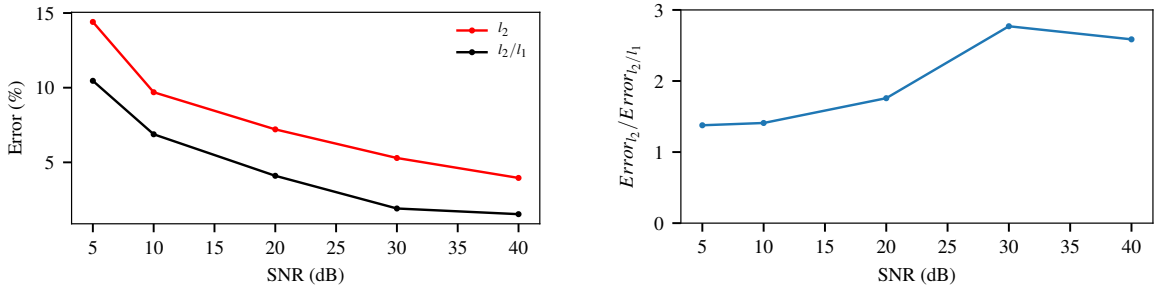
In order to evaluate the influence of the noise, we have added different levels of noise (through the SNR) to the multispectral data and then run the reconstruction algorithms. We depict in Figure 5.34 the reconstruction errors obtained for different SNR. As expected, the proposed algorithms are sensitive to the noise like any regularization method. We observe a decrease of reconstruction errors for both algorithms as the SNR increases.

From the first column of Figure 5.34, we obtained reconstruction error below 5% for an SNR $> 5\,\mathrm{dB}$ for the *HorseHead nebula* object in Figure 5.34.(a), 15% for the object $\mathrm{Synthetic}_1$ and an error below 16% for the object $\mathrm{Synthetic}_2$. An SNR $< 10\,\mathrm{dB}$ show an important increase of reconstruction error for both algorithms of two synthetic objects, i.e. object with high spatial gradients, Although we did not display the multispectral data at SNR $= 5\,\mathrm{dB}$, these later correspond to a complete domination of the noise.

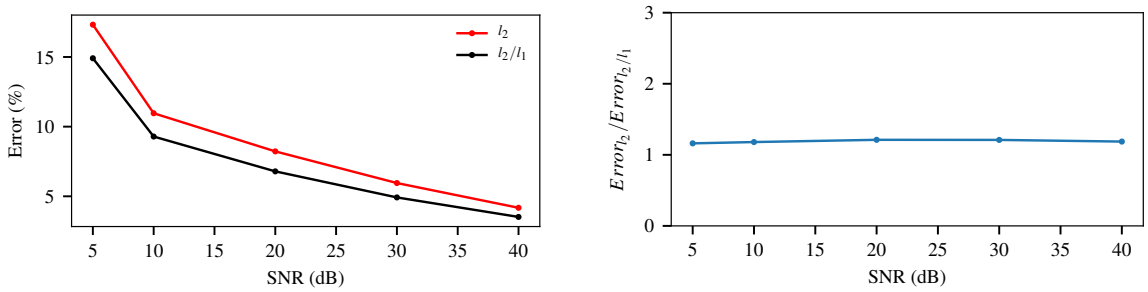
A comparison of the influence of the noise on the improvement factor of the l_2/l_1 algorithm over the l_2 algorithm is shown in the second column of the Figure 5.34. We observe that the improvement factor is not sensitive to the noise for the object *HorseHead nebula* and $\mathrm{Synthetic}_2$ objects. This is because the two objects are simulated with smooth mixture coefficients. However, the improvement factor increases with a high SNR for the object $\mathrm{Synthetic}_1$. This is because high gradients or sharp-edges in the multispectral data become less affected by the noise.



(a) *HorseHead nebula*



(b) Synthetic$_1$



(c) Synthetic$_2$

Figure 5.34: Influence of the noise level on the reconstruction of the three spatio-spectral objects using [left column] quadratic (l_2-norm) and half-quadratic (l_2/l_1-norm) regularization. [right column] Influence of the noise on the improvement factor of the l_2/l_1 algorithm over the l_2 algorithm.

5.7 Conclusion

This chapter presents a high-resolution reconstruction method of a 2D+λ spatio-spectral object observed by a multispectral imaging system from a few low-resolution data.

We first modeled the object with a set of M mixture coefficients and spectral components using a linear mixing model, with $M < P$. This allowed us to reduce the complexity of the reconstruction compared to Chapter 4. Next, we developed an accurate linear for-

ward model where multispectral data are the sum of 2D convolutions between mixture coefficients and observation matrix. The observation matrix preserves the spectral information of the object and accounts for the instrument parameters such as spectral-variant PSF, spectral response of the instrument. The reconstruction of the spatio-spectral object is done through the estimation of the mixture coefficients. We tackled this ill-posed problem using the regularized least-squares method and proposed two multichannel reconstruction methods. The first method uses a multichannel quadratic regularization in order to enforce spatial smoothness to the solution, while the second method uses a multichannel half-quadratic regularization to enforce high spatial gradient to the solution.

We were able to validate the performance of the reconstruction algorithm using three spatio-spectral objects, one model of an astrophysical object and two synthetic data. The execution time to compute the multichannel quadratic solution is about 1 seconds compared to 20 seconds (for 50 iterations) for the computation of the multichannel half-quadratic solution. In addition, an acceleration time of ten is obtained compared the conjugate gradient algorithm thanks to the implementation in the *Fourier* domain. The reconstruction results are very satisfactory, with reconstructed objects free from noise and blur. In all experiments we obtained a relative error below 5% for an SNR = 30 dB.

However, obviously the reconstruction error increases for a lower SNR, especially for objects with high spatial gradients. Our results are based on the hypothesis that a set of high-resolution spectral components are known which not always true.

Chapter 6

General Conclusion

6.1 Summary

The first stage of the data processing chain of our inverse problem framework is the instrument model described in Chapter 3. In this chapter, we focused in modeling the response of the multispectral imaging system, particularly the mid-infrared imager (MIRI) on board the James Webb Space Telescope (JWST). This multispectral imaging system is a diffraction-limited instrument. It consists of an optical system that collects the photon flux and focus them on the detector of the imager. The detector integrates and samples the multispectral data.

The developed instrument model aims to simulate the multispectral data based on the instrument response. It relates the continuous original object at the input of the imaging system to the discrete multispectral data at the output. The specificity of the instrument model is that it accounts the non-stationarity of the optical response (or PSF) of the JWST. It blurs the spatial distribution of the original spatio-spectral object according to the wavelength through a 2D spatial convolution. The PSF linearly depends on the wavelength, i.e. the larger the wavelength the wider the PSF as expected from the diffraction theory. Moreover, the blurred object is spectrally integrated over a set of P broad bands (or spectral responses) by the detector within a broad spectral range (wavelength ratio of 5) and a very low spectral resolving power ($\lambda/\triangle\lambda$) from 3.5 to 16. These degradations severely limit the spatial and spectral resolutions of the multispectral data and strongly challenge the data analysis.

Finally, the instrument model allows us to simulate the multispectral data by taking into account the response of the imaging system. It is also used to develop the next stages of the inverse problem framework such as the forward model and the reconstruction.

In Chapter 4 we tackled the reconstruction of a spatio-spectral object by jointly exploiting and processing the whole set of low-resolution multispectral data. Due to the lack of knowledge of the real PSF, a common approach in the literature consists of considering a broadband PSF to process the multispectral data issued from individual bands separately. However, this approach neglects the spectral variability of the PSF within a spectral band and introduces considerable inaccuracy to the model. Therefore we proposed to use simulated monochromatic PSFs and model the spectral distribution of the object with a piecewise linear function instead of using a non-accurate predefined spectrum.

Next, we have developed a forward model where the discrete stack of unknown spectral channels $\boldsymbol{x}$ is related to the discrete multispectral dataset $\boldsymbol{y}$ through the observation matrix $\boldsymbol{H}$. It represents degradations occurring within and between channels. The forward model is linear and is formulated by a sum of N_λ 2D convolutions between spectral channels and

the observation matrix, where each multispectral data depends on the whole unknown vector of the object $\boldsymbol{x}$. Moreover the spectral distribution is carried out by the channels and weighted by the blocks of the observation matrix. Concerning the reconstruction method, it is based on the regularized least-squares method where the solution is obtained as the minimizer of a cost function. In fact the problem is ill-posed because of the ill-conditioning of the matrix $\boldsymbol{H}^T\boldsymbol{H}$ and the under-determination of the forward model. To correct this ill-posedness we have added prior information about the solution to the cost function $\mathcal{J}(\boldsymbol{x})$ by penalizing the spatial and spectral gradients to enforce smoothness to the solution. This choice is mainly motivated for obtaining a differentiable cost function which allows us to implement a fast algorithm.

The obtained multichannel quadratic solution is linear and consists mainly of inverting the Hessian matrix $\boldsymbol{Q}$. However, this matrix is very large and cannot be inverted in the spatio-spectral domain. Thus we proposed to compute the solution with two different approaches. The first one consists of using an iterative optimization algorithm such as the conjugate gradient, thus no matrix inversion is required. The second one is based on the matrix inversion using diagonalization in the Fourier domain by considering the circulant approximation of the block matrices of the Hessian matrix.

Simulated data of the MIRI imager of the JWST have been used to highlight the complexity of the instrument response, and both algorithms have been tested on a simplified model of the astrophysical object *HorseHead nebula* of size $1000 \times 256 \times 256$ pixels. The exact quadratic solution is computed in the Fourier domain. It is ten times faster than the iterative one. Even though a significant increase in spatial and spectral resolutions is obtained compared to the broadband method, the proposed method has some shortcomings originating mainly from the multispectral dataset and the object model. Firstly, it is difficult to choose the number of channels N_λ; a high number implies a high dimensional observation matrix but increases the between channels degradations. Secondly, the spectral bands of the JWST/MIRI imager are broad and their number is limited (less than 9), meaning that the spectral resolution of the multispectral data is very poor. Therefore additional prior knowledge about the spectral distribution of the object is required in order to improve the obtained results.

In Chapter 5 we aimed to overcome the encountered limitations of Chapter 4 such as the lack of spectral information in the multispectral dataset about the original object. The idea is based on using a linear mixing model to model the spatial and spectral distributions of the object. We proposed a hypothesis wherein we considered the set of high-resolution spectral components or templates to be known. This allows us to reduce the complexity of the previous problem and to represent more accurately the object. Therefore the spatio-spectral object is modeled by a sum of M high-resolution spectral components (supposed to be known) weighted by M mixture coefficients for each spatial position.

Same as in Chapter 4 we developed a linear forward model formulated by a sum of M discrete convolutions between mixture coefficients and observation matrix. Here the spectral information is *completely embedded in the blocks of the observation matrix* $\boldsymbol{H}$. With respect to the reconstruction of the spatio-spectral object, it is achieved through the estimation of the mixture coefficients. The reconstruction method is based on the regularized least-squares methods where the solution is obtained as a minimizer of a con-

vex cost function. We considered two types of regularization in order to correct the ill-posedness of the problem caused by the ill-conditioning of the matrix $\boldsymbol{H}^T\boldsymbol{H}$ and the over-determination of the forward model. The first one enforces spatial smoothness to the solution, whereas the second one enforces a high spatial gradient for the purpose of reconstructing the object with further spatial distribution compared to Chapter 4. This is assured by using multichannel quadratic regularization and multichannel half-quadratic regularization, respectively. The quadratic solution $\hat{\boldsymbol{x}}_{l_2}$ is linear and consists mainly of inverting the Hessian matrix $\boldsymbol{Q}_{l_2}$ in the Fourier domain, whereas the half-quadratic solution $\hat{\boldsymbol{x}}_{l_2/l_1}$ is updated iteratively with auxiliary variables that catch high spatial gradients and prevent the algorithm from smoothing them.

The two algorithms l_2 and l_2/l_1 have been implemented on simulated data using three objects having different spatial and spectral distributions. The reconstruction results of both algorithms are very satisfactory. The three reconstructed objects show a clear denoising and deconvolution of the multispectral data. We obtained a relative error that does not exceed 5% for a noise level of SNR = 30 dB. The algorithm l_2/l_1 gives better reconstruction results than the l_2 algorithm especially for objects with high spatial gradients. It computes the solution in 20 seconds (for 50 iterations) compared to 1 second for the l_2 algorithm thanks to the matrix inversion in the Fourier domain. On the other hand, the shortcomings of the proposed reconstruction methods mainly come from the hypothesis that high-resolution spectral components are known which is not always valid.

6.1.1 Papers Published in International Conferences

1. [Hadj-Youcef *et al.* 2018]: M. A. Hadj-Youcef, F. Orieux, A. Fraysse and A. Abergel. Spatio-Spectral Multichannel Reconstruction from few Low-Resolution Multispectral Data. In 2018 26th European Signal Processing Conference (EUSIPCO), Sept 2018.

2. [Hadj-Youcef *et al.* 2017a]: M. A. Hadj-Youcef, F. Orieux, A. Fraysse and A. Abergel. Restoration from multispectral blurred data with non-stationary instrument response. In 2017 25th European Signal Processing Conference (EUSIPCO), pages 503–507, Aug 2017.

6.1.2 Papers Published in National Conference

1. [Hadj-Youcef *et al.* 2017b]: Mohamed Amine Hadj-Youcef, François Orieux, Aurélia Fraysse and Alain Abergel. Restauration d'objets astrophysiques à partir de données multispectrales floues et d'une réponse instrument non-stationnaire. In 26eme Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017, Juan Les Pins, France, September 2017.

6.1.3 Articles Published in Journal

1. [Boucaud *et al.* 2016]: Boucaud, A., Bocchio, M., Abergel, A., Orieux, F., Dole, H., Hadj-Youcef, M. A. (2016). Convolution kernels for multi-wavelength imaging.

Astronomy & Astrophysics, 596, A63.

6.2 Perspectives

Several aspects are to be improved in the contributions of this thesis. The instrument model developed in Chapter 3 was used to simulate observations of the JWST for the imager MIRI. In fact, an evaluation of this model is necessary, for instance, by performing a comparison with the observations taken from another simulator, such as the MIRI Simulator software (MIRISim) released in April 2018[1]. This tool is developed by the MIRI european team, and in France in a collaboration between DAP/AIM/ CEA (Commissariat d'Energie Atomique)[2] and the IAS (Institut d'Astrophysique Spatiale)[3] with a purpose to simulate observations of the imager or spectrometer of MIRI.

In Chapter 4, the object model depends on a parameter N_λ which set the number of channels to represent the object. This parameter is very important for object representation and reconstruction. It defines the dimension of the problem, i.e. the number of unknown, and has a direct link with the spectral sampling of the reconstructed object. It also controls the spectral regularization added to stabilize the regularized least-squares solution as well as its computation. Thus, this parameter needs to be estimated, for instance, according to the number of multispectral data and the width of the spectral bands.

We have shown in Chapter 5 that considering a set of a few high-resolution spectral components helps a lot in the reconstruction of a high-resolution spatio-spectral object from a small number of low-resolution multispectral data. For instance, these spectral components can be extracted from hyperspectral data using source separation techniques such as the principal components analysis. The instrument MIRI of the JWST can be exploited perfectly in this manner because it delivers two types of data : (1) the imager delivers multispectral data, i.e. data with high spatial-resolution and limited spectral-resolution, and (2) the spectrometer delivers hyperspectral data, i.e. data with high spectral-resolution and limited spatial-resolution. This allows us to make use of complementary advantages of both data in order to reconstruct a spatio-spectral object with high-resolution. Moreover, the threshold parameter s of the Huber function (in the l_2/l_1 algorithm) depends on the gradients of the object. Thus, finding a way to estimate it will reduce from the data will reduce the number of parameters of the algorithm.

Finally, tests and validations of the reconstruction algorithms in Chapters 4 and 5 have been performed on multispectral data that were simulated using the developed instrument model. Thus it is necessary to evaluate the performance of our algorithms and analyze how they perform on real multispectral data. We expect this evaluation to reveal further strong/weak points of the algorithm which are needed to improve. For instance, we intend to consider a noise other than white Gaussian (e.g. Poissonian noise) and account for

[1] https://jwst.fr/wp/?p=2172
[2] http://www.cea.fr/
[3] https://www.ias.u-psud.fr/

photon noise which varies according to the wavelength.

Appendix A

Author's Publications

Spatio-Spectral Multichannel Reconstruction from few Low-Resolution Multispectral Data

M.A. Hadj-Youcef*†, F. Orieux*†, A. Fraysse*, A. Abergel†
* Laboratoire des Signaux et Systèmes, Univ. Paris-Sud, CNRS, CentraleSupélec, Université Paris-Saclay
3 rue Joliot-Curie, 91 192 Gif-sur-Yvette, France
Email: amine.hadjyoucef@ias.u-psud.fr
† Institut d'Astrophysique Spatiale, CNRS, UMR 8617, Univ. Paris-Sud, Université Paris-Saclay
Univ. Paris-Sud, 91405 Orsay, France

***Abstract*—This paper deals with the reconstruction of a 3-D spatio-spectral object observed by a multispectral imaging system, where the original object is blurred with a spectral-variant PSF (Point Spread Function) and integrated over few broad spectral bands. In order to tackle this ill-posed problem, we propose a linear forward model that accounts for direct (or auto) channels and between (or cross) channels degradation, by modeling the imaging system response and the spectral distribution of the object with a piecewise linear function. Reconstruction based on regularization method is proposed, by enforcing spatial and spectral smoothness of the object. We test our approach on simulated data of the Mid-InfraRed Instrument (MIRI) Imager of the James Webb Space Telescope (JWST). Results on simulated multispectral data show a significant improvement over the conventional multichannel method.**

***Index Terms*—Inverse problems, Image reconstruction, Deconvolution, System modeling, Multispectral restoration**

I. Introduction

Multispectral imaging systems are used in many fields, e.g. astrophysics [1], remote sensing [2], medicine [3] or microscopy [4]. This paper deals with the inverse problem of joint restoration. Our goal is to reconstruct a discrete 3-D spatio-spectral object from a small number of 2-D Multi-Spectral (MS) observed data when this continuous 3-D object is degraded by the instrument that suffers the diffraction due to the limited size of the optical system. This physical degradation affects its spatial resolution (in the form of blur) accordingly to the wavelength. Moreover, before its spatial sampling, the blurred object is integrated by the detector over the different wide spectral bands, which results in low spectral resolution multispectral data. Therefore, the multispectral data are severely degraded and contain limited spectral information about the original object.

Multichannel restoration has been extensively studied in the literature. Multichannel forward models have been proposed in [5], [6], where the system response is a block-diagonal matrix with circulant blocks. For instance, [7], [8], [9] address multichannel 2-D deconvolution problem for hyperspectral image deconvolution. They take into account the within-channel degradation, but not the between channel (or cross-channel) degradation. Hence, this approach is not suitable for MS imaging, especially if spectral bands are broad and overlapping, which implies a strong correlation between channels. In [10] a model is proposed that reduces these limitations since the system response is represented by a block matrix corresponding to within and between channel degradations. However, this model is mostly used when the number of channels and observations is the same, e.g. color image restoration [11], [12], [13].

In this paper we propose a multispectral forward model that accounts for within and between channels degradation (or auto and cross-channel), where (1) the number of multispectral data is much lower than the number of spectral channels and (2) a set of low-resolution multispectral data are degraded by a spectral-variant PSF and integrated over broad spectral bands. Reconstruction of a spatio-spectral object is performed using regularization method, by accounting for spatial and spectral quadratic regularization. Simulated results are provided with a comparison to multichannel 2-D deconvolution for an application to the MIRI Imager on board the JWST[1].

This paper is organized as follows. In Sec. II we present the problem formulation. The imaging system response and the forward model are described in Sec. III. The reconstruction method is presented in Sec. IV. Simulation and results are presented in Sec. V including a brief description of the JWST/MIRI Imager. Conclusions and perspectives are provided in Sec. VI.

II. Problem formulation

The general form of the multispectral problem we are considering is the one proposed in [10], [11]. It follows the discrete linear forward model

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{x} = \left[\boldsymbol{x}^{(1)T}, \boldsymbol{x}^{(2)T}, \ldots, \boldsymbol{x}^{(M)T}\right]^T \in \mathbb{R}^{MN_kN_l}$ is the stack of M spectral channels represented in a vector form, each channel containing $N_k \times N_l$ pixels, where $x_{k,l}^{(m)}$ denotes the $(k,l)^{th}$ spatial position in the $(m)^{th}$ wavelength. The vector $\boldsymbol{y} = \left[\boldsymbol{y}^{(1)T}, \boldsymbol{y}^{(2)T}, \ldots, \boldsymbol{y}^{(P)T}\right]^T \in \mathbb{R}^{PN_iN_j}$ is the stack of multispectral observed data acquired via P broad spectral bands of the imaging system. $\boldsymbol{n} = \left[\boldsymbol{n}^{(1)T}, \boldsymbol{n}^{(2)T}, \ldots, \boldsymbol{n}^{(P)T}\right]^T \in \mathbb{R}^{PN_iN_j}$ represents an additive

[1] https://jwst.nasa.gov/

unknown noise. The full system response is a PN_iN_j by MN_kN_l block matrix,

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}^{1,1} & \boldsymbol{H}^{1,2} & \dots & \boldsymbol{H}^{1,M} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}^{P,1} & \boldsymbol{H}^{P,2} & \dots & \boldsymbol{H}^{P,M} \end{bmatrix}, \quad P \ll M, \quad (2)$$

defined by a set of $P \times M$ Toeplitz sub-matrices $\boldsymbol{H}^{p,m} \in \mathbb{R}^{N_iN_j \times N_kN_l}$, which are approximated for computational ease by circulant blocks.

However, all sub-matrices are ill-conditioned, meaning that $\boldsymbol{H}$ is also ill-conditioned, which leads to an ill-posed problem. Each multispectral observation $\boldsymbol{y}^{(p)}$ depends on all spectral channels through the blocks of $\boldsymbol{H}$, where $\boldsymbol{H}^{p,t}, t = m$ represents the direct (or auto) observation, whereas the block $\boldsymbol{H}^{p,t}, t \neq m$ accounts for between (or cross) degradation occurring between channels. We are interested in the particular case where we have few low-resolution multispectral observed data compared to spectral channels, i.e. $P \ll M$, which means there is a lack of spectral information in the data. For instance, for the multispectral data observed by the JWST/MIRI imager, $P = 9$ and $M = 1000$, and the size of $\boldsymbol{H}$ is 9×256^2 by 1000×256^2 for a 256×256 pixel detector.

