Unmixing and deconvolution methods of objects in images : applications in astronomical high-contrast imaging

Samuel Thé

THESE de DOCTORAT DE
L'UNIVERSITE CLAUDE BERNARD LYON 1

Ecole Doctorale ED52
Physique et Astrophysique (PHAST)

Discipline : Astrophysique

Soutenue publiquement le 07/03/2023, par :
Samuel Thé

Méthodes de démélange et de déconvolution d'objets dans des images
Applications en imagerie astronomique à haut-contraste

Devant le jury composé de :

Quantin-Nataf, Cathy	Professeure des Universités, Université Lyon 1	Présidente
Bourguignon, Sébastien	Maître de Conférences, Centrale Nantes	Rapporteur
Mugnier, Laurent	Chercheur, ONERA Lille	Rapporteur
Boccaletti, Anthony	Directeur de Recherche, CNRS Paris	Examinateur
Ferrari, André	Professeur des Universités, Université Côte d'Azur	Examinateur
Meillier, Céline	Maître de Conférences, Université de Strasbourg	Examinatrice
Langlois, Maud	Directrice de Recherche, CNRS Lyon	Directrice de thèse
Thiébaut, Éric	Astronome Adjoint, Université Lyon 1	Co-directeur de thèse
Denis, Loïc	Professeur des Universités, Université Saint-Etienne	Invité

Remerciements

J'aimerais tout d'abord remercier Éric, sans qui ces trois ans (et quelques mois) n'auraient pas été possible. Merci à lui pour sa patience à toute épreuve, son encadrement et sa confiance qui m'ont guidé tout au long de ma thèse. Merci aussi à Loïc et Maud pour leur soutien sans faille et leurs conseils éclairés. Je m'estime extrêmement chanceux d'avoir eu à mes côtés de tels directeurs de thèse. Leurs enseignements m'ont permis d'apprendre énormément et j'espère en apporter la preuve dans ce manuscrit. Je remercie également Sébastien Bourguignon et Laurent Mugnier d'avoir accepté de faire le rapport de ma thèse, ainsi que Cathy Quantin-Nataf, Céline Meillier, Anthony Boccaletti et André Ferrari pour leur participation à mon jury de soutenance. Un grand merci à Ferréol, Gil, Isabelle et Michel, pour leur accueil, les discussions très enrichissantes que nous avons partagé, leurs remarques et conseils avisés durant ces années au CRAL. J'aimerais aussi particulièrement remercier les doctorants, stagiaires, post-docs et ingénieur de l'équipe qui ont rendu ces années aussi amusante : Laurence pour nos discussions animées (scientifiques et autres) et les sortis jeux de sociétés dont j'ai eu besoin à certains moments pour tenir le coup, Clément et ces blagues plus que réconfortantes, Olivier pour son aide précieuse peu importe l'heure, Jules et Camille pour leurs réflexions presque philosophiques autour d'une bière, Antoine toujours près pour un thé bien chaud et Théo, co-bureau de cette dernière année et sa bonne humeur débordante. Enfin, j'aimerais remercier ma femme qui m'a toujours encouragé, ma sœur et mes amis pour leur soutien tout au long de ma thèse. Ma dernière pensée de gratitude est pour ma mère qui a tant fait et tant sacrifié pour moi et mes projets. Merci maman.

La nuit étoilée de Vincent van Gogh.

Résumé par chapitre

Dans cette section sont résumés en français les différents chapitres de ce manuscrit.

Introduction

L'imagerie directe d'environnements circumstellaires joue un rôle clé dans notre compréhension de la formation et de l'évolution des exo-planètes. Mesurer directement la distribution de lumière des objets se trouvant dans ces environnements peut nous donner des informations essentielles sur les phénomènes de formations et la diversité des exo-planètes. L'instrument SPHERE installé sur le VLT a été spécifiquement conçu dans ce but. Grâce à une optique adaptative de très haut niveau, cet instrument est capable d'imager ce type d'environnement avec une résolution angulaire se rapprochant de la limite de diffraction du télescope.

Ce manuscrit se consacre à l'imagerie directe de ces environnements et aux principaux problèmes dont souffre cette méthode d'observation. Outre le flou apporté par l'instrument et le bruit provenant de l'acquisition des données, le signal provenant de l'objet d'intérêt est le plus souvent bien plus faible que l'intensité lumineuse de son étoile hôte. Afin de limiter la pollution lumineuse venant de l'étoile, SPHERE est composé d'un puissant coronographe, permettant de réduire d'un facteur 10 la lumière provenant de l'étoile. Cependant, certaines aberrations dans le chemin optique forment sur le détecteur des fuites stellaires qui sont encore jusqu'à $\sim 10^5$ fois plus brillantes que les signaux provenant d'objets d'intérêts (*e.g.* compagnon, disque). Afin de parvenir à observer l'objet d'intérêt, il est donc nécessaire de faire appel à des méthodes numériques de traitement des données.

Dans ce manuscrit, je détaille de telles méthodes visant à extraire l'objet d'intérêt malgré les composantes de nuisances que sont le bruit, le flou et les fuites stellaires.

Chapitre 1. Les problèmes inverses

Ce chapitre introductif traite des problèmes inverses et vise à donner les outils nécessaires à la bonne compréhension des autres chapitres.

Une première partie est consacrée à la formation des données. Partant des différents phénomènes mis en jeux lors de l'acquisition de données par un détecteur (*e.g.* interaction lumière/matière, courant d'obscurité), je présente un modèle de correction des données brutes. La calibration de la composante instrumentale de ce modèle permet de corriger les effets de l'instrument. Pour utiliser au mieux l'approche problèmes inverses, il est aussi nécessaire de développer un modèle de la statistique (espérances et variances) des données pré-traitées.

Avec ce modèle statistique, il est possible de développer le modèle direct des données dépendant des différents effets de l'instrument et de l'objet d'intérêt. Le problème de déconvolution est présenté en tant qu'illustration. Dans ce contexte, l'objet d'intérêt (à retrouver) est convolué/flouté par la fonction d'étalement du point (PSF en anglais).

Je présente différentes méthodes permettant de retrouver l'objet d'intérêt comme le filtre de Wienner et l'estimateur du maximum *a posteriori*. Ce dernier favorise certaines propriétés dans le résultat de l'estimation via des contraintes sur l'objet à estimer et via des fonctions de régularisations. Le recours à ces fonctions de régularisations n'est cependant pas anodin et ajoute des hyper-paramètres à régler au problème de reconstruction. Des résultats sur données simulées sont présentés afin de comparer ces différentes méthodes.

Le reste du manuscrit reposant sur l'estimation au sens du maximum *a posteriori*, je détaille, dans une autre partie de ce chapitre, certaines méthodes permettant le réglage des hyper-paramètres dans nos problèmes d'estimation.

Enfin, certaines méthodes d'optimisation sont présentées afin de justifier mes choix dans les problèmes présentés le long de ce manuscrit.

Chapitre 2. AMORS: Réglage facilité des hyper-paramètres et amélioration de la convergence pour les modèles bi-linéaires

Poursuivant sur le problème de reconstruction d'objets d'intérêt dans des données floutées et bruitées, se pose la question de l'estimation de la PSF. En effet, cette fonction qui caractérise la réponse de l'instrument est souvent difficile à calibrer précisément. En utilisant l'approche problèmes inverses, il est cependant possible de l'estimer conjointement avec l'objet d'intérêt : c'est la déconvolution aveugle. Le modèle direct résultant est un modèle bi-linéaire en l'objet d'intérêt et en la PSF, ayant pour propriété une ambiguïté d'échelle. En d'autres termes, multiplier par un facteur l'un des deux paramètres et diviser l'autre par ce même facteur donne le même résultat, quelle que soit la valeur du facteur.

De nombreuses méthodes existent afin d'estimer l'objet d'intérêt et la PSF. Dans ce chapitre, je présente une nouvelle méthode reposant sur l'estimation alternée de l'objet, de la PSF et du facteur d'échelle optimal au sens du maximum *a posteriori*. En effet, ce dernier peut être estimé en minimisant le même critère et a une forme analytique. En injectant dans le problème ce facteur d'échelle, plusieurs propriétés intéressantes apparaissent dans le cas de fonctions de régularisations homogènes:

- Mettre à l'échelle les paramètres revient à changer le réglage des hyper-paramètres des termes de régularisations.
- Les régularisations sont de ce fait liées et un seul hyper-paramètre règle en réalité l'importance des deux régularisations.

De ces observations résultent l'algorithme Amors qui estime les paramètres du modèle bi-linéaire (objet d'intérêt et PSF), à un facteur d'échelle près. Cet algorithme permet de n'avoir à régler qu'un seul paramètre, l'autre ayant une valeur arbitraire fixe. Une attention toute particulière est apportée au facteur d'échelle initial du modèle. Utilisant des outils précédemment introduits pour la résolution de ce genre de problème, Amors peut être considéré comme une boite noire, facile à régler, et qui peut être incorporée dans une stratégie non supervisée. La méthode est généralisée à tous les modèles bi-linéaires, donc pas seulement au contexte de la déconvolution aveugle. Amors est en particulier applicable dans les problèmes de dé-mélange et de séparation de sources.

Des résultats obtenus sur données simulées sont présentés afin de valider Amors. Pour ce faire, je présente tout d'abord une version homogène de la régularisation de préservation des bords (*edge-preserving smoothness*). Les résultats sont très similaires à ceux obtenus dans le cas d'une simple déconvolution (donc à PSF connue). Des tests pour différentes échelles de l'initialisation sont aussi effectués, prouvant que la prise en compte du facteur d'échelle optimal et l'étape initiale de raffinement de ce facteur permet une convergence beaucoup plus rapide que les méthodes d'estimation alternées standards.

Chapitre 3. Caractérisation spectrale d'objets sub-stellaires à l'aide de la spectroscopie longue fente

Ce chapitre présente une méthode de traitement de données de spectrographie longue fente, permettant l'extraction du spectre d'un compagnon (*e.g.* exo-planète) orbitant autour de son étoile. L'instrument SPHERE/IRDIS permet ce type d'observation dans l'infra-rouge. Les données sont cependant fortement perturbées par les fuites stellaires. En effet, ces dernières sont bien plus intenses (jusqu'à $\sim 10^5$ fois plus intenses) que le signal du compagnon. Il s'agit donc de distinguer la composante venant du compagnon de la composante venant de l'étoile dans les données.

Pour ce faire, l'état de l'art se base sur une approximation de faible rang utilisant la décomposition en valeurs singulières tronquée des données. Ce faisant (et via une transformation des données pour aligner les fuites stellaires) les méthodes communément utilisées déduisent un modèle des fuites stellaires en négligeant le bruit et le signal du compagnon dans les données. Le signal du compagnon est reconstruit en moyennant le signal dans le voisinage de sa position dans les résidus formés par la soustraction de ce modèle aux données. Ces méthodes souffrent cependant du phénomène d'auto-soustraction, puisqu'une partie du compagnon est prise en compte dans le modèle des fuites stellaires et est donc soustraite aux données. Par ailleurs, la décomposition en valeurs singulières tronquée (sur laquelle est construit le modèle des fuites stellaires) ne prend pas en compte la non-uniformité du bruit dans les données et est donc statistiquement sous-optimale. Enfin, pour obtenir le modèle de la composante de l'étoile, les données doivent être transformées afin d'aligner les fuites stellaires. Cette transformation dépend fortement de la calibration géométrique du détecteur, c'est-à-dire de la relation existant entre les pixels et leurs coordonnées spatiales et spectrales.

Pour remédier à ces limitations, je présente dans la suite de ce chapitre la méthode Exospeco. Afin de convenablement séparer la composante du compagnon de celle de l'étoile, un modèle des données est conçu explicitant ces deux composantes. L'algorithme estime de manière alternée les deux composantes jusqu'à convergence. Le modèle direct dépendant fortement des coordonnées spatio-spectrales, il est essentiel de les estimer précisément. Afin de traiter de véritables données, une étape nécessaire de calibration géométrique a été développée afin de déduire via des fichiers spécifiques les lois de distributions spectrale et spatiale le long du détecteur.

Les résultats sur de vraies données réalisées par SPHERE/IRDIS montre l'importance de la calibration. Un effet de cisaillement de la loi de dispersion spatiale est visible dans les données. Exospeco est simple d'utilisation grâce à Amors. En effet, le modèle des fuites stellaires étant bi-linéaire, Amors est exploité pour estimer les deux composantes (spatiale et spectrale) de la composante stellaire. Ainsi, Exospeco nécessitant le réglage de seulement deux hyper-paramètres : ceux des régularisations appliquées à la densité spectrale d'énergie de l'étoile et du compagnon. Les tests effectués sur données réelles montrent la facilité de réglage de ces deux hyper-paramètres. Afin de contrôler la qualité de reconstruction de l'algorithme, j'ai effectué des tests sur des spectres synthétiques injectés dans de vraies données. Ces résultats montrent une bonne qualité de reconstruction pour des compagnons 10^{-4} fois moins brillants que leur étoile et situés de quelques $0.5''$ à plus de $1.8''$. Une comparaison à l'état de l'art montre bien les bénéfices d'Exospeco.

Chapitre 4. Reconstruction d'un objet étendu à partir d'images haut contraste

Les fuites stellaires ne limitent pas seulement l'étude spectrale de compagnon. SPHERE/IRDIS possède aussi un mode d'observation permettant l'imagerie des milieux circumstellaires. Ce mode d'observation permet l'étude de la morphologie de disques de grains et poussières en orbite autour de l'étoile et qui permet de comprendre la formation des exo-planètes. Les fuites stellaires sont fortement visibles dans ce type de données étant bien plus brillantes que l'objet d'intérêt ($\sim 10^4$ fois plus brillantes, en général).

Afin de pouvoir séparer la composante venant de l'étoile de la composante venant de l'objet d'intérêt, SPHERE/IRDIS utilise l'imagerie angulaire différentielle (ADI). Cette méthode d'observation permet d'apporter de la diversité pour distinguer les fuites stellaires de l'objet en capturant l'environnement autour de l'étoile plusieurs fois au cours de la nuit, permettant au champ de vision de tourner (rotation de la Terre) alors que la pupille de l'instrument reste fixe dans le champ.

Dans la suite est développé le modèle direct des données ADI prenant en compte la rotation du champ de vision, la transmission du masque de coronographie, la PSF de l'instrument et le bruit présent dans les données. Après une rapide étude des différentes méthodes de l'état de l'art, je présente l'approche problèmes inverses permettant la séparation des deux composantes dans les données. Le problème peut être séparé en deux étapes:

- L'estimation de la statistique des données (moyenne et covariance) connaissant le modèle de l'objet d'intérêt.
- La reconstruction de l'objet d'intérêt, sachant un modèle statistique des données.

La stratégie d'estimation alternée des différentes inconnues du problème est une solution simple pour séparer ces deux composantes. Cependant, là où le problème de reconstruction est proche des autres problèmes détaillés dans ce manuscrit, l'estimation de la covariance demande quant à elle une discussion plus détaillée.

En effet, au vu du nombre d'inconnues ($\sim 10^9$ éléments) et du nombre de mesures ($\sim 10^5$), estimer complètement la covariance n'est pas possible. Il faut donc limiter la covariance à une structure de plus faible complexité. Je présente dans ce chapitre une étude de l'impact de différentes approximations de la matrice de covariance et sur leur capacité à décrire les données. Pour ce faire, je développe une étude basée sur la validation croisée *leave-one-out* (LOOCV) des données centrées/résidus (*i.e.* la moyenne a été soustraite). Cette étude valide la prise en compte des corrélations spatiales dans les données permettant une meilleure description des résidus. Plus particulièrement, elle permet de valider la méthode ASAP d'approximation de la matrice de précision (inverse de la covariance), basée sur l'estimation parcimonieuse de la factorisation de Cholesky de la précision. La LOOCV m'a permise aussi de développer une méthode non supervisée de sélection du niveau de parcimonie de l'approximation faite par ASAP.

Au vu des résultats sur données ADI ne présentant aucun objet connu, la prise en compte des corrélations permet une meilleure description des fuites stellaires ce qui permet une meilleure distinction entre l'objet d'intérêt et ces fuites. Suivant l'exemple de REXPACO, la suite logique est de tester cette approximation dans le problème de reconstruction sur données avec un objet synthétique injecté et sur données réelles.

Conclusion

Dans ce manuscrit sont proposés de nouvelles méthodes permettant la séparation entre les objets d'intérêts et les différentes sources polluant les données. Pour améliorer ces méthodes et afin de s'approcher davantage du masque de coronographie et visualiser des objets plus faibles, il est possible de travailler sur le modèle des fuites stellaires afin d'expliquer ces dernières, proche du masque. Travailler sur le réglage automatique des hyper-paramètres permettrait aussi une utilisation simplifiée de ces méthodes tout en garantissant un réglage optimal des régularisations. Bien entendu, ces avancées doivent être conjointement développées avec de nouveaux instruments plus performants.

Unmixing and deconvolution methods of objects in images

Applications in astronomical high-contrast imaging

Abstract

High-contrast imaging of circumstellar environments plays a key role in understanding the formation and evolution of exoplanets. Instruments such as SPHERE have been dedicated to imaging such environments thanks to the coupling of extreme adaptive optics and powerful coronagraphy. However, aberrations in the light path result in strong artifacts in the acquired data. These "stellar leakages", combined with blurring effects and noises from the instrument limit greatly the study of such environments. Indeed, the stellar leakages are strongly preponderant in the data and make the recovery of faint objects of interest (exoplanets spectra, protoplanetary disks) difficult. During my PhD thesis, I developed several methods to overcome the challenging problem of disentangling objects of interest from perturbed data. Using realistic simulations and raw data from the SPHERE/IRDIS instrument, I discuss the modeling of the stellar leakages component in high-contrast imaging. Throughout the chapters, I model the data in terms of statistics and in terms of the different components that form them. Thanks to these models and following the inverse problems framework, I develop algorithms that unmix the objects of interest and the other sources of nuisance occurring in the data. Using the maximum a *posteriori*, to estimate the parameters of interest, I focused my work on three applications in astronomical imaging. Addressing the unavoidable blurring of optical instruments, I first worked on the **blind deconvolution** problem. Stating the problem as the estimation of both the object and the instrumental effects, characterized by its point spread function, I discussed the scaling ambiguity of the bi-linear model formed by these two components. Indeed scaling one component by a factor and the other by the inverse factor yields the same model. I proposed a strategy to estimate both components using this scaling ambiguity. Called AMORS, the algorithm uses the optimal scaling in the sense of the maximum *a posteriori* to reduce the number of hyper-parameters when estimating parameters of a bi-linear model. Noting that problems such as blind deconvolution depend greatly on the starting point, I also considered the starting point of the algorithm. Including the optimal scaling in the estimation process improves the convergence speed whatever the scaling of the initialization. I also focused my work on the SPHERE/IRDIS long-slit spectroscopy (LSS) mode. To **extract a companion spectrum** from LSS data, I developed the EXOSPECO algorithm. It uses a combined model of the contribution of the companion and the stellar leakages. Jointly with EXOSPECO, I developed a flexible calibration method that takes into account any distortions due to the optics of the instrument (*e.g.* shear effects). I demonstrated the importance of this step by confronting the algorithm with real LSS data. Thanks to its alternating estimation of the two components, and using only a single frame of LSS data, EXOSPECO reduces effectively the self-subtraction bias that affects state-of-the-art methods. The results on injected companions in real data obtained by the method show a faithful recovery of the companion SED with a relative

error between the synthetic and extracted SED of less than 20%, for companions 10^4 fainter than their host star at angular separations until 0.4″. Building upon the AMORS algorithm, EXOSPECO is easy to use by limiting the number of tuning parameters. In the last part of my work, I discussed the impact of the statistical model of these stellar leakages when imaging extended objects such as protoplanetary disks around young stars. Noting that the structures present in angular differential imaging (ADI) data are spatially correlated, I discuss **several approximations of the covariance matrix** of the noise component. This component of the data accounts for fluctuations of the stellar leakages, photon noise, as well as other sources of uncertainties. I validated on a real ADI dataset the ASAP approximation which builds a sparse approximation of the inverse of the covariance matrix. Finally, I developed an unsupervised strategy that selects automatically the level of sparsity of the approximation of the covariance.

Table des Matières

Contents

List of Symbols and Notations

ADC Analog to Digital Converter

ADI Angular Differential Imaging

ADU Analog to Digital Unit

ENDoF Effective Number of Degrees of Freedom

ExAO Extreme Adaptive Optics

IRDIS Infra-Red Dual Band Imager and Spectrograph

LSS Long-Slit Spectroscopy

MAP Maximum *A Posteriori*

mas milliarcseconds

MSE Mean Squared Error

NIR Near Infra-Red

PCA Principle Component Analysis

PSF Point Spread Function

PSNR Peak Signal-to-Noise Ratio

RDI Reference Differential Imaging

SED Spectral Energy Density

SNR Signal-to-Noise Ratio

SVD Singular Value Decomposition

TSVD Truncated Singular Value Decomposition

Instruments

IRDIS InfraRed Dual Imager Spectrograph

SPHERE Spectro Polarimetric High contrast Exoplanet REsearch

VLT Very Large Telescope

Mathematics

$\star$ Convolution product

$\odot$ Hadamard product (componentwise multiplication)

$\boldsymbol{v}$ Vector

$\mathbf{M}$ Matrix-operator

$\mathbf{I}$ Identity matrix

$\|\mathbf{M}\|_{\mathrm{F}}^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} |M_{i,j}|^2$ Squared Frobenius norm of $\mathbf{M}$

$\|\boldsymbol{v}\|^2 = \|\boldsymbol{v}\|_2^2 = \boldsymbol{v}^{\mathsf{T}} \boldsymbol{v}$ squared ℓ_2 norm of $\boldsymbol{v} \in \mathbb{R}^N$

$\|\boldsymbol{v}\|_{\mathbf{M}}^2 = \boldsymbol{v}^{\mathsf{T}} \mathbf{M} \boldsymbol{v}$ Squared Mahalanobis norm of $\boldsymbol{v} \in \mathbb{R}^N$

$\|\boldsymbol{v}\|_1 = \sum_{n=1}^{N} |v_n|$ ℓ_1 norm of $\boldsymbol{v} \in \mathbb{R}^N$

Statistics

$\hat{\boldsymbol{v}}$ Estimator of vector $\boldsymbol{v}$

$\mathrm{E}[\boldsymbol{v}]$ Statistical expectation of vector $\boldsymbol{v}$

$\mathrm{p}(\boldsymbol{v})$ Probability density distribution of vector $\boldsymbol{v}$

Introduction

The study of the surroundings of stars (circumstellar environments) is very important for our understanding of the formation processes and the evolution of solar systems. The physical phenomena involved in these environments are not fully understood yet, in particular in the formation of planets from dust particles and gas around the star. Studying these environments can also inform us about the vast diversity of companions that can form (*e.g.* super-earths, gas giants).

To study such environments, many methods have been developed throughout the years. They usually rely on indirect observations of the companion. The measurement of radial velocities is a common method of detection of sub-stellar companions (Latham et al., 1989). A companion orbiting around its host has a gravitational effect of *push-pull* on the star. This movement induces an oscillating Doppler shifting of the stellar light along the line of sight, which can be detected by high-resolution spectroscopy. However, this method does not work for face-on orbiting companions. In other cases, it only gives lower-bound information on the mass of the companion with the strongest Doppler shifting for edge-on orbits. This method also requires observations of at least one full orbit period to confirm the detection of a companion. The photometry method is another method to detect companions (Charbonneau et al., 2000). The companion reduces the apparent light of its host star during its transits. When allied with radial velocities, one can deduce numerous information on the companion such as its orbiting parameters, its mass, its temperature and the composition of its atmosphere. This method is however limited to companions with edge-on orbits and requires numerous observations over time.

Direct imaging is a method that provides unique information to characterize circumstellar environments. It is particularly adapted to the study of hot exoplanets (emitting in the infrared) around young stars, which coincides with the early stages of planetary systems. Because direct imaging produces an actual image of the surrounding of a star, it is also vastly used in the study of protoplanetary disks (disks of dust and gas that can form exoplanets) and debris disks (remaining disks after exoplanets formation). The measurement of the light distribution of the objects orbiting the star also allows for characterizing these objects spectrally. Direct imaging requires only a short observation time (a few hours) to have information on the companion's age, its mass and its composition.

The study of these stellar surroundings requires however the close observation of stars that are located tens to hundreds of light-years from us. Separating the light from companions or disks from that of the star is very challenging: their *angular separation*, *i.e.* the angle between the object, the system of observation and the star, is indeed very small, from 100 milliarcseconds (mas) to a few arcseconds. The angular resolution of telescopes, *i.e.* their capacity to distinguish two objects seen under a small angle, is at least limited by diffraction to λ/D with λ the wavelength at which we observe and D the diameter of the telescope's aperture. The larger the telescope the better the resolution. In ground-based astronomy, the turbulence of the atmosphere, caused by wind and temperature changes, limits further this resolution. Moreover, the signal from the objects of interest is dominated by the much brighter star. The *high contrast* between the companion intensity and its star is usually below 10^{-6} in infrared. Finally, the instrument itself adds perturbations to the scene it

observes as noises coming from the detector used to acquire data.

Telescopes have been designed to overcome these many limitations, like the Very Large Telescope (VLT) in the Atacama desert of Chili. Operated by the European Southern Observatory (ESO), it is composed of four units (UT) telescopes of 8.2m diameter and four auxiliary telescopes of 1.8m. On the third unit (UT3) is installed the SPHERE instrument (Spectro-Polarimetric High-contrast Exoplanet REsearch) (Beuzit et al., 2019). It is designed specifically for the study of circumstellar environments (see Fig. 1). It is composed of an Extreme Adaptive Optics (ExAO) system that can correct in real-time the effects of atmospheric turbulence, thus increasing greatly the resolution of the telescope. SPHERE also has a powerful coronagraph (Carbillet et al., 2011) which is used to hide the stellar light in order to reduce the contrast between the star and the object of interest.

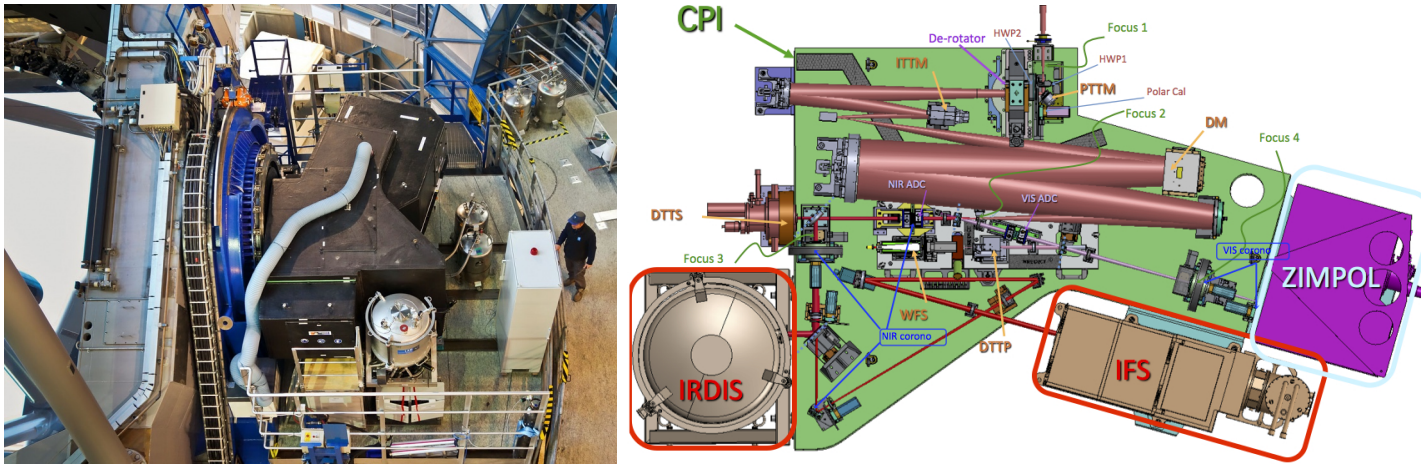

Figure 1 – The SPHERE instrument (Beuzit et al., 2019), pictured in the top-left image, is composed of an ExAO system, powerful coronagraphy and three instruments: the IFS, the ZIMPOL and the IRDIS instruments (seen in the top-right figure)(source: de Boer, J. et al., 2020).

SPHERE is composed of three scientific instruments: the Zurich IMaging POLarimeter (ZIMPOL) used for polarimetric observations in the visible and the Integral Field Instrument (IFS) and Infra-Red Dual Band Imager and Spectrograph (IRDIS), both dedicated to Near Infra-Red (NIR) observations. The IRDIS instrument (particularly discussed in this manuscript) (Dohlen et al., 2008), covers a 0.95 to 2.4 μm wavelength range. It possesses different observation modes such as the Long-Slit Spectroscopy (LSS) mode that is used to characterize circumstellar environments or the dual-band imaging mode. The latter mode images circumstellar environments at two different wavelengths at the same time thanks to a beam splitter and two chromatic filters on the resulting light paths. The LSS mode consists of placing a slit on the light path, followed by a dispersive element that decomposes the light (see Chapter 3).

Because of the atmospheric distortions in the field of view, the resolution of the telescope is limited to $\sim 0.6''$ in usual observation conditions. Thanks to its ExAO, SPHERE can reach an angular resolution close to the diffraction limit at $\sim$ 10mas on average. There remain residuals of the atmospheric perturbations as the correction is not perfect. This limitation has a blurring effect on the data taken by the instrument, characterized by the Point Spread Function (PSF) of the instrument (see Chapters 1 and 2). The strong coronagraphy of SPHERE also allows reaching contrast values of 10^{-3} to 10^{-5} between a companion and its host star, depending on the angular separation and wavelength. The masking of the star is still imperfect and with the limits imposed by diffraction, optical aberrations in the light path form a nuisance component in the data, called on-axis PSF, that mixes with the signal of interest (companion, disk). As it images in the Near Infra Red (NIR), the detector of IRDIS is very sensitive to temperature. Even with a cryostat protective enclosure, the acquired data is still corrupted by different sources of noise. Note that the SPHERE instrument has a derotator in the light's path which can play an important role in the acquisition of data. Indeed,

thanks to this device, the instrument's optical response can be stabilized while the field of view rotates throughout the night. This mode called *stabilized pupil* can help develop observation methods to improve the extraction of exoplanet or disk signals. Indeed, by taking multiple acquisitions throughout the same night of an object, the resulting data is an ensemble of images where the field of view has rotated (due to the Earth's rotation), while the response of the instrument (on-axis PSF) stayed quasi-static. This method of observation, called Angular Differential Imaging (ADI), allows using post-processing methods such as statistical modeling to separate/un-mix the object of interest from other components in the data (see Chapter 4).

All in all, the data produced by instruments like SPHERE are still corrupted by different phenomena which makes the recovery of the signal of interest (companions, disks) difficult. While taking into account the unavoidable noises of any acquisition method, data post-processing methods need to be developed to increase the performance of the instrument. Indeed, we want to observe areas usually close to the star, and so, we want to push the resolution capabilities of the telescope to the diffraction limit, while not being limited in angular separation between the object and its host star by the so-called on-axis PSF and the contrast. In this manuscript, we follow the inverse problems approach to recover objects from noisy and corrupted astronomical data. This problem is however not exclusively found in high-contrast imaging and so, the methods we propose in this manuscript can be adapted to numerous signal-processing contexts.

Chapter 1 details the formation of data by an optical instrument and the recovery of the signal of interest from such data in a general inverse problems framework. While illustrating these methods on a deconvolution problem, this chapter aims at giving all the tools necessary for the rest of the manuscript. Chapter 2 discusses the solving of bilinear models and more precisely the blind-deconvolution problem by a new approach, called AMORS, aiming to simplify the recovery of the objects of interest in such problems. Chapter 3 is focused on the design of EXOSPECO, a method for the extraction of a companion's spectrum in high-contrast imaging using the long-slit spectroscopy mode of the IRDIS instrument. Using the dual-band imaging mode of IRDIS and the ADI observation method, Chapter 4 is about recovering extended objects such as protoplanetary disks from high-contrast data thanks to a careful analysis of the noise in such data and particularly focusing on the impact of correlations in the data.

Chapter 1

The inverse problems framework

Separating several components present in the signal and compensating for instrument effects (*e.g.* noise, blur, *etc.*) is a problem found in many scientific contexts. Physics-based models of the instrument and of the observed scene can help to significantly reduce these nuisances. The inverse problems approach is a powerful formalism used in image processing to reconstruct images and retrieve parameters that characterize the objects from imperfect measurements. While physics-based models can describe the measurements corresponding to a given image/object thanks to a *direct* (or *forward*) model, the inverse problems approach uses measurements to recover the parameters of such model usually by means of the optimization of some criterion.

This first chapter is an introduction to the inverse problems approaches followed throughout this manuscript. It aims at providing the necessary tools used in the following chapters to process simulated and real astronomical data. To put in context and illustrate the problems and commonly used methods in image processing, we propose to consider throughout this chapter the problem of deconvolution. Section 1.1 describes the important step of carefully modeling the detector's effects on the data and how to take them into account. Section 1.2 describes the effects of the instrument and how they can be modeled. Section 1.3 gives examples of commonly used methods to reconstruct the object of interest while Section 1.4 details unsupervised strategies to tune the hyper-parameters introduced in the problem. Finally, Section 1.5 describes the main optimization methods used in this manuscript to solve the problems at hand.

In This Chapter

Topics:

- Careful model of the formation of data by a detector, including its statistical moments.
- General formalism of the inverse problems direct model and commonly used estimation methods.
- Description of some unsupervised methods useful for tuning hyper-parameters.
- Details on the optimization methods used in the manuscript.

1.1 Formation of data

In the visible and near-infrared astronomy context, the imaging system can be decomposed into two major components. One is the instrument responsible for the collection of light and image formation.

The other is the detector placed in the image plane of the instrument to record the distribution of the light intensity. In this section, we investigate the imaging chain which produces the observations we use in the inverse problem approach. Indeed, multiple physical phenomena occur during the measurement of the light flux in the detector's plane.

In charge of the acquisition, the detector samples the light intensity in the image plane over a grid of N pixels by counting the number of photo-electrons captured by each pixel's well during the exposition time of the acquisition. However, due to the quantum nature of light, an integer number of photons are collected which creates a photon noise or shot noise. It is particularly visible at a low Signal to Noise Ratio (SNR). Due to their thermal random motion, electrons in the substrate are passing the potential barrier of the pixels' well, adding to the effective count of electrons. This *dark current* - because it is not produced by light - increases with time and temperature. Although these sources of noise are considered as following a Poisson distribution, at high-intensity levels, these distributions can however be approximated by Gaussian distributions, as a result of the central limit theorem. Finally, Analog to Digital Converters (ADCs) convert the number of electrons into an integer numerical value. The digital conversion is responsible for adding rounding errors to the acquisition. Read-out noise describes the imperfect counting of electrons by these devices. ADCs also add up bias to the measurements. Taking into account these considerations yields the level given by the detector for the n-th pixel, as a function of the number of captured electrons (from light and dark current) e_n:

$$r_n = \frac{e_n}{g_n} + z_n \tag{1.1}$$

with g_n the ADC's gain[1] in electrons per Analog to Digital Unit (ADU) and z_n a random variable representing the ADC's bias, read-out noise and rounding errors that occur during the conversion[2]. Given the model of Eq. (1.1), the statistics of the raw data can be summarized by its statistical moments:

$$\mathrm{E}\big[r_n\big] = \frac{\mathrm{E}\big[e_n\big]}{g_n} + \mathrm{E}\big[z_n\big] \tag{1.2}$$

$$\mathrm{Var}(r_n) = \frac{\mathrm{Var}(e_n)}{g_n^2} + \mathrm{Var}(z_n) \tag{1.3}$$

The number of electrons e_n detected by the pixel following a Poisson distribution, its expectation can be decomposed in two terms:

$$\mathrm{E}\big[e_n\big] = \underbrace{q_n\,\phi_n^{[\mathrm{obj}]}}_{\mathrm{E}\left[e_n^{[\mathrm{obj}]}\right]} + \mathrm{E}\big[e_n^{[\mathrm{bg}]}\big] \tag{1.4}$$

where $e_n^{[\mathrm{obj}]}$ is the number of electrons resulting from the conversion of the flux $\phi_n^{[\mathrm{obj}]}$ coming from the image of interest on the detector's plane, via the uneven factor q_n. This factor represents the quantum efficiency as well as local modifications of intensity such as vignetting. $e_n^{[\mathrm{bg}]}$ contains all the other sources of electrons, *e.g.* the dark current and stray light (*e.g.* the instrument's thermal background) that do not interest us.

Given the assumed statistics of the random variables in Eq. (1.1) (see App. A for more details),

[1] The term "gain" used here represents the number of electrons necessary to pass from a quantified level of ADU to the next one.

[2] Note that due to their very different origins, z_n and e_n can be considered as mutually independent.

we deduce the statistical model of the raw data:

$$\mathrm{E}\big[r_n\big] = \underbrace{\frac{q_n}{g_n}}_{\propto 1/a_n} \phi_n^{[\mathrm{obj}]} + \underbrace{\frac{\mathrm{E}\big[e_n^{[\mathrm{bg}]}\big]}{g_n} + \mathrm{E}\big[z_n\big]}_{c_n} \tag{1.5}$$

$$\mathrm{Var}(r_n) = \frac{q_n}{g_n^2}\,\phi_n^{[\mathrm{obj}]} + \frac{\mathrm{E}\big[e_n^{[\mathrm{bg}]}\big]}{g_n^2} + \mathrm{Var}(z_n) \tag{1.6}$$

where a_n is a coefficient defined in Eq. (1.7). The expectation of the raw data in Eq. (1.5) is modeled as affine with respect to the flux of the object $\phi_n^{[\mathrm{obj}]}$ (Berdeu, Anthony et al., 2020), (Mugnier, Fusco, and Conan, 2004). Noting also the possibly uneven behavior of the detector, we choose to correct the raw data by:

$$d_n = a_n\,(r_n - c_n) \tag{1.7}$$

with a_n and c_n correcting factors to take into account the non-uniform behavior of the detector throughout its pixels grid and de-bias the data from signals other than one of the objects of interest. Appendix A summarizes a calibration methodology to estimate the different parameters in Eqs. (1.1) and (1.7). With this pre-processing step, we can work on data that are no longer biased by the detector and so, are simpler to handle.

Given this corrected model and simplifications described in Appendix A, each pixel d_n of the pre-processed data can be approximated as a Gaussian variable with the following expectation and variance:

$$\mathrm{E}\big[d_n\big] = \frac{a_n\,q_n}{g_n}\,\phi_n^{[\mathrm{obj}]} = \frac{\phi_n^{[\mathrm{obj}]}}{\phi^{[\mathrm{flat}]}} \tag{1.8}$$

$$\mathrm{Var}(d_n) = f_n\,\mathrm{E}\big[d_n\big] + \sigma_n^2 \tag{1.9}$$

with the expressions of σ_n^2 and f_n given in Eq. (A.14) of App. A and $\phi^{[\mathrm{flat}]}$ the constant flux of a diffuse source illuminating the entire detector, used to calibrate the non-uniformity of the response of the detector. Eq. (1.8) shows that the pre-processed data is uniformly proportional to the flux of the detector's image plane and is expressed in arbitrary units. It is however possible to convert to given units if the calibration flux $\phi^{[\mathrm{flat}]}$ is known.

The pre-processing of the raw data $\boldsymbol{r} \in \mathbb{R}^N$ of N pixels gives us the statistical model of the pre-processed data $\boldsymbol{d} \in \mathbb{R}^N$. These sufficient statistics propagate all information from the raw data and are very useful in the inverse problems framework (see Section 1.3). That is why we decide to work on the simpler pre-processed $\boldsymbol{d}$, called "data" throughout the manuscript.

1.2 The instrumental response and direct model

From the statistical model of Eqs. (1.8) and (1.9), the pre-processed data $\boldsymbol{d}$ can be decomposed as:

$$\boldsymbol{d} = \boldsymbol{m}(\boldsymbol{x}) + \boldsymbol{b} \tag{1.10}$$

with $\boldsymbol{m}(\boldsymbol{x})$, the so-called direct model, representing the expectation of $\boldsymbol{d}$ for a given object of interest $\boldsymbol{x}$ and instrument and $\boldsymbol{b} \sim \mathcal{N}(\boldsymbol{0}, \mathbf{C})$ a centered multivariate Gaussian variable of covariance matrix $\mathbf{C}$ which accounts for the different sources of noise described in Section 1.1 (and also possibly modeling errors in $\boldsymbol{m}(\boldsymbol{x})$). Equation (1.9) gives the variance of $\boldsymbol{d}$, *i.e.* the diagonal elements of $\mathbf{C}$.

Commonly used processing pipelines (*e.g.* (Pavlov, Feldt, and Henning, 2008)) generally consider simplified models, in particular, that the pre-processed data variance is identical for all pixels, *i.e.* it

is independent and identically distributed (i.i.d.). On the contrary, we use in the following manuscript a Julia (Bezanson et al., 2017) code[3] which gives as output the pre-processed data $\boldsymbol{d}$ and the value of the variance for each pixel. In that case, Eq. (1.9) can be rewritten:

$$\mathrm{Var}(d_n) = f_n \left[\boldsymbol{m}(\boldsymbol{x})\right]_n + \sigma_n^2 \tag{1.11}$$

where $[\boldsymbol{m}(\boldsymbol{x})]_n$ designates the n-th element of $\boldsymbol{m}(\boldsymbol{x})$. However, this model of the variance makes the estimation of the parameters $\boldsymbol{x}$ quite challenging. To avoid this problem, we choose the following approximation:

$$\mathrm{Var}(d_n) \simeq f_n \max(0, d_n) + \sigma_n^2 \tag{1.12}$$

which is accurate at high intensity, the $\max(0, d_n)$ errors in low intensity being negligible compared to σ_n^2. It is similar to the approximation made in Mistral (Mugnier, Fusco, and Conan, 2004) while also taking into account the non-uniformity of the detector's response.

In Eq. (1.8) our detector's model assumes a response that is linear with respect to the intensity of light. There are however some pixels that do not follow this rule. They usually appear as dead (dark) or saturated (bright) in the data and are not correctly described by the model of Eq.(1.8). Called "bad pixels", they need to be identified beforehand to be discarded by any subsequent processing. The inverse problem framework is particularly well suited to deal with these pixels: the corresponding variance can be set to infinity. Doing so avoids the need for interpolation of bad pixels that would introduce correlations in the pre-processed data.

Note also that due to cross-talk between pixel wells and other phenomena, the pixels of $\boldsymbol{d}$ may not be independent. Furthermore, extra processing steps such as the interpolation of bad pixels commonly done in the standard pipelines induce correlations to the pre-processed data. This limitation of the final statistical model assumed by Eqs. (1.8) and (1.9) is discussed in Chapter 4. This chapter details different possible approximations (i.i.d., independent, correlated) of the statistics of $\boldsymbol{d}$ and their impact on the description of real noisy data.

For our direct model to be valid, the acquisition needs also to plan the exposition time so that enough photons are seen by the detector. The choice of the exposition time needs to reach a tradeoff between having enough time to collect photons which will ensure a reasonable SNR while avoiding the saturation of the pixel wells.

Imperfections from the instrument and the observation technics such as misalignments or aberrations in the optics, turbulences in the light's path and diffraction phenomena have various effects that produce distortions. Considering the distribution of light from the object $\boldsymbol{x}$ is incoherent, the instrumental response is linear in intensity and the resulting pre-processed data can be modeled by:

$$\boldsymbol{d} = \mathbf{H}\,\boldsymbol{x} + \boldsymbol{b} \tag{1.13}$$

with $\mathbf{H}$ a linear operator implementing the instrument effects, including for instance the rotation of the field of view and blur effects.

Direct model in the deconvolution problem context

When the instrument capturing the data has a small field of view, the effects modeled by $\mathbf{H}$ can be assumed to be shift invariant[4]. In other words, the various phenomena caused by the instrument do not depend on the absolute position in the image plane. In the deconvolution problem, such effects are represented by a spatial Point Spread Function (PSF), $\boldsymbol{h}$, of the optical system. It is the image captured by the instrument of a point-like source (Dirac distribution) in the field of view. This PSF is usually spatially extended and has a blurring effect on the observation. Being shift-invariant, this

[3] https://git-cral.univ-lyon1.fr/emmt/scientificdetectors/

[4] Approximation that holds in high-contrast imaging, except close or under the coronagraphic mask.

effect is modeled as the convolution of the scene $\boldsymbol{x}$ by the PSF of the instrument, rewriting the direct model in Eq. (1.13) as (see Fig. 1.1):

$$\boldsymbol{m} = \mathbf{H}\,\boldsymbol{x} \simeq \boldsymbol{h} \star \boldsymbol{x} \tag{1.14}$$

with $\star$ the spatial convolution operator and the approximations errors noted by $\simeq$ can be taken into account by the term $\boldsymbol{b}$. This convolution model is flexible enough to take into account many different blurring processes coming from the observation. The diffraction is the most common limitation as it concerns every instrument. Given a circular pupil of diameter D, it limits the angular resolution to λ/D with λ the wavelength of the light. Even ideal optical systems (*i.e.* aberrations-free) are limited by diffraction and cannot resolve point-like sources perfectly. Their response in such conditions has been modeled by Airy (1835) as an Airy function that limits its capacity to distinguish details in the scene. In our astronomical context, turbulence in the atmosphere is severely modifying the wavefront of the light we want to observe. Adaptive Optics (AO) systems correct the wavefront "on the fly" trying to reach the diffraction limit. This system has however its limitations and still leaves residual aberrations. Other blurring effects can be observed such as motion blur which is produced by movements of the instrument or object moving significantly during the integration time of the acquisition.