III. Modeling of within and between channel degradations

A. Object Model

We first define the 3-D spatio-spectral object of interest with $\phi(\alpha, \beta, \lambda) : \mathbb{R}^3 \to \mathbb{R}$, having two spatial parameters $(\alpha, \beta) \in \mathbb{R}^2$ and one spectral parameter $\lambda \in \mathbb{R}_+$. In order to handle the lack of spectral information in the data, we propose to perform approximation of the M spectral channels by M' channels, with $M' < M$, by modeling the object spectral distribution with a piecewise linear function [14]. Moreover, we are interested in reconstructing a discrete version of the object, hence, we define two basis functions for spatial and spectral discretization, $b_s()$ and $b_\lambda()$, respectively. They are defined upon two grids, $\mathcal{G}_s = \{\alpha_k, \beta_l\}_{k,l=1}^{N_k,N_l}$ and $\mathcal{G}_\lambda = \{\lambda^{(m)}\}_{m=1}^{M'}$, respectively. Thus, the object is modeled by

$$\phi(\alpha, \beta, \lambda) = \sum_{m=1}^{M'} \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} x_{k,l}^{(m)} \, b_s^{(k,l)}(\alpha, \beta) b_\lambda^{(m)}(\lambda), \quad (3)$$

where $b_\lambda^{(m)}(\lambda)$ is a uniform piecewise linear function, for instance first-order B-spline function [15]. The parameter M' compromises between the sharpness of the spectral sampling of the modeled object and the unknown spectral channels to reconstruct $x^{(m)}$.

B. Imaging System Response

In this section we provide the multispectral imaging system response by establishing an imaging system model that relates the input to the output. The instrument we are considering is composed of an optical system and a detector. Due to light diffraction of ϕ, the optical system response is modeled by a 2-D spatial convolution [16] with a spectral variant optical response, known as Point Spread Function (PSF) $h(\alpha, \beta, \lambda)$. This blurs the object accordingly to the wavelength and limits its spatial resolution (as illustrated in Sec. V). The blurred object is integrated over broad spectral bands $\tau_p(\lambda)$ and sampled pixel-by-pixel on the 2-D detector grid, $\mathcal{G}_{\text{samp}} = \{\alpha_i, \beta_j\}_{i,j=1}^{N_i,N_j}$, with α_i, β_j being the 2-D angular positions of pixels (i, j) and N_i, N_j are the total number of pixel according dimensions α and β. We introduce a basis function $b_{\text{samp}}^{(i,j)}(\alpha, \beta)$ to carry out spatial sampling. It is defined on the pixel sensitive surface Ω_{pix}. Moreover, a noise term $n_{i,j}^{(p)}$ is added for each pixel (i, j) and band p, e.g. readout noise of the detector. Finally, the imaging system model is given by

$$y_{i,j}^{(p)} = \int_{\mathbb{R}_+} \tau_p(\lambda) \Bigg(\iint_{\Omega_{\text{pix}}} \Bigg(\iint_{\mathbb{R}^2} \phi(\alpha', \beta', \lambda) \quad (4)$$
$$h(\alpha - \alpha', \beta - \beta', \lambda) d\alpha' d\beta' \Bigg) b_{\text{samp}}^{(i,j)}(\alpha, \beta) d\alpha d\beta \Bigg) d\lambda + n_{i,j}^{(p)}$$

this model links the 3-D continuous input $\phi(\alpha, \beta, \lambda)$ to the 2-D discrete output $y_{i,j}^{(p)}$ through a complex instrument response, which includes spectral windowing and five sums, two for spatial 2-D convolutions, two for spatial sums and one for spectral integration. Note that the above model does not include any non-ideal characteristics of the detector, which are assumed to be corrected upstream.

C. Forward Model and definition of $\boldsymbol{H}^{p,m}$

The discrete forward model links the discrete spectral channels to the discrete multispectral data. It is obtained by substituting equation (3) in (4). This yields

$$y_{i,j}^{(p)} = \sum_{m=1}^{M'} \sum_{k=1}^{N_k} \sum_{l=1}^{N_l} H_{i,j;k,l}^{p,m} \, x_{k,l}^{(m)} + n_{i,j}^{(p)}, \quad (5)$$

with

$$H_{i,j;k,l}^{p,m} = \int_{\mathbb{R}_+} \tau_p(\lambda) b_\lambda^{(m)}(\lambda) \Bigg(\iint_{\Omega_{\text{pix}}} \Big(\iint_{\mathbb{R}^2} b_s^{(k,l)}(\alpha', \beta') \quad (6)$$
$$h(\alpha - \alpha', \beta - \beta', \lambda) d\alpha' d\beta' \Big) b_{\text{samp}}^{(i,j)}(\alpha, \beta) d\alpha d\beta \Bigg) d\lambda.$$

In addition, we consider for instance a rectangular impulse function [17] for the sampling function. i.e. $b_{\text{samp}}^{(i,j)}(\alpha, \beta) = \frac{1}{\triangle\alpha\triangle\beta} \Pi_{\triangle\alpha,\triangle\beta}(\alpha - \alpha_i, \beta - \beta_j)$, with $\triangle\alpha, \triangle\beta$ are the sampling steps according to dimensions α and β, respectively. Thus, the system response becomes a convolution matrix $H_{i,j;k,l}^{p,m} = H_{i-k;j-l}^{p,m}$. Therefore, the vector-matrix representation of (5) is

$$\boldsymbol{y}^{(p)} = \sum_{m=1}^{M'} \boldsymbol{H}^{p,m} \boldsymbol{x}^{(m)} + \boldsymbol{n}^{(p)}, \quad (7)$$

where the $p\text{-}th$ multispectral data $\boldsymbol{y}^{(p)}$ is a sum of M' discrete 2-D spatial convolutions between spectral channels and convolution matrices $\boldsymbol{H}^{p,m}$ (blocks of the matrix $\boldsymbol{H}$ in (2)). Thus, it accounts for within and between channels degradation. The discrete multispectral forward model with the

full imaging system $\boldsymbol{H}$ response takes the form in (1). Without loss of generality, we consider $N_i = N_j = N_k = N_l = N$.

IV. Reconstruction

The reconstruction of the object of interest ϕ relies on the reconstruction of its spectral channels $\boldsymbol{x}^{(m)} : m = 1, 2, \ldots, M'$ using a regularization method. The solution $\hat{\boldsymbol{x}}$ is obtained by minimizing an objective function $\mathcal{J}(\boldsymbol{x})$,

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \ \{\mathcal{J}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{x}, \boldsymbol{y})\}, \tag{8}$$

where $\mathcal{Q}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2$ is the data fidelity that enforces agreement of the solution with the data.

Therefore, the solution is $\hat{\boldsymbol{x}} = \left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{y}$, called the Least-Squares solution. However, this solution is unstable because of the ill-conditioning of the matrix $\boldsymbol{H}$, hence the problem is ill-posed. To correct this ill-posedness we add regularization terms to $\mathcal{J}(\boldsymbol{x})$, this method is called Regularized Least-Squares [18]. The objective function becomes $\mathcal{J}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{x}, \boldsymbol{y}) + \mathcal{R}_s(\boldsymbol{x}) + \mathcal{R}_\lambda(\boldsymbol{x})$, where $\mathcal{R}_s(\boldsymbol{x}) = \mu_s \|\boldsymbol{D}_s\boldsymbol{x}\|_2^2$ is a spatial regularization which enforces spatial smoothness between neighboring pixels of $\boldsymbol{x}$, with $\boldsymbol{D}_s \in \mathbb{R}^{M'N^2 \times M'N^2}$ is a second-order finite difference operator along the spatial dimension. $\mathcal{R}_\lambda(\boldsymbol{x}) = \mu_\lambda \|\boldsymbol{D}_\lambda\boldsymbol{x}\|_2^2$ refers to spectral regularization. It enforces the similarity between intensity values of corresponding pixels in neighboring channels, with $\boldsymbol{D}_\lambda \in \mathbb{R}^{M'N^2 \times M'N^2}$ is a first-order finite difference operator along the spectral direction. μ_s and μ_λ are regularization parameters, they are set to compromise between fidelity to the data and spatial smoothness, and spectral smoothness across channel, respectively. Therefore the objective function, $\mathcal{J}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \mu_s \|\boldsymbol{D}_s\boldsymbol{x}\|_2^2 + \mu_\lambda \|\boldsymbol{D}_\lambda\boldsymbol{x}\|_2^2$. is a sum of quadratic, linear and differentiable terms. Thus we obtain the solution

$$\hat{\boldsymbol{x}} = \left(\boldsymbol{H}^T\boldsymbol{H} + \mu_s\boldsymbol{D}_s^T\boldsymbol{D}_s + \mu_\lambda\boldsymbol{D}_\lambda^T\boldsymbol{D}_\lambda\right)^{-1}\boldsymbol{H}^T\boldsymbol{y}, \tag{9}$$

where $\boldsymbol{Q} = \boldsymbol{H}^T\boldsymbol{H} + \mu_s\boldsymbol{D}_s^T\boldsymbol{D}_s + \mu_\lambda\boldsymbol{D}_\lambda^T\boldsymbol{D}_\lambda$ contains Toepltiz blocks $\boldsymbol{Q}^{i,j} : i, j = 1, \ldots, M'$ of size $N^2 \times N^2$. However, $\boldsymbol{Q}^{i,j} \neq \boldsymbol{Q}^{i+t,j+t}$, hence $\boldsymbol{Q}$ is not a Toeplitz matrix. We propose to compute the solution without inverting $\boldsymbol{Q}$, but by computing the solution iteratively using the following form:

$$\hat{\boldsymbol{x}}^{k+1} = \hat{\boldsymbol{x}}^k - a\left[\boldsymbol{Q}\hat{\boldsymbol{x}}^k - \boldsymbol{H}^T\boldsymbol{y}\right], \tag{10}$$

with $\hat{\boldsymbol{x}}^{k=0} = \boldsymbol{0}$ corresponds to the initialization and a is a convergence parameter of the algorithm. A conjugated gradient (CG) algorithm [19] is implemented.

V. Simulation Results

A. JWST/MIRI Imager

We apply the proposed reconstruction algorithm to multispectral data, simulated using the model in (4), for the Mid-InfraRed Instrument (MIRI) Imager [20] on-board the James Webb Space Telescope (JWST), the next flagship space telescope of NASA, ESA and the Canadian Space Agency (CSA) to be launched in 2020. This imager provides nine multispectral observations ($P = 9$) integrated over a broad range of spectral bands, from 5 μm to 28 μm [21]. The nine bands are shown in Fig. 1. Note that overlapping of the spectral bands increases the between channels degradation. The MIRI Imager detector has a pixel pitch of 0.11 arcsecond, i.e. $\Omega_{\text{pix}} = 0.11 \times 0.11$ arcsecond^2. We use the official PSF simulator of the JWST mission, *WebbPSF* [22], [23], to simulate realistic PSF images at different wavelengths, as shown in Fig. 2. The PSF is complex due to the segmented mirror of the JWST. We clearly observe an enlargement of the PSF according to the wavelength, i.e. the longer the wavelength the wider the PSF, as expected from diffraction theory [16].

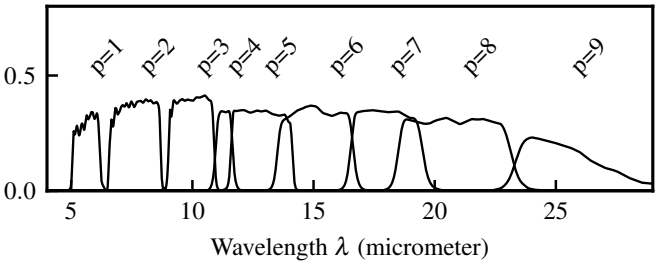


Fig. 1. The nine broad bands of the JWST/MIRI Imager [21] covering the mid-infrared wavelength range from 5 to 28 μm.

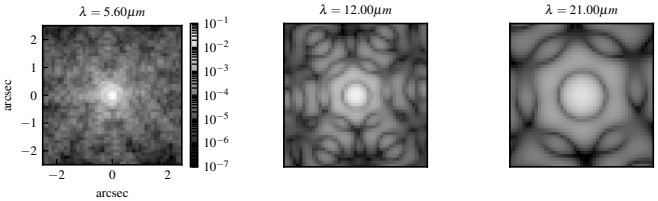


Fig. 2. Monochromatic PSF of the JWST/MIRI imager simulated using *WebbPSF* [22], [23] and displayed in the same logarithmic scale. We clearly observe the dependency of the PSF to the wavelength as expected from diffraction theory.

B. Setup of the Experiment

The original 3D object ϕ is a simplified spatio-spectral model of the Horsehead nebula [24]. A spatial region of 256×256 pixels ($N = 256$) is taken for the simulation with $M' = 1000$ spectral samples uniformly distributed within $4-28$ μm. Nine multispectral data are simulated using (4) with a zero-mean white Gaussian noise added in order to obtain a global Signal-to-Noise Ratio (SNR) of $30, 20, 10$ dB.

$$SNR_{(dB)} = 10\log_{10}\left(\frac{\frac{1}{PN^2}\|\boldsymbol{y}\|_2^2}{\sigma_n^2}\right), \tag{11}$$

where σ_n is the standard deviation of the noise, P is the number of MS data and N^2 is the total number of pixels in the MS data.

Reconstruction results are summarized in Table I together with a comparison between the proposed method and the

multichannel 2-D deconvolution method (MDec) [6] (independent channel restoration using an averaged PSF per channel). The regularization parameters μ_s and μ_λ are adjusted by running the code for different values in a range $[10^{-4}, 10^{-2}]$ and keeping the pair that minimizes the objective function $\mathcal{J}(\hat{\boldsymbol{x}}(\hat{\mu_s}, \hat{\mu_\lambda}))$. For a quantitative comparison between the original object $\boldsymbol{f}_{\text{orig}}$ and the reconstructed $\boldsymbol{f}_{\text{rec}}$, we compute the relative reconstruction error as defined by

$$Error(\%) = 100 \times \left\| \boldsymbol{f}_{\text{orig}} - \boldsymbol{f}_{\text{rec}} \right\|_2 / \left\| \boldsymbol{f}_{\text{orig}} \right\|_2 .$$

C. Discussion

Fig. 3 shows the spectral reconstruction result of one single pixel $(100, 150)$, comparing the original spectrum ϕ_{orig}, the reconstructed spectrum using our method ϕ_{rec}, and the reconstructed spectrum using multichannel 2-D deconvolution ϕ_{MDec}. The original spectral distribution is complex with spectral features at short wavelengths $(4 - 13\,\mu\text{m})$ and continuum. Therefore, an accurate reconstruction using a few MS data is difficult, if not impossible, without using a strong prior knowledge of the spectrum of the object. The reconstructed spectrum computed with our method (using piecewise linear model) allows us to reconstruct an envelope-like spectral distribution which significantly increases the spectral resolution compared to multichannel 2-D deconvolution. Several values of $M' = \{20, 40, 60\}$ have been tested, and the reconstruction results for three wavelengths, 7.8, 16 and 21 μm, are reported in Table I. Increasing M' improves the spectral resolution of the object model, but increases the between channels degradation and the number of unknowns. Moreover, we find that there is not much error improvement for $M' > 60$. In any case, the proposed reconstruction shows smaller reconstruction errors compared to the multichannel 2-D deconvolution ; this is due to our model accounting for within and between channel degradations. Spatial reconstruction results at different wavelengths are illustrated in Fig. 4. As anticipated, a better reconstruction is obtained at $\lambda = 16\,\mu$m and $\lambda = 21\,\mu$m than at $\lambda = 7.8\,\mu$m (see the fourth row of the figure) since within the integration windows at long wavelengths the spectrum of the object does not contain any feature.

TABLE I
RELATIVE RECONSTRUCTION ERRORS (SEE THE TEXT) FOR THE HORSEHEAD NEBULA [24] OF SIZE $1000 \times 256 \times 256$

SNR (dB)	λ (μm)	Error (%) Proposed Reconstruction			MDec
		M'=20	M'=40	M'=60	
30	7,8	49,44	42,37	41,42	52,85
	16,0	2,44	4,11	4,80	7,89
	21,0	1,87	3,82	4,26	11,92
20	7,8	49,50	43,07	41,46	52,84
	16,0	7,41	7,66	8,98	8,02
	21,0	4,42	5,40	5,77	11,97
10	7,8	50,71	43,71	42,38	52,84
	16,0	19,67	21,25	25,51	8,56
	21,0	10,85	11,31	13,38	12,13

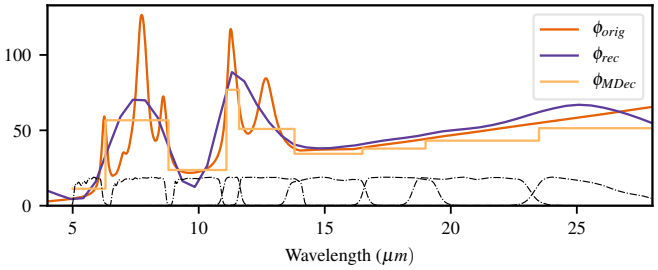


Fig. 3. Comparison between one single pixel $(100, 150)$ spectrum from the original object ϕ_{orig}, the proposed reconstruction ϕ_{rec} (with $M' = 60$) and the multichannel 2-D deconvolution ϕ_{MDec}. The nine MS data ($P = 9$) were corrupted with zero-mean Gaussian noise of 30 dB.

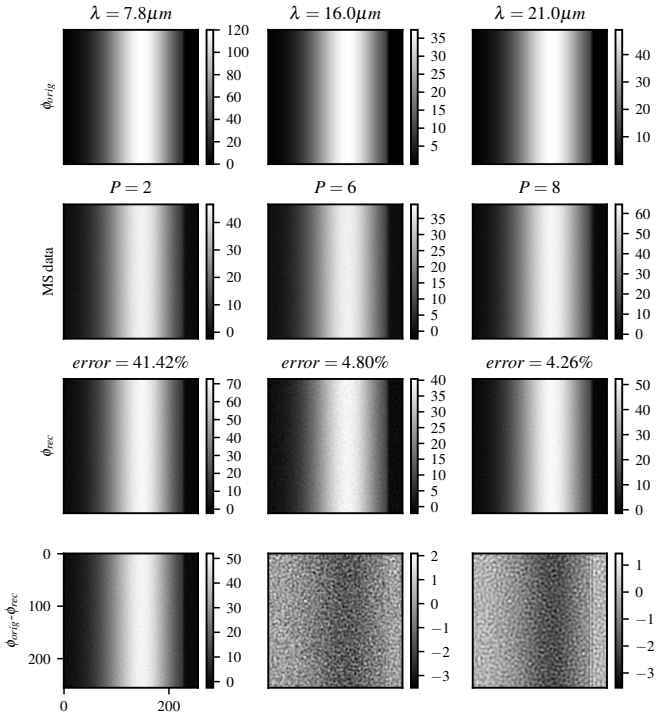


Fig. 4. [First row] Original spectral channel of the Horse Head nebula. [Second row] Simulated MS data with 30 dB corresponding to the bands that include wavelengths of the first row (see Fig. 1). [Third row] Proposed reconstruction. [Fourth row] Difference between the original and reconstructed spectral channels. The original and reconstructed objects are in physical units, whereas MS data are in detector units.

VI. Conclusion

In this paper we address the reconstruction of 3-D spatio-spectral object observed by a multispectral imaging system from a few low-resolution data. A discrete forward model is defined accounting for within and between channel degradations using a piecewise linear function to model the spectral distribution of the sought object. A quadratic reconstruction is proposed by considering spatial and spectral regularization terms. Results on simulated data applied to the JWST/MIRI Imager highlights the complexity of the problem. A clear increase of spatial and spectral distribution is achieved compared to multichannel 2-D deconvolution method.

REFERENCES

[1] Roberto Lionello, Jon A Linker, and Zoran Mikić, "Multispectral emission of the sun during the first whole sun month: Magnetohydrodynamic simulations," *The Astrophysical Journal*, vol. 690, no. 1, pp. 902, 2008.

[2] David A Landgrebe, *Signal theory methods in multispectral remote sensing*, vol. 29, John Wiley & Sons, 2005.

[3] Troy O McBride, Brian W Pogue, Steven Poplack, Sandra Soho, Wendy A Wells, Shudong Jiang, Keith D Paulsen, et al., "Multispectral near-infrared tomography: a case study in compensating for water and lipid content in hemoglobin imaging of the breast," *Journal of biomedical optics*, vol. 7, no. 1, pp. 72–79, 2002.

[4] ME Dickinson, G Bearman, S Tille, R Lansford, and SE Fraser, "Multispectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy," *Biotechniques*, vol. 31, no. 6, pp. 1272–1279, 2001.

[5] B Hunt and Olaf Kubler, "Karhunen-loeve multispectral image restoration, part i: Theory," *IEEE transactions on acoustics, speech, and signal processing*, vol. 32, no. 3, pp. 592–600, 1984.

[6] Nikolas P Galatsanos and Roland T Chin, "Digital restoration of multichannel images," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 3, pp. 415–421, 1989.

[7] Simon Henrot, Charles Soussen, and David Brie, "Fast positive deconvolution of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 828–833, 2013.

[8] S Bongard, F Soulez, Éric Thiébaut, and É Pecontal, "3d deconvolution of hyper-spectral astronomical data," *Monthly Notices of the Royal Astronomical Society*, vol. 418, no. 1, pp. 258–270, 2011.

[9] Yingying Song, David Brie, El-Hadi Djermoune, and Simon Henrot, "Regularization parameter estimation for non-negative hyperspectral image deconvolution," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5316–5330, 2016.

[10] Nikolas P Galatsanos, Aggelos K Katsaggelos, Roland T Chin, and Allen D Hillery, "Least squares restoration of multichannel images," *IEEE Transactions on Signal Processing*, vol. 39, no. 10, pp. 2222–2236, 1991.

[11] Richard R Schultz and Robert L Stevenson, "Stochastic modeling and estimation of multispectral image data," *IEEE Transactions on Image Processing*, vol. 4, no. 8, pp. 1109–1119, 1995.

[12] You-Wei Wen, Michael K Ng, and Yu-Mei Huang, "Efficient total variation minimization methods for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2081–2088, 2008.

[13] Junfeng Yang, Wotao Yin, Yin Zhang, and Yilun Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 569–592, 2009.