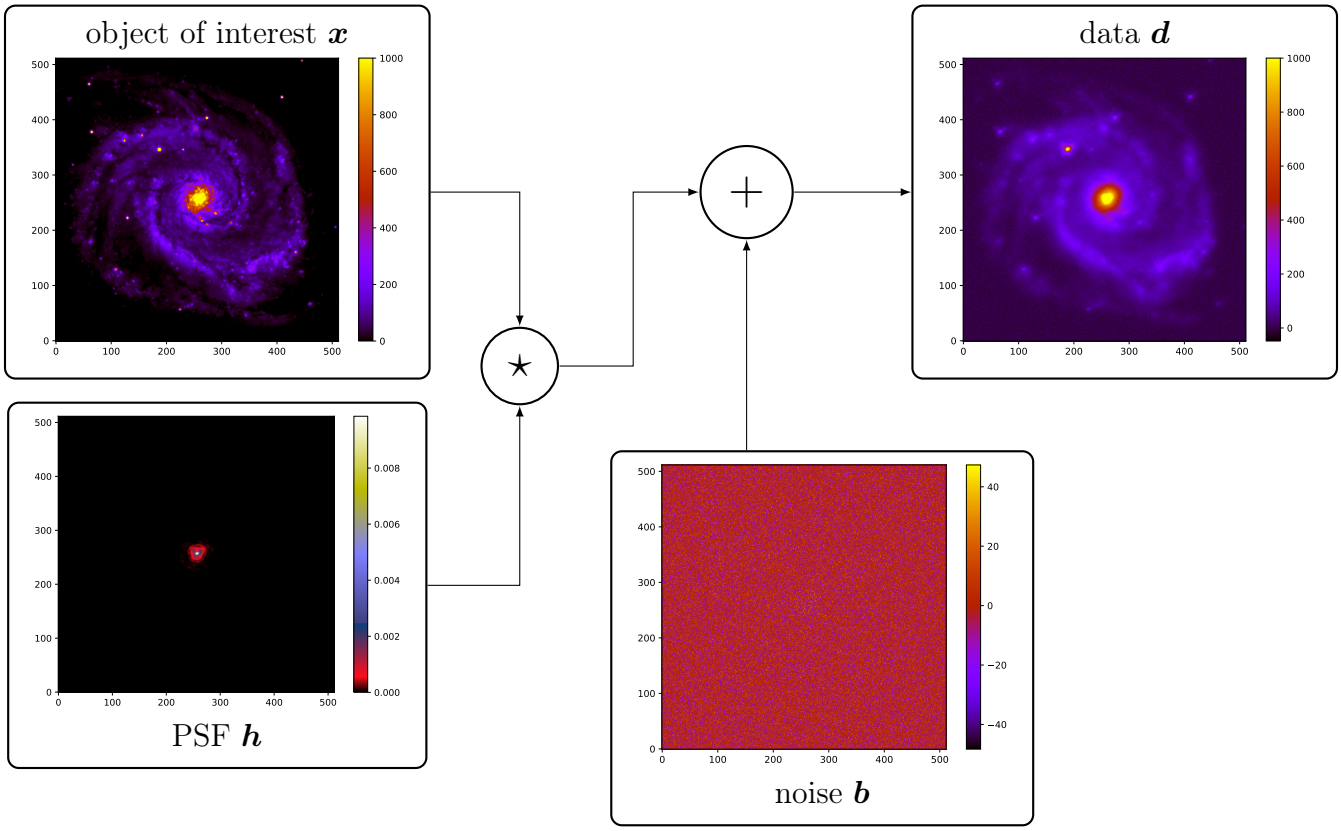

Figure 1.1 – Model of the data $\boldsymbol{d}$ as the convolution of an object of interest $\boldsymbol{x}$ and a PSF $\boldsymbol{h}$ and corrupted by a non-uniform Gaussian noise $\boldsymbol{b}$.

Illustration on simulated deconvolution problem

Following the direct model of Eq. (1.14), we simulate astronomical data by convolving an image of the galaxy M51 (shown in Fig. 1.2) by an experimental PSF from the first AO system at the 3.6m Canada-France Hawaï telescope[5]. To illustrate the deconvolution problem on realistic simulated data,

[5]Courtesy of François Roddier

we decided to use a mixture of Poisson and Gaussian noise, representing respectively the photon noise and read-out noise contributions. Going back to Eq. (1.12), we consider in our simulations a uniform response of the detector leading to the noise model $\boldsymbol{b} \sim \mathcal{N}(\mathbf{0}, \mathbf{C} = \mathrm{diag}(\boldsymbol{v}))$[6], where:

$$\boldsymbol{v} = \frac{\boldsymbol{m}}{g} + \sigma^2 \tag{1.15}$$

with $\boldsymbol{m} = \boldsymbol{h} \star \boldsymbol{x}$ the blurred image, g the detector's ADCs gain and σ the read-out noise standard deviation. Tests in the following chapter are done on such simulated data of Peak Signal to Noise Ratio (PSNR):

$$\mathrm{PSNR} = 10 \log\Big(\frac{\big(\max(\boldsymbol{m}) - \min(\boldsymbol{m})\big)^2}{\frac{1}{N}\sum_{n=1}^{N}(d_n - m_n)^2} \Big) \simeq 40 \tag{1.16}$$

with N the number of pixels. Note that we consider here the detector's response as uniform to simplify the simulation process.

1.3 Estimation of the model's parameters

1.3.1 Straightforward inversion

A first straightforward approach to recover an estimation $\hat{\boldsymbol{x}}^{[\mathrm{inv}]}$ of the scene $\boldsymbol{x}$ would be given by:

$$\hat{\boldsymbol{x}}^{[\mathrm{inv}]} = \mathbf{H}^{\dagger}\, \boldsymbol{d} \tag{1.17}$$

with $\boldsymbol{H}^{\dagger}$ the pseudo-inverse of $\mathbf{H}$. $\mathbf{H}^{\dagger}$ can be expressed differently depending on the problem (dimensions, invertibility of $\mathbf{H}$). For instance, the Moore-Penrose pseudo-inverse is used when the problem is ill-posed (more unknowns than measurements, $\mathbf{H}$ is a wide matrix). In this case, $\mathbf{H}$ is not a square matrix and so is not directly invertible. When $\mathbf{H}^{\mathsf{T}}\mathbf{H}$ is invertible, the pseudo-inverse is:

$$\mathbf{H}^{\dagger} = (\mathbf{H}^{\mathsf{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathsf{T}}\,. \tag{1.18}$$

Note that if $\mathbf{H}^{-1}$ exists, the pseudo-inverse solution amounts to:

$$\hat{\boldsymbol{x}}^{[\mathrm{inv}]} = \mathbf{H}^{-1}\,\boldsymbol{d}\,. \tag{1.19}$$

The result of this straightforward inversion on simulated data is shown in Fig. 1.2 where no features of the ground truth are recognizable. This heavily perturbed image can be explained by the fact that dividing $\boldsymbol{d}$ by the PSF (or its pseudo-inverse) rewrites Eq. (1.17) as:

$$\hat{\boldsymbol{x}}^{[\mathrm{inv}]} = \boldsymbol{x} + \mathbf{H}^{\dagger}\,\boldsymbol{b} \tag{1.20}$$

with the first term explained by the fact that we consider the ideal (but usually not true) case where all modes of the ground truth $\boldsymbol{x}$ are seen in the data. The Fourier transform of the instrument operator is usually low at high frequencies (*i.e.* it has a smooth shape) whereas the noise $\boldsymbol{b}$ is significant at all frequencies. The last term of Eq. (1.20) then leads to an amplification of the noise $\boldsymbol{b}$ as it is divided at high frequencies by low levels of the MTF. This illustrates that finding $\boldsymbol{x}$ from $\boldsymbol{d}$ requires proper reconstruction methods.

[6] with $\mathrm{diag}(\boldsymbol{v})$ an operator forming a diagonal matrix with diagonal elements given by the vector $\boldsymbol{v}$

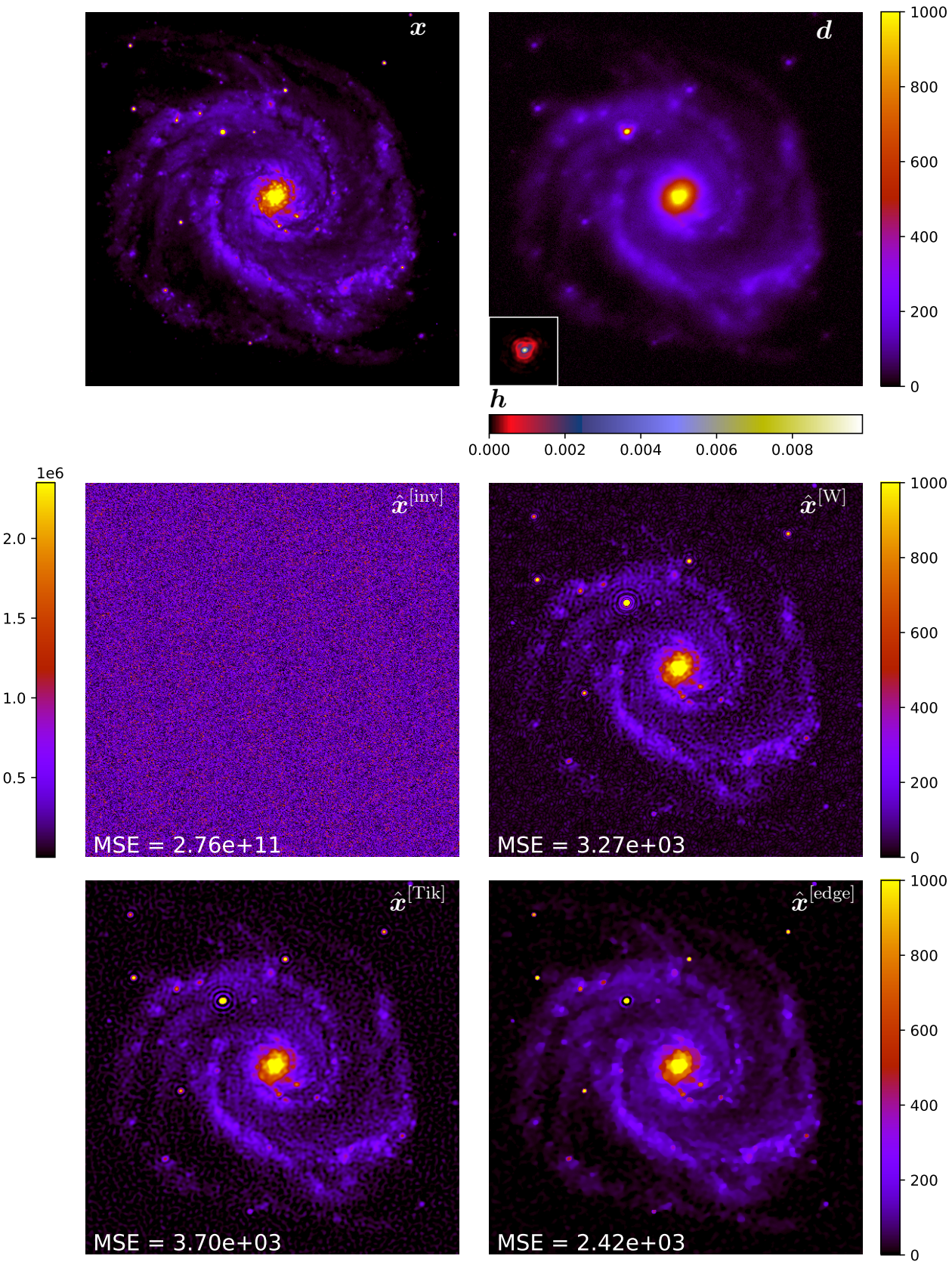

Figure 1.2 – Ground truth object $\boldsymbol{x}$ giving the simulated data $\boldsymbol{d}$ when convoluted to PSF $\boldsymbol{h}$ (overlapped to the top-right figure), as discussed in Section 1.2. The results of the straightforward inversion $\boldsymbol{x}^{[\mathrm{inv}]}$ (color bar on the left), the Wiener filter $\boldsymbol{x}^{[\mathrm{W}]}$, the MAP estimator with quadratic regularization $\boldsymbol{x}^{[\mathrm{Tik}]}$ and with the edge-preserving regularization $\boldsymbol{x}^{[\mathrm{edge}]}$ are presented. The MSE between the ground truth $\boldsymbol{x}$ and the different estimators are given.

1.3.2 The Wiener filter

To retrieve an estimation of the scene and avoid noise amplification, one may want to minimize the Mean Squared Error (MSE). Indeed, this criterion measures the discrepancies between the ground truth and the estimate. In that spirit, Wiener, 1949 developed an affine transformation of the data, relying on the minimization of this criterion which gives the estimator:

$$\hat{\boldsymbol{x}}^{[\mathrm{W}]} = \mathbf{F}^{[\mathrm{W}]}\,\boldsymbol{d} + \boldsymbol{p}^{[\mathrm{W}]} \tag{1.21}$$

with $\boldsymbol{d} = \mathbf{H}\,\boldsymbol{x} + \boldsymbol{b}$ the model of the data of Eq. (1.13). The parameters of this transform are deduced by:

$$(\mathbf{F}^{[\mathrm{W}]}, \boldsymbol{p}^{[\mathrm{W}]}) = \underset{\mathbf{F},\boldsymbol{p}}{\arg\min}\,\mathrm{MSE} \tag{1.22}$$

where

$$\mathrm{MSE} = \mathrm{E}_{\boldsymbol{x},\boldsymbol{b}}\big[\|\boldsymbol{x}^{[\mathrm{W}]} - \boldsymbol{x}\|^2\big] = \mathrm{E}_{\boldsymbol{x},\boldsymbol{b}}\big[\|\mathbf{F}\,\boldsymbol{d} + \boldsymbol{p} - \boldsymbol{x}\|^2\big] \tag{1.23}$$

with $\mathrm{E}_{\boldsymbol{u}}[\boldsymbol{u}]$ the expectation of a random variable $\boldsymbol{u}$, noted $\bar{\boldsymbol{u}}$ in the following. The minimization gives the estimates:

$$\boldsymbol{p}^{[\mathrm{W}]} = \underbrace{\mathrm{E}_{\boldsymbol{x}}[\boldsymbol{x}]}_{\bar{\boldsymbol{x}}} - \mathbf{F}^{[\mathrm{W}]}\,\underbrace{\mathrm{E}_{\boldsymbol{b}}[\boldsymbol{d}]}_{\bar{\boldsymbol{d}} = \mathbf{H}\,\bar{\boldsymbol{x}} + \bar{\boldsymbol{b}}} \tag{1.24}$$

and

$$\mathbf{F}^{[\mathrm{W}]} = \mathrm{Cov}(\boldsymbol{x})\,\mathbf{H}^{\mathsf{T}}\Big(\mathbf{H}\,\mathrm{Cov}(\boldsymbol{x})\,\mathbf{H}^{\mathsf{T}} + \mathrm{Cov}(\boldsymbol{b})\Big)^{-1} \tag{1.25}$$

Note that the effect of $\boldsymbol{p}^{[\mathrm{W}]}$ is to recenter the estimator $\boldsymbol{x}^{[\mathrm{W}]}$. Using the hypothesis that $\mathrm{E}[\boldsymbol{x}] = \mathbf{0}$ and $\mathrm{E}[\boldsymbol{b}] = \mathbf{0}$, the transformation's parameter $\boldsymbol{p}^{[\mathrm{W}]} = \mathbf{0}$ and the Wiener estimator of the scene is simply:

$$\hat{\boldsymbol{x}}^{[\mathrm{W}]} = \mathbf{F}^{[\mathrm{W}]}\,\boldsymbol{d} \tag{1.26}$$

Proof: Taking the derivative of the MSE with respect to $\boldsymbol{p}$

$$\begin{aligned}\frac{\partial \mathrm{MSE}}{\partial \boldsymbol{p}} &= 2\,\mathrm{E}_{\boldsymbol{x},\boldsymbol{b}}\big[\mathbf{F}\,\boldsymbol{d} + \boldsymbol{p} - \boldsymbol{x}\big] \\ &= 2\big(\mathbf{F}\,\mathrm{E}_{\boldsymbol{b}}[\boldsymbol{d}] + \boldsymbol{p} - \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{x}]\big)\end{aligned} \tag{1.27}$$

the estimate $\boldsymbol{p}^{[\mathrm{W}]}$ is obtained for:

$$\begin{aligned}\frac{\partial \mathrm{MSE}}{\partial \boldsymbol{p}}\bigg|_{\boldsymbol{p}=\boldsymbol{p}^{[\mathrm{W}]}} &= \mathbf{0} \\ \Leftrightarrow \boldsymbol{p}^{[\mathrm{W}]} &= \underbrace{\mathrm{E}_{\boldsymbol{x}}[\boldsymbol{x}]}_{\bar{\boldsymbol{x}}} - \mathbf{F}\,\underbrace{\mathrm{E}_{\boldsymbol{b}}[\boldsymbol{d}]}_{\bar{\boldsymbol{d}} = \mathbf{H}\,\bar{\boldsymbol{x}} + \bar{\boldsymbol{b}}}\end{aligned} \tag{1.28}$$

which when injected in Eq. (1.23) yields:

$$\begin{aligned}\mathrm{MSE} &= \mathrm{E}_{\boldsymbol{x},\boldsymbol{b}}\Big[\big\|\mathbf{F}\,\boldsymbol{d} + \big(\bar{\boldsymbol{x}} - \mathbf{F}\,\bar{\boldsymbol{d}}\big) - \boldsymbol{x}\big\|^2\Big] \\ &= \mathrm{E}_{\boldsymbol{x},\boldsymbol{b}}\Big[\big\|\mathbf{F}\,\big(\boldsymbol{d} - \bar{\boldsymbol{d}}\big) - \big(\boldsymbol{x} - \bar{\boldsymbol{x}}\big)\big\|^2\Big]\end{aligned} \tag{1.29}$$

Taking the derivative of this MSE[7] with respect to $\mathbf{F}$:

$$\begin{aligned}\frac{\partial \text{MSE}}{\partial \mathbf{F}} &= 2\,\text{E}_{\boldsymbol{x},\boldsymbol{b}}\left[\left(\mathbf{F}\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)-\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)\right)\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)^{\mathsf{T}}\right] \\ &= 2\,\mathbf{F}\,\underbrace{\text{E}_{\boldsymbol{b}}\left[\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)^{\mathsf{T}}\right]}_{=\,\text{Cov}(\boldsymbol{d})} - 2\,\text{E}_{\boldsymbol{x},\boldsymbol{b}}\left[\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)^{\mathsf{T}}\right]\end{aligned} \tag{1.30}$$

with $\text{Cov}(\boldsymbol{d})$ the covariance matrix of $\boldsymbol{d}$, the estimate can be deduced in the same manner:

$$\begin{aligned}&\left.\frac{\partial \text{MSE}}{\partial \mathbf{F}}\right|_{\mathbf{F}=\mathbf{F}^{[\text{W}]}} = \mathbf{0} \\ \Leftrightarrow\ &\mathbf{F}^{[\text{W}]}\,\text{Cov}(\boldsymbol{d}) = \text{E}_{\boldsymbol{x},\boldsymbol{b}}\left[\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)^{\mathsf{T}}\right]\end{aligned} \tag{1.31}$$

Given that $\boldsymbol{d} = \mathbf{H}\,\boldsymbol{x} + \boldsymbol{b}$ and $\bar{\boldsymbol{d}} = \text{E}[\boldsymbol{d}] = \mathbf{H}\,\text{E}[\boldsymbol{x}] + \text{E}[\boldsymbol{b}] = \mathbf{H}\,\bar{\boldsymbol{x}} + \bar{\boldsymbol{b}}$,

$$\boldsymbol{d}-\bar{\boldsymbol{d}} = \mathbf{H}\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right) - \left(\boldsymbol{b}-\bar{\boldsymbol{b}}\right) \tag{1.32}$$

Considering $\boldsymbol{x}$ and $\boldsymbol{b}$ as mutually independent, the right-hand term of Eq. (1.31) can then be rewritten:

$$\begin{aligned}\text{E}\left[\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)^{\mathsf{T}}\right] &= \text{E}\left[\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)\left(\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)^{\mathsf{T}}\mathbf{H}^{\mathsf{T}} - \left(\boldsymbol{b}-\bar{\boldsymbol{b}}\right)^{\mathsf{T}}\right)\right] \\ &= \underbrace{\text{E}\left[\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)^{\mathsf{T}}\right]}_{=\,\text{Cov}(\boldsymbol{x})}\mathbf{H}^{\mathsf{T}} - \underbrace{\text{E}\left[\boldsymbol{x}-\bar{\boldsymbol{x}}\right]}_{=\,\mathbf{0}}\underbrace{\text{E}\left[\boldsymbol{b}-\bar{\boldsymbol{b}}\right]^{\mathsf{T}}}_{=\,\mathbf{0}^{\mathsf{T}}}\end{aligned} \tag{1.33}$$

In the same way:

$$\begin{aligned}\text{Cov}(\boldsymbol{d}) &= \text{E}\left[\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)\left(\boldsymbol{d}-\bar{\boldsymbol{d}}\right)^{\mathsf{T}}\right] \\ &= \text{E}\left[\left(\mathbf{H}\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)-\left(\boldsymbol{b}-\bar{\boldsymbol{b}}\right)\right)\left(\mathbf{H}\left(\boldsymbol{x}-\bar{\boldsymbol{x}}\right)-\left(\boldsymbol{b}-\bar{\boldsymbol{b}}\right)\right)^{\mathsf{T}}\right] \\ &= \mathbf{H}\,\text{Cov}(\boldsymbol{x})\,\mathbf{H}^{\mathsf{T}} + \text{Cov}(\boldsymbol{b})\end{aligned} \tag{1.34}$$

Considering these two last results, we can deduce from Eq. (1.31):

$$\mathbf{F}^{[\text{W}]} = \text{Cov}(\boldsymbol{x})\,\mathbf{H}^{\mathsf{T}}\left(\mathbf{H}\,\text{Cov}(\boldsymbol{x})\,\mathbf{H}^{\mathsf{T}} + \text{Cov}(\boldsymbol{b})\right)^{-1} \tag{1.35}$$

■

Going back to the deconvolution problem explained in Section 1.2, the operator $\mathbf{H}$ is approximating a convolution. Using the Fourier transform $\breve{\boldsymbol{x}}$ of $\boldsymbol{x}$, it is possible to redefine Eq. (1.26) in the Fourier space, where each spectral channel of the estimate ℓ is given by:

$$\breve{\hat{x}}_\ell = \breve{F}_\ell^{[\text{W}]}\,\breve{d}_\ell \tag{1.36}$$

with $\breve{d}_\ell = \breve{h}_\ell\,\breve{x}_\ell + \breve{b}_\ell$ the model of the Fourier transformed data (Eq. (1.14)). When considering that $\text{Cov}(\boldsymbol{x})$ and $\text{Cov}(\boldsymbol{b})$ are stationary, the Wiener filter is then defined by:

$$\breve{F}_\ell^{[\text{W}]} = \frac{\breve{h}_\ell^*}{|\breve{h}_\ell|^2 + \underbrace{\text{E}\left[|\breve{b}_\ell|^2\right]/\text{E}\left[|\,\breve{x}_\ell|^2\right]}_{1/\text{SNR}^2}} \tag{1.37}$$

[7] Noting that $\frac{\partial}{\partial \mathbf{A}}\left(\|\mathbf{A}\,\boldsymbol{x}-\boldsymbol{s}\|^2\right) = 2\left(\mathbf{A}\,\boldsymbol{x}-\boldsymbol{s}\right)\boldsymbol{x}^{\mathsf{T}}$

with the Signal-to-Noise Ratio (SNR) defined as the power spectral density of the scene over the variance of the noise. The Wiener filter yields good results because of how it is handling the noise component. The higher, compared to $|\breve{h}_\ell|^2$, the SNR at the spectral channel ℓ (*i.e.* the noise is low), the closer the Wiener filter is to the direct inversion of the operator, discussed in Section 1.3.1. At frequencies with a low SNR compared to $|\breve{h}_\ell|^2$ (the noise is preponderant), the filter will be close to zero, thus cutting off the corresponding Fourier frequency information.

In practice, the power spectral densities of the scene and of the noise are not known and need to be estimated. As the scene is a natural image, its spectral distribution can be modeled as a parametric power model. For instance, Conan et al., 1998 developed a custom model in the astronomical context, based on the power model of Kattnig and Primot, 1997, where:

$$\mathrm{E}\left[|\breve{b}_\ell|^2\right] = \text{cste} \tag{1.38}$$

and

$$\mathrm{E}\left[|\breve{x}_\ell|^2\right] = \frac{\mu_{\boldsymbol{d}}^2}{1 + \left(\nu_\ell / \nu_{\text{cut}}\right)^p} \tag{1.39}$$

with $\mu_{\boldsymbol{d}}$ the average total flux, ν_ℓ the length of the spatial frequency of the spectral channel ℓ, ν_{cut} a cutoff frequency, and p a hyper-parameter to estimate. The parameters can be estimated thanks to an iterative minimization method (*e.g.* Newuoa (Powell, 2006), see Section 1.5.3) applied to an estimation of the MSE criterion[8]. Thanks to this tuning of hyper-parameters, artifacts in the reconstruction due to bad tuning can be avoided. Fig. 1.2 gives the results of this fitted Wiener filter. As can be seen, this estimation of the Wiener filter gives reasonably good results, reducing by far the MSE compared to the straightforward inversion. It is however difficult to estimate the hyper-parameters of the model as it needs to precisely describe the power spectral density. Other methods exist that can enforce more general prior information (see Section 1.3.4).

1.3.3 The Maximum Likelihood

Given the probability density function of the data, it is possible to estimate the scene by an inverse problem approach. That is, under the Gaussian noise hypothesis:

$$\mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x}) = \left|\frac{\mathbf{W}}{2\pi}\right|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\left\|\boldsymbol{d} - \mathbf{H}\,\boldsymbol{x}\right\|_{\mathbf{W}}^2\right\} \tag{1.40}$$

with $\|\boldsymbol{x}\|_{\mathbf{W}}^2 = \boldsymbol{x}^\mathsf{T}\,\mathbf{W}\,\boldsymbol{x}$ the squared Mahalanobis norm of $\boldsymbol{x}$ and $\mathbf{W} = \mathbf{C}^{-1}$ the precision matrix of $\boldsymbol{d}$ or an approximation that can take into account bad pixels, approximation errors such as in Eq. (1.14). We want to find the parameter $\boldsymbol{x}$ which maximizes Eq. (1.40). The Maximum Likelihood (ML) estimator is defined by:

$$\begin{aligned}
\hat{\boldsymbol{x}}^{[\mathrm{ML}]} &= \arg\max_{\boldsymbol{x}} \mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x}) \\
&= \arg\min_{\boldsymbol{x}} -\log \mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x}) \\
&= \arg\min_{\boldsymbol{x}} \|\boldsymbol{d} - \mathbf{H}\,\boldsymbol{x}\|_{\mathbf{W}}^2
\end{aligned} \tag{1.41}$$

With this estimator, we want to minimize the discrepancies between the data and our model, measured by the Mahalanobis norm.

This estimator has a closed-form solution which can be deduced by zeroing the gradient of the function to minimize:

$$\frac{\partial}{\partial \boldsymbol{x}}\left(\|\boldsymbol{d} - \mathbf{H}\,\boldsymbol{x}\|_{\mathbf{W}}^2\right)\Big|_{\boldsymbol{x}=\hat{\boldsymbol{x}}^{[\mathrm{ML}]}} = 2\,\mathbf{H}^\mathsf{T}\,\mathbf{W}\,(\boldsymbol{d} - \mathbf{H}\,\hat{\boldsymbol{x}}^{[\mathrm{ML}]}) = \mathbf{0} \tag{1.42}$$

[8] See Section 1.4 for a more detailed explanation on the estimation of hyper-parameters.

giving the well-known system of *normal equations*:

$$(\mathbf{H}^\mathsf{T}\,\mathbf{W}\,\mathbf{H})\,\hat{\boldsymbol{x}}^{[\mathrm{ML}]} = \mathbf{H}^\mathsf{T}\,\mathbf{W}\,\boldsymbol{d} \tag{1.43}$$

which, in this form, takes into account the correlations and variances of the data.

Note that for an independent and identically distributed noise, *i.e.* $\mathbf{C} = a\,\mathbf{I}$ with $\mathbf{I}$ the identity matrix and $a \in \mathbb{R}_+$ the variance of $\boldsymbol{b}$, the ML estimator amounts to the straightforward inversion of Eq. (1.17):

$$\begin{aligned}\hat{\boldsymbol{x}}^{[\mathrm{ML}]} &= (\mathbf{H}^\mathsf{T}\,\mathbf{H})^{-1}\,\mathbf{H}^\mathsf{T}\,\boldsymbol{d} \\ &= \hat{\boldsymbol{x}}^{[\mathrm{inv}]}\end{aligned} \tag{1.44}$$

when $\mathbf{H}^\mathsf{T}\,\mathbf{H}$ is invertible.

The problem defined in Eq. (1.41) is usually ill-conditioned and so the ML estimator is not suitable. Indeed, in these kinds of problems, small perturbations in the data (*e.g.* due to noise) lead to large changes in the solution, because the estimator attempts to fit the noise in the data. The ML estimator is an unbiased estimator but with a high variance.

1.3.4 Taking *a priori* knowledge into account - The Maximum *A Posteriori*

To reduce the variance of the estimator, it is necessary to introduce some additional knowledge in its design. The Maximum *A Posteriori* (MAP) method allows taking into account *a priori* information. It is expressed by the maximization of the object's probability given the data:

$$\hat{\boldsymbol{x}}^{[\mathrm{MAP}]} = \arg\max_{\boldsymbol{x}} \mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{d}, \boldsymbol{\theta}) \tag{1.45}$$

with $\boldsymbol{\theta}$ representing some extra parameters. $\boldsymbol{x}$ not being observed, it is possible to rewrite the problem using the Bayes rule from which:

$$\mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{d}, \boldsymbol{\theta}) \propto \mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x})\,\mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{\theta}) \tag{1.46}$$

with $\mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{\theta})$ an assumed prior distribution of the scene parametrized by $\boldsymbol{\theta}$ and considering the statistics of $\boldsymbol{d}$ does not depend on $\boldsymbol{\theta}$. It is possible to simplify the problem, by expressing this prior probability as being a Gibbs distribution, *i.e.* $\mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{\theta}) \propto \exp\big\{-\mathcal{J}(\boldsymbol{x};\boldsymbol{\theta})\big\}$. In the Gaussian hypothesis, Eq. (1.45) becomes with this rule:

$$\begin{aligned}\hat{\boldsymbol{x}}^{[\mathrm{MAP}]} &= \arg\max_{\boldsymbol{x}} \mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x})\,\mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{\theta}) \\ &= \arg\min_{\boldsymbol{x}} \big(-\log(\mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x})) - \log(\mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{\theta}))\big) \\ &= \arg\min_{\boldsymbol{x}} \big(\underbrace{\|\boldsymbol{d} - \mathbf{H}\,\boldsymbol{x}\|^2_{\mathbf{W}}}_{\text{data fidelity}} + \underbrace{\mathcal{J}(\boldsymbol{x};\boldsymbol{\theta})}_{\textit{prior}}\big)\end{aligned} \tag{1.47}$$

with $\mathcal{J}(\boldsymbol{x};\boldsymbol{\theta})$ a *regularization* term. Its role is to drive the estimator toward the prior, reducing in return the variance of the estimator. Such a model usually favors simple solutions presenting continuity or sparsity of the estimation. In $\boldsymbol{\theta}$ are contained *hyper-parameters* that control the level of this regularization compared to the squared Mahalanobis norm. In other words, tuning these hyper-parameters controls the level of *a priori* information we want to enforce in our solution with respect to the fidelity it must have to the data. Hence, the value of $\boldsymbol{\theta}$ is of uttermost importance as it tunes the bias-variance tradeoff of the estimator.

Note that, depending on the form of the regularization, a closed-form solution of Eq. (1.47) does not always exist, leading to the need for iterative methods of minimization such as gradient descent or quasi-Newton methods (see Section 1.5).

Link with the Wiener filter

With a Gaussian prior (*e.g.* quadratic smoothness), we can rewrite Eq. (1.47) as:

$$\hat{\boldsymbol{x}}^{[\mathrm{MAP}]} = \arg\min_{\boldsymbol{x}} \left\{ \mathcal{L} = \|\boldsymbol{d} - \mathbf{H}\,\boldsymbol{x}\|^2_{\mathbf{W}} + \mu\,\|\boldsymbol{x}\|^2_{\mathbf{M}} \right\} \tag{1.48}$$

with $\mathbf{M}$ playing the role of precision matrix of $\boldsymbol{x}$[9] and assuming $\mathrm{E}[\boldsymbol{x}] = \mathbf{0}$. In that case, a closed-form solution can be derived as:

$$\begin{aligned} \left.\frac{\partial \mathcal{L}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=-\hat{\boldsymbol{x}}^{[\mathrm{MAP}]}} &= 2\mathbf{H}^\mathsf{T}\,\mathbf{W}\left(\boldsymbol{d} - \mathbf{H}\,\hat{\boldsymbol{x}}^{[\mathrm{MAP}]}\right) + 2\mu\,\mathbf{M}\,\hat{\boldsymbol{x}}^{[\mathrm{MAP}]} = \mathbf{0} \\ \Leftrightarrow\ \hat{\boldsymbol{x}}^{[\mathrm{MAP}]} &= \left(\mathbf{H}^\mathsf{T}\,\mathbf{W}\,\mathbf{H} + \mu\,\mathbf{M}\right)^{-1}\mathbf{H}^\mathsf{T}\,\mathbf{W}\,\boldsymbol{d} \end{aligned} \tag{1.49}$$

Recalling the Wiener filter expression of Eq. (1.25):

$$\begin{aligned} \hat{\boldsymbol{x}}^{\mathrm{Wiener}} &= \mathrm{Cov}(\boldsymbol{x})\,\mathbf{H}^\mathsf{T}\left(\mathbf{H}\,\mathrm{Cov}(\boldsymbol{x})\,\mathbf{H}^\mathsf{T} + \mathrm{Cov}(\boldsymbol{b})\right)^{-1}\boldsymbol{d} \\ &= \left(\mathrm{Cov}(\boldsymbol{x}) + \mathbf{H}^\mathsf{T}\,\mathrm{Cov}(\boldsymbol{b})^{-1}\,\mathbf{H}\right)^{-1}\mathbf{H}^\mathsf{T}\,\mathrm{Cov}(\boldsymbol{b})^{-1}\,\boldsymbol{d} \end{aligned} \tag{1.50}$$

which holds for $\mathrm{Cov}(\boldsymbol{x})$ and $\mathrm{Cov}(\boldsymbol{b})$ definite positive matrices[10]. We find back the MAP estimator by identifying $\mathrm{Cov}(\boldsymbol{x}) = \mu\,\mathbf{M}$ and $\mathrm{Cov}(\boldsymbol{b})^{-1} = \mathbf{W}$ of Eq. (1.49). Although in these conditions the Wiener filter and the MAP estimator are the same, it is interesting to remark that the two methods result from different hypotheses. While they both assume a linear model of the data $\boldsymbol{d}$ with respect to $\boldsymbol{x}$, the Wiener filter only assumes that the object of interest $\boldsymbol{x}$ and the noise $\boldsymbol{b}$ are independent and centered. Whereas the MAP estimator supposes $\boldsymbol{x}$ and $\boldsymbol{b}$ as following centered multivariate Gaussian distributions.

1.3.5 Regularization functions

By adding *a priori* knowledge to the estimation problem, we want to favor specific properties in the resulting estimates. It is therefore important to define adequately such information. In practice, prior knowledge can be added to the estimation problem as support constraints and/or as regularization functions. Some commonly used constraints and regularizations in the imaging context are:

- Positivity, *i.e.* $\boldsymbol{x} \geq \mathbf{0}$ whenever $\boldsymbol{x}$ represents a non-negative physical quantity (*e.g.* light distribution in the imaging context).
- Support in the spatial domain: the number of non-zero pixels of $\boldsymbol{x}$ can be reduced to improve the conditioning of the problem.
- Reducing the support in the spectral domain when estimating the Fourier transform $\breve{\boldsymbol{x}}$ instead of $\boldsymbol{x}$. This amounts to cutting off the high frequencies which correspond to a smoothing effect on the solution.
- Sparse *a priori* such as implemented for basis pursuit (Chen et al., 1996) and LASSO methods (Tibshirani, 1996):

$$\mathcal{J}_{\ell 1}(\boldsymbol{x};\mu) = \mu\,\|\boldsymbol{x}\|_1 = \mu\sum_n |x_n| \tag{1.51}$$

  which limits the number of non-zero elements in $\boldsymbol{x}$ and with μ a hyper-parameter controlling the levels of the regularization (contained in $\boldsymbol{\theta}$ in Eq. (1.47)).

[9] $\mathbf{M} = \sum_n \mathbf{D}_n^\mathsf{T}\,\mathbf{D}_n$ in the case of the quadratic smoothness of Eq. (1.52).

[10] And where the second line holds thanks to the identity which states that considering $\mathbf{P}$ and $\mathbf{R}$ two positive definite matrices $\mathbf{P}\,\mathbf{B}^\mathsf{T}\,(\mathbf{B}\,\mathbf{P}\,\mathbf{B}^\mathsf{T} + \mathbf{R})^{-1} = (\mathbf{P}^{-1} + \mathbf{B}^\mathsf{T}\,\mathbf{R}^{-1}\,\mathbf{B})^{-1}\,\mathbf{B}^\mathsf{T}\,\mathbf{R}^{-1}$

Continuity *a priori* on the shape of the estimator $\boldsymbol{x} \in \mathbb{R}^N$ is usually enforced on its gradient by regularization functions. Instances of such functions are:

- Quadratic smooth *a priori* (Tikhonov and Arsenin, 1977):

$$\mathcal{J}_{\text{Tik}}(\boldsymbol{x}; \mu) = \mu \sum_n \|\mathbf{D}_n \boldsymbol{x}\|^2 \tag{1.52}$$

  with $\mathbf{D}_n : \mathbb{R}^N \to \mathbb{R}^p$ an approximation of the p-dimensional gradient of the n-th element of $\boldsymbol{x}$. For instance, first order finite differences on a 2D image $\boldsymbol{x} \in \mathbb{R}^{I \times J}$ give the gradient at the (i, j) element coordinates:

$$\mathbf{D}_{i,j}\, \boldsymbol{x} = \begin{pmatrix} x_{i+1,j} - x_{i,j} \\ x_{i,j+1} - x_{i,j} \end{pmatrix} \tag{1.53}$$

  where (i, j) represents the row and column coordinates of the element linked to the lexicographic index $n = j\, I + i$ in column-major order and the *a priori* writes:

$$\mathcal{J}_{\text{Tik}}(\boldsymbol{x}; \mu) = \mu \Big(\sum_{i,j} (x_{i+1,j} - x_{i,j})^2 + \sum_{i,j} (x_{i,j+1} - x_{i,j})^2 \Big) \tag{1.54}$$

- Total variation (Rudin, Osher, and Fatemi, 1992):

$$\mathcal{J}_{\text{TV}}(\boldsymbol{x}; \mu) = \mu \sum_n \|\mathbf{D}_n \boldsymbol{x}\| \tag{1.55}$$

- Edge-preserving regularization:

$$\mathcal{J}_{\text{edge}}(\boldsymbol{x}; \mu, \epsilon) = \mu \sum_n \left(\sqrt{\|\mathbf{D}_n \boldsymbol{x}\|^2 + \epsilon^2} - \epsilon \right) \tag{1.56}$$

  with $\epsilon \geqslant 0$ and which corresponds to the total variation when $\epsilon \to 0^+$.

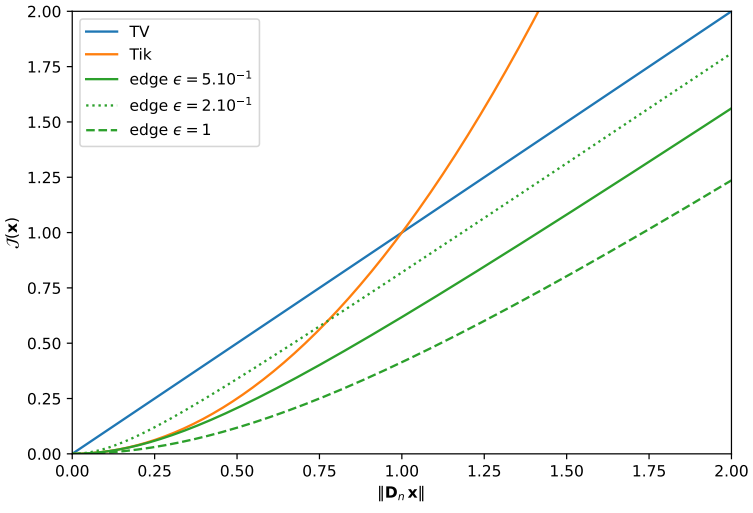

Figure 1.3 – Smooth regularization functions compared to the norm of the gradient.

Fig. 1.3 shows the behavior of these three smoothing functions with respect to the norm of the gradient. It highlights the advantages of each of the regularization functions. Quadratic smoothness is usually the most used regularization to favor continuity as it is easy to implement. It will penalize much more high gradients than small ones. This behavior will ensure smooth solutions without erasing small spatial variations of intensity. It however also smooths straight edges and can create

ripples in the solution. Total variation (plain blue line) favors sparse gradient, that is solutions that are piecewise constant functions with straight edges. It is well indicated when imaging objects with straight edges (*e.g.* an object in front of a uniform background). It can however erase small details in the image as it penalizes oscillating functions and creates so-called cartoon effects (Strong and Chan, 2003). The last regularization, called edge-preserving, is a trade-off between the total variation and quadratic smoothness. Tuned by ϵ, its behavior adapts to the level of the gradient at the n-th pixel:

$$\sqrt{\|\mathbf{D}_n\,\boldsymbol{x}\|^2+\epsilon^2}-\epsilon \simeq \begin{cases} \frac{1}{2\epsilon}\,\|\mathbf{D}_n\,\boldsymbol{x}\|^2 & \text{for } \|\mathbf{D}_n\,\boldsymbol{x}\| \ll \epsilon \\ \|\mathbf{D}_n\,\boldsymbol{x}\| & \text{for } \|\mathbf{D}_n\,\boldsymbol{x}\| \gg \epsilon \end{cases} \tag{1.57}$$

as a quadratic smoothing regularization or a total variation depending on the level of $\|\mathbf{D}_n\,\boldsymbol{x}\|$ compared to ϵ.

Proof: For $\|\mathbf{D}_n\,\boldsymbol{x}\| \ll \epsilon$:

$$\begin{aligned} \sqrt{\|\mathbf{D}_n\,\boldsymbol{x}\|^2+\epsilon^2}-\epsilon &= \epsilon\sqrt{1+\frac{\|\mathbf{D}_n\,\boldsymbol{x}\|^2}{\epsilon^2}}-\epsilon \\ &\simeq \epsilon\left(1+\frac{\|\mathbf{D}_n\,\boldsymbol{x}\|^2}{2\epsilon^2}\right)-\epsilon \\ &\simeq \frac{1}{2\epsilon}\,\|\mathbf{D}_n\,\boldsymbol{x}\|^2 \end{aligned} \tag{1.58}$$

the function behaves like a quadratic smoothing regularization. For $\|\mathbf{D}_n\,\boldsymbol{x}\| \gg \epsilon$:

$$\begin{aligned} \sqrt{\|\mathbf{D}_n\,\boldsymbol{x}\|^2+\epsilon^2}-\epsilon &= \|\mathbf{D}_n\,\boldsymbol{x}\|\sqrt{1+\frac{\epsilon^2}{\|\mathbf{D}_n\,\boldsymbol{x}\|^2}}-\epsilon \\ &\simeq \|\mathbf{D}_n\,\boldsymbol{x}\|\left(1+\frac{\epsilon^2}{2\|\mathbf{D}_n\,\boldsymbol{x}\|^2}\right)-\epsilon \\ &\simeq \|\mathbf{D}_n\,\boldsymbol{x}\| \end{aligned} \tag{1.59}$$

■

Instances of this function compared to the regularization presented in Eqs. (1.55) and (1.52) are shown in Fig. 1.3. This double behavior allows straightening edges in the image (hence its name) while gently smoothing small variations of intensity. These properties are very interesting in astronomical images where the objects we want to image are usually smooth and contrasted with a low-level background. The $\|\mathbf{D}_n\,\boldsymbol{x}\| \ll \epsilon$, usually true for most pixels in the astronomical context, depends however on the ϵ parameter. In the following, we choose to decouple the effects of this hyper-parameter with μ, the one tuning the level of the regularization in Eq. (1.47):

$$\mathcal{J}^{\text{v2}}_{\text{edge}}(\boldsymbol{x};\mu,\epsilon) = 2\mu\,\epsilon\sum_{n=1}^{N}\left(\sqrt{\|\mathbf{D}_n\,\boldsymbol{x}\|^2+\epsilon^2}-\epsilon\right) \tag{1.60}$$

which will ensure for $\epsilon > 0$ that $\lim\limits_{\|\mathbf{D}_n\,\boldsymbol{x}\|\to 0}\mathcal{J}^{\text{v2}}_{\text{edge}}(\boldsymbol{x};\mu,\epsilon) = \mu\,\|\mathbf{D}_n\,\boldsymbol{x}\|^2$ (Denneulin, L. et al., 2021).

Results of the MAP estimator with different regularizations

To demonstrate the efficiency of the edge-preserving regularization in the astronomical context, Fig. 1.2 shows the result of the MAP estimation with the quadratic smoothness of Eq. (1.52), $\boldsymbol{x}^{[\text{Tik}]}$,

(bottom-left) and with the edge-preserving, $\boldsymbol{x}^{[\text{edge}]}$ (bottom right). To evaluate the quality of the MAP estimator, the hyper-parameters μ and ϵ were tuned to minimize the true MSE between the scene estimator and the ground truth. This allows us to remove any artifact due to bad regularization tuning and see the best reconstruction possible by the MAP estimator. The quadratic smoothness result is an estimator quite close to the Wiener filter with noticeable ripples, common for quadratic regularization. with the edge-preserving regularization (bottom-right of Fig. 1.2) the ripples that were visible in the Wiener solution and MAP solution with quadratic regularization are no longer part of the result, while the edges of the arms of the galaxy are better defined. In astronomical imaging, the scene that we want to de-blur is usually composed of an object well contrasted with respect to its background. Using edge-preserving regularization in these contexts is then very useful as it favors images containing objects with sharp edges.