[14] MA Hadj-Youcef, François Orieux, Aurélia Fraysse, and Alain Abergel, "Restoration from multispectral blurred data with non-stationary instrument response," in *Signal Processing Conference (EUSIPCO), 2017 25th European*. IEEE, 2017, pp. 503–507.

[15] Philippe Thévenaz, Thierry Blu, and Michael Unser, "Interpolation revisited [medical images application]," *IEEE Transactions on medical imaging*, vol. 19, no. 7, pp. 739–758, 2000.

[16] Joseph W Goodman, *Introduction to Fourier optics*, Roberts and Company Publishers, 2005.

[17] Leonid Yaroslavsky, *Digital holography and digital image processing: principles, methods, algorithms*, Springer Science & Business Media, 2013.

[18] Guy Demoment, "Image reconstruction and restoration: Overview of common estimation structures and problems," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 2024–2036, 1989.

[19] Jonathan Richard Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," 1994.

[20] Patrice Bouchet, Macarena García-Marín, P-O Lagage, Jérome Amiaux, J-L Auguéres, Eva Bauwens, JADL Blommaert, CH Chen, ÖH Detre, Dan Dicken, et al., "The Mid-Infrared Instrument for the James Webb Space Telescope, III: MIRIM, The MIRI Imager," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 953, pp. 612, 2015.

[21] Alistair Glasse, GH Rieke, Eva Bauwens, Macarena García-Marín, ME Ressler, Steffen Rost, Tuomo Ville Tikkanen, Bart Vandenbussche, and GS Wright, "The Mid-Infrared Instrument for the James Webb Space Telescope, IX: Predicted Sensitivity," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 953, pp. 686, 2015.

[22] Marshall D Perrin, Rémi Soummer, Erin M Elliott, Matthew D Lallo, and Anand Sivaramakrishnan, "Simulating point spread functions for the James Webb Space Telescope with WebbPSF," International Society for Optics and Photonics, 2012, pp. 84423D–84423D.

[23] Marshall D Perrin, Anand Sivaramakrishnan, Charles-Philippe Lajoie, Erin Elliott, Laurent Pueyo, Swara Ravindranath, and Loïc Albert, "Updated point spread function simulations for jwst with webbpsf," in *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*. International Society for Optics and Photonics, 2014, vol. 9143, p. 91433X.

[24] A Abergel, D Teyssier, JP Bernard, F Boulanger, A Coulais, D Fosse, E Falgarone, M Gerin, M Perault, J-L Puget, et al., "Isocam and molecular observations of the edge of the horsehead nebula," *Astronomy & Astrophysics*, vol. 410, no. 2, pp. 577–585, 2003.

Restoration from Multispectral Blurred Data with Non-Stationary Instrument Response

M.A. Hadj-Youcef*†, F. Orieux*†, A. Fraysse*, A. Abergel†
* Laboratoire des Signaux et Systèmes, Univ. Paris-Sud, CNRS, CentraleSupélec, Université Paris-Saclay
3 rue Joliot-Curie, 91 192 Gif-sur-Yvette, France
Email: amine.hadjyoucef@ias.u-psud.fr
† Institut d'Astrophysique Spatiale, CNRS, UMR 8617, Univ. Paris-Sud, Université Paris-Saclay
Univ. Paris-Sud, 91405 Orsay, France

***Abstract*—In this paper we propose an approach of image restoration from multispectral data provided by an imaging system. We specifically address two topics: (i) Development of a multi-wavelength direct model for non-stationary instrument response that includes a spatial convolution and a spectral integration, (ii) Implementation of multispectral image restoration using a regularized least-square, based on a quadratic criterion and minimized by a gradient algorithm. We test our approach on simulated data of the Mid-InfraRed Instrument IMager (MIRIM) of the James Webb Space Telescope (JWST). Our method shows a clear increase of spatial resolution compare to conventional methods.**

***Index Terms*—Direct Model, Multispectral Imaging, Inverse Problems, Image Restoration**

I. Introduction

Multispectral imaging instrument are used in many fields, e.g. remote sensing [1], medicine [2], astrophysics [3]. Imaging system on board space telescopes are specific subject of this work, especially the Mid-InfraRed Instrument IMager (MIRIM) [4] on board the James Webb Space Telescope (JWST)[1] which is the next space telescope of NASA, ESA and the Canadian Space Agency (CSA) to be launched at the end of 2018. Such a system is mainly composed of an optical system (or optic) that gathers source's light (or object) from space and provides it to the imaging instrument, which contains a spectral filter (or photometric band) to select the wavelength band of interest and an infrared detector that integrates and discretizes the light, providing a 2D discrete image data.

Due to its passage through the optics, photon flux of the object of interest is diffracted on the focal plane of the telescope. The result is a spatial 2D-convolution of the object with the optic response known as Point Spread Function (PSF) [5]. On the other hand, the PSF vary accordingly to the wavelength, blurring the object of interest and limiting its spatial resolution. The second problem occurs during the integration of the object by the detector, where the continuous 2D+λ object is integrated spectrally over the filter bandpass, the 2D discrete image data has thus no spectral dimension.

Conventional approaches consider measured PSFs [6], [4] but only PSFs at specific wavelengths are measured. Other approaches assume a broadband PSF, by averaging available monochromatic PSFs, weighted by the filter+detector spectral response and the object spectrum to observe [7]. Broadband PSF smooths the structure of the PSF and makes it dependent on the object spectrum to observe, especially for filters with wide bands. In addition, a monochromatic object spectrum is assumed for unknown sources, which results in a stationary PSF and leads to inaccurate optics response and direct model. Spatially-variant PSF has also been studied in [8] which approximates the spatial-variant PSF using PSF-interpolation. The same idea was used in [9] to approximate the spectral-variant PSF. More recently, [10] reviews and provides models of spatial-variant PSF. Other works treated deconvolution problem with unknown parameter of the PSF or unknown PSF, known as myope or blind deconvolution [11]. In the other hand, the spectral integration is mostly approached by a spectral convolution, such as in [12] where the 3D-PSF is separated into spectral and spatial-invariant PSF, and [9] by using PSF-interpolation so that the direct model is a 2D+λ convolution. In this paper we deal with the continuous spectral integration by taking into account informations from available multispectral data of multi-filter, and preserve the spectral non-stationarity of PSF using PSF simulator. This work aims to restore a discrete version of the original spatio-spectral object from available multispectral discrete data. Our main contribution is the development of an instrument model and a direct model for an imaging system with a non-stationary instrument response.

The paper is organized as follows. In Sec. II, we develop the instrument model. Then we present the direct model along with the restoration in Sec. III. Experimental results on simulated data with an application to the instrument MIRIM are presented and discussed in Sec. IV. Finally, we conclude our work and provide perspectives in Sec. V.

II. Instrument Model

In this section we develop the instrument model for the imager. The block diagram of the instrument is represented in Fig. 1, and is composed of an optic response, spectral filtering and detector integration. The object of interest at the entrance of the acquisition system is a 3D object, $\phi(\alpha, \beta, \lambda) : \mathbb{R}^3 \rightarrow \mathbb{R}$, having two spatial dimensions $(\alpha, \beta) \in \mathbb{R}^2$ (angles in radian)

[1] https://jwst.nasa.gov/

and one spectral dimension $\lambda \in \mathbb{R}^+$ (in microns). During the observation process with the filter f ($f \in [1, n_f]$ where n_f stands for total number of filters), the object is modified by the instrument components, providing 2D-discrete data $\boldsymbol{y}^{(f)}$.

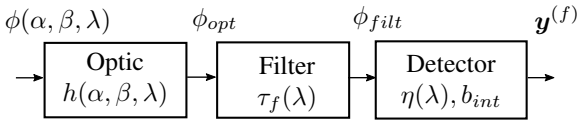


Fig. 1. Block diagram of the instrument model of the imager.

A. Instrument Response

1) Optic: The effect of the optic is carried out by the (α, β)-convolution of the input object $\phi(\alpha, \beta, \lambda)$ with the spectral-variant PSF $h(\alpha, \beta, \lambda)$ as follows

$$\begin{aligned} \phi_{opt}(\alpha, \beta, \lambda) &= \phi(\alpha, \beta, \lambda) \underset{(\alpha,\beta)}{*} h(\alpha, \beta, \lambda) \\ &= \iint_{\mathbb{R}^2} \phi(\alpha', \beta', \lambda)\, h(\alpha - \alpha', \beta - \beta', \lambda)\, d\alpha' d\beta', \end{aligned} \tag{1}$$

where $\underset{(\alpha,\beta)}{*}$ stands for 2D spatial convolution.

Fig. 2 illustrates the non-stationarity of the PSF. We display two simulated PSFs of the instrument JWST/MIRIM at different wavelengths from the mid-infrared range, $\lambda_1 = 7.7\mu m$ and $\lambda_2 = 25.25\mu m$, PSF_{λ_2} is clearly larger than PSF_{λ_1}.

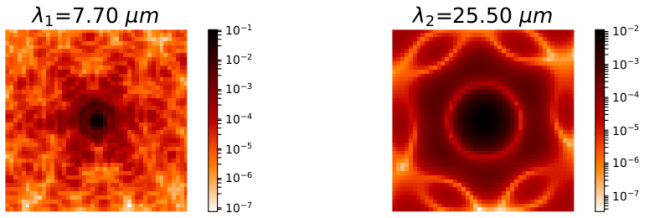


Fig. 2. Monochromatic PSFs simulated at $\lambda_1 = 7.7\,\mu m$ and $\lambda_2 = 25.25\,\mu m$, using the simulation tool WebbPSF [13] for the instrument MIRIM on board the JWST.

2) Filter: The object passes through the filter f in order to select the wavelength range of interest. This operation is performed through the filter response $\tau_f(\lambda)$ as follows

$$\phi_{filt}^{(f)}(\alpha, \beta, \lambda) = \tau_f(\lambda)\, \phi_{opt}(\alpha, \beta, \lambda). \tag{2}$$

3) Detector: The detector integrates the filtered object and performs sampling. We model its response in two steps :

- The spectral integration is weighted by the spectral response of the detector (or quantum efficiency) $\eta(\lambda)$,

$$y^{(f)}(\alpha, \beta) = \int_{\mathbb{R}_+} \eta(\lambda) \phi_{filt}^{(f)}(\alpha, \beta, \lambda) d\lambda. \tag{3}$$

- We introduce a basis function b_{int} to perform spatial sampling within spatial integration, e.g. indicator function of a sensitive area of the pixel Ω_{pix}. The discrete data of the filter f and pixel (i, j) is

$$y_{i,j}^{(f)} = \iint_{\Omega_{pix}} y^{(f)}(\alpha, \beta) b_{\text{int}}(\alpha - \alpha_{i,j}, \beta - \beta_{i,j})\, d\alpha d\beta. \tag{4}$$

with $\alpha_{i,j}$ and $\beta_{i,j}$ the angular directions of the pixel position (i, j) defined on the detector grid. We denote N the total number of pixels.

B. Complete Formulation

The complete formulation of the instrument model is obtained by substituting Eqs.(1)-(2) and (3) in Eq.(4). This yields

$$\begin{aligned} y_{i,j}^{(f)} = \int_{\mathbb{R}_+} \eta(\lambda)\tau_f(\lambda) \Big(\iint_{\Omega_{pix}} \phi(\alpha, \beta, \lambda) \underset{(\alpha,\beta)}{*} h(\alpha, \beta, \lambda) \\ b_{\text{int}}(\alpha - \alpha_{i,j}, \beta - \beta_{i,j})\, d\alpha d\beta \Big) d\lambda. \end{aligned} \tag{5}$$

This model links the continuous 3D object $\phi(\alpha, \beta, \lambda)$ at the entrance of the imaging system to the 2D discrete data $\boldsymbol{y}^{(f)}$ through the instrument response, which includes spectral windowing and five integrations, two for spatial 2D-convolution, two for spatial integration and one for spectral integration. The developed instrument model takes into account the non-stationarity of the instrument response through the spectral-variant PSF. Moreover, the above model does not include non-ideal characteristics of the detector [14]. All these effects are assumed to be corrected through the pipeline stages of the data reduction plan [15]. In this paper, we consider the same sampling grid of the object as for the data, and that all pixels are regularly disposed on the detector grid, having the same area Ω_{pix}.

In order to simplify the instrument model, we consider a constant object over pixel area. Thus, we define b_{int} as a rectangular function over the pixel area. The impact of this approximation is slightly important for detectors with good resolution, such as MIRIM detector [16]. Thus, Eq.(5) becomes

$$\begin{aligned} y_{i,j}^{(f)} = \Omega_{pix} \int_{\mathbb{R}_+} \eta(\lambda)\tau_f(\lambda) \\ \phi(\alpha_{i,j}, \beta_{i,j}, \lambda) \underset{(\alpha_{i,j},\beta_{i,j})}{*} h(\alpha_{i,j}, \beta_{i,j}, \lambda)\, d\lambda \end{aligned} \tag{6}$$

where $\underset{(\alpha_{i,j},\beta_{i,j})}{*}$ stands for discrete 2D-convolution.

III. Direct Model and Inverse Problem

In this section we first model the object spectrum, then we develop the direct model, thereafter we tackle the inverse problem of the object.

A. Continuous Object Spectrum

In this paper we model the spectrum object with a continuous piecewise linear function, as shown in Fig.(3). This choice allows us to obtain a simple model that preserves the spectral distribution of the object with less complexity, whereas conventional approaches do not. They generally consider a

model with broadband PSF, defined as a spectral integration of monochromatic PSF weighted by the filter + detector + object spectrum window [7]. Hence, one can only attempt to restore a spatial distribution with such a model [17]. The object is modeled as follows

$$\phi\left(\alpha_{i,j},\beta_{i,j},\lambda\right)=\sum_{b=1}^{n_b}\Big(x^{(b)}\left(\alpha_{i,j},\beta_{i,j}\right)g_{+}^{(b)}(\lambda)+ x^{(b-1)}\left(\alpha_{i,j},\beta_{i,j}\right)g_{-}^{(b)}(\lambda)\Big)\mathbb{1}_{\left[\lambda^{(b-1)},\lambda^{(b)}\right]}(\lambda), \tag{7}$$

where $b\in[1,...,n_b]$ is the index of the bandwidth, $\mathbb{1}(\lambda)$ the indicator function, which is equal to 1 for $\lambda\in\left[\lambda^{(b-1)},\lambda^{(b)}\right]$ and 0 otherwise. The discrete parameters $x_{i,j}^{(b-1)}$ and $x_{i,j}^{(b)}$ are intensities of the object at $\lambda^{(b-1)}$ and $\lambda^{(b)}$, respectively. $g_{-}^{(b)}(\lambda)=\frac{1}{2}-\frac{\lambda-\lambda_c^{(b)}}{\lambda^{(b)}-\lambda^{(b-1)}}$ and $g_{+}^{(b)}(\lambda)=\frac{1}{2}+\frac{\lambda-\lambda_c^{(b)}}{\lambda^{(b)}-\lambda^{(b-1)}}$ express the linear slope of the object spectrum.

Thanks to the parameterization in Eq. (7), the object spectrum is continuous in λ and the positivity constraint is naturally fulfilled.

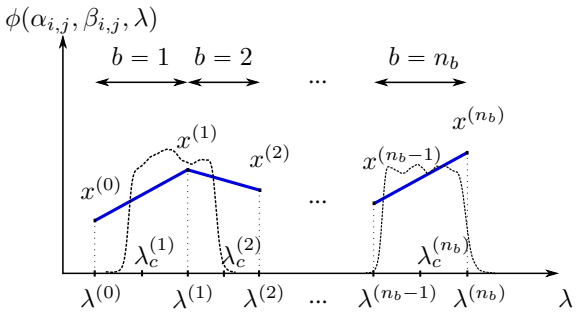


Fig. 3. Representation of a piecewise linear spectrum, for a single position over a wavelength range of $[\lambda^{(0)},\lambda^{(n_b)}]$. (in dots) Filters transmission to illustrate the multi-filter analysis.

B. Direct model

The discrete data $y^{(f)}$ contains information about the object of interest within the filter bandpass. In fact, we use a multi-filter analysis to develop a direct model that accounts for informations of the whole wavelength range of the instrument. By substituting Eq.(7) in Eq.(6) and rearranging similar terms, only wavelength-dependent terms are left in the spectral integration without the object parameter x. This yields

$$y_{i,j}^{(f)}=\sum_{b=0}^{n_b}h_{\text{int}}^{(f,b)}(\alpha_{i,j},\beta_{i,j})\underset{(\alpha_{i,j},\beta_{i,j})}{*}x^{(b)}(\alpha_{i,j},\beta_{i,j}) \tag{8}$$

where $h_{\text{int}}^{(f,b)}$ is a 2D integrated PSF

$$h_{\text{int}}^{(f,b)}(\alpha_{i,j},\beta_{i,j})=\Omega_{pix}\int_{\mathbb{R}_+}\eta(\lambda)\tau_f(\lambda)\left[g_0^{(b+1)}(\lambda)+g_1^{(b)}(\lambda)\right]h(\alpha_{i,j},\beta_{i,j},\lambda)d\lambda, \tag{9}$$

with the boundary conditions $g_{-}^{(n_b+1)}(\lambda)=g_{+}^{(0)}(\lambda)=0$.

Eq. (8) can be formulated in a vector form as follows

$$\boldsymbol{y}^{(f)}=\sum_{b=0}^{n_b}\boldsymbol{H}_{\text{int}}{}^{(f,b)}\boldsymbol{x}^{(b)}+\boldsymbol{\epsilon}^{(f)}, \tag{10}$$

where $\boldsymbol{H}_{\text{int}}^{(f,b)}\in\mathbb{R}^{N\times N}$ is a convolution matrix. $\boldsymbol{x}^{(b)}\in\mathbb{R}^N$ and $\boldsymbol{y}^{(f)}\in\mathbb{R}^N$ are vector representations of $y^{(f)}$ and $x^{(b)}$. $\boldsymbol{\epsilon}^{(f)}\in\mathbb{R}^N$ corresponds to readout noise and modeling error.

The linear direct model in Eq. (10) is a sum of (n_b+1) spatial convolutions, between the 2D object parameters $\boldsymbol{x}^{(b)}$ and 2D system response $\boldsymbol{H}_{\text{int}}^{(f,b)}$. Moreover, the observation $\boldsymbol{y}^{(f)}$ contains information about the unknown object over the whole spectral range, weighted by the spectral windows $\eta(\lambda)\tau_f(\lambda)$. By combining all available multispectral data in one vector $\boldsymbol{y}$, Eq. (10) becomes

$$\boldsymbol{y}=\boldsymbol{H}\boldsymbol{x}+\boldsymbol{\epsilon}, \tag{11}$$

where $\boldsymbol{H}=\left\{\boldsymbol{H}_{\text{int}}^{(f,b)}\right\}_{\substack{f=1:n_f\\b=0:n_b}}$ is a block matrix containing $n_f\times(n_b+1)$ convolution matrices corresponding to the integrated PSFs $h_{\text{int}}^{(f,b)}$. We denote by $\boldsymbol{x}=\left[\boldsymbol{x}^{(0)}\dots\boldsymbol{x}^{(n_b)}\right]^t$ the vector of (n_b+1) unknowns object parameters, $\boldsymbol{y}=\left[\boldsymbol{y}^{(1)}\dots\boldsymbol{y}^{(n_f)}\right]^t$ vector of all available data, and $\boldsymbol{\epsilon}=\left[\boldsymbol{\epsilon}^{(1)}\dots\boldsymbol{\epsilon}^{(n_f)}\right]^t$ vector of additive noise associated to data.

C. Restoration

Restoration of $\boldsymbol{x}$ relies on regularized least-square approach in inverse problems [18], where the solution is obtained by minimizing a criterion $\mathcal{J}$ as follows

$$\hat{\boldsymbol{x}}=\underset{\boldsymbol{x}}{\text{argmin}}\ \mathcal{J}\left(\boldsymbol{x}\right). \tag{12}$$

The criterion of least-square (called data fidelity) is $\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|_2^2$. It enforces agreement of the solution $\hat{\boldsymbol{x}}$ with the data $\boldsymbol{y}$. Least-square problem is ill-posed because of the ill-conditioning of $\boldsymbol{H}$. In fact, regularized least-square estimator aims of correcting this ill-posedness by adding a regularization term $\|\boldsymbol{C}\boldsymbol{x}\|_2^2$ to the criterion, with $\boldsymbol{C}$ a well-conditioned matrix. We are particularly interested in a quadratic-regularization (l_2-norm) in order to obtain a differentiable criterion. This allow us to implement a fast calculation of the solution. Therefore, the regularized least-square criterion is

$$\mathcal{J}(\boldsymbol{x})=\|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|_2^2+\|\boldsymbol{C}\boldsymbol{x}\|_2^2, \tag{13}$$

the matrix $\boldsymbol{C}=\text{diag}\left\{\sqrt{\mu_0}\boldsymbol{D}_{\alpha,\beta},\dots,\sqrt{\mu_{n_b}}\boldsymbol{D}_{\alpha,\beta}\right\}$, where $\boldsymbol{D}_{\alpha,\beta}$ is a 2D spatial constraint operator, e.g. discrete Laplacian operator in case of a spatially smooth object. Moreover, regularization parameters $\mu_0,\dots,\mu_{n_b}$ are set to compromise between the two terms of the criterion, data fidelity and regularization.

As $\mathcal{J}$ is a quadratic form, solution of the problem $\hat{\boldsymbol{x}}$ is explicit and obtained by canceling the gradient of the criterion defined in Eq.(13) :

$$\hat{\boldsymbol{x}} = (\underbrace{\boldsymbol{H}^t\boldsymbol{H} + \boldsymbol{C}^t\boldsymbol{C}}_{\boldsymbol{Q}})^{-1}\boldsymbol{H}^t\boldsymbol{y} \tag{14}$$

where $\boldsymbol{Q}$ is a Hessian matrix of size $(n_b+1)N \times (n_b+1)N$.

We first attempt to compute the solution by inverting $\boldsymbol{Q}$ using diagonalization of circular matrices in Fourier space, therefore, computing $\boldsymbol{Q}^{-1}$ means inverting N square matrices of size $(n_b + 1) \times (n_b + 1)$. This calculation turns out heavy for $n_b > 4$. We propose instead to compute the solution without inverting $\boldsymbol{Q}$, i.e. by solving a linear system $(\boldsymbol{H}^t\boldsymbol{H} + \boldsymbol{C}^t\boldsymbol{C})\hat{\boldsymbol{x}} = \boldsymbol{H}^t\boldsymbol{y}$ through an iterative scheme, e.g. using numerical optimization algorithm such as the conjugated gradient (CG) algorithm [19]. Moreover, we consider circular convolution and compute the solution in Fourier space for efficient computation.