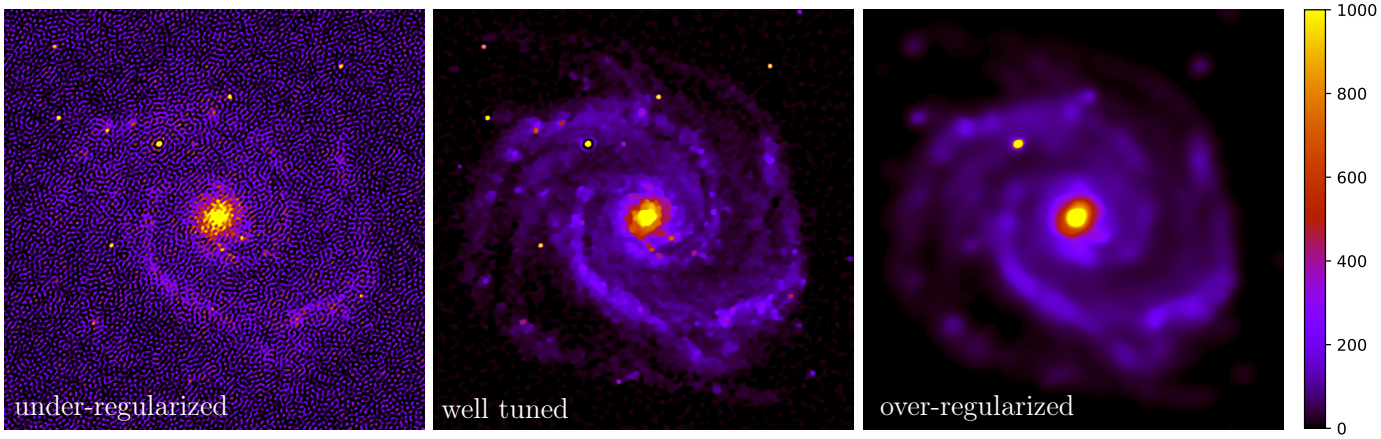

Figure 1.4 – Effects of a mistuning of the hyper-parameter μ on the results of the MAP estimator with an edge-preserving regularization. The center figure recalls the result of Fig. 1.2 with the optimal value of the hyper-parameter $\mu^{[\text{opt}]}$ according to the MSE. The left figure shows the result when choosing the hyper-parameter as $\mu^{[\text{opt}]}/100$, while the right figure when $100\,\mu^{[\text{opt}]}$.

The choice of the regularization function is important as it influences greatly the estimation. The tuning of the regularization level with respect to the data fidelity term must also be done properly. Indeed, the hyper-parameter μ is driving the regularity we want to enforce to the solution. As presented in Fig. 1.4, for a very high value of μ, the *a priori* information will be much more important than the fidelity to the data and so the resulting solution will be over-regularized, resulting in our case in a blurred solution. For a really low value, the fidelity to the data will, however, be preponderant, *i.e.* the noise will be fitted in the solution and the result will be under-regularized.

1.4 Unsupervised methods

Let's recall the MAP estimation problem of Eq. (1.47) in a more general formalism:

$$\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}) = \arg\min_{\boldsymbol{x}} \Big\{ \mathcal{L} = \mathcal{D}(\boldsymbol{d}, \boldsymbol{x}) + \mathcal{R}(\boldsymbol{x}; \boldsymbol{\theta}) \Big\} \tag{1.61}$$

with $\mathcal{D}(\boldsymbol{d}, \boldsymbol{x})$ measuring the discrepancies between the data $\boldsymbol{d}$ and the estimate $\boldsymbol{x}$ via the direct model and $\mathcal{R}(\boldsymbol{x}; \boldsymbol{\theta})$ the regularization function which brings *a priori* information to the estimation. The solution of Eq. (1.61) depends on the set of hyper-parameters $\boldsymbol{\theta} > \mathbf{0}$ (*e.g.* μ and ϵ for the edge-preserving regularization, or μ and λ when different regularization are involved (see Chapter 2)). Choosing the value of these hyper-parameters is usually difficult and can lead to very different solutions. To avoid hand-tuning them via trial and error, which can be time-consuming and lack

objectivity, many approaches have been developed to automatically choose hyper-parameters according to a criterion. The L-curve criterion (Hansen, 2001), the Generalized Cross-Validation (GCV) (Wahba et al., 1985) or the SURE minimization (Stein, 1981) are examples of commonly used methods to perform an unsupervised estimation of $\boldsymbol{\theta}$. The rather direct method to do so is to consider the estimation of the parameters of interest (here $\boldsymbol{x}$) as a *black box* and solve the following hierarchical problem:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}>0}{\arg\min}\, \mathcal{C}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}); \boldsymbol{\theta}) \tag{1.62}$$

with $\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})$ the solution of Eq. (1.61) for given data $\boldsymbol{d}$ and hyper-parameters $\boldsymbol{\theta}$ and $\mathcal{C}$ a criterion.

Ideally, we would like to tune the hyper-parameters so that the Mean Squared Error (MSE) is the lowest possible. When developing this criterion ($\mathcal{C}$ = MSE):

$$\begin{aligned}
\text{MSE} &= \mathrm{E}[\|\boldsymbol{x} - \hat{\boldsymbol{x}}_{\boldsymbol{\theta}}\|^2] = \sum_n \mathrm{E}[(x_n - \hat{x}_{\boldsymbol{\theta}\,n})^2] \\
&= \sum_n \Big(x_n^2 - 2\, x_n\, \mathrm{E}[\hat{x}_{\boldsymbol{\theta}\,n}] + \mathrm{E}[\hat{x}^2_{\boldsymbol{\theta}\,n}] \Big) \\
&= \sum_n \Big(x_n^2 - 2\, x_n\, \mathrm{E}[\hat{x}_{\boldsymbol{\theta}\,n}] + \mathrm{E}[\hat{x}_{\boldsymbol{\theta}\,n}]^2 \Big) + \sum_n \Big(-\mathrm{E}[\hat{x}_{\boldsymbol{\theta}\,n}]^2 + \mathrm{E}[\hat{x}^2_{\boldsymbol{\theta}\,n}] \Big) \\
&= \underbrace{\sum_n \Big(x_n - \mathrm{E}[\hat{x}_{\boldsymbol{\theta}\,n}] \Big)^2}_{\text{Bias}^2} + \underbrace{\sum_n \Big(\mathrm{E}[\hat{x}^2_{\boldsymbol{\theta}\,n}] - \mathrm{E}[\hat{x}_{\boldsymbol{\theta}\,n}]^2 \Big)}_{\text{Variance}}
\end{aligned} \tag{1.63}$$

a bias term and a variance term appear, which depend on the level of $\boldsymbol{\theta}$. Looking at the μ parameter of the regularization functions defined in Section 1.3.5, it is directly responsible for the level of *a priori* in the estimator. When μ increases, the term Bias^2 of Eq. (1.63) becomes preponderant in the MSE, thus reducing the variance of the estimator. On the contrary, when μ decreases, it is the Variance term that dominates in the MSE. The hyper-parameter then plays the role of a trade-off between *a priori* information and variance.

As we are not measuring $\boldsymbol{x}$ directly but via the direct model, Eldar, 2009 discusses the possibility to minimize the Generalized MSE (GMSE) instead of the MSE. In the Gaussian hypothesis, *i.e.* $\boldsymbol{d} \sim \mathcal{N}(\boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})), \mathbf{W}^{-1})$, the problem writes:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Omega_{\boldsymbol{\theta}}}{\arg\min} \Big\{ \text{GMSE}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})) = \mathrm{E}[\|\boldsymbol{m}(\boldsymbol{x}) - \boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}))\|^2_{\mathbf{W}}] \Big\} \tag{1.64}$$

with $\Omega_{\boldsymbol{\theta}}$ the set of possible hyper-parameters, $\boldsymbol{m}(\boldsymbol{x})$ the ground truth direct model and $\boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}))$ the model deduced by the estimation of $\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})$ at given data $\boldsymbol{d}$, and respectively noted $\boldsymbol{m} = \boldsymbol{m}(\boldsymbol{x})$ and $\hat{\boldsymbol{m}}(\boldsymbol{\theta}) = \boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}))$ in the following to simplify the notation. Note however that this criterion cannot take into account discrepancies between $\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}$ and the ground truth $\boldsymbol{x}$ that are not measured by the direct model. The GMSE also needs the ground truth model $\boldsymbol{m}$ which is unknown in practice.

1.4.1 Generalized Stein's Unbiased Risk Estimator (GSURE)

Following the work of Stein, 1981, Eldar, 2009 generalizes his method to estimate the GMSE criterion:

$$\begin{aligned}
\text{GMSE}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})) &= \mathrm{E}[\|\boldsymbol{m} - \hat{\boldsymbol{m}}(\boldsymbol{\theta})\|^2_{\mathbf{W}}] \\
&= \mathrm{E}\Big[\|(\hat{\boldsymbol{m}}(\boldsymbol{\theta}) - \boldsymbol{d}) + (\boldsymbol{d} - \boldsymbol{m})\|^2_{\mathbf{W}} \Big] \\
&= \mathrm{E}\Big[\|\hat{\boldsymbol{m}}(\boldsymbol{\theta}) - \boldsymbol{d}\|^2_{\mathbf{W}} \Big] + \mathrm{E}\Big[\|\boldsymbol{d} - \boldsymbol{m}\|^2_{\mathbf{W}} \Big] + 2\, \mathrm{E}\big[(\hat{\boldsymbol{m}}(\boldsymbol{\theta}) - \boldsymbol{d})^{\mathsf{T}}\, \mathbf{W}\, (\boldsymbol{d} - \boldsymbol{m}) \big]
\end{aligned} \tag{1.65}$$

Since $\boldsymbol{m} = \mathrm{E}\big[\boldsymbol{d}\big]$ and $\mathbf{W}$ is the inverse of the covariance matrix of $\boldsymbol{d}$, the second term of the criterion is the number of valid measurements (number of degrees of freedom):

$$\mathrm{E}\Big[\|\boldsymbol{d}-\boldsymbol{m}\|^2_{\mathbf{W}}\Big] = \mathrm{E}\Big[\|\boldsymbol{b}\|^2_{\mathbf{W}}\Big] = N \tag{1.66}$$

To simplify the GMSE expression, Stein's lemma indicates that for any weakly differentiable N-variate function $\boldsymbol{f}$ and for Gaussian data $\boldsymbol{d} \sim \mathcal{N}(\boldsymbol{m}, \mathbf{W}^{-1})$:

$$\mathrm{E}\big[\mathbf{f}(\boldsymbol{d})^\mathsf{T}\,\mathbf{W}\,(\boldsymbol{d}-\boldsymbol{m})\big] = \mathrm{E}\Big[\mathrm{tr}\Big(\frac{\partial \mathbf{f}(\boldsymbol{d})}{\partial \boldsymbol{d}}\Big)\Big] \tag{1.67}$$

Proof:

$$\begin{aligned}
\mathrm{E}\big[\mathbf{f}(\boldsymbol{d})^\mathsf{T}\,\mathbf{W}\,(\boldsymbol{d}-\boldsymbol{m})\big] &= \int \mathbf{f}(\boldsymbol{d})^\mathsf{T}\,\mathbf{W}\,(\boldsymbol{d}-\boldsymbol{m})\,\mathrm{p}(\boldsymbol{d})\,\mathrm{d}\boldsymbol{d} \\
&= -\int \mathbf{f}(\boldsymbol{d})^\mathsf{T}\,\frac{\partial \mathrm{p}(\boldsymbol{d}\,|\,\boldsymbol{x})}{\partial \boldsymbol{d}}\,\mathrm{d}\boldsymbol{d} \\
&= -\sum_n \int f_n(\boldsymbol{d})\,\frac{\partial \mathrm{p}(d_n\,|\,\boldsymbol{x})}{\partial d_n}\,\mathrm{d}d_n \\
&= \sum_n \int \frac{\partial f_n(\boldsymbol{d})}{\partial d_n}\,\mathrm{p}(d_n\,|\,\boldsymbol{x})\,\mathrm{d}d_n \\
&= \sum_n \mathrm{E}\Big[\frac{\partial f_n(\boldsymbol{d})}{\partial d_n}\Big] \\
&= \mathrm{E}\Big[\mathrm{tr}\Big(\frac{\partial \mathbf{f}(\boldsymbol{d})}{\partial \boldsymbol{d}}\Big)\Big]
\end{aligned} \tag{1.68}$$

where $f_n(\boldsymbol{d})$ is the n-th element of $\boldsymbol{f}(\boldsymbol{d})$ and the fourth right-hand side results from integration by parts and so, requires that $\mathbf{f}$ be weakly differentiable. ■

Setting $\mathbf{f}(\boldsymbol{d}) = \hat{\boldsymbol{m}}(\boldsymbol{\theta}) - \boldsymbol{d}$ leads to:

$$\mathrm{E}\big[(\hat{\boldsymbol{m}}(\boldsymbol{\theta})-\boldsymbol{d})^\mathsf{T}\,\mathbf{W}\,(\boldsymbol{d}-\boldsymbol{m})\big] = \mathrm{E}\Big[\mathrm{tr}\Big(\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{\theta})}{\partial \boldsymbol{d}} - \mathbf{I}\Big)\Big] \tag{1.69}$$

$$= \mathrm{E}\Big[\mathrm{tr}\Big(\frac{\partial \hat{\boldsymbol{m}}(\boldsymbol{\theta})}{\partial \boldsymbol{d}}\Big)\Big] - N \tag{1.70}$$

with $\mathbf{I}$ the identity matrix.

Going back to Eq. (1.65), the GMSE can then be written as:

$$\mathrm{GMSE}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})) = \mathrm{E}\Big[\|\hat{\boldsymbol{m}}(\boldsymbol{\theta})-\boldsymbol{d}\|^2_{\mathbf{W}}\Big] + 2\,\mathrm{E}\Big[\mathrm{tr}\big(\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\big)\Big] - N \tag{1.71}$$

with $\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}$ the Jacobian matrix of $\hat{\boldsymbol{m}}(\boldsymbol{\theta})$ evaluated at $\boldsymbol{d}$. The GSURE criterion is an unbiased estimator of the GMSE:

$$\mathrm{GSURE}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})) = \|\hat{\boldsymbol{m}}(\boldsymbol{\theta})-\boldsymbol{d}\|^2_{\mathbf{W}} + 2\,\mathrm{tr}\Big(\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\Big) - N \tag{1.72}$$

with $\mathrm{E}\Big[\mathrm{GSURE}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}))\Big] = \mathrm{GMSE}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}))$, and the estimation of the hyper-parameters is done by:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\Omega_{\boldsymbol{\theta}}}{\arg\min}\,\mathrm{GSURE}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})) \tag{1.73}$$

Note that as for the GMSE, the GSURE criterion only takes into account what is measured, and so, if some modes of the ground truth $\boldsymbol{x}$ are not measured or are attenuated by the direct model, they are not taken into account by the tuning of $\hat{\boldsymbol{\theta}}$. To compute the GSURE criterion, we need the trace term which is usually difficult to compute.

1.4.2 Computing the GSURE criterion

Girard, 1989 and Ramani, Blu, and Unser, 2008 demonstrate a simple way to estimate the trace term. Using a perturbation method by a centered perturbation $\boldsymbol{\delta}$ of covariance the identity matrix, the second order of the Taylor expansion of the direct model writes:

$$\boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}+\epsilon\,\boldsymbol{\delta})) = \boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})) + \epsilon\,\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}(\boldsymbol{d})\,\boldsymbol{\delta} + \epsilon^2\,\boldsymbol{l}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}(\boldsymbol{d}) \tag{1.74}$$

with $\boldsymbol{l}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}$ the Lagrange remainder of the decomposition.

Then, subtracting by $\boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}))$, multiplying by $\boldsymbol{\delta}^{\mathsf{T}}$ and taking the expectation yields:

$$\mathrm{E}\Big[\boldsymbol{\delta}^{\mathsf{T}}\left(\boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}+\epsilon\,\boldsymbol{\delta})) - \boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}))\right)\Big] = \epsilon\,\mathrm{E}\big[\boldsymbol{\delta}^{\mathsf{T}}\,\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\,\boldsymbol{\delta}\big] + \epsilon^2\,\mathrm{E}\big[\boldsymbol{\delta}^{\mathsf{T}}\,\boldsymbol{l}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\big] \tag{1.75}$$

with

$$\begin{aligned}
\mathrm{E}\big[\boldsymbol{\delta}^{\mathsf{T}}\,\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\,\boldsymbol{\delta}\big] &= \mathrm{E}\Big[\sum_{i,j}\delta_i\,J_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})\,i,j}\,\delta_j\Big] = \sum_{i,j} J_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})\,i,j}\mathrm{E}[\delta_i\,\delta_j] \\
&= \sum_i J_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})\,i,i}\mathrm{E}[\delta_i^2] \qquad \text{since the entries of } \boldsymbol{\delta} \text{ are mutually independent} \\
&= \sum_i J_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})\,i,i}\,I_{i,i} = \mathrm{tr}\big(\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\big)
\end{aligned} \tag{1.76}$$

and $\mathrm{E}\big[\boldsymbol{\delta}^{\mathsf{T}}\,\boldsymbol{l}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\big] = \text{cst}$. Then

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathrm{E}\Big[\boldsymbol{\delta}^{\mathsf{T}}\big(\boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}+\epsilon\,\boldsymbol{\delta})) - \boldsymbol{m}(\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d}))\big)\Big] = \mathrm{tr}\big(\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\big) \tag{1.77}$$

Note that this method works as a Monte-Carlo strategy because each pixel of the perturbed data is considered independent and so, gives enough realization of the perturbation $\boldsymbol{\delta}$. ϵ here plays the role of the standard deviation of the perturbation. In practice, we choose $\epsilon^2 = \frac{1}{10N}\,\|\boldsymbol{d}\|^2$ to scale the perturbation compared to the data.

1.4.3 Other strategies for unsupervised methods

Generalized Cross-Validation

The Generalized Cross-Validation (GCV) criterion, as introduced by Craven and Wahba, 1978, discusses the case when the precision matrix of the data is not reliably known. In this case, the GSURE criterion cannot be used as it depends on $\mathbf{W}$. They introduced the minimization problem:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{(1/N)\,\|\mathbf{I} - \mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\,\boldsymbol{d}\|^2}{\Big((1/N)\;\mathrm{tr}\big(\mathbf{I} - \mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}\big)\Big)^2} \tag{1.78}$$

where $\mathbf{J}_{\hat{\boldsymbol{m}}(\boldsymbol{\theta})}$ is the Jacobian matrix and $\hat{\boldsymbol{m}}(\boldsymbol{\theta})$ the direct model deduced by the estimation of $\hat{\boldsymbol{x}}_{\boldsymbol{\theta}}(\boldsymbol{d})$ in Eq. (1.61).

Hierarchical Bayesian strategy

To counter the difficulty of tuning the hyper-parameters, a solution is to consider them also as random variables. Recalling Eq. (1.45), it amounts to maximizing:

$$\mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{d}, \boldsymbol{\theta}) \propto \mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x})\, \mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{\theta})\, \mathrm{p}(\boldsymbol{\theta}) \tag{1.79}$$

with $\mathrm{p}(\boldsymbol{\theta})$ a usually non-informative distribution of the hyper-parameters $\boldsymbol{\theta}$, or *hyper-priors*, which needs to be as general as possible as we generally don't know well $\boldsymbol{\theta}$. An example of the hyper-parameter μ responsible for the level of the regularization is the Jeffreys prior, one of the most used hyper-prior:

$$\mathrm{p}(\mu) \propto \mu^{-\beta}, \;\; \beta \geqslant 0 \tag{1.80}$$

as it does not assume anything on the form of the hyper-parameter except that it is positive. This prior can even be improper and amounts to $\mathrm{p}(\mu) \propto \text{cste}$.

With this approach, the problem defined in Eq. (1.45) can then be rewritten by marginalizing over $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{x}}^{[\mathrm{MAP}]} = \arg\max_{\boldsymbol{x}} \int_{\mathbb{R}^+} \mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x})\, \mathrm{p}(\boldsymbol{x} \,|\, \boldsymbol{\theta})\, \mathrm{p}(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta} \tag{1.81}$$

thus removing the hyper-parameters from the estimation process. Hierarchical Bayesian can also be used when multiple parameters of interest need to be estimated (see Section 2.1.3).

1.5 Optimization problems and their resolution

1.5.1 Finding the minimum of smooth cost functions

In Sections 1.3 and 1.4, we discussed the different estimation problems at the core of this manuscript. They are all expressed as an optimization problem consisting of minimizing functions. In this section, we detail the different methods that can be used to solve these problems. Indeed, the convexity of the function to minimize, whether it is differentiable or not, or the high dimension of the parameters to estimate, are properties that constrain the choice of the optimization method.

With the inverse problem approach, the estimation problem can be written in a general formalism. We look for a local minimizer (or an approached value) $\hat{\boldsymbol{x}}$ of a criterion $\mathcal{L}$:

$$\hat{\boldsymbol{x}} \simeq \arg\min_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}) \tag{1.82}$$

Provided that $\mathcal{L}$ is twice differentiable, a local minimum is obtained for:

$$\left.\frac{\partial \mathcal{L}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}} = \boldsymbol{g}(\hat{\boldsymbol{x}}) = \mathbf{0} \quad \text{and} \quad \left.\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{x}^2}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}} = \mathbf{B}(\hat{\boldsymbol{x}}) \succ \mathbf{0} \tag{1.83}$$

with $\boldsymbol{g}(\boldsymbol{x})$ the gradient of the criterion with respect to $\boldsymbol{x}$ and $\mathbf{B}(\boldsymbol{x})$ its Hessian defined as definite positive. The minima of the criteria in the problems of Eq. (1.61) and (1.62) do not necessarily have a closed-form. They, therefore, require iterative methods which will approach the minimum of the criterion at convergence. Finding $\hat{\boldsymbol{x}}$ can be done by building a sequence of estimates $\{\boldsymbol{x}_{[k]}\}_{k\in\mathbb{N}}$ following an iterative estimation strategy. Beginning with an initial value $\boldsymbol{x}_{[0]}$, the sequence must ensure that for each iteration k, $\mathcal{L}(\boldsymbol{x}_{[k+1]}) < \mathcal{L}(\boldsymbol{x}_{[k]})$ and such that $\lim_{k\to+\infty} \boldsymbol{g}(\boldsymbol{x}_{[k]}) = \mathbf{0}$. The sequence can then be built by:

$$\boldsymbol{x}_{[k+1]} = \boldsymbol{x}_{[k]} + \boldsymbol{s}_{[k]} \tag{1.84}$$

with $\boldsymbol{s}_{[k]}$ called the stepped direction of minimization and chosen so that the iterates go towards a local minimum of $\mathcal{L}$.