IV. Simulation Results

A. Application to JWST/MIRIM

We apply the proposed approach to the instrument MIRIM on board the space telescope JWST. The optical system of the telescope is mainly composed of a 6.5-meters primary mirror, made up of 18-hexagonal segmented mirror. The imaging instrument MIRIM has nine photometric bands that cover the mid-infrared range of 5 to $28\mu m$. Their transmission profiles are given in [4]. The object of interest is guided through mirrors to the MIRIM detector [16], where it will be integrated and sampled. MIRIM detector has a pixel scale of 0.11 arcsec/pixel ($\Omega_{pix} = 0.11^2$ arcsecond^2) and its spectral window $\eta(\lambda)\tau_f(\lambda)$ is known as Photon Conversion Efficiency (PCE) [20], see Fig. 4. Moreover, non-stationary PSF is a set of monochromatic PSFs simulated on a discrete wavelength grid that covers the whole spectral range of the instrument, simulated with WebbPSF [13] the official PSF simulator of mission JWST.

B. Setup

In order to validate the proposed approach, all tests were done on simulated data, where we simulated the original object as a cube of size $64\times64\times9$ ($N = 64, n_b = 8$) with a Gaussian spatial distribution ($\sigma = 3$) and a linear spectrum over the instrument spectral range. We implemented the direct model in Eq.(11) by computing the nine integrated PSFs and performing convolutions with original object parameters. As a result, nine multispectral data are simulated corresponding to nine MIRIM photometric bands. Signal-to-noise ratio (SNR) [21] is used to evaluate the level of the white Gaussian zero-mean noise added to the data. We implement the conjugate gradient algorithm to restore the object parameter $\hat{\boldsymbol{x}}$ and compare it to the conventional approach given in Sec. III-A, referenced as as "standard". Object restoration in case of the standard approach consists of deconvolution of nine data by considering computed broadband PSFs. We use the unsupervised deconvolution method of [22], with a quadratic-regularization and same prior. We obtain deconvolved observations as well as estimation of regularization parameters, these latter are used for both approaches, proposed and standard ones. Finally, the restored object of interest $\hat{\boldsymbol{\phi}}$ is deducted from the restored object parameter $\hat{\boldsymbol{x}}$ using Eq.(7). Restoration quality is measured with a relative error $\left\|\boldsymbol{\phi}_{\text{original}} - \hat{\boldsymbol{\phi}}\right\|_2 / \left\|\boldsymbol{\phi}_{\text{original}}\right\|_2$ between the original object of interest $\boldsymbol{\phi}_{\text{original}}$ and the restored object $\hat{\boldsymbol{\phi}}$.

C. Discussion

Fig. 4 illustrates restoration results in a spectral dimension, where we display the spectrum of the central pixel of the original object, proposed and standard restorations. All data are corrupted by an additive zero-mean Gaussian noise of 30 dB. We observe that the restored spectrum using the proposed approach fits very well the original spectrum along the whole wavelength range. However, the standard approach fails to do so, because of the inaccuracy of the instrument response considered in the standard approach. In addition, Fig. 5 shows observed data using filter $f = 8$ (named F2100W), the original object at $\lambda = 18.7\mu m$ and the spatial restoration with both approaches. The restoration result using our approach shows a good restoration of the dynamic and the spatial details, the blur caused by the PSF is also removed, with an error of 6.27% compared to 21.29% obtained with the standard approach. Moreover, we notice that the wider the filter band, the better the results of our approach compared to the conventional approach, which is the case for the instrument MIRIM.

V. Conclusion

In this paper we present the problem of image restoration from multispectral data acquired by an imager, where we restored a spatio-spectral object. An instrument model with a complex non-stationary response is detailed, including a spatial convolution with a spectral-variant PSF and a spectral integration. Then, we developed a multi-wavelength direct model, which is the sum of spatial 2D-convolution of the object parameter with an integrated PSF. The idea relies on using a multi-filter processing and an approximation of a continuous piecewise linear spectrum. Multispectral restoration is implemented using regularized least-square based on quadratic criterion. Preliminary restoration results are provided on simulated data, where we obtained a clear improvement of restoration quality compared to classical approach.

Several aspects are to improve in our approach, such as the estimation of the (n_b+1) regularization parameters. As future work, generalization of the direct model by introducing a set of basis functions of the object discretization, and increasing the spectral sampling of the sought object by adding spectral prior of the object. Finally, validation of direct model and restoration of real astrophysical object.

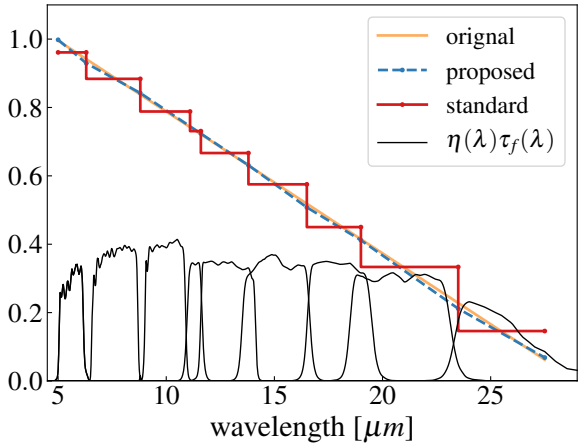


Fig. 4. Illustration of the spectral restoration. All intensities were normalized to 1 for better visualization and comparison.

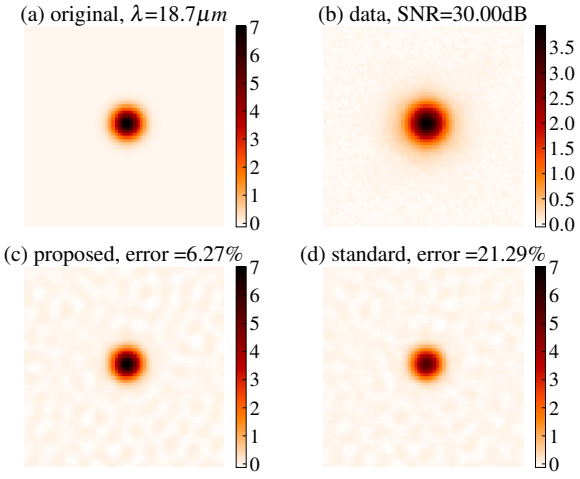


Fig. 5. Illustration of the spatial restoration : (a) original object at $\lambda = 18.7\mu m$. (b) data for the filter F2100W. (c) proposed restoration, (d) standard restoration. Parameters used are : SNR = 30 dB, $\sigma_{\text{Gauss}} = 3$

REFERENCES

[1] D. A. Landgrebe, *Signal theory methods in multispectral remote sensing*. John Wiley & Sons, 2005, vol. 29.

[2] S. Ozer, D. L. Langer, X. Liu, M. A. Haider, T. H. van der Kwast, A. J. Evans, Y. Yang, M. N. Wernick, and I. S. Yetik, "Supervised and unsupervised methods for prostate cancer segmentation with multispectral mri," *Medical physics*, vol. 37, no. 4, pp. 1873–1883, 2010.

[3] R. Lionello, J. A. Linker, and Z. Mikić, "Multispectral emission of the sun during the first whole sun month: Magnetohydrodynamic simulations," *The Astrophysical Journal*, vol. 690, no. 1, p. 902, 2008.

[4] P. Bouchet, M. García-Marín, P.-O. Lagage, J. Amiaux, J.-L. Augueres, E. Bauwens, J. Blommaert, C. Chen, O. Detre, D. Dicken *et al.*, "The Mid-Infrared Instrument for the James Webb Space Telescope, III: MIRIM, The MIRI Imager," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 953, p. 612, 2015.

[5] J. W. Goodman, *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[6] P. Guillard, T. Rodet, S. Ronayette, J. Amiaux, A. Abergel, V. Moreau, J. Augueres, A. Bensalem, T. Orduna, C. Nehmé *et al.*, "Optical performance of the JWST/MIRI flight model: characterization of the point spread function at high resolution," in *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics, 2010, pp. 77 310J–77 310J.

[7] N. Geis and D. Lutzn, "Herschel/PACS modelled point-spread functions."

[8] L. Denis, É. Thiébaut, and F. Soulez, "Fast model of space-variant blurring and its application to deconvolution in astronomy," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 2817–2820.

[9] F. Soulez, E. Thiébaut, and L. Denis, "Restoration of hyperspectral astronomical data with spectrally varying blur," *EAS Publications Series*, vol. 59, pp. 403–416, 2013.

[10] É. Thiébaut, L. Dénis, F. Soulez, and R. Mourya, "Spatially variant psf modeling and image deblurring," in *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics, 2016, pp. 99 097N–99 097N.

[11] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3715–3727, 2006.

[12] E. Villeneuve and H. Carfantan, "Nonlinear deconvolution of hyperspectral data with mcmc for studying the kinematics of galaxies," *IEEE Transactions on Image Processing*.

[13] M. D. Perrin, R. Soummer, E. M. Elliott, M. D. Lallo, and A. Sivaramakrishnan, "Simulating point spread functions for the James Webb Space Telescope with WebbPSF," in *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics, 2012, pp. 84 423D–84 423D.

[14] G. Wright, D. Wright, G. Goodson, G. H. Rieke, G. Aitink-Kroes, J. Amiaux, A. Aricha-Yanguas, R. Azzollini, K. Banks, D. Barrado-Navascues *et al.*, "The mid-infrared instrument for the james webb space telescope, ii: Design and build," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 953, p. 595, 2015.

[15] K. D. Gordon, C. Chen, R. E. Anderson, R. Azzollini, L. Bergeron, P. Bouchet, J. Bouwman, M. Cracraft, S. Fischer, S. D. Friedman *et al.*, "The Mid-Infrared Instrument for the James Webb Space Telescope, X: Operations and Data Reduction," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 953, p. 696, 2015.

[16] G. Rieke, M. Ressler, J. E. Morrison, L. Bergeron, P. Bouchet, M. García-Marín, T. Greene, M. Regan, K. Sukhatme, and H. Walker, "The Mid-Infrared Instrument for the James Webb Space Telescope, VII: The MIRI Detectors," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 953, p. 665, 2015.

[17] G. Aniano, B. Draine, K. Gordon, and K. Sandstrom, "Common-resolution convolution kernels for space-and ground-based telescopes," *Publications of the Astronomical Society of the Pacific*, vol. 123, no. 908, p. 1218, 2011.

[18] J. Idier, *Bayesian approach to inverse problems*.

[19] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," 1994.

[20] A. Glasse, G. Rieke, E. Bauwens, M. García-Marín, M. Ressler, S. Rost, T. V. Tikkanen, B. Vandenbussche, and G. Wright, "The Mid-Infrared Instrument for the James Webb Space Telescope, IX: Predicted Sensitivity," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 953, p. 686, 2015.

[21] R. C. Gonzalez and R. E. Woods, "Image processing," *Digital image processing*, vol. 2, 2007.

[22] F. Orieux, J.-F. Giovannelli, and T. Rodet, "Bayesian estimation of regularization and point spread function parameters for wiener–hunt deconvolution," *JOSA A*, vol. 27, no. 7, pp. 1593–1607, 2010.

Restauration d'objets astrophysiques à partir de données multispectrales floues et d'une réponse instrument non-stationnaire

MA HADJ-YOUCEF[1,2], F ORIEUX[1,2], A FRAYSSE[1], A ABERGEL[2]

[1]Laboratoire des Signaux et Systèmes, Univ. Paris-Sud,
CNRS, CentraleSupélec, Université Paris-Saclay
3 rue Joliot-Curie, 91 192 Gif-sur-Yvette, France

[2]Institut d'Astrophysique Spatiale, CNRS, UMR
8617, Univ. Paris-Sud, Université Paris-Saclay
Univ. Paris-Sud, 91405 Orsay, France

amine.hadjyoucef@ias.u-psud.fr, orieux@l2s.centralesupelec.fr,
fraysse@l2s.centralesupelec.fr, alain.abergel@ias.u-psud.fr

Résumé – Cet article traite de la restauration du flux lumineux à partir de données multispectrales fournies par un imageur à bord d'un télescope spatial. Les problèmes abordés concernent la limitation de la résolution spatiale causée par la réponse optique variant spectralement et l'intégration spectrale de l'objet sur une large bande. Nous avons développé un modèle instrument prenant en compte ces effets et proposé un modèle direct qui exploite conjointement l'ensemble des données à différentes bandes spectrales. Nous avons mis en œuvre la restauration de l'objet inconnu en utilisant la méthode des moindres carrés régularisés et la solution est calculée par l'algorithme du gradient conjugué. Nous avons testé notre approche sur des données simulées de l'imageur de Mid-InfraRed Instrument (MIRI) à bord du futur télescope spatial James Webb (JWST). Notre méthode montre une nette augmentation des résolutions spatiale et spectrale par rapport aux méthodes conventionnelles.

Abstract – In this paper we deal with the restoration of astrophysical objects from multispectral data acquired by an imager on board of a space telescope. The problems we address are the spatial resolution limitation caused by the spectral varying optical response and the broadband spectral integration of the object. Indeed, we develop an instrument model that takes into account these effects, and we propose a direct model by exploiting all the data from different spectral band jointly. The restoration of the object is implemented using the regularized least squares method, and the solution is computed by the conjugate gradient algorithm. We test our approach on simulated imagery data from the Mid-InfraRed Instrument (MIRI) on board the future James Webb Space Telescope (JWST). Our method shows a significant increase in spatial and spectral resolution compared to conventional methods.

1 Introduction

Les instruments d'imagerie multispectrale sont utilisés dans de nombreux domaines, comme la télédétection, l'imagerie médicale ou l'astrophysique. Dans ce travail nous nous intéressons à l'imageur du Mid-InfraRed Instrument (MIRI)[1], à bord du futur télescope spatial James Webb (JWST) [1] de la NASA (en collaboration avec l'ESA), dont le lancement est prévu fin 2018. Le faisceau de lumière 3D (2 dimensions spatiales et 1 spectrale) à l'entrée du télescope est diffracté et focalisé sur le détecteur. Le résultat est modélisé par un produit de convolution spatiale 2D de l'objet avec la réponse impulsionnelle, ou PSF (Point Spread Function) [2]. Cette PSF floute l'objet différemment en fonction de la longueur d'onde, ce qui entraîne une réduction de résolution spatiale. De plus, la distribution spectrale de l'objet disparaît lors de l'intégration sur de larges fenêtres spectrales.

Les approches conventionnelles généralement négligent les variations spectrales de PSF. Certains travaux considèrent une PSF 2D mesurée, mais uniquement à des longueurs d'onde spécifiques [1, 3]. D'autres utilisent une PSF 2D à large bande en calculant la moyenne pondérée de PSFs monochromatiques [4, 5], mais la PSF est alors supposée spectralement invariante, conduisant à un modèle instrument inexact. D'autre part, [6] a traité un cas où la variation spectrale de PSF est faible et la PSF 3D approchée par deux PSFs spatialement et spectralement invariantes ; l'intégration spectrale du filtre + détecteur revient à une convolution spectrale. De façon comparable dans [7], l'interpolation de PSF est utilisée pour prendre en compte la variation spectrale de PSF, ce qui conduit à un modèle de convolution 3D.

Nous proposons une nouvelle approche pour reconstruire l'objet original 2D+λ à partir d'un modèle direct, prenant en compte l'intégration spectrale continue et préservant la variation en longueur d'onde exacte de la PSF. Notre contribution réside dans le développement du modèle instrument et du modèle direct et d'une première méthode de restauration. Nous suppo-

1. https ://jwst.nasa.gov/

sons que le spectre de l'objet est linéaire par morceaux, afin d'extraire son contenu spectral, mais également de conserver sans approximation les variations spectrales de PSF. Par ailleurs, nous traitons l'ensemble des données multispectrales disponibles issues de plusieurs filtres pour restaurer l'information sur la totalité de la bande spectrale de l'instrument. L'objet 3D est alors représenté par des paramètres 2D spatiaux, calculés en utilisant la méthode des moindres carrés régularisés.

Notre approche est testée sur des données simulées de MIRI. Les résultats de la restauration montrent un gain significatif en résolution spatiale et fournissent plus d'information spectrale que les méthodes utilisant une PSF à large bande.

2 Modèle instrument

L'objet d'intérêt à l'entrée du système optique est un flux lumineux $\phi(\alpha, \beta, \lambda) : \mathbb{R}^3 \rightarrow \mathbb{R}$ ayant deux dimensions spatiales $(\alpha, \beta) \in \mathbb{R}^2$ et une dimension spectrale $\lambda \in \mathbb{R}^+$. Durant le processus d'observation, l'objet est modifié par la réponse instrument, fournissant une sortie discrète 2D $\boldsymbol{y}^{(f)}$ pour chaque filtre $f \in [1, n_f]$ (n_f est le nombre total des filtres). On notera dans la suite N le nombre total de pixels de données.

Dans ce qui suit, nous allons modéliser la réponse instrument présentée sur le schéma en blocs dans la figure 1. Elle est composée de la réponse du système optique et du système de mesure de l'imageur, comprenant le filtre et le détecteur.

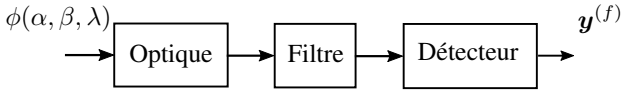


FIGURE 1 – Diagramme en blocs du modèle instrument

2.1 Réponse du système optique

Elle est modélisée par une convolution spatiale 2D de l'objet $\phi(\alpha, \beta, \lambda)$ avec une PSF variant en longueur d'onde :

$$\phi_{opt}(\alpha, \beta, \lambda) = \phi(\alpha, \beta, \lambda) \underset{(\alpha,\beta)}{*} h(\alpha, \beta, \lambda). \tag{1}$$

Afin d'illustrer la non-stationnarité spectrale de la PSF, la figure 2 présente des PSFs à des longueurs d'onde différentes du spectre moyen-infrarouge (en utilisant l'outil WebbPSF [8]). On observe clairement la variation spectrale de PSF, avec une largeur à mi-hauteur proportionnelle à λ.

2.2 Réponse du filtre

L'objet $\phi_{opt}(\alpha, \beta, \lambda)$ est ensuite spectralement filtré. Cette opération est modélisée par un produit avec la transmission du filtre $\tau_f(\lambda)$

$$\phi^{(f)}_{filt}(\alpha, \beta, \lambda) = \tau_f(\lambda)\ \phi_{opt}(\alpha, \beta, \lambda). \tag{2}$$

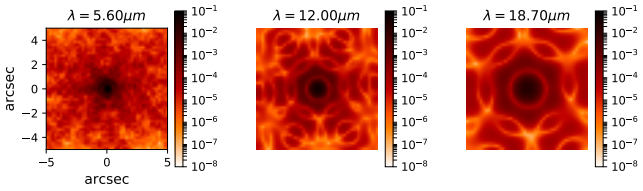


FIGURE 2 – Simulation de PSF monochromatique du JWST/MIRI

2.3 Réponse du détecteur

Le détecteur effectue une intégration spectrale et spatiale. L'intégration spectrale est pondérée par la réponse spectrale du détecteur (ou efficacité quantique) $\eta(\lambda)$. L'objet est intégré spatialement sur la matrice du détecteur 2D, la réponse pour un pixel est modélisé par

$$y^{(f)}(i,j) = \iint \int_{\mathbb{R}_+} \eta(\lambda)\phi^{(f)}_{filt}(\alpha, \beta, \lambda)d\lambda \\ B_{\text{int}}\left(\frac{\alpha - i\triangle_i}{\triangle_i}, \frac{\beta - j\triangle_j}{\triangle_j}\right) d\alpha\, d\beta, \tag{3}$$

où $\triangle_i \times \triangle_j$ est la surface du pixel (i, j). B_{int} est une fonction d'intégration sur la surface du pixel.

Le modèle complet de l'instrument est obtenu en substituant les équations (1)-(2) dans (3) :

$$y^{(f)}(i,j) = \iint \int_{\mathbb{R}_+} \eta(\lambda)\tau_f(\lambda)\, \phi(\alpha, \beta, \lambda) \underset{(\alpha,\beta)}{*} h(\alpha, \beta, \lambda) \\ B_{\text{int}}\left(\frac{\alpha - i\triangle_i}{\triangle_i}, \frac{\beta - j\triangle_j}{\triangle_j}\right) d\lambda d\alpha d\beta. \tag{4}$$

Ce modèle relie l'objet d'intérêt continu aux données discrètes à travers une réponse système complexe, incluant un fenêtrage spectral et cinq intégrations. Dans la suite nous avons fait deux hypothèses simplificatrices : (1) les pixels sont régulièrement disposés et de même surface $\triangle_i \times \triangle_j = \triangle^2$. (2) la constance de l'objet sur la surface des pixels, en choisissant pour fonction d'intégration B_{int} une fonction rectangle, ce qui permet de passer d'une convolution continue à une convolution discrète. Finalement, le modèle instrument devient

$$y^{(f)}(i,j) = \triangle^2 \int_{\mathbb{R}_+} \eta(\lambda)\tau_f(\lambda)\phi(i, j, \lambda) \underset{(i,j)}{*} h(i, j, \lambda)\, d\lambda. \tag{5}$$

3 Modèle direct et inversion

3.1 Modèle pour le spectre de l'objet

L'objectif est de reconstruire l'objet 2D+λ sur la totalité de la bande spectrale de l'instrument à partir de données multispectrales obtenues avec différents filtres. Pour cela, la paramétrisation de l'objet doit permettre des liens entre les filtres. Nous avons choisi de prendre pour chaque pixel un spectre

linéaire par morceaux avec une contrainte de continuité, tel qu'illustré sur la figure 3. Ce spectre s'écrit comme

$$\phi_{i,j}(\lambda) = \sum_{b=1}^{n_b} x_{i,j}^{(b)}\ g_+^{(b)}(\lambda) + x_{i,j}^{(b-1)}\ g_-^{(b)}(\lambda)\mathbb{1}_{\left[\lambda^{(b-1)},\lambda^{(b)}\right]}(\lambda) \tag{6}$$

avec $b \in [1, n_b]$ l'indice de la bande et $\mathbb{1}(\lambda)$ la fonction indicatrice sur $[\lambda^{(b-1)}, \lambda^{(b)}]$.

La fonction $\phi_{i,j}^{(b)}(\lambda)$ est linéaire

$$\phi_{i,j}^{(b)}(\lambda) = x_{i,j}^{(b)}\ g_+^{(b)}(\lambda) + x_{i,j}^{(b-1)}\ g_-^{(b)}(\lambda) \tag{7}$$

où $x_{i,j}^{(b)}$ est l'intensité du pixel (i,j) à $\lambda^{(b)}$ et $g_\pm^{(b)}(\lambda)$ la fonction contenant la variation linéaire sur la bande b :

$$g_\pm^{(b)}(\lambda) = \frac{1}{2} \pm \frac{\lambda - \lambda_c^{(b)}}{\lambda^{(b)} - \lambda^{(b-1)}}.$$

Ce modèle présente plusieurs avantages. Tout d'abord il permet une description linéaire du spectre plutôt que constante sans ajout de degré de liberté. La positivité sur toute la largeur de bande est respectée avec la positivité des intensités $x^{(b)}$. Enfin, plusieurs jeux de données correspondant à différents filtres dépendent des mêmes inconnues.

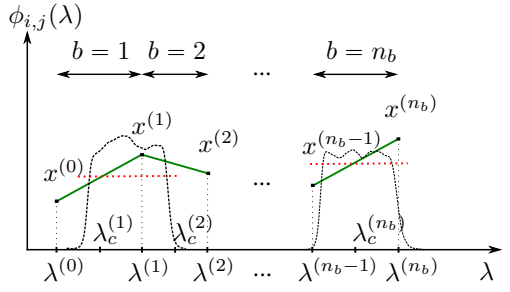


FIGURE 3 – Illustration du modèle linéaire par morceaux du spectre de l'objet (en continues). Spectre d'objet constant sur la bande spectrale du filtre (en pointillés).

3.2 Modèle direct

Le modèle direct pour un filtre f est obtenu en substituant l'équation (6) dans (5). En réarrangeant les termes de l'équation résultante, les termes dépendant de la longueur d'onde restent dans l'intégration spectrale :

$$y^{(f)}(i,j) = \sum_{b=0}^{n_b} h_{\text{int}}^{(f,b)}(i,j) \underset{(i,j)}{*} x^{(b)}(i,j) \tag{8}$$

avec

$$h_{\text{int}}^{(f,b)}(i,j) = \int \eta(\lambda)\tau_f(\lambda)\Big(g_+^{(b+1)}(\lambda) + g_-^{(b)}(\lambda)\Big) h(i,j,\lambda)\, d\lambda. \tag{9}$$

L'équation (8) montre que l'image obtenue pour le filtre f est la somme de $n_b + 1$ convolutions 2D où chaque image d'intensités $x^{(b)}$ est convoluée avec une réponse 2D, elle même somme pondérée des PSFs monochromatiques $h(\lambda)$.

La représentation matrice-vecteur de l'équation (8) s'écrit

$$\boldsymbol{y}^{(f)} = \sum_{b=0}^{n_b} \boldsymbol{H}_{\text{int}}^{(f,b)} \boldsymbol{x}^{(b)}, \tag{10}$$

où $\boldsymbol{y}^{(f)} \in \mathbb{R}^N$ est le vecteur de données, $\boldsymbol{x}^{(b)} \in \mathbb{R}^N$ l'ensemble des inconnues et $\boldsymbol{H}_{\text{int}}^{(f,b)} \in \mathbb{R}^N \times \mathbb{R}^N$ une matrice de convolution de réponse impulsionnelle $h_{\text{int}}^{(f,b)}(i,j)$. En définitive, le modèle direct de l'ensemble des données est linéaire et s'écrit

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\epsilon}, \tag{11}$$

où $\boldsymbol{\epsilon} \in \mathbb{R}^{n_f N}$ contient les erreurs liés à chaque observation. La matrice $\boldsymbol{H}$ est une matrice bloc de taille $n_f N \times (n_b+1)N$ où chaque bloc est une matrice de convolution.

3.3 Inversion

Nous estimons les intensités $\boldsymbol{x}$ dans l'équation (11) en minimisant le critère

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}}\, J(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \underbrace{\sum_{b=0}^{n_b} \mu_b \|\boldsymbol{D}\boldsymbol{x}\|_2^2}_{\|\boldsymbol{C}\boldsymbol{x}\|_2^2}. \tag{12}$$

Il s'agit d'une méthode de moindres carrés régularisés [9]. Le terme de régularisation $\|\boldsymbol{C}\boldsymbol{x}\|_2^2$ permet d'ajouter une information a priori sur l'objet à restaurer, dans le but de compenser le mauvais conditionnement de la matrice $\boldsymbol{H}$ et de stabiliser la solution des moindres carrés. Nous choisissons une norme quadratique pour obtenir une solution linéaire et tirer profit des moyens de calculs rapides de la solution.

Nous avons choisi un opérateur différentiel 2D pour $\boldsymbol{D}$ afin de promouvoir les solutions lisses, $\boldsymbol{C} = \text{diag}\{\mu_0\boldsymbol{D}, \ldots, \mu_{n_b}\boldsymbol{D}\}$ où $\mu_0, \ldots, \mu_b$ sont les paramètres de régularisation. Le critère $J(\boldsymbol{x})$ est quadratique et la solution $\hat{\boldsymbol{x}}$ est obtenue en annulant son gradient :

$$\hat{\boldsymbol{x}} = (\underbrace{\boldsymbol{H}^t\boldsymbol{H} + \boldsymbol{C}^t\boldsymbol{C}}_{\boldsymbol{Q}})^{-1}\boldsymbol{H}^t\boldsymbol{y} \tag{13}$$

où $\boldsymbol{Q}$ est une matrice de taille $(n_b+1)N \times (n_b+1)N$.

En fixant $n_b + 1 = n_f$ la matrice $\boldsymbol{H}$ devient carrée. Par conséquent, l'inversion $\boldsymbol{Q}^{-1}$ peut se faire par diagonalisation dans l'espace de Fourier, en faisant une approximation circulante des matrices de convolution et en inversant N matrices carrées de taille $n_b \times n_b$. En revanche, ce calcul s'avère lourd pour $n_b > 4$. Nous proposons plutôt le calcul de la solution par résolution du système linéaire $\boldsymbol{Q}\hat{\boldsymbol{x}} = \boldsymbol{H}^t\boldsymbol{y}$ à l'aide d'un algorithme itératif, sans inversion de matrice, comme le gradient conjugué [10].

4 Simulations et résultats

4.1 Application à l'imageur JWST/MIRI

Nous considérons l'imageur MIRI, et nous évaluons l'inversion sur des données simulées. Les PSF du JWST sont simulées

avec l'outil WebbPSF. MIRI comporte neuf bandes photométriques dans l'infrarouge moyen ($5\,\mu m - 28\,\mu m$) ; leurs profils spectraux $\eta(\lambda)\tau_f(\lambda)$ sont fournis dans [1].

4.2 Simulation

Afin de valider notre approche, nous avons effectué des simulations sur des données synthétiques. Nous avons donc simulé un cube pour l'objet original de taille $64 \times 64 \times 9$, ayant une source spatialement gaussienne (σ =3) et un spectre linéaire. Ensuite, neuf observations sont simulées pour les neuf bandes de MIRI. Toutes les données sont corrompues avec un bruit blanc gaussien centré avec un rapport signal sur bruit SNR $\approx$ 30 dB (défini dans [11] page 376). Nous comparons nos résultats aux travaux utilisant une PSF 2D à large bande [4, 5], appelés large bande ou LB dans la suite. Nous considérons donc neuf PSFs à large bande ne variant pas spectralement. Par conséquent, la restauration LB consiste à déconvoluer les données filtre par filtre indépendamment. Nous avons utilisé la méthode de déconvolution proposé dans [12]. Cette méthode est non-supervisée et permet d'estimer les paramètres de régularisations pour les deux approches de restaurations.

4.3 Résultats

Nous illustrons la restauration du contenu spectral à la figure 4 en affichant le spectre du pixel central de la source (normalisé à 1). Le spectre restauré par notre approche coïncide parfaitement au spectre original sur la totalité de la bande spectrale de l'instrument, tandis que la restauration LB ne permet pas de restaurer le spectre original. La figure 5 présente le contenu spatial restauré de l'objet original à la longueur d'onde $\lambda = 18.7\,\mu m$ contenue dans le filtre $f = 8$. Nous observons un effet de déconvolution de données et une restauration des intensités spatiales, avec une erreur relative de 6% au lieu des 21% pour l'approche LB. Cette différence apparaît bien dans les résidus affichés figure 5 (c) et (f).

5 Conclusion

Nous avons présenté le problème de restauration d'objet 2D+λ à partir de l'ensemble de données multispectrales 2D floues. Nous avons développé un modèle instrument de l'imageur prenant en compte la variation en longueur d'onde de la PSF et l'intégration spectrale sur de larges fenêtres spectrales. Nous avons ensuite développé un modèle direct linéaire, en modélisant le spectre de l'objet par une fonction continue, choisie linéaire par morceaux, ce qui nous a permis d'effectuer un traitement multi-filtre. Finalement, la solution du problème est obtenue en utilisant la méthode des moindres carrés régularisés. Les résultats de restauration préliminaires sont obtenus à partir de données simulées. Par rapport aux approches conventionnelles utilisant une PSF 2D à large bande, nous avons obtenu une augmentation significative de résolution spatiale et une meilleure reconstruction de l'information spectrale.

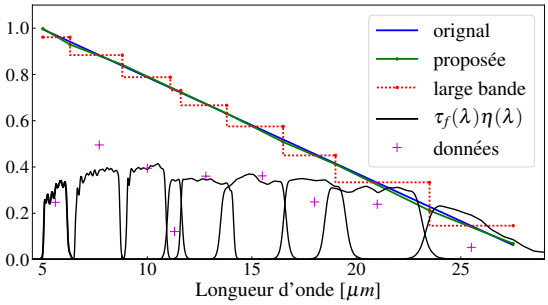


FIGURE 4 – Illustration de la reconstruction spectrale pour le pixel central de la source.

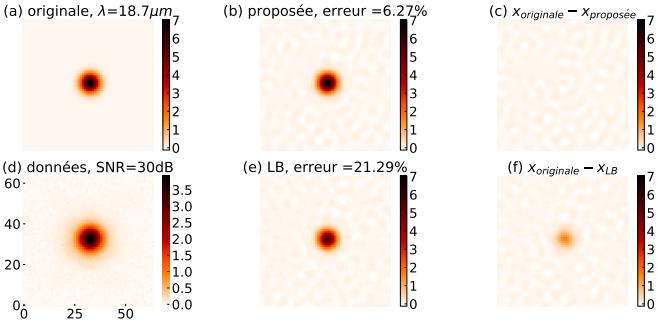


FIGURE 5 – Illustration de la restauration spatiale sur des coupes à $\lambda = 18.7\mu m$

Références

[1] P. Bouchet, M. García-Marín, P.-O. Lagage, J. Amiaux, J.-L. Auguéres, E. Bauwens, J. Blommaert, C. Chen, Ö. Detre, D. Dicken *et al.*, "The Mid-Infrared Instrument for the James Webb Space Telescope, III : MIRIM, The MIRI Imager," *Publications of the Astronomical Society of the Pacific*, vol. 127, no. 953, p. 612, 2015.

[2] J. W. Goodman, *Introduction to Fourier optics*. Roberts and Company Pub., 2005.

[3] P. Guillard, T. Rodet, S. Ronayette, J. Amiaux, A. Abergel, V. Moreau, J. Augueres, A. Bensalem, T. Orduna, C. Nehmé *et al.*, "Optical performance of the jwst/miri flight model : characterization of the point spread function at high resolution," in *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics, 2010, pp. 77 310J–77 310J.

[4] G. Aniano, B. Draine, K. Gordon, and K. Sandstrom, "Common-resolution convolution kernels for space-and ground-based telescopes," *Publications of the Astronomical Society of the Pacific*, vol. 123, no. 908, p. 1218, 2011.

[5] N. Geis and D. Lutzn, "Herschel/PACS modelled point-spread functions," 2010.

[6] E. Villeneuve and H. Carfantan, "Nonlinear deconvolution of hyperspectral data with mcmc for studying the kinematics of galaxies," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4322–4335, Oct 2014.

[7] F. Soulez, E. Thiébaut, and L. Denis, "Restoration of hyperspectral astronomical data with spectrally varying blur," *EAS Publications Series*, vol. 59, pp. 403–416, 2013.

[8] M. D. Perrin, R. Soummer, E. M. Elliott, M. D. Lallo, and A. Sivaramakrishnan, "Simulating point spread functions for the James Webb Space Telescope with WebbPSF," in *SPIE Astronomical Telescopes+ Instrumentation*. International Society for Optics and Photonics, 2012, pp. 84 423D–84 423D.

[9] J. Idier, *Bayesian approach to inverse problems*. John Wiley & Sons, 2013.

[10] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," 1994.

[11] R. C. Gonzalez and R. E. Woods, "Image processing," *Digital image processing*, vol. 2, 2007.

[12] F. Orieux, J.-F. Giovannelli, and T. Rodet, "Bayesian estimation of regularization and point spread function parameters for wiener–hunt deconvolution," *JOSA A*, vol. 27, no. 7, pp. 1593–1607, 2010.

A&A 596, A63 (2016)
DOI: 10.1051/0004-6361/201629080

Astronomy & Astrophysics

Convolution kernels for multi-wavelength imaging

A. Boucaud[1,2], M. Bocchio[1], A. Abergel[1], F. Orieux[1,3], H. Dole[1], and M. A. Hadj-Youcef[1,3]

[1] Institut d'Astrophysique Spatiale, CNRS, UMR 8617, Univ. Paris-Sud, Université Paris-Saclay, IAS, Bât. 121, Univ. Paris-Sud, 91405 Orsay, France
e-mail: alexandre.boucaud@ias.u-psud.fr
[2] Sorbonne Universités, UPMC Univ Paris 6 et CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98bis bd Arago, 75014 Paris, France
[3] Laboratoire des Signaux et Systèmes (Univ. Paris-Sud, CNRS, CentraleSupélec, Université Paris-Saclay), 91192 Gif-sur-Yvette, France

Received 9 June 2016 / Accepted 5 September 2016

ABSTRACT

Astrophysical images issued from different instruments and/or spectral bands often require to be processed together, either for fitting or comparison purposes. However each image is affected by an instrumental response, also known as point-spread function (PSF), that depends on the characteristics of the instrument as well as the wavelength and the observing strategy. Given the knowledge of the PSF in each band, a straightforward way of processing images is to homogenise them all to a target PSF using convolution kernels, so that they appear as if they had been acquired by the same instrument. We propose an algorithm that generates such PSF-matching kernels, based on Wiener filtering with a tunable regularisation parameter. This method ensures all anisotropic features in the PSFs to be taken into account. We compare our method to existing procedures using measured *Herschel*/PACS and SPIRE PSFs and simulated JWST/MIRI PSFs. Significant gains up to two orders of magnitude are obtained with respect to the use of kernels computed assuming Gaussian or circularised PSFs. A software to compute these kernels is available at https://github.com/aboucaud/pypher

Key words. methods: observational – techniques: image processing – telescopes – techniques: photometric

1. Introduction

The point-spread function (PSF), also known as *beam*, is one of the main characteristics of any astronomical imager. It is a model of the diffraction pattern resulting from the interaction between the electromagnetic radiation and the instrument optics and detectors at every wavelength. Since most instruments operate on a single or a series of bandpasses (through e.g. filters), the resulting effective PSF is an integral of the monochromatic PSFs over the wavelength range, weighted by the instrumental throughput and the source energy distribution of a given astronomical object. A more accurate model can even include convolutional effects such as guiding errors, trailing effects from a scanning mode, smearing by the detector response, or even non-convolutional effects like the brighter-fatter effect. Once imaged, these model PSFs exhibit a complex shape, including anisotropy, wings, and spikes that extend far from the centre. Another classic feature of the PSF derived from the laws of optics is the radially oscillating pattern of the response, especially in the monochromatic case, creating a series of peaks and valleys. These secondary peaks can account for a non-negligible amount of the total power of the PSF. For ground-based astronomy, however, the atmospheric turbulence creates a smearing that redistributes the power of these peaks and valleys and enables the PSF to be modelled by simple analytic profiles such as two-dimensional (2D) Gaussians. On the contrary, space telescopes can benefit from a much higher resolution at the expense of a full complexity of the PSF. To cite a few examples, the effective PSF of IRAS maps was elliptical owing to the scanning strategy, so the angular resolution was strongly anisotropic, with ratios up to 1:6 (e.g. $0.75' \times 4.6'$ at $25\,\mu$m, from Wheelock et al. 1994). More recently, the effective PSFs of the *Planck*/HFI[1] maps appeared to have an ellipticity in the range 1.04 to 1.4, depending on the spectral band (Ade et al. 2011); the PSF of the PACS[2] photometer (Poglitsch et al. 2010) on board the *Herschel* satellite, characterised by Lutz (2012), showed a narrow core, a tri-lobe pattern and knotty structured diffraction rings. As we push the boundaries of both optical performances and detector capabilities of future missions, optical designs highly increase in complexity. Hence, for upcoming space surveys (Euclid[3], WFIRST[4]) or observatories (Athena[5] or JWST[6]), the characterisation and processing of elaborated PSFs become a crucial task.

Most astrophysical studies necessitate multi-wavelength observations, either from multiple bands or filters within an instrument or from various instruments and telescopes. The different maps are affected by a particular PSF and the pixel-based data comparison cannot be straightforward. However, a technique widely used in multi-band photometry is to perform the measurements on PSF homogenised data, that is to select a dataset as reference (usually the one with the worst resolution, or wider PSF) and transform the other datasets so they are PSF-matched with the reference PSF; this technique is called PSF homogenisation or PSF matching (see e.g. Bertin et al. 2002; Gordon et al. 2008; Darnell et al. 2009; Desai et al. 2012; Hildebrandt et al. 2012). Usually, PSF homogenisation is achieved by convolving the

[1] http://planck.esac.esa.int
[2] http://herschel.esac.esa.int
[3] http://www.euclid-ec.org/
[4] http://wfirst.gsfc.nasa.gov/
[5] http://www.the-athena-x-ray-observatory.eu/
[6] http://http://www.jwst.nasa.gov/

image with a kernel that is generated from the PSF corresponding to the image and the reference PSF. In the literature, one can distinguish between parametric kernels and non-parametric methods. Parametric methods use a fit of an analytic model to each PSF (Moffat, multiple Gaussians, etc.) or their decomposition on a proper basis (e.g. Gauss-Hermite polynomials or shapelets), and result in an analytic expression of the kernel (e.g. Kuijken 2008; Hildebrandt et al. 2012; Liu et al. 2015); non-parametric methods use pixel information from the image (e.g. Alard 2000) or adopt effective PSF images (e.g. Gordon et al. 2008; Aniano et al. 2011) to compute the kernels.

With the purpose of taking the full complexity and angular extension of the PSFs of space instruments into account, we address the creation of PSF-matching kernels for multi-wavelength studies. We then present two use cases for these kernels: one based on the *Herschel* satellite data and a second on simulations for the Mid-InfraRed Instrument Imager (MIRI) of the *James Webb* Space Telescope (JWST). We also deliver a programme called `pypher` that computes the kernels given two PSF images (see Appendix A). This code has initially been developed in preparation for the Euclid mission (Laureijs et al. 2010).

In Sect. 2, we describe the algorithm for the generation of convolution kernels used to match the resolution of images. In Sect. 3, we assess the improvement brought by these kernels on the multi-wavelength study of dust properties with the *Herschel* satellite, and show in Sect. 4 the reconstruction power of such kernels on PSF simulations of the future JWST satellite, before summarising this work in Sect. 5.

2. Kernel generation

2.1. Data model

We first consider an astrophysical image $\boldsymbol{y}$, observed with an instrument modelled as a linear invariant system,

$$\boldsymbol{y} = \boldsymbol{h} * \boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{h}$ is the PSF convolved with the unknown sky $\boldsymbol{x}$, $\boldsymbol{n}$ is the image noise and $*$ stands for the discrete convolution (see e.g. Gonzalez & Woods 2008).

Given two PSF models $\boldsymbol{h}_a$ and $\boldsymbol{h}_b$, where a and b refer to different frequency bands from the same or various instruments, the process we are interested in, referred to as PSF-matching, is to transform the image $\boldsymbol{y}_a$ acquired at the angular resolution of $\boldsymbol{h}_a$

$$\boldsymbol{y}_a = \boldsymbol{h}_a * \boldsymbol{x} + \boldsymbol{n}_a, \tag{2}$$

into an image $\boldsymbol{y}_{a,b}$ at the angular resolution of $\boldsymbol{h}_b$

$$\begin{aligned} \boldsymbol{y}_{a,b} &= \boldsymbol{k}_{a,b} * \boldsymbol{y}_a, & (3)\\ &\simeq \boldsymbol{h}_b * \boldsymbol{x} & (4) \end{aligned}$$

where $\boldsymbol{k}_{a,b}$ is the matching kernel from $\boldsymbol{h}_a$ to $\boldsymbol{h}_b$.

This paper presents a linear algorithm that computes the kernel $\boldsymbol{k}_{a,b}$ to produce the image $\boldsymbol{y}_{a,b}$ from the original image $\boldsymbol{y}_a$ through a convolution. To this end, we need to construct the kernel $\boldsymbol{k}_{a,b}$ such that

$$\boldsymbol{h}_b = \boldsymbol{h}_a * \boldsymbol{k}_{a,b}. \tag{5}$$

2.2. Kernel generation

For such linear systems as Eq. (5), one can seek an estimate of $\boldsymbol{k}_{a,b}$, denoted $\hat{\boldsymbol{k}}_{a,b}$, which minimises the least squares criterion J

$$\begin{aligned} \hat{\boldsymbol{k}}_{a,b} &= \arg\min_{\boldsymbol{k}_{a,b}} J(\boldsymbol{k}_{a,b}), & (6)\\ &= \arg\min_{\boldsymbol{k}_{a,b}} \left\| \boldsymbol{h}_b - \boldsymbol{h}_a * \boldsymbol{k}_{a,b} \right\|^2 . & (7) \end{aligned}$$

However, the presence of the convolution makes the system ill-posed for the inversion, hence the solution to Eq. (7) is not stable. The only way to stabilise the solution is to add information. For the considered system, we use a technique called *regularisation*. We choose a ℓ_2 norm to have a linear estimator and use Fourier filtering; and penalise the high frequencies in which we expect the noise to dominate, using a high-pass filter $\boldsymbol{d}$. This corresponds to adding a relative degree of smoothness between values of neighbouring pixels.

$$J(\boldsymbol{k}_{a,b}) = \left\| \boldsymbol{h}_b - \boldsymbol{h}_a * \boldsymbol{k}_{a,b} \right\|^2 + \mu \left\| \boldsymbol{d} * \boldsymbol{k}_{a,b} \right\|^2 , \tag{8}$$

where $\boldsymbol{d}$ is the second-order differential operator (i.e. 2D Laplacian matrix)

$$\boldsymbol{d} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{9}$$

and μ the regularisation parameter, which tunes the balance between the data fidelity and the penalisation. Other norms, such as ℓ_1, $\ell_2\ell_1$ or TV (Total Variations) are known to better preserve the image details, but produce non linear estimators that require iterative algorithms to solve.