In order to find the direction of minimization, the main methods used in this manuscript rely on a quadratic model of the criterion. For the k-th iteration, the Taylor expression of $\mathcal{L}$ around $\boldsymbol{x}_{[k]}$ gives the quadratic model:

$$\mathcal{L}(\boldsymbol{x}_{[k]} + \boldsymbol{s}) \approx \mathcal{M}_{[k]}(\boldsymbol{s}) = \mathcal{L}(\boldsymbol{x}_{[k]}) + \boldsymbol{g}(\boldsymbol{x}_{[k]})^\mathsf{T}\, \boldsymbol{s} + \frac{1}{2}\, \boldsymbol{s}^\mathsf{T}\, \mathbf{B}_k\, \boldsymbol{s} \tag{1.85}$$

with $\mathbf{B}_k$ the Hessian $\mathbf{B}(x_{[k]})$ or an approximation of it. This model $\mathcal{M}$ is however not valid everywhere and we need to limit the length of the stepped direction $\boldsymbol{s}$ of minimization. Two families of methods provide this control.

Line search methods

From Eq. (1.85), it is possible to deduce the direction to follow at each iteration k in the Newton method sense:

$$\begin{aligned} \boldsymbol{q}_{[k]} &= \arg\min_{\boldsymbol{s}} \mathcal{M}_{[k]}(\boldsymbol{s}) \\ &= -\mathbf{B}_k^{-1}\, \boldsymbol{g}(\boldsymbol{x}_{[k]}) \end{aligned} \tag{1.86}$$

Line search methods rely on the estimation of a step length $\lambda_{[k]}$ which limits the update of $\boldsymbol{x}$ by defining the stepped direction as $\boldsymbol{s}_{[k]} = \lambda_{[k]}\, \boldsymbol{q}_{[k]}$.

The step must ensure that $\mathcal{L}(\boldsymbol{x}_{[k]} + \lambda_{[k]}\, \boldsymbol{s}_{[k]}) < \mathcal{L}(\boldsymbol{x}_{[k]})$. Its value is, therefore, important as it tunes how much the next estimator will have decreased the criterion. On one hand, the quadratic approximation is only true locally so we need to limit the step's length. On the other hand, infinitely small steps would stop the convergence of the method. The Wolfe conditions (Wolfe, 1969) are a pair of conditions on the sufficient decrease and curvature of the criterion:

$$\mathcal{L}(\boldsymbol{x}_{[k]} + \lambda_{[k]}\, \boldsymbol{s}_{[k]}) \leqslant \mathcal{L}(\boldsymbol{x}_{[k]}) + c_1\, \lambda_{[k]}\, \boldsymbol{g}(\boldsymbol{x}_{[k]})^\mathsf{T}\, \boldsymbol{s}_{[k]} \tag{1.87a}$$

$$\boldsymbol{g}(\boldsymbol{x}_{[k]} + \lambda_{[k]}\, \boldsymbol{s}_{[k]})^\mathsf{T}\, \boldsymbol{s}_{[k]} \geqslant c_2\, \boldsymbol{g}(\boldsymbol{x}_{[k]})^\mathsf{T}\, \boldsymbol{s}_{[k]} \tag{1.87b}$$

with $c_1 \in]0, 1[$ often chosen as really low (*e.g.* 10^{-4}) and $c_2 \in]c_1, 1[$.

Trust-region methods

Contrary to the line search, trust-region methods (Nocedal and Wright, 1999) use a constrained estimation of the direction of minimization:

$$\boldsymbol{s}_{[k]} = \arg\min_{\boldsymbol{s}} \mathcal{M}_{[k]}(\boldsymbol{s}) \qquad \text{s.t.} \quad \|\boldsymbol{s}\| < \Delta_{[k]} \tag{1.88}$$

with $\Delta_{[k]}$ defining the radius of a region where we trust the quadratic model of Eq. (1.85) to be valid around the estimate $\boldsymbol{x}_{[k]}$ at iteration k (the *trust-region*). Such methods adjust the radius of the region at iteration $k+1$ by computing the ratio:

$$\rho_{[k]} = \frac{\mathcal{L}(\boldsymbol{x}_{[k]}) - \mathcal{L}(\boldsymbol{x}_{[k]} + \boldsymbol{s}_{[k]})}{\mathcal{M}_{[k]}(\mathbf{0}) - \mathcal{M}_{[k]}(\boldsymbol{s}_{[k]})} \tag{1.89}$$

where the numerator is the actual reduction while the denominator is the predicted reduction. If $\rho_{[k]} \ll 1$, the model is not representative of the criterion in the region. Its radius needs to decrease so that the approximation holds. If $\rho_{[k]}$ is close to 1, the reduction of both the criterion and its approximation is close. The region can be extended so that at the next iteration, the direction can be less constrained.

1.5.2 Newton/quasi-Newton methods

When solving problems such as the one in Eq. (1.61), we usually look for a large number of parameters, *e.g.* the image $\boldsymbol{x}$ we want to reconstruct in the deconvolution problem. To ease up the estimation process, we usually choose a differentiable criterion allowing us to use the methods described above. Following the development of the line search, the Newton and quasi-Newton methods rely on the direction defined in Eq. (1.86) to find the next estimate at each iteration. However, when the parameters to estimate are of large size, the computation of the Hessian of the criterion $\mathcal{L}$ can be costly in convergence time and difficult. Methods have been developed to approximate the inverse of $\mathbf{B}_k$ efficiently such as the BFGS method (Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970). It uses the hypothesis that the model at an iteration $k+1$ should explain the gradient of $\mathcal{L}$ at iterations k and $k+1$. Thanks to the information of the gradient at previous iterations, and constraining the Hessian to be similar between consecutive iterations, the BFGS method estimates efficiently $\mathbf{B}_k$.

Considering that, for a large number of variables, the rank of $\mathbf{B}_{[k]}$ is high, the matrix is heavy to compute. Nocedal, 1980 came up with a lighter method, called the L-BFGS algorithm, using an iterative strategy and remembering a few previous iterations so that the direction $\boldsymbol{s}_{[k]} = -\mathbf{B}_{[k]}\,\boldsymbol{g}(\boldsymbol{x}_{[k]})$ is computed without storing the approximation $\mathbf{B}_{[k]}$.

In his paper, Thiébaut, 2002 developed a bounded L-BFGS method, called VMLM-B, to include bound constraints ($\forall n,\, l_{\boldsymbol{x}\,n} \leqslant x_n \leqslant u_{\boldsymbol{x}\,n}$). In the following manuscript, VMLM-B is used extensively as a method of optimization to solve estimation problems such as Eq. (1.61). Indeed, it allows us to enforce positivity constraint ($l_{\boldsymbol{x}\,n} = 0$ and $u_{\boldsymbol{x}\,n} = +\infty$) on the solutions (see Section 1.3.5), thus improving the estimation process compared to an unconstrained estimation.

1.5.3 Derivative-free methods of minimization

While quasi-Newton methods are very efficient, they require the function to minimize to be differentiable and need its expression and the one of its derivative. The latter can however be difficult to express. To counteract this difficulty, derivative-free methods have been developed that only need the computation of the criterion. Solving the problem in Eq. (1.62), *i.e.* estimating hyper-parameters, is usually the type of problem that can be solved by such methods as they involve only a few unknowns.

Univariate function minimization

When there is only one parameter μ to find, the problem:

$$\hat{\mu} = \arg\min_{\mu} \mathcal{C}(\mu) \tag{1.90}$$

can be solved simply by derivative-free methods. For instance, the golden search method (Kiefer, 1953) is an iterative method that determines at each iteration a smaller interval around the minimum following the golden ratio and with the help of three previously computed values. This method ensures the convergence towards a local minimum at a constant reduction rate throughout the iterations.

To speed up the convergence, Brent's method (Brent, 2013) bases its iterative strategy on another consideration. Provided the criterion is smooth enough in an interval $[a, b]$, Brent's method uses at iteration k a parabolic model $\mathcal{Q}_{[k]}$ of the criterion $\mathcal{C}$ interpolated on the three previously computed values, as seen in Fig. 1.5. The value of the criterion $\mathcal{C}$ corresponding to the minimum of this model $\mathcal{Q}$ is then stored and a new parabolic model is interpolated at iteration $k+1$. If the minimum of $\mathcal{Q}_{[k]}$ is located outside of the interval $[a, b]$, a golden search is applied to ensure a reduction of the interval.

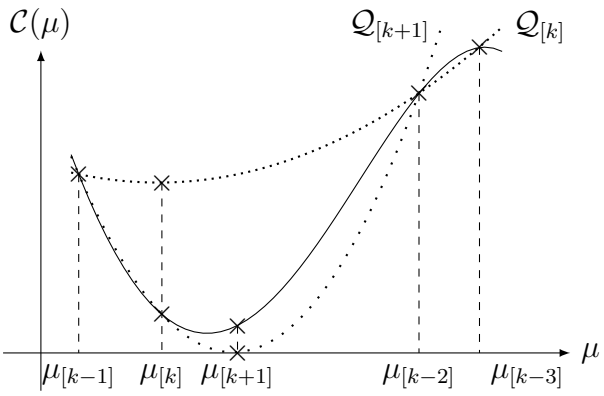

Figure 1.5 – Principle of Brent's methods.

Multivariate function minimization

Let $\mathcal{C}(\boldsymbol{\theta})$ be a function to minimize in order to find the estimate of multiple parameters $\hat{\boldsymbol{\theta}}$:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathcal{C}(\boldsymbol{\theta}) \tag{1.91}$$

Inspired by Newton's approach, many non-derivative algorithms have been developed for multiple parameter estimation. Powell's Newuoa (Powell, 2006) is an example of methods mixing a parabolic model with a trust-region approach. The Newuoa method differs from a trust-region method by not needing the gradient and Hessian of the criterion. Instead of the quadratic model derived from the Taylor development (Eq. (1.85)), it takes an interpolated parabolic model, *i.e.* $\mathcal{M}_{[k]} = \mathcal{Q}_{[k]}$, as defined in Brent's method. Such a model needs $(N+1)(N+2)/2$ function values for N parameters, which can be costly in terms of computation time. To reduce the number of evaluations of the criterion, the method also uses an iterative updating strategy by only interpolating on $m < (N+1)(N+2)/2$ function values[11] and estimating the other parameters by an update of the parabolic model. This update consists in minimizing the Frobenius norm $\left\| \frac{\partial^2 \mathcal{Q}_{[k+1]}}{\partial \boldsymbol{\theta}^2} - \frac{\partial^2 \mathcal{Q}_{[k]}}{\partial \boldsymbol{\theta}^2} \right\|_{\mathrm{F}}^2$.

All these methods have however limitations as they optimize only locally the criterion. In the case of possible poor local minima, it is necessary to choose good initial parameters.

1.6 Conclusions

We presented in this chapter the main methodologies used in the rest of the manuscript. Going from physical phenomena taking place in the detector, we discussed the pre-processing strategy of raw data by a careful model of the pre-processed data $\boldsymbol{d}$ (Eq. (1.7)). Simpler to handle while summarized by sufficient statistics (Eqs. (1.8) and (1.9)), we use this pre-processed data throughout the rest of the manuscript. The basic principles of the inverse problem are also presented and illustrated thanks to the deconvolution problem. The problem of disentangling the object of interest from spurious signals and effects of the instrument, recurrent in this manuscript, is discussed in a general formalism. Insisting on the Maximum *a posteriori* and the necessary regularization functions, we also detailed an unsupervised strategy based on the GSURE criterion which allows us to properly tune hyper-parameters of our problems in the rest of the manuscript. Finally, we discussed the choice of optimization methods to solve the different problems faced in the manuscript. Due to the high dimensionality of our unknowns, the quasi-Newton framework is chosen as it achieves a good trade-off between complexity and convergence speed, whereas the estimation of hyper-parameters is left to

[11] Powell recommends $m = 2N + 1$ to have enough samples for the model while reducing the number of evaluations of the criterion.

derivative-free methods which are well-suited for fewer unknowns while not requiring the derivative of the criterion to minimize.

Chapter 2

AMORS: Easier tuning of hyper-parameters and improved convergence for bi-linear models

Inverse problems framework can tackle many estimation problems but relies on knowledge of the acquisition method: the instrument and the detector used to measure the data (see Sections 1.1 and 1.2). Priors are also playing an important role in the estimation process and need to be leveled appropriately compared to the data fidelity measure. The tuning of the regularization hyper-parameters is therefore something delicate that can be fastidious when done manually.

Another source of difficulty is the potential error that we can do building the direct model. Taking back the example of the deconvolution problem discussed throughout Chapter 1, a faithful model of the Point Spread Function (PSF) of the instrument, $\boldsymbol{h}$, is necessary to recover the object of interest $\boldsymbol{x}$ from blurred data $\boldsymbol{d}$. When it is not well known, this assumption raises the question of how we can recover the PSF as well as the object of interest.

In this chapter, we decide to investigate further these problematic by studying the blind deconvolution problem. That is the estimation problem where neither the object of interest nor the PSF is known and both need to be estimated. Section 2.1 deals with the different strategies used by the community to solve this difficult problem. Analyzing the problem in a more general context, we develop in Section 2.2 the AMORS algorithm: a strategy that robustly and efficiently estimates both components in the direct model while being unsupervised. Section 2.3 presents results on simulated data to validate the advantages of our method. Finally, Section 2.4 concludes the chapter.

In This Chapter

Prerequisite:

- The formation of the direct model (Section 1.2).
- Maximum *a posteriori* formalism (Sections 1.3.4 and 1.3.5).

Contributions:

- Exploit the scaling ambiguity in bi-linear models.
- Development of an alternating estimation procedure AMORS.
- Presentations at EUSIPCO 2020 (Thé et al., 2021) and SPIE 2022 (Thé et al., 2022a).

2.1 The blind deconvolution problem

If the PSF is perfectly known and spatially invariant, image deblurring amounts to a deconvolution problem. Most of the time, this PSF is however difficult to obtain. Calibration processes can be difficult to perform and theoretical models are usually not flexible enough to take into account all phenomena described in 1.2. Estimating both the object of interest $\boldsymbol{x}$ and the PSF $\boldsymbol{h}$ is then necessary to deblur properly the data.

This problem is however much more difficult: it is even more ill-posed than the deconvolution problem (there are not enough measurements for the number of parameters) and the model of the blurring process has many ambiguities which complicate the un-mixing process. A spatial shift of one of the components can for example be compensated by an opposite shift in the other. The PSF and the object of interest have the same mathematical role and can be exchanged. It is also possible that one of the components (or both) is in itself a convolution of other functions. Another degeneracy called the scaling ambiguity is responsible for an infinite number of couples of equivalent solutions to the deconvolution problem:

$$\forall \alpha > 0,\ (\boldsymbol{h}/\alpha) \star (\alpha\, \boldsymbol{x}) = \boldsymbol{h} \star \boldsymbol{x}\,. \tag{2.1}$$

Many methods have been proposed in the literature to disentangle the two components and recover the object of interest $\boldsymbol{x}$, we briefly review some of them in the next sub-sections.

2.1.1 Joint estimation of components in the data

Early methods of blind deconvolution already relied on the alternated estimation of the two components by different sub-methods (*e.g.* with the Wiener filter as in Ayers and Dainty, 1988). As Yu-Li You and Kaveh, 1996 describes, a regularized joint estimation approach can be developed for $\boldsymbol{x}$ and $\boldsymbol{h}$:

$$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{h}}) = \underset{\boldsymbol{x},\boldsymbol{h}}{\arg\min} \Big\{ \mathcal{L}_{\text{bdec}}(\boldsymbol{x}, \boldsymbol{h}; \mu, \lambda) = \|\boldsymbol{d} - \boldsymbol{h} \star \boldsymbol{x}\|^2_{\mathbf{W}} + \mu\, \mathcal{J}(\boldsymbol{x}) + \lambda\, \mathcal{K}(\boldsymbol{h}) \Big\} \tag{2.2}$$

Note that this optimization problem requires regularization of both components (see section 2.2.2 for a formal justification). Because we need to estimate the PSF as well, the regularization choices are important for the method to distinguish the object of interest from the PSF. Choosing a restricted spatial domain support (Section 1.3.5) avoids for instance solutions like swapping the PSF and the image ($\hat{\boldsymbol{x}} = \boldsymbol{h}$ and $\hat{\boldsymbol{h}} = \boldsymbol{x}$). Many methods perform the estimation of the Fourier transform $\breve{\boldsymbol{h}}$ of the PSF $\boldsymbol{h}$. Doing so can help to enforce constraints on the spectral domain (cutoff frequency) and prevent trivial solutions $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{h}}) = (\boldsymbol{d}, \boldsymbol{\delta})$, with $\boldsymbol{\delta}$ the Dirac distribution. Making the hypothesis that the total flux of $\boldsymbol{d}$ is similar to the one of $\boldsymbol{x}$, it is possible to normalize the PSF, *i.e.* $\sum_n h_n = 1$ (Chan and Wong, 1998,You and Kaveh, 1996). Perrone and Favaro, 2014 emphasis that regularizing only the object (*i.e.* $\lambda = 0$) with a quadratic smoothness or edge-preserving regularization (Eq. (1.52) or (1.56)) leads to favor either the solution of interest $(\boldsymbol{x}, \boldsymbol{h})$ or the blurred solution $(\boldsymbol{h} \star \boldsymbol{x}, \boldsymbol{\delta})$, with $\boldsymbol{\delta}$ the Dirac function. Indeed, they prove for $\mathcal{J}$ as in Eq. (1.52) or (1.56) that:

$$\begin{aligned} \mathcal{L}_{\text{bdec}}(\boldsymbol{x} \star \boldsymbol{h}, \boldsymbol{\delta}; \mu, \lambda) &= \|\boldsymbol{d} - (\boldsymbol{h} \star \boldsymbol{x}) \star \boldsymbol{\delta}\|^2_{\mathbf{W}} + \mu\, \mathcal{J}(\boldsymbol{h} \star \boldsymbol{x}) + \lambda\, \mathcal{K}(\boldsymbol{\delta}) \\ &= \|\boldsymbol{d} - \boldsymbol{h} \star \boldsymbol{x}\|^2_{\mathbf{W}} + \mu\, \mathcal{J}(\boldsymbol{h} \star \boldsymbol{x}) + \lambda\, \mathcal{K}(\boldsymbol{\delta}) \\ &\leqslant \mathcal{L}_{\text{bdec}}(\boldsymbol{x}, \boldsymbol{h}; \mu, \lambda) \end{aligned} \tag{2.3}$$

where $\mathcal{J}(\boldsymbol{h} \star \boldsymbol{x}) \leqslant \mathcal{J}(\boldsymbol{x})$, as the blurred solution $\boldsymbol{h} \star \boldsymbol{x}$ has usually smoother gradients than the sharp solution $\boldsymbol{x}$. Another example is given by Benichoux, Vincent, and Gribonval, 2013 observing that with a sparse *prior* or normalization constraints, the global minimum of the function in Eq. (2.2) is the solution $(\boldsymbol{\delta}, \boldsymbol{h} \star \boldsymbol{x})$.

2.1.2 Alternated estimation strategy for blind deconvolution

To resolve these sensitivities to regularization and constraints and ensure quick convergence towards interesting solutions, Yu-Li You and Kaveh, 1996 developed an Alternated Minimization (AM) algorithm. It consists in decomposing the optimization problem into two deconvolution sub-problems that are much easier to solve and for which existing algorithms may be re-used:

$$\hat{\boldsymbol{x}}(\boldsymbol{h};\mu) = \underset{\boldsymbol{x}\in\mathbb{X}}{\arg\min} \|\boldsymbol{d} - \boldsymbol{h} * \boldsymbol{x}\|_2^2 + \mu\, \mathcal{J}(\boldsymbol{x}) \tag{2.4a}$$

$$\hat{\boldsymbol{h}}(\boldsymbol{x};\lambda) = \underset{\boldsymbol{h}\in\mathbb{H}}{\arg\min} \|\boldsymbol{d} - \boldsymbol{x} * \boldsymbol{h}\|_2^2 + \lambda\, \mathcal{K}(\boldsymbol{h}) \tag{2.4b}$$

with $\mathbb{X}$ and $\mathbb{H}$, two sub-sets that depend on the constraints applied. Nonetheless, this method is still sensitive to initialization and the global convergence is only towards a local minimum. Chan and Wong, 2000 developed the Projected AM (PAM) algorithm which includes a normalization step of the PSF, thus leading to more interesting solutions.

2.1.3 Marginalization methods

An alternative to the alternated estimation strategy is to marginalize the criterion. Rather than minimizing the joint posterior distribution of $\boldsymbol{x}$ and $\boldsymbol{h}$ (generalizing the joint estimation method of Eq. (2.2)):

$$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{h}}) = \underset{\boldsymbol{x},\boldsymbol{h}}{\arg\max}\, \mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x}, \boldsymbol{h})\, \mathrm{p}(\boldsymbol{x} \,|\, \mu)\, \mathrm{p}(\boldsymbol{h} \,|\, \lambda)\,, \tag{2.5}$$

it is possible to marginalize this distribution to drop the dependency on the object of interest $\boldsymbol{x}$:

$$\hat{\boldsymbol{h}} = \underset{\boldsymbol{h}}{\arg\max} \int_{\boldsymbol{x}} \mathrm{p}(\boldsymbol{d} \,|\, \boldsymbol{x}, \boldsymbol{h})\, \mathrm{p}(\boldsymbol{x} \,|\, \mu)\, \mathrm{p}(\boldsymbol{h} \,|\, \lambda)\, \mathrm{d}\boldsymbol{x} \tag{2.6}$$

The estimated PSF can then be used in a similar strategy or embedded in other alternated strategies such as the Expectation-Minimization algorithm (Levin et al., 2011), or directly applied, for instance, as combined with a Wiener filter (Blanco and Mugnier, 2011), to retrieve the un-blurred object of interest. However, marginalizing $\boldsymbol{x}$ from the criterion assumes that the prior model and the noise are well-known and valid on the entire set of possible solutions of $\boldsymbol{h}$. This is not necessarily true at least for the prior as discussed by Molina, 1994.

2.1.4 Deep learning methods

Many architectures have been considered in deep neural networks to tackle the blind deconvolution problem (Zhang et al., 2022). Deep Auto-Encoders are usually used in this context as they are easy to implement. They rely on two structures, an encoder which will extract the features, or modes, of the data to a latent space of fewer dimensions, and a decoder which takes parameters in the latent space and reconstructs the data. The two modules are trained thanks to an extensive learning database with blurred data and their sharp counterpart. Another architecture commonly used is the Generative Adversarial Network which generates sharp images from blurred data and compares them to a database containing real sharp images via a discriminator.

Training these networks is however difficult. The choice of a loss to minimize, the training computational cost, the number of examples in the learning database, as well as their diversity are making their use complicated in our context.

2.2 Using the scaling ambiguity to our advantage

2.2.1 Generalizing the problem to bi-linear models

Blind deconvolution is only a sub-problem of a more general estimation problem that applies to many different contexts and can be summarized as an un-mixing problem between two components. Other examples of applications are blind source separations, low-rank approximations, or separation of spectral and spatial contributions (see Chapter 3).