Denoting the Fourier transform of any two-dimensional vector $\boldsymbol{u}$ by $\tilde{\boldsymbol{u}}$, the convolution theorem states that the real-space convolution is equivalent to a termwise product in Fourier space

$$\boldsymbol{h} * \boldsymbol{k} \Leftrightarrow \tilde{\boldsymbol{h}} \odot \tilde{\boldsymbol{k}}, \tag{10}$$

where $\odot$ symbolises the termwise product between vectors/matrices.

Under these assumptions, the cancellation of the first gradient of the criterion (8) leads to the classical regularised mean square solution of Eq. (7) in Fourier space

$$\tilde{\hat{\boldsymbol{k}}}_{a,b} = \boldsymbol{w} \odot \tilde{\boldsymbol{h}}_b \tag{11}$$

where $\boldsymbol{w}$ is a Wiener filter with high-frequency penalisation

$$\boldsymbol{w}(\mu) = \frac{\tilde{\boldsymbol{h}}_a^{\dagger}}{|\tilde{\boldsymbol{h}}_a|^2 + \mu |\tilde{\boldsymbol{d}}|^2}; \quad \mu \neq 0 \tag{12}$$

and $\tilde{\boldsymbol{h}}_a^{\dagger}$ stands for the complex conjugate of matrix $\tilde{\boldsymbol{h}}_a$.

The real-space convolution kernel $\boldsymbol{k}_{a,b}$, is eventually obtained via the inverse Fourier transform of Eq. (11). For two instruments a and b, this kernel is thus only parametrised by the regularisation parameter μ. The optimal balance between the data and the penalisation is found by setting μ to the inverse of the signal-to-noise ratio (S/N) of the homogenised image, $\boldsymbol{y}_a$ in this case.

We provide with this paper the `pypher` programme (see Appendix A); an implementation of the algorithm 1 presented below, which computes $\boldsymbol{k}_{a,b}$. We note that the optical transfer function (OTF) that appears in the algorithm is the discrete Fourier transform of a signal that has been translated so that its peak value is the first vector entry (i.e. Im[0, 0] for an image).

3. Impact on *Herschel* data analysis

The `pypher` kernels allow us to convolve multiple astronomical images to a common angular resolution. In order to preserve all the information during this process, a good knowledge of the effective PSFs of the instruments used is required. We focus here on the particular analysis of *Herschel* photometric images that have been widely used in the last few years to measure the dust temperature and the spectral index β across many astronomical objects from multi-band imaging with PACS and SPIRE.

Given the uncertainties on the PSFs of PACS and SPIRE instruments, since the beginning of the operational period of *Herschel* and in recent years, convolutions from PACS to SPIRE angular resolution have been performed assuming Gaussian PSFs with a given FWHM estimated from dedicated observations of asteroids. In 2011, a better characterisation of the PSF of the instruments allowed Aniano et al. (2011, ADGS11) to develop a method to construct convolution kernels assuming circular PSFs. While a full analysis of the PSF of the SPIRE instrument has been finalised recently by the Instrument Control Centre (ICC, Schultz 2015), a final PSF characterisation for the PACS instrument has not been released yet by the ICC (Lutz 2015). Parallel work has been performed by Bocchio et al. (2016), who computed PACS effective PSFs[7] from the combination of Vesta and Mars dedicated observations, which will be used hereafter. The current knowledge of both PACS and SPIRE PSFs, and the `pypher` code allow us to construct effective convolution kernels with an unprecedented precision. In this section we show how and to what extent the use of different convolution kernels can affect the results.

3.1. Herschel PSFs and kernels

For each *Herschel* band, λ, we define four PSF images:

1. The effective PSF (E_λ), taken from Bocchio et al. (2016) for PACS and Schultz (2015) for SPIRE.
2. The Gaussian PSF (G_λ), computed fitting a 2D Gaussian profile to E_λ. The PSF G_λ has the same FWHM as E_λ but does not account for secondary lobes and faint structures.
3. The circular PSF (C_λ), computed from E_λ by averaging the image intensity in annular bins. All information on the asymmetry of the PSF is then lost.
4. The Aniano PSF (A_λ), a circular PSF from the ADGS11 paper.

For each of the first three types of PSF, we define the corresponding matching kernels between the bands $\lambda_1$and λ_2 as

$$K^X_{\lambda_1,\lambda_2} \quad \text{with} \quad X \in \{E, G, C\} \tag{13}$$

and compute them with `pypher`. We also consider the Aniano kernels defined here as $K^A_{\lambda_1,\lambda_2}$ and computed in ADGS11 via a different method than that used in this work.

The homogenised images can therefore be denoted by

$$X_{70,350} = E_{70} * K^X_{70,350}. \tag{14}$$

[7] These PSFs are publicly available at `http://idoc-herschel.ias.u-psud.fr/sitools/client-user/Herschel/project-index.html`.

Algorithm 1: Matching kernel generation recipe

inputs: $\boldsymbol{h}_a$ 2D array of size $N_a \times N_a$ and pixel scale p_a,
$\boldsymbol{h}_b$ 2D array of size $N_b \times N_b$ and pixel scale p_b,
angles α_a and α_b (see Appendix A)
regularisation factor μ.
output: $\boldsymbol{k}_{a,b}$ 2D array of size $N \times N$ and pixel scale p.

`/* PSF warping: */`

$N = N_b$; $p = p_b$

for *i in {a, b}* **do**
  if $\alpha_i \neq 0$ **then**
    rotate $\boldsymbol{h}_i$ through an angle α_i
  end
end

if $p_a \neq p$ **then**
  rescale $\boldsymbol{h}_a$ to the pixel scale p
end

for $\boldsymbol{u}$ *in {*$\boldsymbol{h}_a$, $\boldsymbol{d}$*}* **do**
  if *size(*$\boldsymbol{u}$*)* $< N \times N$ **then**
    pad $\boldsymbol{u}$ with zeros to a size of $N \times N$
  else
    trim $\boldsymbol{u}$ to a size of $N \times N$
  end
end

`/* Wiener filter: */`

for $\boldsymbol{u}$ *in {*$\boldsymbol{h}_a$, $\boldsymbol{d}$*}* **do**
  compute the OTF[7] of $\boldsymbol{u}$: $\tilde{\boldsymbol{u}}$
end
compute $\boldsymbol{w}(\mu)$ following Eq. (12)

`/* Kernel: */`

compute the discrete Fourier transform of $\boldsymbol{h}_b$: $\tilde{\boldsymbol{h}}_b$
compute $\tilde{\tilde{\boldsymbol{k}}}_{a,b}$ following Eq. (11) via a termwise product
inverse Fourier transform $\tilde{\tilde{\boldsymbol{k}}}_{a,b}$ to obtain the kernel $\boldsymbol{k}_{a,b}$

3.2. Kernel comparison

In this paragraph we assess the impact of using approximations of effective PSFs in the homogenisation process, which is directly related to the creation of the convolution kernel. Because the effective kernel $K^E_{\lambda_1,\lambda_2}$ is the target of our kernel generation algorithm, it is expected to produce better results than the other types described above.

To compare the different types of kernels, we chose to measure the difference between the effective PSF at λ_1 matched to λ_2 and the effective PSF at λ_2. We define the relative residuals R^X as

$$R^X_{\lambda_1,\lambda_2} = \frac{|E_{\lambda_2} - E_{\lambda_1} * K^X_{\lambda_1,\lambda_2}|}{E_{\lambda_2}} \tag{15}$$

for each type of kernel.

For this comparison test, we consider the matching of the PACS 70 μm PSF to the resolution of SPIRE 350 μm. In the remainder of this paragraph, we refer to these bands as 70 and 350.

The two images on the left of Fig. 1 represent the effective PSFs of PACS 70 μm E_{70} and SPIRE 350 μm E_{350}. The central column shows the kernels computed with `pypher` from Gaussian (top) and circular (middle) approximations of the effective PSFs, as described in Sect. 3.1, and those computed directly from the

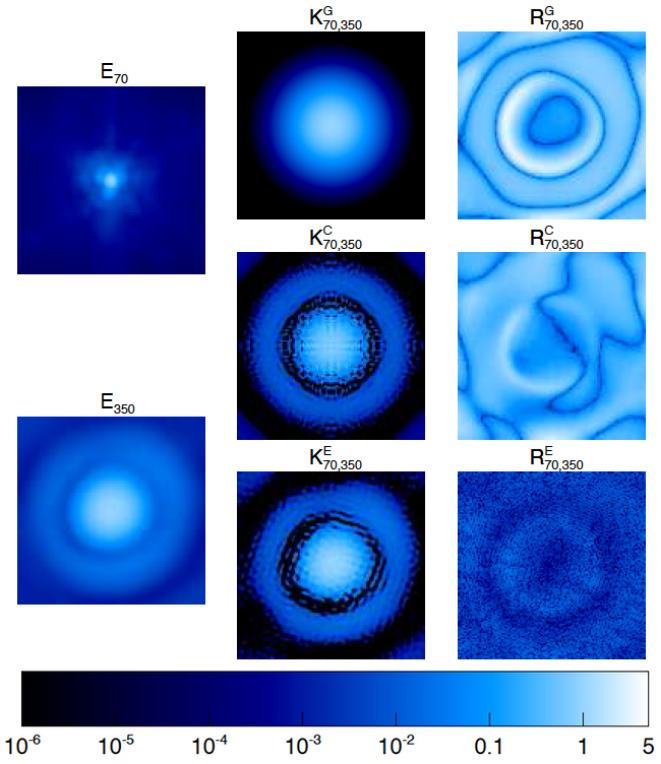


Fig. 1. Effective PACS 70 μm and SPIRE 350 μm PSFs on the *left*. The *central column* shows Gaussian, circular, and effective kernels for these PSFs computed following the procedure described in Sect. 3.1. The corresponding relative residual images (see Eq. (15)) are shown in the *right column*. The *bottom* kernel and residual image correspond to this work. All images are $120'' \times 120''$.

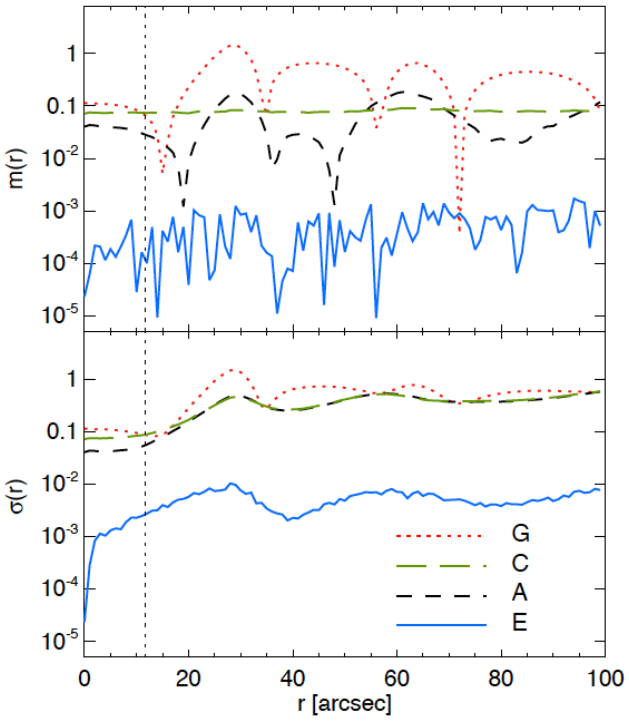


Fig. 2. First ($m(r)$, *top panel*) and second ($\sigma(r)$, *bottom panel*) polar moments of the residuals (right column images of Fig. 1 + $R^A_{70,350}$) as a function of the distance r to the image centre. The different lines represent the kernel types used for the homogenisation: Gaussian (red dotted), circular (green long dashed), ADGS11 (black dashed) and effective (blue solid, this work). The vertical black dotted line indicates the HWHM of the SPIRE 350 μm PSF.

left-side PSFs (bottom). On these kernel images, one can see the characteristics of the input PSFs: the Gaussian approximation has a single lobe, the circular approximation is axisymmetrical and presents a second lobe, and the last approximation has two lobes and the general shape of E_{350}. Next to these kernels, the associated homogenisation residual images are shown (see Eq. (15) for computation).

Both $R^G_{70,350}$ and $R^C_{70,350}$ have residuals of the order of 10% within the first lobe of E_{350} (central region of radius equal to the half width at half maximum (HWHM) $\simeq 12''$). Outside of that region, the reconstruction from these two kernels is even worse. In particular, the extinction ring that marks the transition between the first and second lobe does not exist in the Gaussian case and is slightly shifted owing to azimuthal averaging in the circular case, which leads to a big residual error in both cases (white circle on $R^G_{70,350}$ and $R^C_{70,350}$). Both kernel and residual images from ADGS11, which are not shown in Fig. 1, present very similar behaviour to the circular approximation. Using the kernel constructed with effective PSFs, $R^E_{70,350}$, we obtain very homogeneous residuals of the order of 0.1%.

To analyse these residuals in more detail, we introduce the first m and second σ polar moments of the residuals as follows:

$$m(r) = \frac{1}{n_{\rm ang}} \sum_{i=1}^{n_{\rm ang}} R^X(r, \theta_i),$$
$$\sigma(r) = \sqrt{\frac{1}{n_{\rm ang}} \sum_{i=1}^{n_{\rm ang}} [R^X(r, \theta_i)]^2}, \quad (16)$$

where $R^X(r, \theta_i)$ is the intensity of the residual image (computed using the kernel X) at a distance r from the image centre and at an angle $\theta_i = 2\pi i / n_{\rm ang}$, with $n_{\rm ang} = 100$. The moments express the intensity and dispersion of the residuals along the PSF radius.

These two values, computed on the residual images for the four kernel types, are shown in Fig. 2. As previously stated, the Gaussian case (dotted red lines) is only stable within a circle of radius equal to the HWHM of E_{350}. At further distance it shows very high first and second moments close to unity, and establishes very poor matching. The circular case (long-dashed green) exhibits a constant first moment below 10% at all distances, which is better than the Gaussian case. This is mainly due to the computation of the first moment that is very similar to the circularising process and averages out the measurements. However, the asymmetry and local structures of E_{70} and E_{350} are lost and the second moment is comparable to that of $R^G_{70,350}$. To produce the kernels, ADGS11 used narrower versions of PACS and SPIRE PSFs than those used in this work. Both first and second moments (dashed black) therefore present bumps at the position of the lobes. Finally, the effective case (blue) best matches E_{350}. In amplitude, the first moment is $\lesssim$0.1% and the second moment $\lesssim$1% at all distances, even if we note that the second moment presents the two bumps observed earlier. By comparison, these moments are two orders of magnitude lower than those of the other three cases.

This concludes in a major improvement in using `pypher` kernels with effective PSFs with respect to Gaussian, circular or ADGS11 kernels from an image processing standpoint. Next we test again these kernels on the determination of meaningful parameters from data or simulations.

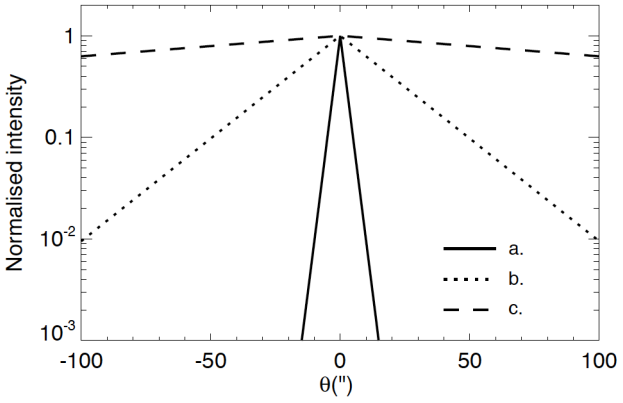


Fig. 3. Intrinsic dust profiles (see Eq. (17)) for scenarios *a.*, *b.* and *c.* listed in Table 1, corresponding to characteristic scale heights of z_d = 0.1, 1 and 10 kpc, respectively. These profiles are then convolved with *Herschel* PSFs to simulate an edge-on galaxy observed in different bands.

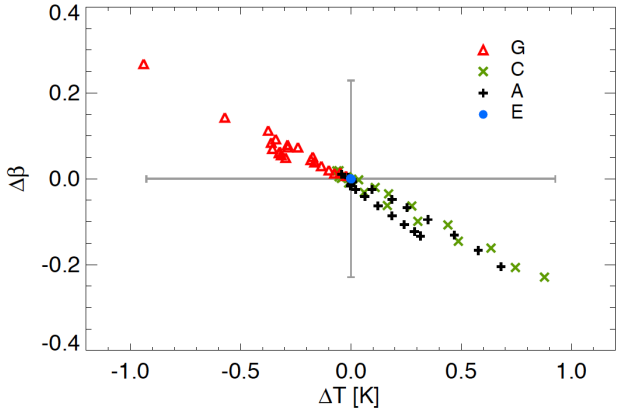


Fig. 4. Parameter $\Delta\beta$ as a function of ΔT for the four kernel types, using an edge-on galaxy with the dust profile *b.* (z_d = 1 kpc, $S/N = 10^4$). For a given kernel type, there are 21 data points, each of which correspond to a measurement on a single pixel along the galaxy profile. The spread thus represents the systematic error on the $\beta - T$ measurement for this kernel type. Error bars on both axes at the centre indicate the statistical errors on β and T obtained from the χ^2-fitting routine.

Table 1. Scale height, z_d, and signal-to-noise ratio (S/N) adopted for the simulation of dust profiles of an edge-on galaxy.

Scenario	z_d (kpc)	S/N
a.	0.1	10^4
b.	1.0	10^4
c.	10.0	10^4
b.1	1.0	10^2
b.2	1.0	10^1

3.3. Dust properties study

Herschel observations are often used to retrieve information about dust properties in our Galaxy and in local galaxies. In this section we show how the choice of PSFs to construct the convolution kernels can affect pixel-by-pixel measurements.

We consider an edge-on galaxy at $D \sim 10$ Mpc from us with an intrinsic vertical profile given by

$$I_d(z) = I_d(0) \exp\left(\frac{-z}{z_d}\right), \quad (17)$$

where z_d is the scale height of the vertical dust distribution. We introduce the angular distance $\theta \simeq z/D$ and its characteristic value $\theta_d \simeq z_d/D$. Three scale heights are examined, z_d = 0.1, 1 and 10 kpc (scenarios *a.*, *b.* and *c.* from Table 1), corresponding to $\theta_d \simeq 2, 20$ and $200''$. Their intrinsic dust profile is shown in Fig. 3.

We convolve the intrinsic dust profile to the effective PSFs of PACS 70, 100 and 160 μm and SPIRE 250 and 350 μm, while keeping the pixel size at $1''$. We then make the basic assumption that the dust in the whole galaxy has a temperature of T_d = 20 K and spectral index of β = 1.6 (following Planck Collaboration XI 2014), and rescale the convolved models accordingly.

In order to simulate real data, we add Gaussian statistical noise to the models and consider three different values of S/N, 10^4, 10^2, 10 (scenarios *b.*, *b.1* and *b.2* from Table 1), with respect to the dust emission at the peak position (z = 0). These images are then homogenised to the resolution of SPIRE 350 μm using the four kernel types described in Sect. 3.1 and resampled with a common pixel size of $10''$.

Finally, using a minimum χ^2 method, we fit the multi-wavelength data on a pixel-by-pixel basis to a modified blackbody

$$I_\nu = \tau_{\nu_0} (\nu/\nu_0)^\beta B_\nu(T), \quad (18)$$

where τ_{ν_0} is the optical depth at the reference frequency ν_0 and $B_\nu(T)$ is the blackbody radiation for a grain at temperature T.

We define $\Delta\beta$ and ΔT as the deviations from the reference dust spectral index β and temperature T_d, respectively. Figure 4 shows $\Delta\beta$ as a function of ΔT for scenario *b.* where z_d = 1 kpc and $S/N = 10^4$, for the different kernel types. Except for the case where effective PSFs are used, systematic discrepancies are present (up to $\Delta T \pm 1$ K and $\Delta\beta \pm 0.3$), which are comparable in amplitude to the statistical errors (horizontal and vertical bars in Fig. 4), and a spurious strong negative correlation appears between dust temperature and β parameter.

In order to show where this effect is most significant within the galaxy profile, we illustrate in Fig. 5 the quantities ΔT and $\Delta\beta$ at various angular distances from the galaxy centre and for different kernel types (G: Gaussian; C: circular; A: ADGS11; E: effective). Each column thus represents a vertical cut of the modelled galaxy. Dashed lines indicate the characteristic scale height of the intrinsic dust abundance profile. The top three panels show scenarios *a.*, *b.* and *c.* where the scale height varies and the S/N is kept constant at 10^4. The two bottom panels show scenarios *b.1* and *b.2* where the scale height is fixed at 1 kpc and we vary the S/N (see Table 1 for a summary). Each panel of Fig. 5 shows the relationship between $\Delta\beta$ and ΔT just as in Fig. 4, and adds the spatial information to the data points.

Depending on the considered dust scale height, deviations from the reference dust temperature and β parameter reach ± 2 K and ± 0.6, respectively, with higher deviations for shorter scale heights and at higher galactic latitudes. Regardless of the convolution kernel adopted, the negative correlation is clearly measured for all the pixels in the vertical cut.