Let us consider the following model of the data $\boldsymbol{d} \in \mathbb{R}^N$:

$$\boldsymbol{d} = \boldsymbol{h} \circledast \boldsymbol{x} + \boldsymbol{b} \tag{2.7}$$

with $\boldsymbol{h} \in \mathbb{H}$ and $\boldsymbol{x} \in \mathbb{X}$ forming a bi-linear model thanks to the operator $\circledast$ (*e.g.* convolution, element-wise multiplication), *i.e.*:

$$\forall(\alpha, \beta) \in \mathbb{R}^2, (\alpha\boldsymbol{h}) \circledast (\beta\,\boldsymbol{x}) = (\alpha\beta)\,(\boldsymbol{h} \circledast \boldsymbol{x}) \tag{2.8a}$$
$$\forall(\boldsymbol{h}_1, \boldsymbol{h}_2) \in \mathcal{H}, (\boldsymbol{h}_1 + \boldsymbol{h}_2) \circledast \boldsymbol{x} = (\boldsymbol{h}_1 \circledast \boldsymbol{x}) + (\boldsymbol{h}_2 \circledast \boldsymbol{x}) \tag{2.8b}$$
$$\forall(\boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathcal{H}, \boldsymbol{h} \circledast (\boldsymbol{x}_1 + \boldsymbol{x}_2) = (\boldsymbol{h} \circledast \boldsymbol{x}_1) + (\boldsymbol{h} \circledast \boldsymbol{x}_2)\,. \tag{2.8c}$$

$\mathbb{H}$ and $\mathbb{X}$ are subsets of $\mathbb{R}^N$ that can restrict the solutions to enforce constraints such as normalization, positivity, *etc.*. $\boldsymbol{b} \in \mathbb{R}^N$ is a spurious signal coming from phenomena described in 1.1 (not necessarily Gaussian).

In such cases, the joint MAP estimation problem can be written more generally as:

$$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{h}}) = \underset{\boldsymbol{x}\in\mathbb{X}, \boldsymbol{h}\in\mathbb{H}}{\arg\min} \Big\{ \mathcal{L}(\boldsymbol{x}, \boldsymbol{h}; \mu, \lambda) = \mathcal{G}(\boldsymbol{x} \circledast \boldsymbol{h}) + \mu\,\mathcal{J}(\boldsymbol{x}) + \lambda\,\mathcal{K}(\boldsymbol{h}) \Big\} \tag{2.9}$$

where $\mathcal{G}(\boldsymbol{x} \circledast \boldsymbol{h})$ measures the fidelity of the model compared to the data and $\mathcal{J}(\boldsymbol{x})$ and $\mathcal{K}(\boldsymbol{h})$ are regularization terms.

2.2.2 Taking into account the scaling ambiguity

As is generally the case in image reconstruction, we only consider homogeneous regularization functions. In this case, $\mathcal{J}$ and $\mathcal{K}$ follow:

$$\exists q \in \mathbb{R},\ \forall\alpha \geqslant 0,\ \forall\boldsymbol{x},\ \ \mathcal{J}(\alpha\,\boldsymbol{x}) = \alpha^q\,\mathcal{J}(\boldsymbol{x}) \geqslant 0 \tag{2.10a}$$
$$\exists r \in \mathbb{R},\ \forall\alpha \geqslant 0,\ \forall\boldsymbol{h},\ \ \mathcal{K}(\alpha\,\boldsymbol{h}) = \alpha^r\,\mathcal{K}(\boldsymbol{h}) \geqslant 0 \tag{2.10b}$$

where $\alpha \geqslant 0$ as both $\boldsymbol{x}$ and $\boldsymbol{h}$ represent physical quantities and q and r are called the degrees of the homogeneous functions. For example, the Smooth Tikhonov regularization of Eq. (1.52) corresponds to a homogeneous function of degree 2 while the total variation regularization of Eq. (1.55) is homogeneous of degree 1.

Taking the homogeneous property of the regularizations, the scaling invariance of Eq. (2.1) can be written explicitly in the function to minimize in Eq. (2.9):

$$\begin{aligned} \forall\alpha > 0,\ \mathcal{L}(\alpha\,\boldsymbol{x}, \boldsymbol{h}/\alpha; \mu, \lambda) &= \mathcal{G}((\alpha\,\boldsymbol{x}) \circledast (\boldsymbol{h}/\alpha)) + \mu\,\mathcal{J}(\alpha\,\boldsymbol{x}) + \lambda\,\mathcal{K}(\boldsymbol{h}/\alpha) \\ &= \mathcal{G}(\boldsymbol{x} \circledast \boldsymbol{h}) + \mu\alpha^q\,\mathcal{J}(\boldsymbol{x}) + \lambda\alpha^{-r}\,\mathcal{K}(\boldsymbol{h}) \\ &= \mathcal{L}(\boldsymbol{x}, \boldsymbol{h}; \alpha^q\mu, \lambda\alpha^{-r}) \end{aligned} \tag{2.11}$$

This means that by scaling the two parameters, one by α and the other one by $1/\alpha$, the regularization weights are modified while the bilinear model (and hence $\mathcal{G}$) is not. From this statement, one can

wonder about the optimal scaling of the problem. An estimator of this factor can be found by solving:

$$\begin{aligned}\hat{\alpha} &= \arg\min_{\alpha>0} \mathcal{L}(\boldsymbol{x}, \boldsymbol{h}; \alpha^q \mu, \lambda \alpha^{-r}) \\ &= \arg\min_{\alpha>0} \Big(\mu \alpha^q \, \mathcal{J}(\boldsymbol{x}) + \lambda \alpha^{-r} \, \mathcal{K}(\boldsymbol{h}) \Big)\end{aligned} \tag{2.12}$$

Taking the first and second derivatives with respect to α yields:

$$\frac{\partial \mathcal{L}(\boldsymbol{x}, \boldsymbol{h}; \alpha^q \mu, \lambda \alpha^{-r})}{\partial \alpha} = q \mu \alpha^{q-1} \, \mathcal{J}(\boldsymbol{x}) - r \lambda \alpha^{-r-1} \, \mathcal{K}(\boldsymbol{h}) \tag{2.13}$$

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{x}, \boldsymbol{h}; \alpha^q \mu, \lambda \alpha^{-r})}{\partial \alpha^2} = q(q-1) \mu \alpha^{q-2} \, \mathcal{J}(\boldsymbol{x}) + r(r+1) \lambda \alpha^{-r-2} \, \mathcal{K}(\boldsymbol{h}) \tag{2.14}$$

showing that, provided $\mathcal{J}(\boldsymbol{x}) > 0$ and $\mathcal{K}(\boldsymbol{h}) > 0$, the problem is strictly convex ($\frac{\partial^2 \mathcal{L}}{\partial \alpha^2} > 0$) for $\alpha \in \mathbb{R}_*^+$, $r > 0$ and $q \geqslant 1$[1]. In other words, there is a closed-form solution to the optimal scaling factor, which depends on $\boldsymbol{x}$ and $\boldsymbol{h}$ and the hyper-parameters μ and λ:

$$\hat{\alpha}(\boldsymbol{x}, \boldsymbol{h}; \mu, \lambda) = \left(\frac{r \lambda \, \mathcal{K}(\boldsymbol{h})}{q \mu \, \mathcal{J}(\boldsymbol{x})} \right)^{\frac{1}{q+r}} \tag{2.15}$$

Using this expression, it is possible to write a new function to minimize by substituting α with its optimal value in Eq. (2.11):

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{h}; \eta) = \mathcal{G}(\boldsymbol{x} \circledast \boldsymbol{h}) + \eta \, \mathcal{J}(\boldsymbol{x})^{\frac{r}{q+r}} \, \mathcal{K}(\boldsymbol{h})^{\frac{q}{q+r}} \tag{2.16}$$

where taking

$$\eta = \left(\left(\left(\frac{r}{q} \right)^q + \left(\frac{q}{r} \right)^r \right) \mu^r \, \lambda^q \right)^{\frac{1}{q+r}} \tag{2.17}$$

ensures that

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{h}; \eta) = \min_{\alpha>0} \mathcal{L}(\boldsymbol{x}, \boldsymbol{h}; \alpha^q \mu, \lambda \alpha^{-r}) = \min_{\alpha>0} \mathcal{L}(\alpha \, \boldsymbol{x}, \boldsymbol{h}/\alpha; \mu, \lambda) \tag{2.18}$$

[1] These two former conditions are usually met using common regularization functions.

Proof:

$$
\begin{aligned}
\mathcal{H}(\boldsymbol{x},\boldsymbol{h};\eta(\mu,\lambda)) &= \min_{\alpha>0} \mathcal{L}(\boldsymbol{x},\boldsymbol{h};\alpha^q\mu,\lambda\alpha^{-r}) \\
&= \mathcal{L}(\boldsymbol{x},\boldsymbol{h};\alpha^q\mu,\lambda\alpha^{-r}) \mid_{\alpha=\hat{\alpha}(\boldsymbol{x},\boldsymbol{h};\mu,\lambda)} \\
&= \mathcal{G}(\boldsymbol{x}\circledast\boldsymbol{h}) + \mu\Big(\frac{r\lambda\,\mathcal{K}(\boldsymbol{h})}{q\mu\,\mathcal{J}(\boldsymbol{x})}\Big)^{\frac{q}{q+r}}\mathcal{J}(\boldsymbol{x}) + \lambda\Big(\frac{r\lambda\,\mathcal{K}(\boldsymbol{h})}{q\mu\,\mathcal{J}(\boldsymbol{x})}\Big)^{-\frac{r}{q+r}}\mathcal{K}(\boldsymbol{h}) \\
&= \mathcal{G}(\boldsymbol{x}\circledast\boldsymbol{h}) + \Bigg(\Big(\frac{r}{q}\Big)^{\frac{q}{q+r}}\mu\Big(\frac{\lambda}{\mu}\Big)^{\frac{q}{q+r}}\mathcal{J}(\boldsymbol{x})\Big(\frac{\mathcal{K}(\boldsymbol{h})}{\mathcal{J}(\boldsymbol{x})}\Big)^{\frac{q}{q+r}} \\
&\qquad + \Big(\frac{r}{q}\Big)^{\frac{-r}{q+r}}\lambda\Big(\frac{\lambda}{\mu}\Big)^{\frac{-r}{q+r}}\mathcal{K}(\boldsymbol{h})\Big(\frac{\mathcal{K}(\boldsymbol{h})}{\mathcal{J}(\boldsymbol{x})}\Big)^{\frac{-r}{q+r}}\Bigg) \\
&= \mathcal{G}(\boldsymbol{x}\circledast\boldsymbol{h}) + \Bigg(\Big(\frac{r}{q}\Big)^{q}\mu^r\lambda^q\,\mathcal{J}(\boldsymbol{x})^r\,\mathcal{K}(\boldsymbol{h})^q + \Big(\frac{q}{r}\Big)^{r}\mu^r\lambda^q\,\mathcal{J}(\boldsymbol{x})^r\,\mathcal{K}(\boldsymbol{h})^q\Bigg)^{\frac{1}{q+r}} \\
&= \mathcal{G}(\boldsymbol{x}\circledast\boldsymbol{h}) + \underbrace{\Bigg(\Big(\Big(\frac{r}{q}\Big)^{q} + \Big(\frac{q}{r}\Big)^{r}\Big)\mu^r\lambda^q\Bigg)^{\frac{1}{q+r}}}_{\eta(\mu,\lambda)}\mathcal{J}(\boldsymbol{x})^{\frac{r}{q+r}}\,\mathcal{K}(\boldsymbol{h})^{\frac{q}{q+r}} \qquad (2.19)
\end{aligned}
$$

■

This formulation implies that when considering only the morphology[2] of the pair $(\boldsymbol{x},\boldsymbol{h})$, their regularizations are intrinsically linked. It shows the importance of regularization on the two components of the direct model. In other words, tuning either μ or λ has an impact on η and so acts on the level of the two regularizations. As discussed in section 2.1, Eq. (2.9) has many local minima and the regularizations play a key role in the minima found by the optimization method. Thanks to the scaling ambiguity, we propose to make their tuning easier by only considering one hyper-parameter.

2.2.3 AMORS: Alternated Minimization using Optimal ReScaling

Section 2.2.2 shows that, when dealing with bi-linear direct models, only one hyper-parameter is necessary to tune both the regularization of $\boldsymbol{x}$ and $\boldsymbol{h}$. This statement and the new criterion derived in Eq. (2.16) lead to two possible methods to solve the estimation problem with optimal rescaling.

A first approach would be to choose an hyper-parameter η and to minimize jointly $\mathcal{H}(\boldsymbol{x},\boldsymbol{h};\eta)$ with respect to $\boldsymbol{x}$ and $\boldsymbol{h}$:

$$
(\hat{\boldsymbol{x}},\hat{\boldsymbol{h}}) = \underset{\boldsymbol{x}\geqslant\boldsymbol{0},\boldsymbol{h}\geqslant\boldsymbol{0}}{\arg\min}\Big\{\mathcal{H}(\boldsymbol{x},\boldsymbol{h};\eta) = \mathcal{G}(\boldsymbol{x}\circledast\boldsymbol{h}) + \eta\,\mathcal{J}(\boldsymbol{x})^{\frac{r}{q+r}}\,\mathcal{K}(\boldsymbol{h})^{\frac{q}{q+r}}\Big\} \qquad (2.20)
$$

This solution is however difficult as it requires minimizing a non-linear function and does not re-use existing algorithms.

A second possible approach is to fix one of the hyper-parameters (μ or λ) and minimize the function $\mathcal{L}(\boldsymbol{x},\boldsymbol{h};\hat{\alpha}^q\mu,\lambda\hat{\alpha}^{-r})$ with $\hat{\alpha}$ the optimal scaling factor deduced by Eq. (2.15). As seen in section 2.1.2, using the AM estimation method in this context is very efficient for convex regularizations[3], as it decomposes the problem in two convex sub-problems (Eq. (2.4a) and (2.4b)), much easier to solve thanks to already existing methods. Moreover, Thiébaut, 2002 has shown that optimally scaling the parameters in a joint estimation strategy improves greatly the convergence of an iterative optimization method.

[2] We are interested in the family of solutions $\{(\alpha\,\boldsymbol{x},\boldsymbol{h}/\alpha)\,|\,\alpha>0\}$.

[3] which is a common property as seen in section 2.1.1

Algorithm 2.1: AMORS: Alternated Minimization using the Optimal ReScaling in a bi-linear direct model context

Input: The criterion $\mathcal{L}$, algorithms to solve Eqs. (2.4a) and (2.4b), a well-tuned regularization parameter μ and an arbitrarily fixed hyper-parameter λ, an initial scaling factor $\alpha_{[0]}$ and an initial guess $\boldsymbol{h}_{[0]}$.

Output: $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{h}}$, a local minimum of the function defined in Eq. (2.11).

$k \leftarrow 0$

while *not converged* **do**

  while *true* **do**

    ▶ **Estimation of $\boldsymbol{x}$**

    $\boldsymbol{x}_{[k+1]} \leftarrow \hat{\boldsymbol{x}}(\boldsymbol{h}_{[k]}; \alpha_{[k]}^{q}\mu)$ ◁ Eq. (2.4a)

    ▶ **Update of the scaling factor α**

    $\alpha_{[k+1/2]} \leftarrow \hat{\alpha}(\boldsymbol{x}_{[k+1]}, \boldsymbol{h}_{[k]}; \mu, \lambda)$ ◁ Eq. (2.15)

    ▶ **Re-scaling initialization**

    if $k > 0$ **or** $\alpha_{[k+1/2]} \approx \alpha_{[k]}$ **then Break**

    $\alpha_{[k]} \leftarrow \alpha_{[k+1/2]}$

  ▶ **Estimation of $\boldsymbol{h}$**

  $\boldsymbol{h}_{[k+1]} \leftarrow \hat{\boldsymbol{h}}(\boldsymbol{x}_{[k+1]}; \alpha_{[k+1/2]}^{-r}\lambda)$ ◁ Eq. (2.4b)

  ▶ **Update of the scaling factor α**

  $\alpha_{[k+1]} \leftarrow \hat{\alpha}(\boldsymbol{x}_{[k+1]}, \boldsymbol{h}_{[k+1]}; \mu, \lambda)$ ◁ Eq. (2.15)

  $k \leftarrow k + 1$

$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{h}}) \leftarrow (\boldsymbol{x}_{[k]}, \boldsymbol{h}_{[k]})$

We develop here an Alternated Minimization using the Optimal ReScaling (AMORS) which updates α after each estimation of the parameters $\boldsymbol{x}$ and $\boldsymbol{h}$. The algorithm shown in Alg. 2.1 is the result of these considerations. Note that $\boldsymbol{x}$ and $\boldsymbol{h}$ have the same role in this algorithm, meaning that they can be exchanged. Given that $\mathcal{L}$ is not jointly convex with respect to $\boldsymbol{x}$ and $\boldsymbol{h}$, the initialization $\boldsymbol{h}_{[0]}$ is important in terms of convergence speed and solutions found. To reduce this sensitivity, we propose to initially re-scale the problem by alternatively estimating $\boldsymbol{x}$ and updating α, until the scaling factor converges. This simple strategy makes the algorithm robust to the scaling of the initialization, which also improves the convergence speed. It allows also to use an arbitrary value for the initial scaling factor, say $\alpha_{[0]} = 1$ (which is refined by this process at the beginning of the method). Note that if there are any constraints on the PSF (*e.g.* normalization), they can be enforced at the end of the algorithm.

2.2.4 Unsupervised tuning of remaining hyper-parameters

Given the simplicity of Alg. 2.1, it can be embedded into an unsupervised strategy as a *black box* algorithm. Using a criterion $\mathcal{C}$ that depends on $\boldsymbol{x}$, $\boldsymbol{h}$ and μ such as the GSURE (see Section 1.4), it is possible to derive a hierarchical strategy consisting in solving:

$$\hat{\mu}_{\boldsymbol{x}} = \underset{\mu>0}{\arg\min}\, \mathcal{C}(\hat{\boldsymbol{x}}(\mu), \hat{\boldsymbol{h}}(\mu); \mu) \tag{2.21}$$

with

$$(\hat{\boldsymbol{x}}(\mu), \hat{\boldsymbol{h}}(\mu), \hat{\alpha}(\mu)) = \underset{\boldsymbol{x}\in\mathbb{X}, \boldsymbol{h}\in\mathbb{H}, \alpha>0}{\arg\min}\, \mathcal{L}(\boldsymbol{x}, \boldsymbol{h}; \alpha^{q}\mu, \alpha^{-r}) \tag{2.22}$$

solved by AMORS (Alg.2.1) when fixing λ arbitrarily (say $\lambda = 1$).

2.3 Applying the method to blind deconvolution

Although our method is not restricted to this context (see Chapter 3), we decide here to validate and evaluate the advantages of AMORS in the blind deconvolution context. Using the same simulation process of blurred data as the one explained in Section 1.2, we illustrate the advantages of our method AMORS. The loss function used to recover the object of interest and PSF is defined as Eq. (2.2) under positivity constraint:

$$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{h}}) = \underset{\boldsymbol{x}\geqslant\mathbf{0},\boldsymbol{h}\geqslant\mathbf{0}}{\arg\min} \|\boldsymbol{d} - \mathbf{h} \star \boldsymbol{x}\|^2_{\mathbf{W}} + \mu\, \mathcal{J}(\boldsymbol{x}) + \lambda\, \mathcal{K}(\boldsymbol{h}) \tag{2.23}$$

and each sub-problem of AMORS (Eqs. (2.4a) and (2.4b)) is solved by the quasi-Newton method VMLM-B ((Thiébaut, 2002), see Section 1.5.2).

2.3.1 Proposition of a homogeneous edge-preserving regularization

Taking advantage of the scaling ambiguity requires the regularization to be homogeneous. Commonly used regularizations have this property such as the regularizations defined in Eq. (1.52) and (1.55). Although some specific regularization functions do not have this property, it is sometimes possible to adapt them in order to obtain the homogeneity property.

An important example is the edge-preserving regularization which is not homogeneous in its hyperbolic form, shown in Eq. (1.56). As seen in Section 1.3.5, this function is a middle-ground between the $\ell 1$ and squared $\ell 2$ norm and is very efficient in a MAP estimation context. To respect the necessary hypothesis for using scaling ambiguity, we propose in the following, a homogeneous version of degree 2 of the edge-preserving regularization:

$$\boxed{\mathcal{J}_{\text{hep}}(\boldsymbol{x}) = 2\rho\, \|\boldsymbol{x}\| \sum_n \left(\sqrt{\|\mathbf{D}_n\, \boldsymbol{x}\|^2 + \rho^2\, \|\boldsymbol{x}\|^2} - \rho\, \|\boldsymbol{x}\| \right)} \tag{2.24}$$

with $\rho \geqslant 0$ a threshold which tunes the $\ell 1$-$\ell 2$ behavior. Following the decomposition of Section 1.3.5:

$$2\rho\, \|\boldsymbol{x}\| \left(\sqrt{\|\mathbf{D}_n\, \boldsymbol{x}\|^2 + \rho^2\, \|\boldsymbol{x}\|^2} - \rho\, \|\boldsymbol{x}\| \right) \simeq \begin{cases} \|\mathbf{D}_n\, \boldsymbol{x}\|^2 & \text{for } \|\mathbf{D}_n\, \boldsymbol{x}\| \ll \rho\, \|\boldsymbol{x}\| \\ 2\rho\, \|\boldsymbol{x}\|\, \|\mathbf{D}_n\, \boldsymbol{x}\| & \text{for } \|\mathbf{D}_n\, \boldsymbol{x}\| \gg \rho\, \|\boldsymbol{x}\| \end{cases} \tag{2.25}$$

Proof: When $\|\mathbf{D}_n\, \boldsymbol{x}\| \ll \rho\, \|\boldsymbol{x}\|$:

$$\begin{aligned} 2\rho\, \|\boldsymbol{x}\| \left(\rho\, \|\boldsymbol{x}\| \sqrt{1 + \frac{\|\mathbf{D}_n\, \boldsymbol{x}\|^2}{\rho^2\, \|\boldsymbol{x}\|^2}} - \rho\, \|\boldsymbol{x}\| \right) &\approx 2\rho\, \|\boldsymbol{x}\| \left(\rho\, \|\boldsymbol{x}\| \left(1 + \frac{\|\mathbf{D}_n\, \boldsymbol{x}\|^2}{2\rho^2\, \|\boldsymbol{x}\|^2} \right) - \rho\, \|\boldsymbol{x}\| \right) \\ &\approx 2\rho\, \|\boldsymbol{x}\| \left(\frac{\|\mathbf{D}_n\, \boldsymbol{x}\|^2}{2\, \rho\, \|\boldsymbol{x}\|} \right) \\ &\approx \|\mathbf{D}_n\, \boldsymbol{x}\|^2 \end{aligned} \tag{2.26}$$

the regularization behaves as a quadratic smoothing regularization (see Eq. (1.52)), and when $\|\mathbf{D}_n \boldsymbol{x}\| \gg \rho \|\boldsymbol{x}\|$:

$$
\begin{aligned}
2\rho\|\boldsymbol{x}\| \left(\|\mathbf{D}_n \boldsymbol{x}\| \sqrt{1 + \frac{\rho^2 \|\boldsymbol{x}\|^2}{\|\mathbf{D}_n \boldsymbol{x}\|^2}} - \rho\|\boldsymbol{x}\| \right) &\approx 2\rho\|\boldsymbol{x}\| \left(\|\mathbf{D}_n \boldsymbol{x}\| \left(1 + \frac{\rho^2 \|\boldsymbol{x}\|^2}{2\|\mathbf{D}_n \boldsymbol{x}\|^2}\right) - \rho\|\boldsymbol{x}\| \right) \\
&\approx 2\rho\|\boldsymbol{x}\| \left(\|\mathbf{D}_n \boldsymbol{x}\| + \rho\|\boldsymbol{x}\| \underbrace{\left(\frac{\rho\|\boldsymbol{x}\|}{2\|\mathbf{D}_n \boldsymbol{x}\|} - 1 \right)}_{\simeq -1} \right) \\
&\approx 2\rho\|\boldsymbol{x}\| \underbrace{\left(\|\mathbf{D}_n \boldsymbol{x}\| - \rho\|\boldsymbol{x}\| \right)}_{\simeq \|\mathbf{D}_n \boldsymbol{x}\|} \\
&\approx 2\rho\|\boldsymbol{x}\| \|\mathbf{D}_n \boldsymbol{x}\| \qquad (2.27)
\end{aligned}
$$

it behaves as a non-quadratic norm with respect to the gradient of $\boldsymbol{x}$ (Eq. (1.55)). ■

Note that the expression in Eq. (2.24) allows decoupling the two hyper-parameters as done in Section 1.3.5. By observing eq. (1.56), one can remark that our formulation is similar to the commonly used function when taking $\epsilon = \rho\|\boldsymbol{x}\|$. The tuning of the parameter ρ can be done by studying the distribution of the gradient of the data compared to its $\ell 2$ norm.