Scenarios *a.*, *b.* and *c.* show that very low (<1%) deviations from the reference values are obtained using the effective kernels, while the use of other convolution kernels leads to larger errors for $z \gtrsim z_d$. However, as expected, in decreasing the level of S/N (scenarios *b.1* and *b.2*), the noise starts to dominate over

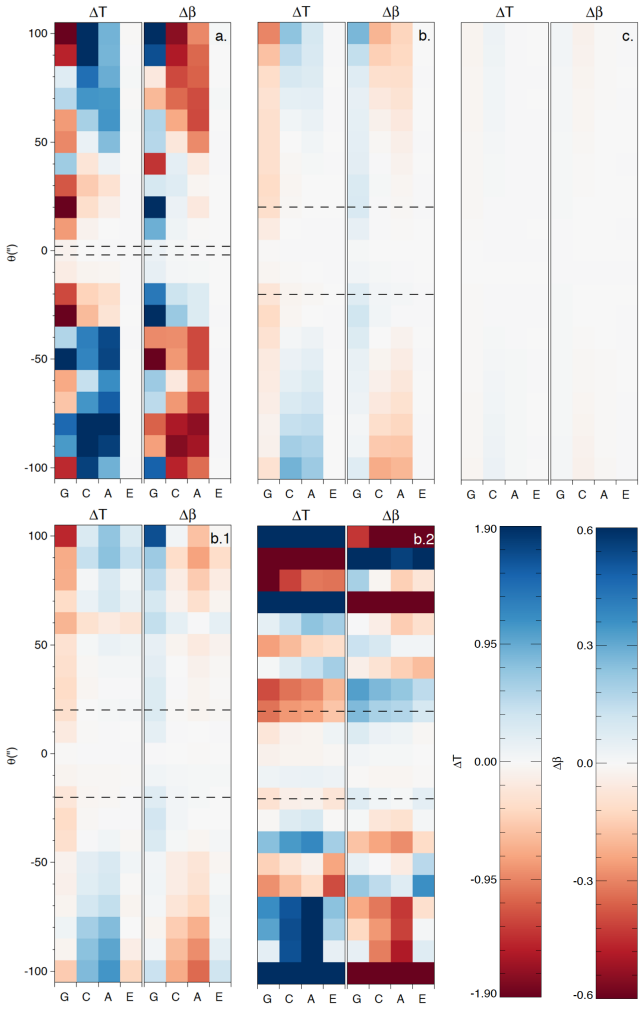


Fig. 5. Heat maps of ΔT and $\Delta\beta$ measured on a simulated edge-on galaxy at various angular distances θ from the centre, as a function of the kernel type used. Each panel represents a scenario from Table 1. Dashed lines indicate the scale height of the dust profile z_d. This figure highlights the very small scatter in both ΔT and $\Delta\beta$ when using the kernels with effective PSFs (column E on each plot, until the S/N becomes to low as in panel **b.2**).

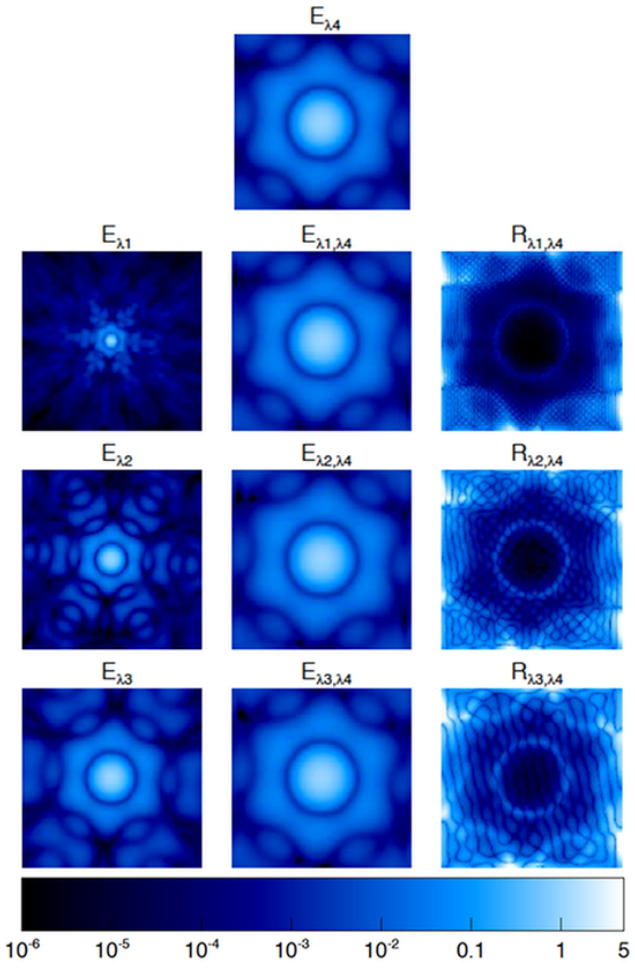


Fig. 6. Proof of concept of PSF homogenisation for the JWST/MIRI instrument. The PSFs from the *first column* at 5.6 μm, 11.3 μm and 18.8 μm, respectively, are homogenised to the PSF on *top of the second column* at 25.5 μm, using `pypher` kernels. The resulting homogenised PSFs are shown in the *second column*. They are visually indistinguishable from the effective one on top. The relative residuals on the *right* confirm that the central part of the PSF is reconstructed at least to 0.1%.

the signal and very large discrepancies are observed in temperature and β, regardless of the adopted convolution kernel.

4. JWST PSF simulations

To show the reliability of our method for complex-shaped PSFs, we now test our algorithm on simulated PSFs from a telescope with an uncommon optical design. For this purpose, we select the Mid-InfraRed Instrument Imager (MIRI) of the *James Webb* Space Telescope (JWST, Bouchet et al. 2015). As the JWST main mirror is made of several hexagonal mirrors, the optical response of this instrument shows highly non-symmetrical features that need to be accounted for in the homogenisation process.

We used WebbPSF (Perrin et al. 2012), the official JWST PSF simulation tool to generate a set of four broadband PSFs of $5'' \times 5''$ centred at $\lambda_1 = 5.6\,\mu$m, $\lambda_2 = 11.3\,\mu$m, $\lambda_3 = 18.0\,\mu$m and $\lambda_4 = 25.5\,\mu$m in order to cover the spectral range of MIRI. These broadband PSFs were generated from a set of 20 monochromatic PSFs assuming flat spectral energy distribution of the source, and oversampled at four times the pixel size of the detector, corresponding to a pixel scale of 0.11 arcsec. They are shown in the first column (and top of the second column) of Fig. 6.

Using `pypher`, we then computed three matching kernels, namely, K_{λ_1,λ_4}, K_{λ_2,λ_4} and K_{λ_3,λ_4} (following the notation from Sect. 3.1) to homogenise the first three PSFs to the angular resolution of E_{λ_4}. Using the same procedure as in Sect. 3.2, we compared these homogenised PSFs to the original PSF using the residual formalism (15) applied to the MIRI bandpasses. The resulting residual images R_{λ_1,λ_4}, R_{λ_2,λ_4} and R_{λ_3,λ_4} are shown on the last column of Fig. 6.

The central region of the residual images, within the two main lobes of the E_{λ_4} PSF, has a low level of residuals ($\sim 10^{-5}$) in the three configurations Along the image borders, there are some non-negligible residual patches. A quick visual comparison with the homogenised PSFs (central column) shows that these patches correspond to extremely faint regions of the PSF ($<10^{-6}$ w.r.t. the peak) and thus have a very low impact in the matching process.

5. Conclusions

In this paper, we propose a new method for the generation of static PSF homogenisation kernels which is applicable for

instruments presenting complex PSFs such as recent or future space-born telescopes. The PSF on such optical systems is hardly ever static over the field-of-view, but we restricted the purpose of this paper to the production of homogenization kernels for the study of regions of interest on the image, where the PSF can be considered non-variable. The treatment of the PSF varying over the whole field-of-view of modern instruments cannot be linearized as in this work and requires a very different approach. It will be the subject of a following paper. The application on *Herschel*/PACS and SPIRE and JWST/MIRI instruments demonstrates the performance of the proposed algorithm in terms of low residuals (better than 10^{-2}–10^{-3} and 10^{-5}–10^{-6} for observed and simulated PSFs, respectively).

To assess the improvement brought by our algorithm for multi-wavelength studies, we address the estimation of dust temperature and spectral index β of astronomical objects using multi-band images taken in the submillimeter spectral range by *Herschel*. This estimation is made via pixel-by-pixel measurements across these images which have different intrinsic angular resolutions. Homogenisation kernels are thus traditionally used to bring all the images to the same angular resolution. Most of the analyses performed so far use either Gaussian kernels, or the circularised kernels produced by Aniano et al. (2011). However, effective PSFs of space imagers are anisotropic, so these methods are not accurate enough therefore to introduce systematic anti-correlation on β and temperature measurements with an amplitude that can be larger than the statistical noise. We have checked that using `pypher` kernels, systematic errors are in any case negligible compared to statistical noise.

Finally, we provide the `pypher` software (Boucaud 2016) to compute homogenisation kernels to be used for current and future instruments.

Acknowledgements. A.B. would like to thank Hacheme Ayasso for useful discussions. We acknowledge the CNES (Centre National d'Études Spatiales) for supporting this work as part of the Euclid SGS (Science Ground Segment) within the Euclid Consortium. We acknowledge the Euclid Consortium, the European Space Agency and agencies and institutes supporting the development of Euclid. Part of this work has received funding from the European Union's Seventh Framework Programme (FP7/2007–2013) for the DustPedia project (grant agreement No. FP7-SPACE-606847). This research has made use of the Herschel Database operated by the Integrated Data and Operation Center (IDOC) at the Institut d'Astrophysique Spatiale (IAS) under contract with CNES and CNRS. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013).

References

Ade, P. A., Aghanim, N., Arnaud, M., et al. 2011, A&A, 536, A7
Alard, C. 2000, A&AS, 144, 363
Aniano, G., Draine, B. T., Gordon, K. D., & Sandstrom, K. 2011, PASP, 123, 1218 (ADGS11)
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Data Analysis Software and Systems XI, 281, 228
Bocchio, M., Bianchi, S., & Abergel, A. 2016, A&A, 591, A117
Boucaud, A. 2016, pypher: Python PSF Homogenization kERnels, http://dx.doi.org/10.5281/zenodo.61392
Bouchet, P., García-Marín, M., Lagage, P.-O., et al. 2015, PASP, 127, 612
Darnell, T., Bertin, E., Gower, M., et al. 2009, in Astronomical Data Analysis Software and Systems XVIII, 411, 18
Desai, S., Armstrong, R., Mohr, J. J., et al. 2012, ApJ, 757, 83
Gonzalez, R. C., & Woods, R. E. 2008, Digital Image Processing, 3rd edn. (Upper Saddle River, NJ: Prentice Hall)
Gordon, K. D., Engelbracht, C. W., Rieke, G. H., et al. 2008, ApJ, 682, 336
Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, MNRAS, 421, 2355
Kuijken, K. 2008, A&A, 482, 1053
Laureijs, R. J., Duvet, L., Sanz, I. E., et al. 2010, in SPIE Astronomical Telescopes+ Instrumentation, International Society for Optics and Photonics, 77311
Liu, J., Henning, C., Desai, S., et al. 2015, MNRAS, 449, 3370
Lutz, D. 2012, PACS photometer point spread function
Lutz, D. 2015, Herschel-PACS document PICC-ME-TN-033, v2.2, http://herschel.esac.esa.int/twiki/pub/Public/PacsCalibrationWeb/bolopsf_22.pdf
Perrin, M. D., Soummer, R., Elliott, E. M., Lallo, M. D., & Sivaramakrishnan, A. 2012, in SPIE Astronomical Telescopes+ Instrumentation, International Society for Optics and Photonics, 84423
Planck Collaboration XI. 2014, A&A, 571, A11
Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2
Schultz, B. 2015, Analysis Details of SPIRE Photometer Beam Profiles, http://herschel.esac.esa.int/twiki/bin/view/Public/SpirePhotometerBeamProfileAnalysis2
Wheelock, S., Gautier, T., Chillemi, J., et al. 1994, IRAS Sky Survey Atlas ISSA, Explanatory Supplement, JPL

Appendix A: The `pypher` code

A Python code called `pypher` (Boucaud 2016), which computes the static PSF homogenisation kernels described in this work has been made publicly available and can be retrieved at:
https://github.com/aboucaud/pypher

Once installed, this programme can be used through a command-line interface taking as input the PSF images (source and target) as fits files, and specifying the output filename for the kernel,

```
$ pypher psf_a.fits psf_b.fits kernel_a_to_b.fits
```

The tunable parameters are, first, the regularisation parameter μ of the Wiener filter (see Eq. (12)) that penalises the high frequencies, and should be set according to the image that will be homogenised, and second, the position angle of both PSFs with respect to their image to accurately take into account the PSF shape in the homogenisation process.

The programme takes less than a second on a single CPU to compute a kernel from two 512×512 PSF images.

Appendix B

Appendix

B.1 Half-Quadratic Regularization: l_2/l_1-norm

Huber's function is defined as

$$\varphi(\delta) = \begin{cases} \delta^2 & \text{if } |\delta| < s \\ 2s|\delta| - s^2 & \text{otherwise} \end{cases} \tag{B.1}$$

and its first-order derivative is

$$\varphi'(\delta) = \begin{cases} 2\delta & \text{if } |\delta| < s \\ 2s \text{ signe}(\delta) & \text{otherwise} \end{cases} \tag{B.2}$$

The auxiliary function associated to *Huber*'s function [Idier 2001] is

$$\xi_\alpha(b) = \alpha \begin{cases} \frac{1}{1-2\alpha} b^2 & \text{if } |b| < (1-2\alpha)s \\ 2s|b| - (1-2\alpha)s^2 & \text{otherwise} \end{cases}. \tag{B.3}$$

with $\alpha \in [0, \alpha_{max} = 1/2]$.

B.2 Principal Component Analysis Method

Algorithm 2.6 Principal Component Analysis (PCA) algorithm

Input $\boldsymbol{X}, n_{\text{comp}}$
$[n_{\text{samp}}, n_{feat}] = \text{size}(\boldsymbol{X})$
Normalize the data set
$\boldsymbol{X}_{\text{std}} = \frac{\boldsymbol{X} - \text{mean}(\boldsymbol{X})}{\text{std}(\boldsymbol{X})}$
Compute the covariance matrix
$\boldsymbol{\Sigma}_x = \frac{1}{n_{\text{samp}}} \boldsymbol{X}^T \boldsymbol{X}$
Sort the eigen-vectors (columns) of $\boldsymbol{\Sigma}_x$ in decreasing order in term of eigen-values
$ev, eig = \text{eig}(\boldsymbol{\Sigma}_x)$
Select the first n_{comp} columns of $\boldsymbol{\Sigma}_x$
$\boldsymbol{\Sigma}_{x_{\text{new}}} = \boldsymbol{\Sigma}_x[:, 0 : n_{\text{comp}}]$
Project the dataset on the new space
$\boldsymbol{X}_{\text{new}} = \boldsymbol{X} \boldsymbol{\Sigma}_{x_{\text{new}}}^T$
return $\boldsymbol{X}_{\text{new}}$

Bibliography

[Abergel *et al.* 2003] A Abergel, D Teyssier, JP Bernard, F Boulanger, A Coulais, D Fosse, E Falgarone, M Gerin, M Perault, J-L Puget*et al. ISOCAM and molecular observations of the edge of the Horsehead nebula.* Astronomy & Astrophysics, vol. 410, no. 2, pages 577–585, 2003. 53, 54, 156

[Adams *et al.* 1986] John B Adams, Milton O Smith and Paul E Johnson. *Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site.* Journal of Geophysical Research: Solid Earth, vol. 91, no. B8, pages 8098–8112, 1986. 17, 68

[Alliney 1992] Stefano Alliney. *Digital filters as absolute norm regularizers.* IEEE Transactions on Signal Processing, vol. 40, no. 6, pages 1548–1562, 1992. 11

[Aniano *et al.* 2011] G Aniano, BT Draine, KD Gordon and K Sandstrom. *Common-resolution convolution kernels for space-and ground-based telescopes.* Publications of the Astronomical Society of the Pacific, vol. 123, no. 908, page 1218, 2011. 18, 38

[Banham & Katsaggelos 1997] Mark R Banham and Aggelos K Katsaggelos. *Digital image restoration.* IEEE signal processing magazine, vol. 14, no. 2, pages 24–41, 1997. 7

[Beck & Teboulle 2009] Amir Beck and Marc Teboulle. *A fast iterative shrinkage-thresholding algorithm for linear inverse problems.* SIAM journal on imaging sciences, vol. 2, no. 1, pages 183–202, 2009. 13

[Benazza-Benyahia & Pesquet 2006] Amel Benazza-Benyahia and J-C Pesquet. *Multichannel image deconvolution in the wavelet transform domain.* In Signal Processing Conference, 2006 14th European, pages 1–5. IEEE, 2006. 9

[Berne *et al.* 2007] Olivier Berne, C Joblin, Y Deville, JD Smith, M Rapacioli, JP Bernard, J Thomas, W Reach and A Abergel. *Analysis of the emission of very small dust particles from Spitzer spectro-imagery data using blind signal separation methods.* Astronomy & Astrophysics, vol. 469, no. 2, pages 575–586, 2007. 17, 82, 83, 84, 157

[Bertalmio *et al.* 2000] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles and Coloma Ballester. *Image inpainting.* In Proceedings of the 27th annual conference on Computer graphics and interactive techniques, pages 417–424. ACM Press/Addison-Wesley Publishing Co., 2000. 7, 8

[Bioucas-Dias *et al.* 2012] José M Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader and Jocelyn Chanussot. *Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches.* IEEE journal of selected topics in applied earth observations and remote sensing, vol. 5, no. 2, pages 354–379, 2012. 17, 68

[Bongard *et al.* 2011] S Bongard, F Soulez, Éric Thiébaut and É Pecontal. *3D deconvolution of hyper-spectral astronomical data.* Monthly Notices of the Royal Astronomical Society, vol. 418, no. 1, pages 258–270, 2011. 9, 16, 39, 42

[Boucaud *et al.* 2016] Alexandre Boucaud, Marco Bocchio, Alain Abergel, François Orieux, Hervé Dole and Mohamed Amine Hadj-Youcef. *Convolution kernels for multi-wavelength imaging.* Astronomy & Astrophysics, vol. 596, page A63, 2016. 18, 38, 117

[Bouchet *et al.* 2015] Patrice Bouchet, Macarena García-Marín, P-O Lagage, Jérome Amiaux, J-L Auguéres, Eva Bauwens, JADL Blommaert, CH Chen, ÖH Detre, Dan Dicken*et al. The Mid-Infrared Instrument for the James Webb Space Telescope, III: MIRIM, The MIRI Imager.* Publications of the Astronomical Society of the Pacific, vol. 127, no. 953, page 612, 2015. 1, 2, 22, 25, 26, 28, 29, 80, 155

[Buades *et al.* 2005] Antoni Buades, Bartomeu Coll and J-M Morel. *A non-local algorithm for image denoising.* In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 2, pages 60–65. IEEE, 2005. 7

[Chambolle & Pock 2011] Antonin Chambolle and Thomas Pock. *A first-order primal-dual algorithm for convex problems with applications to imaging.* Journal of Mathematical Imaging and Vision, vol. 40, no. 1, pages 120–145, 2011. 8, 12

[Chambolle 2004] Antonin Chambolle. *An algorithm for total variation minimization and applications.* Journal of Mathematical imaging and vision, vol. 20, no. 1-2, pages 89–97, 2004. 12, 74

[Champagnat & Idier 2004] Frédéric Champagnat and Jérôme Idier. *A connection between half-quadratic criteria and EM algorithms.* IEEE Signal Processing Letters, vol. 11, no. 9, pages 709–712, 2004. 75

[Chan & Wong 1998] Tony F Chan and Chiu-Kwong Wong. *Total variation blind deconvolution.* IEEE transactions on Image Processing, vol. 7, no. 3, pages 370–375, 1998. 8

[Chappell *et al.* 2009] Michael A Chappell, Adrian R Groves, Brandon Whitcher and Mark W Woolrich. *Variational Bayesian inference for a nonlinear forward model.* IEEE Transactions on Signal Processing, vol. 57, no. 1, pages 223–236, 2009. 7

[Charbonnier *et al.* 1997] Pierre Charbonnier, Laure Blanc-Féraud, Gilles Aubert and Michel Barlaud. *Deterministic edge-preserving regularization in computed imaging.* IEEE Transactions on image processing, vol. 6, no. 2, pages 298–311, 1997. 74

[Cichocki & Amari 2003] Andrzej Cichocki and Shun-ichi Amari. *Adaptive Blind Signal and Image Processing (new revised and improved edition)*, 2003. 68

[Combettes & Pesquet 2004] Patrick L Combettes and J-C Pesquet. *Image restoration subject to a total variation constraint.* IEEE transactions on image processing, vol. 13, no. 9, pages 1213–1222, 2004. 8

[Combettes & Pesquet 2011] Patrick L Combettes and Jean-Christophe Pesquet. *Proximal splitting methods in signal processing.* In Fixed-point algorithms for inverse problems in science and engineering, pages 185–212. Springer, 2011. 13

[Comon & Jutten 2010] Pierre Comon and Christian Jutten. Handbook of blind source separation: Independent component analysis and applications. Academic press, 2010. 68

[Demoment 1989] Guy Demoment. *Image reconstruction and restoration: Overview of common estimation structures and problems.* IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 12, pages 2024–2036, 1989. 11, 44

[Denis *et al.* 2011] Loïc Denis, Éric Thiébaut and Ferréol Soulez. *Fast model of space-variant blurring and its application to deconvolution in astronomy.* In Image Processing (ICIP), 2011 18th IEEE International Conference on, pages 2817–2820. IEEE, 2011. 17, 39

[Denis *et al.* 2015] Loïc Denis, Eric Thiébaut, Ferréol Soulez, Jean-Marie Becker and Rahul Mourya. *Fast approximations of shift-variant blur.* International Journal of Computer Vision, vol. 115, no. 3, pages 253–278, 2015. 17

[Dobigeon *et al.* 2009] Nicolas Dobigeon, Saïd Moussaoui, Martial Coulon, Jean-Yves Tourneret and Alfred O Hero. *Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery.* IEEE Transactions on Signal Processing, vol. 57, no. 11, pages 4355–4368, 2009. 17, 68

[Donoho 1995] David L Donoho. *De-noising by soft-thresholding.* IEEE transactions on information theory, vol. 41, no. 3, pages 613–627, 1995. 7

[Galatsanos & Chin 1989] Nikolas P Galatsanos and Roland T Chin. *Digital restoration of multichannel images.* IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 3, pages 415–421, 1989. 9, 15, 60, 68

[Galatsanos *et al.* 1991] Nikolas P Galatsanos, Aggelos K Katsaggelos, Roland T Chin and Allen D Hillery. *Least squares restoration of multichannel images.* IEEE Transactions on Signal Processing, vol. 39, no. 10, pages 2222–2236, 1991. 10, 15, 41, 47, 51, 73

[Galatsanos *et al.* 2000] Nikolas P Galatsanos, Miles N Wernick, Aggelos K Katsaggelos and Rafael Molina. *Multichannel image recovery.* Handbook of Image and Video Processing, vol. 12, pages 155–168, 2000. 10, 15

[Geis & Lutz 2010] N Geis and D Lutz. *Herschel/PACS modelled point-spread functions*, 2010. 38

[Geman & Reynolds 1992] Donald Geman and George Reynolds. *Constrained restoration and the recovery of discontinuities.* IEEE Transactions on pattern analysis and machine intelligence, vol. 14, no. 3, pages 367–383, 1992. 13, 74

[Geman & Yang 1995] Donald Geman and Chengda Yang. *Nonlinear image recovery with half-quadratic regularization.* IEEE Transactions on Image Processing, vol. 4, no. 7, pages 932–946, 1995. 13, 74, 75, 76, 92

[Giovannelli 2008] Jean-François Giovannelli. *Unsupervised Bayesian convex deconvolution based on a field with an explicit partition function.* IEEE Transactions on Image Processing, vol. 17, no. 1, pages 16–26, 2008. 75

[Glasse *et al.* 2015] Alistair Glasse, GH Rieke, Eva Bauwens, Macarena García-Marín, ME Ressler, Steffen Rost, Tuomo Ville Tikkanen, Bart Vandenbussche and GS Wright. *The Mid-Infrared Instrument for the James Webb Space Telescope, IX: Predicted Sensitivity.* Publications of the Astronomical Society of the Pacific, vol. 127, no. 953, page 686, 2015. 22, 28, 30, 33, 34, 155

[Goldstein & Osher 2009] Tom Goldstein and Stanley Osher. *The split Bregman method for L1-regularized problems.* SIAM journal on imaging sciences, vol. 2, no. 2, pages 323–343, 2009. 12

[Goodman 2005] Joseph W Goodman. Introduction to Fourier optics. Roberts and Company Publishers, 2005. 1, 24, 25, 31

[Gordon *et al.* 2015] Karl D Gordon, CH Chen, Rachel E Anderson, Ruymán Azzollini, L Bergeron, Patrice Bouchet, Jeroen Bouwman, Misty Cracraft, Sebastian Fischer, Scott D Friedman*et al. The Mid-Infrared Instrument for the James Webb Space Telescope, X: Operations and Data Reduction.* Publications of the Astronomical Society of the Pacific, vol. 127, no. 953, page 696, 2015. 32

[Guillard *et al.* 2010] Pierre Guillard, Thomas Rodet, S Ronayette, J Amiaux, Alain Abergel, V Moreau, JL Augueres, A Bensalem, T Orduna, C Nehmé*et al. Optical performance of the JWST/MIRI flight model: characterization of the point spread function at high resolution.* In SPIE Astronomical Telescopes+ Instrumentation, pages 77310J–77310J. International Society for Optics and Photonics, 2010. 38

[Guo *et al.* 2009] Zhaohui Guo, Todd Wittman and Stanley Osher. *L1 unmixing and its application to hyperspectral image enhancement.* In Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV, volume 7334, page 73341M. International Society for Optics and Photonics, 2009. 17, 68

[Hadj-Youcef *et al.* 2017a] M. A. Hadj-Youcef, F. Orieux, A. Fraysse and A. Abergel. *Restoration from multispectral blurred data with non-stationary instrument response.* In 2017 25th European Signal Processing Conference (EUSIPCO), pages 503–507, Aug 2017. 68, 117

[Hadj-Youcef *et al.* 2017b] Mohamed Amine Hadj-Youcef, François Orieux, Aurélia Fraysse and Alain Abergel. *Restauration d'objets astrophysiques à partir de données multispectrales floues et d'une réponse instrument non-stationnaire*. In 26eme Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017, Juan Les Pins, France, September 2017. 117

[Hadj-Youcef *et al.* 2018] M. A. Hadj-Youcef, F. Orieux, A. Fraysse and A. Abergel. *Spatio-Spectral Multichannel Reconstruction from few Low-Resolution Multispectral Data*. In 2018 26th European Signal Processing Conference (EUSIPCO), 2018. 117

[Haertel & Shimabukuro 2004] Victor Haertel and Yosio Edemir Shimabukuro. *Spectral linear mixing model in low spatial resolution image data*. In Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International, volume 4, pages 2546–2549. IEEE, 2004. 17

[Hedjam & Cheriet 2013] Rachid Hedjam and Mohamed Cheriet. *Historical document image restoration using multispectral imaging system*. Pattern Recognition, vol. 46, no. 8, pages 2297–2312, 2013. 15

[Henrot *et al.* 2013] Simon Henrot, Charles Soussen and David Brie. *Fast positive deconvolution of hyperspectral images*. IEEE Transactions on Image Processing, vol. 22, no. 2, pages 828–833, 2013. 9, 16

[Hunt & Kubler 1984] B Hunt and Olaf Kubler. *Karhunen-Loeve multispectral image restoration, part I: Theory*. IEEE transactions on acoustics, speech, and signal processing, vol. 32, no. 3, pages 592–600, 1984. 9, 41

[Hunt 1971] B Hunt. *A matrix theory proof of the discrete convolution theorem*. IEEE Transactions on Audio and Electroacoustics, vol. 19, no. 4, pages 285–288, 1971. 13, 43, 50

[Hyvärinen *et al.* 2004] Aapo Hyvärinen, Juha Karhunen and Erkki Oja. Independent component analysis, volume 46. John Wiley & Sons, 2004. 68

[Idier 2001] Jérôme Idier. *Convex half-quadratic criteria and interacting auxiliary variables for image restoration*. IEEE Transactions on Image Processing, vol. 10, no. 7, pages 1001–1009, 2001. 12, 76, 143

[Jacoby *et al.* 1984] George H Jacoby, Deidre A Hunter and Carol A Christian. *A library of stellar spectra*. The Astrophysical Journal Supplement Series, vol. 56, pages 257–281, 1984. 68

[Jaynes 2003] Edwin T Jaynes. Probability theory: The logic of science. Cambridge university press, 2003. 14

[Jolliffe 1986] Ian T Jolliffe. *Principal Component Analysis and Factor Analysis*. In Principal component analysis, pages 115–128. Springer, 1986. 68, 81, 82, 157

[Kass & Wasserman 1996] Robert E Kass and Larry Wasserman. *The selection of prior distributions by formal rules.* Journal of the American Statistical Association, vol. 91, no. 435, pages 1343–1370, 1996. 15

[Katsaggelos & Paik 1988] Aggelos K Katsaggelos and Joon K Paik. *Iterative color image restoration algorithms.* In Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on, pages 1028–1031. IEEE, 1988. 15

[Keshava & Mustard 2002] Nirmal Keshava and John F Mustard. *Spectral unmixing.* IEEE signal processing magazine, vol. 19, no. 1, pages 44–57, 2002. 17, 68

[Levenson & Mansfield 2006] Richard M Levenson and James R Mansfield. *Multispectral imaging in biology and medicine: slices of life.* Cytometry part A, vol. 69, no. 8, pages 748–758, 2006. 15

[Ljung 1998] Lennart Ljung. *System identification.* In Signal analysis and prediction, pages 163–173. Springer, 1998. 8

[Loncan *et al.* 2015] Laetitia Loncan, Luis B Almeida, José M Bioucas-Dias, Xavier Briottet, Jocelyn Chanussot, Nicolas Dobigeon, Sophie Fabre, Wenzhi Liao, Giorgio A Licciardi, Miguel Simoes*et al.* *Hyperspectral pansharpening: A review.* arXiv preprint arXiv:1504.04531, 2015. 17

[Love *et al.* 2005] Peter J Love, Alan W Hoffman, Nancy A Lum, Ken J Ando, Joe Rosbeck, William D Ritchie, Neil J Therrien, Roger S Holcombe and Elizabeth Corrales. *1024× 1024 Si: As IBC detector arrays for JWST MIRI.* In Focal Plane Arrays for Space Telescopes II, volume 5902, page 590209. International Society for Optics and Photonics, 2005. 27

[Mairal *et al.* 2014] Julien Mairal, Francis Bach, Jean Ponce*et al.* *Sparse modeling for image and vision processing.* Foundations and Trends® in Computer Graphics and Vision, vol. 8, no. 2-3, pages 85–283, 2014. 13

[Makidon *et al.* 2007] RB Makidon, S Casertano, C Cox and R van der Marel. *The JWST point spread function: Calculation methods and expected properties.* NASA Technic Al Report, 2007. 23, 24, 155

[Mast *et al.* 1982] Terry S Mast, Jerry E Nelson and William J Welch. *Effects of primary mirror segmentation on telescope image quality.* In Advanced Technology Optical Telescopes I, volume 332, pages 123–134. International Society for Optics and Photonics, 1982. 24

[Molina *et al.* 2003] Rafael Molina, Javier Mateos, Aggelos K Katsaggelos and Miguel Vega. *Bayesian annel image restoration using compound Gauss-Markov random fields.* IEEE Transactions on Image Processing, vol. 12, no. 12, pages 1642–1654, 2003. 10, 15

[Molina *et al.* 2006] Rafael Molina, Javier Mateos and Aggelos K Katsaggelos. *Blind deconvolution using a variational approach to parameter, image, and blur estimation.* IEEE Transactions on Image Processing, vol. 15, no. 12, pages 3715–3727, 2006. 8

[Mugnier *et al.* 2004] Laurent M Mugnier, Thierry Fusco and Jean-Marc Conan. *MISTRAL: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images.* JOSA A, vol. 21, no. 10, pages 1841–1854, 2004. 8

[Neumaier 1998] Arnold Neumaier. *Solving ill-conditioned and singular linear systems: A tutorial on regularization.* SIAM review, vol. 40, no. 3, pages 636–666, 1998. 44

[Nuzillard & Bijaoui 2000] Danielle Nuzillard and Albert Bijaoui. *Blind source separation and analysis of multispectral astronomical images.* Astronomy and Astrophysics Supplement Series, vol. 147, no. 1, pages 129–138, 2000. 15

[Orieux *et al.* 2010] François Orieux, Thomas Rodet and J-F Giovannelli. *Instrument parameter estimation in bayesian convex deconvolution.* In 2010 IEEE International Conference on Image Processing, pages 1161–1164. IEEE, 2010. 8

[Park *et al.* 2003] Sung Cheol Park, Min Kyu Park and Moon Gi Kang. *Super-resolution image reconstruction: a technical overview.* IEEE signal processing magazine, vol. 20, no. 3, pages 21–36, 2003. 7, 41

[Perona & Malik 1990] Pietro Perona and Jitendra Malik. *Scale-space and edge detection using anisotropic diffusion.* IEEE Transactions on pattern analysis and machine intelligence, vol. 12, no. 7, pages 629–639, 1990. 74

[Perrin *et al.* 2012] Marshall D Perrin, Rémi Soummer, Erin M Elliott, Matthew D Lallo and Anand Sivaramakrishnan. *Simulating point spread functions for the James Webb Space Telescope with WebbPSF.* In Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, volume 8442, page 84423D. International Society for Optics and Photonics, 2012. 25, 26, 155

[Perrin *et al.* 2014] Marshall D Perrin, Anand Sivaramakrishnan, Charles-Philippe Lajoie, Erin Elliott, Laurent Pueyo, Swara Ravindranath and Loïc Albert. *Updated point spread function simulations for JWST with WebbPSF.* In Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, volume 9143, page 91433X. International Society for Optics and Photonics, 2014. 25

[Pickles 1998] AJ Pickles. *A stellar spectral flux library: 1150–25000 Å.* Publications of the Astronomical Society of the Pacific, vol. 110, no. 749, page 863, 1998. 68

[Plaza *et al.* 2004] Antonio Plaza, Pablo Martínez, Rosa Pérez and Javier Plaza. *A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data.* IEEE transactions on geoscience and remote sensing, vol. 42, no. 3, pages 650–663, 2004. 17

[R. Gastaud 2018] A. Coulais R. Gastaud C. Cossou. *MIRISIM Imsim Report.* Technical report, mai 2018. 38

[Rieke *et al.* 2015] GH Rieke, ME Ressler, Jane E Morrison, L Bergeron, Patrice Bouchet, Macarena García-Marín, TP Greene, MW Regan, KG Sukhatme and Helen Walker.

The Mid-Infrared Instrument for the James Webb Space Telescope, VII: The MIRI Detectors. Publications of the Astronomical Society of the Pacific, vol. 127, no. 953, page 665, 2015. 22, 27, 29, 155

[Rigie & La Rivière 2015] David S Rigie and Patrick J La Rivière. *Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization*. Physics in Medicine & Biology, vol. 60, no. 5, page 1741, 2015. 9

[Rodet *et al.* 2009] Thomas Rodet, François Orieux, J-F Giovannelli and Alain Abergel. *Data inversion for hyperspectral objects in astronomy*. In Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009. WHISPERS'09. First Workshop on, pages 1–4. IEEE, 2009. 16

[Rudin *et al.* 1992] Leonid I Rudin, Stanley Osher and Emad Fatemi. *Nonlinear total variation based noise removal algorithms*. Physica D: Nonlinear Phenomena, vol. 60, no. 1, pages 259–268, 1992. 11, 12, 74

[Sabatke *et al.* 2005] Erin Sabatke, James Burge and Derek Sabatke. *Analytic diffraction analysis of a 32-m telescope with hexagonal segments for high-contrast imaging*. Applied optics, vol. 44, no. 8, pages 1360–1365, 2005. 24

[Sawatzky *et al.* 2014] Alex Sawatzky, Qiaofeng Xu, Carsten O Schirra and Mark A Anastasio. *Proximal ADMM for multi-channel image reconstruction in spectral X-ray CT*. IEEE transactions on medical imaging, vol. 33, no. 8, pages 1657–1668, 2014. 10

[Schultz & Stevenson 1995] Richard R Schultz and Robert L Stevenson. *Stochastic modeling and estimation of multispectral image data*. IEEE Transactions on Image Processing, vol. 4, no. 8, pages 1109–1119, 1995. 10, 15, 16, 41

[Settle & Drake 1993] JJ Settle and NA Drake. *Linear mixing and the estimation of ground cover proportions*. International Journal of Remote Sensing, vol. 14, no. 6, pages 1159–1177, 1993. 17

[Shewchuk 1994] Jonathan Richard Shewchuk. *An introduction to the conjugate gradient method without the agonizing pain*, 1994. 46, 47, 163

[Song *et al.* 2016] Yingying Song, David Brie, El-Hadi Djermoune and Simon Henrot. *Regularization parameter estimation for non-negative hyperspectral image deconvolution*. IEEE Transactions on Image Processing, vol. 25, no. 11, pages 5316–5330, 2016. 9

[Soulez *et al.* 2013] Ferréol Soulez, Eric Thiébaut and Loıc Denis. *Restoration of hyperspectral astronomical data with spectrally varying blur*. EAS Publications Series, vol. 59, pages 403–416, 2013. 16, 39

[Tarabalka *et al.* 2009] Yuliya Tarabalka, Jón Atli Benediktsson and Jocelyn Chanussot. *Spectral–spatial classification of hyperspectral imagery based on partitional clustering techniques*. IEEE Transactions on Geoscience and Remote Sensing, vol. 47, no. 8, pages 2973–2987, 2009. 17, 68

[Thévenaz *et al.* 2000] Philippe Thévenaz, Thierry Blu and Michael Unser. *Interpolation revisited [medical images application].* IEEE Transactions on medical imaging, vol. 19, no. 7, pages 739–758, 2000. 40

[Thiébaut *et al.* 2016] Éric Thiébaut, Loïc Dénis, Ferréol Soulez and Rahul Mourya. *Spatially variant PSF modeling and image deblurring.* In SPIE Astronomical Telescopes+ Instrumentation, pages 99097N–99097N. International Society for Optics and Photonics, 2016. 17, 39

[Tikhonov & Arsenin 1977] AN Tikhonov and Vasili Ya Arsenin. Methods for solving ill-posed problems. John Wiley and Sons, Inc, 1977. 11

[Wells *et al.* 2015] Martyn Wells, J-W Pel, Alistair Glasse, GS Wright, Gabby Aitink-Kroes, Ruymán Azzollini, Steven Beard, BR Brandl, Angus Gallie, VC Geers*et al. The mid-infrared instrument for the James Webb Space telescope, VI: the medium resolution spectrometer.* Publications of the Astronomical Society of the Pacific, vol. 127, no. 953, page 646, 2015. 25, 68

[Wen *et al.* 2008] You-Wei Wen, Michael K Ng and Yu-Mei Huang. *Efficient total variation minimization methods for color image restoration.* IEEE Transactions on Image Processing, vol. 17, no. 11, pages 2081–2088, 2008. 10, 15

[Yang *et al.* 2009a] Junfeng Yang, Wotao Yin, Yin Zhang and Yilun Wang. *A fast algorithm for edge-preserving variational multichannel image restoration.* SIAM Journal on Imaging Sciences, vol. 2, no. 2, pages 569–592, 2009. 10, 15

[Yang *et al.* 2009b] Junfeng Yang, Yin Zhang and Wotao Yin. *An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise.* SIAM Journal on Scientific Computing, vol. 31, no. 4, pages 2842–2865, 2009. 10, 12

[Yaroslavsky 2012] Leonid P Yaroslavsky. Theoretical foundations of digital imaging using matlab®. CRC Press, 2012. 42

[Yaroslavsky 2013] Leonid Yaroslavsky. Digital holography and digital image processing: principles, methods, algorithms. Springer Science & Business Media, 2013. 42, 69

List of Figures

List of Tables

List of Algorithms

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Reconstruction spatio-spectrale à partir de données multispectrales basse résolution. Application à l'instrument infrarouge moyen du Télescope spatial James Webb.

Mots clés : Imagerie multispectrale, Problèmes inverses, Déconvolution, Modélisation du système, Multi-longueur d'onde, Modèle direct, Restauration multispectrale, Reconstruction d'image

Résumé : Cette thèse traite un problème inverse en astronomie. L'objectif est de reconstruire un objet spatio-spectral, ayant une distribution spatiale et spectrale, à partir d'un ensemble de données multispectrales de basse résolution fournies par l'imageur MIRI (Mid-InfraRed Instrument), qui est à bord du prochain télescope spatial James Webb Space Telescope (JWST). Les données multispectrales observées souffrent d'un flou spatial qui dépend de la longueur d'onde. Cet effet est dû à la convolution par la réponse optique (PSF). De plus, les données multispectrales souffrent également d'une sévère dégradation spectrale en raison du filtrage spectral et de l'intégration par le détecteur sur de larges bandes. La reconstruction de l'objet original est un problème mal posé en raison du manque important d'informations spectrales dans l'ensemble des données multispectrales. La difficulté se pose alors dans le choix d'une représentation de l'objet permettant la reconstruction de l'information spectrale. Un modèle classique utilisé jusqu'à présent considère une PSF invariante spectralement par bande, ce qui néglige la variation spectrale de la PSF. Cependant, ce modèle simpliste est convenable que dans le cas d'instrument à une bande spectrale très étroite, ce qui n'est pas le cas pour l'imageur de MIRI. Notre approche consiste à développer une méthode pour l'inversion qui se résume en quatre étapes : (1) concevoir un modèle de l'instrument reproduisant les données multispectrales observées, (2) proposer un modèle adapté pour représenter l'objet à reconstruire, (3) exploiter conjointement l'ensemble des données multispectrales, et enfin (4) développer une méthode de reconstruction basée sur la régularisation en introduisant des priori à la solution.
Les résultats de reconstruction d'objets spatio-spectral à partir de neuf images multispectrales simulées de l'imageur de MIRI montrent une augmentation significative des résolutions spatiale et spectrale de l'objet par rapport à des méthodes conventionnelles. L'objet reconstruit montre l'effet de débruitage et de déconvolution des données multispectrales. Nous avons obtenu une erreur relative n'excédant pas 5% à 30 dB et un temps d'exécution de 1 seconde pour l'algorithme de norm-l_2 et 20 secondes avec 50 itérations pour l'algorithme norm-l_2/l_1. C'est 10 fois plus rapide que la solution itérative calculée par l'algorithme de gradient conjugué.

Title : Spatio spectral reconstruction from low resolution multispectral data. Application to the Mid-Infrared instrument of the James Webb Space Telescope

Keywords : Multipectral imaging, Inverse problems, Deconvolution, System modeling, Multi-wavelength, Forward Model, Multispectral Restoration, Image Reconstruction

Abstract : This thesis deals with an inverse problem in astronomy. The objective is to reconstruct a spatio-spectral object, having spatial and spectral distributions, from a set of low-resolution multispectral data taken by the imager MIRI (Mid-InfraRed Instrument), which is on board the next space telescope James Webb Space Telescope (JWST). The observed multispectral data suffers from a spatial blur that varies according to the wavelength due to the spatial convolution with a shift-variant optical response (PSF). In addition the multispectral data also suffers from severe spectral degradations because of the spectral filtering and the integration by the detector over broad bands. The reconstruction of the original object is an ill-posed problem because of the severe lack of spectral information in the multispectral dataset. The difficulty then arises in choosing a representation of the object that allows the reconstruction of this spectral information. A common model used so far considers a spectral shift-invariant PSF per band, which neglects the spectral variation of the PSF. This simplistic model is only suitable for instruments with a narrow spectral band, which is not the case for the imager of MIRI. Our approach consists of developing an inverse problem framework that is summarized in four steps : (1) designing an instrument model that reproduces the observed multispectral data, (2) proposing an adapted model to represent the sought object, (3) exploiting all multispectral dataset jointly, and finally (4) developing a reconstruction method based on regularization methods by enforcing prior information to the solution.
The overall reconstruction results obtained on simulated data of the JWST/MIRI imager show a significant increase of spatial and spectral resolutions of the reconstructed object compared to conventional methods. The reconstructed object shows a clear denoising and deconvolution of the multispectral data. We obtained a relative error below 5% at 30 dB, and an execution time of 1 second for the l_2-norm algorithm and 20 seconds with 50 iterations for the l_2/l_1-norm algorithm. This is 10 times faster than the iterative solution computed by conjugate gradients.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France