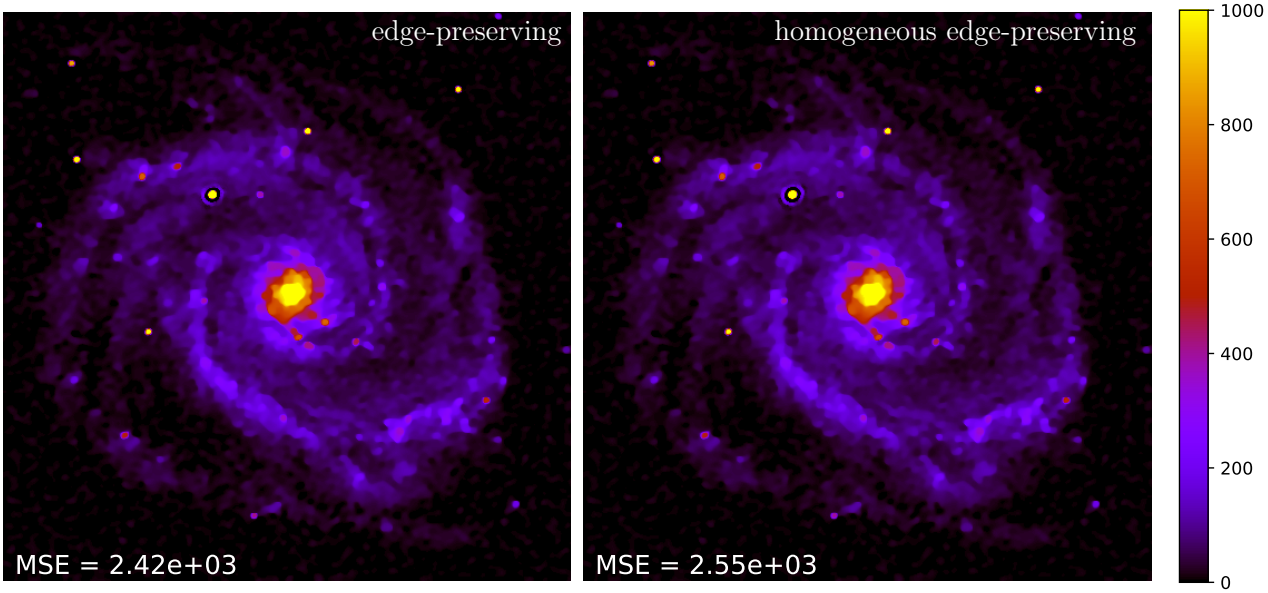

Figure 2.1 – Solution of the deconvolution problem defined in Eq. (1.47) with edge-preserving regularization (Eq. (1.56)) on the left and our homogeneous version on the right (Eq. (2.24))

Taking the simulated data of Section 1.2 of PSNR $\simeq 40$, we test this homogeneous version of the edge-preserving on the deconvolution problem defined in Eq. (1.47). Fig. 2.1 recalls the solution estimated in Section 1.3.5 (left) and compares it to the homogeneous version (right). The results are quite similar with close MSE values demonstrating the equivalence of the two regularizations.

2.3.2 Inverse-crime validation

We consider the blind-deconvolution case on simulated data following Section 1.2. We use, to do so, the 3.6m Canada-France Hawaï telescope experimental PSF and the ground truth of the M-51 galaxy as presented in Fig. 2.2 to simulate data with PSNR $\simeq 40$. At the initial iteration, the algorithm needs an initial PSF for solving Eq. (2.4a). In the following, we choose the Airy function corresponding to

the theoretical PSF and far from the ground truth. To solve the optimization problem, we choose a flat image of ones to start the iterative process (see Section 1.5.2).

Results of AMORS for a quadratic smoothness regularization for both the object and the PSF, $\hat{\boldsymbol{x}}^{[\text{bdec,Tik}]}$ and $\hat{\boldsymbol{h}}^{[\text{bdec,Tik}]}$, are shown in the right-middle Figure of Fig. 2.2. For comparison purposes, the result of the deconvolution with the quadratic smoothing regularization, $\hat{\boldsymbol{x}}^{[\text{dec,Tik}]}$, is given on the left-middle figure. The bottom line gives the results of AMORS with an edge-preserving regularization, $\hat{\boldsymbol{x}}^{[\text{bdec,hep}]}$ and $\hat{\boldsymbol{h}}^{[\text{bdec,hep}]}$, on the right figure and gives the deconvolution results $\hat{\boldsymbol{h}}^{[\text{dec,hep}]}$ for comparison purposes on the left figure (hep = homogeneous edge-preserving). As in the other studies, the hyper-parameters are tuned to minimize the true MSE between the ground truth object and its estimate.

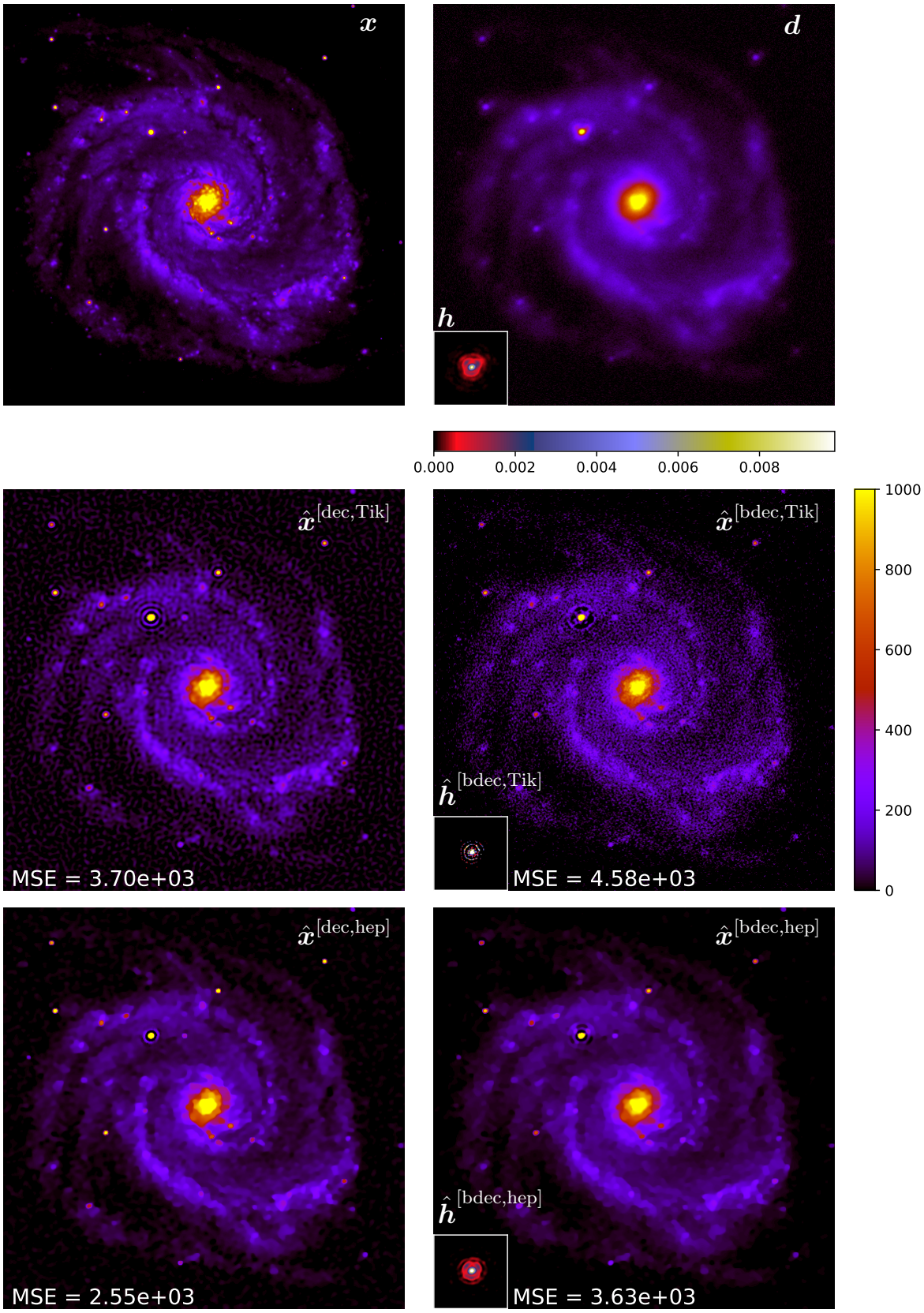

Figure 2.2 – Results on simulated data. The top row presents the simulated data (right) of PSNR $\simeq 40$, corresponding to the convolution of the ground truth object (left) by the PSF (overlapped with the top-right figure). The middle row figures show the results of the deconvolution (left) and of the blind deconvolution (right) problems with a quadratic smoothness regularization for both the object and the PSF. The bottom row shows the same results but with our homogeneous edge-preserving as the object regularization and a quadratic smoothness regularization for the PSF.

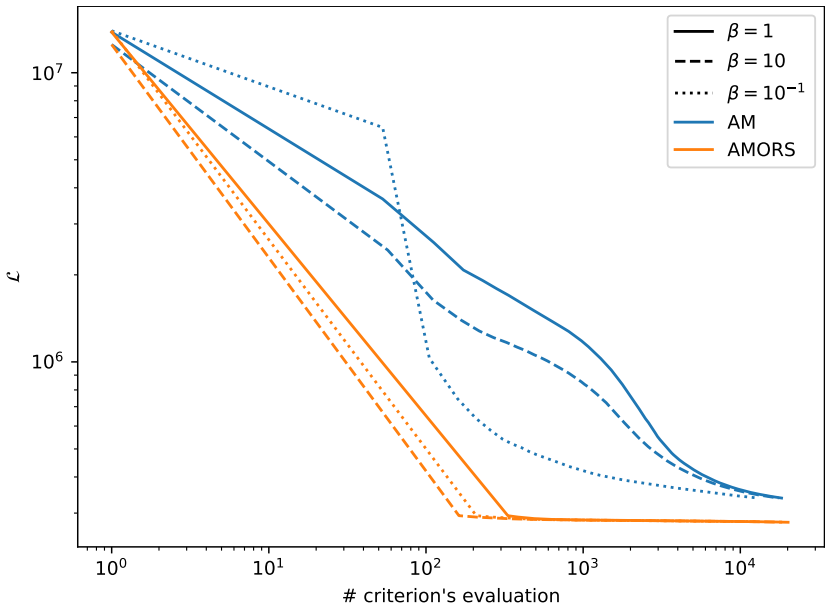

Figure 2.3 – Evolution of the loss function as defined in Eq. (2.2) when using the AMORS algorithm (orange) for different scaled initialization $\boldsymbol{h}_{[0]} = \beta\,\boldsymbol{h}_{\text{init}}$, with $\beta = 10^{-1}$ (dotted lines), $\beta = 1$ (plain lines) and $\beta = 10$ (dashed lines). The evolution of the loss function is also shown in the case of an AM algorithm (blue) for the same initializations.

We can see that the AMORS algorithm gives reasonably good results even for low PSNR of $\simeq 40$. The object and PSF regularization are tuned correctly by our method as they give results close to the ground truth. The edge-preserving *a priori* for the object allied with quadratic smoothness regularization for the PSF shows better results than when considering a quadratic smoothness regularization for the object, as expected and found in the deconvolution case. The object having a lot of sharp edges (arms of the galaxies) and point-like sources (background stars), the edge-preserving *a priori* favors a cleaner reconstruction. Note also that the PSF shape and in particular the variations of its second lobe are well reconstructed, especially considering the symmetry of the initial starting PSF (Airy function).

Studying the impact of the initialization on the results, we choose to use AMORS on the data presented in Fig. 2.2 with a PSNR $\simeq 40$. The algorithm is initialized with $\boldsymbol{h}_{[0]} = \beta\,\boldsymbol{h}_{\text{init}}$, with β a scaling factor of 10^{-1}, 1 or 10 and $\boldsymbol{h}_{\text{init}}$ the initial Airy PSF. In order to validate further our algorithm, Fig. 2.3 shows the evolution of the loss function that is alternatively minimized (Eq. (2.2)) with respect to the number of calls used by the optimization method (VMLM-B here) for a thousand outer-iterations of AMORS. The same loss function is presented with the commonly used AM algorithm (Eqs. (2.4a) and (2.4b)). The different hyper-parameters[4] are tuned objectively by the MSE criterion, using the derivative-free methods Newuoa (see Section 1.5.3) to avoid any artifacts due to an incorrect value of the regularizations. As can be seen, by taking into account the optimal scaling in the estimation process, especially at the beginning of the procedure, AMORS is quite insensitive to the scaling of the initialization and converges quickly towards the same solution, despite an increase or a decrease by a factor of ten of the initial PSF scaling. Comparatively, we can see that the commonly used alternative minimization depends greatly on the scaling of $\boldsymbol{h}_{[0]}$.

[4]the scaling of the object regularization μ and the threshold ρ of the edge-preserving for AMORS, μ, ρ and λ, the scaling of the PSF regularization, for the AM method

2.4 Conclusions and perspectives

In this chapter, we discussed the bilinear model difficulties in the inverse problem context. But more than a problem, we showed that one of the properties of such a model, the scaling ambiguity, can be used to simplify greatly the tuning of hyper-parameters. Thanks to the optimal scaling of the model, we can reduce the number of hyper-parameters while ensuring stability and quick convergence against the scaling of the initialization. Results are shown in a blind-deconvolution context, although, AMORS can be applied to any bilinear direct model. After introducing a custom homogeneous edge-preserving, we demonstrated the efficiency of AMORS in a blind deconvolution context on simulated data. There are still improvements possible to AMORS such as an *on the fly* tuning of the remaining hyper-parameters which is currently being developed, possibly inspired by hierarchical Bayesian methods (Molina, 1994) or by a SURE update (Ammanouil et al., 2019). More tests and validations can be done on AMORS, especially on the convergence speed and the local minimum it converges to. This work is part of an ongoing article that will soon be submitted to a peer-reviewed journal.

Chapter 3

Spectral characterization of sub-stellar objects using long-slit spectroscopy

Studying exoplanets through the measurement of their spectrum is an important part of understanding their formation process. Thanks to direct imaging, many types of sub-stellar companions have been discovered and characterized these last years (Langlois, M. et al., 2021). Beyond the blur phenomenon[1], other perturbations, coming from this instrument and the observation methods, can occur in the data. Remnants of the stellar light are usually still in the data, despite the coronagraphic mask. Disentangling this spurious component from the companion's signal can be challenging.

This chapter focuses on the Long-Slit Spectroscopy (LSS) observation method allied with the powerful coronagraphy necessary to characterize close-by companions of stars. After presenting the principles of such a method of observation, we discuss in section 3.1 some strategies to extract the companion spectrum from LSS data. Choosing the inverse problem framework, a direct model is developed in Section 3.2 relying on a geometric calibration procedure detailed in section 3.3. We present in Section 3.4 our Exospeco algorithm to extract the companion's information in the data. Section 3.5 shows results on data with synthetic companions, in order to validate our method. We conclude on the next steps in Section 3.6.

In This Chapter

Prerequisite:

- Maximum *a posteriori* formalism (Sections 1.3.4 and 1.3.5).
- Amors method (Section 2.2).

Contributions:

- Proposal of a direct model of the long-slit spectroscopy data.
- Building a calibration strategy of the spatial and spectral dispersion laws.
- Development of the Exospeco algorithm to extract the companion SED.
- Accepted paper at A&A journal (Thé et al., 2023) and presentations at Lyot 2022 (Thé et al., 2022b) and Spie 2022 (Thé et al., 2022c).

[1] discussed in Chap. 1 and 2.

3.1 Long-slit spectroscopy in high contrast imaging

3.1.1 Observation principle

Long-slit spectroscopy (LSS) is a method consisting in observing an object through a narrow slit, isolating its light from other sources. The signal is then decomposed spectrally by a dispersive element before being recorded on a detector.

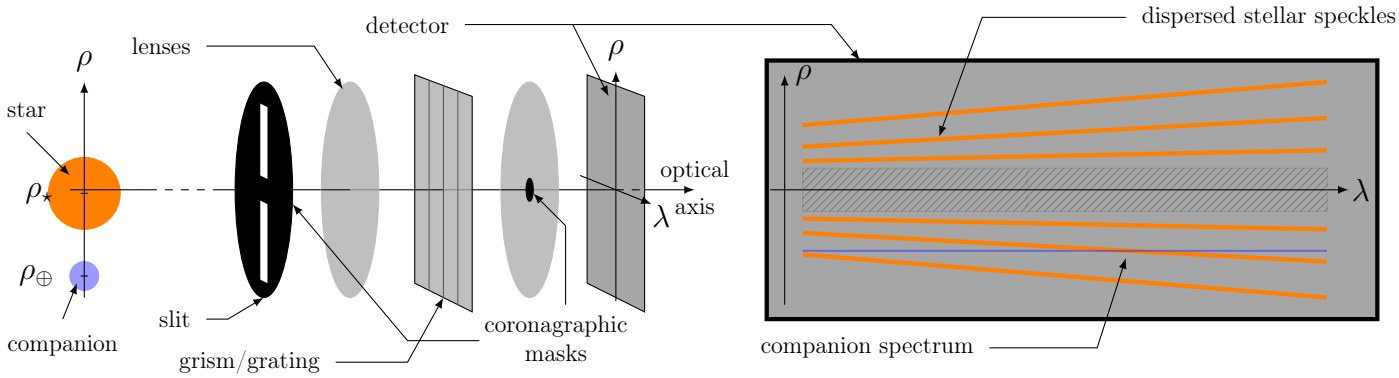

Figure 3.1 – Schematic of an LSS instrument (left) and the data it's collecting on the detector (right), in high-contrast imaging. Wavelength (λ) and angular separation to the coronagraphic mask's center (ρ) are indicated.

In the context of direct imaging of companions, LSS is used by aligning the slit on the companion-star axis, as shown in Fig. 3.1. The IRDIS instrument[2] (Dohlen et al., 2008), was designed for exoplanets study and has an LSS mode in J ($\sim$ 1220 nm), H ($\sim$ 1630 nm) and K ($\sim$ 2190 nm) bands (Vigan et al., 2008). Thanks to Extreme Adaptive Optics (ExAO), it can image stars' surroundings in Low (LRS) or Medium (MRS) Resolution Mode ($\lambda/\Delta\lambda = 35$ or $\lambda/\Delta\lambda = 400$) by correcting the atmospheric distortions of the incoming wavefront. The star being much brighter than the companion, the LSS is combined with coronagraphy to strongly attenuate the star's signal on the detector. The coronagraphic mask produces on the detector a dark area centered on the optical axis, shown by the hatched area on the right side of Fig. 3.1.

The diffraction of starlight creates residuals in the detector plane which, due to aberrations in the optical path, appear in the form of dispersed speckles (represented as orange oblique lines in Fig. 3.1). Lower by a factor of ~ 10 from the star intensity, these so-called dispersed speckles are still 10^3 to 10^5 brighter than the companion's signal. Fig. 3.2 presents real data from IRDIS with the speckles visible as bright oblique lines and a bright companion (between arrowheads). In addition to this spurious component, the data is corrupted by noise (see Section 1.1). The bright and dark spots visible in the data are so-called *bad pixels*. They are usually static and can be detected once and for all through calibration steps. As we want to observe faint companions of stars, the data in itself is not usable and needs numerical processing methods to disentangle the companion signal from the star component.

3.1.2 Continuous model of the data

Given the wavelength λ and the angular separation, ρ, from the coronagraphic mask, the LSS data can be expressed for each pixel n of the detector as:

$$d_n = m(\rho_n, \lambda_n) + b_n \tag{3.1}$$

with b_n a nuisance term following a centered multivariate Gaussian distribution as explained in Section 1.2 and $m(\rho_n, \lambda_n)$ the direct model of the light distribution on the detector's plane at the pixel of

[2]for InfraRed Dual Imaging Spectrograph.

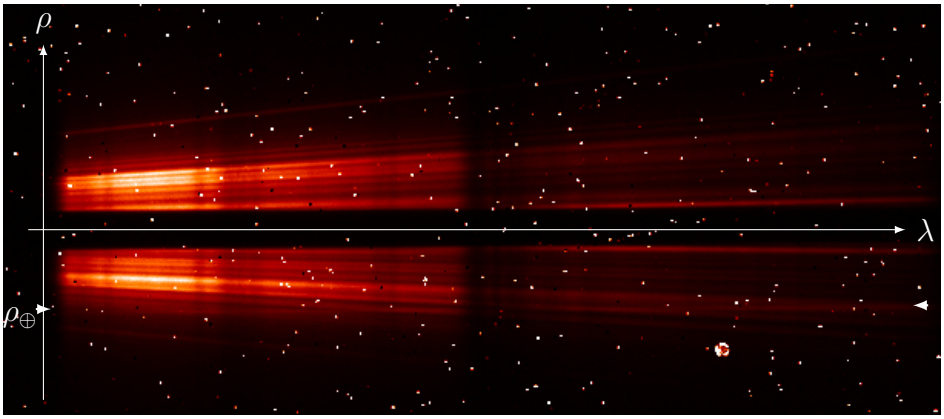

Figure 3.2 – Long-slit spectroscopy data of HR3549 target taken by IRDIS

coordinate (ρ_n, λ_n). Distinguishing the spatial and spectral distribution of the star and companion, the direct model writes:

$$m(\rho,\lambda) = \underbrace{f_\star(\lambda)\, h_\star(\rho,\lambda)}_{\text{star component}} + \overbrace{f_\oplus(\lambda)\, h_\oplus(\rho,\lambda)}^{\text{companion component}} \tag{3.2}$$

where $f_\star$ and $f_\oplus$ are the Spectral Energy Distributions (SEDs) of the star and the companion, while $h_\star$ and $h_\oplus$ represent the on-axis and off-axis PSFs, as presented in Figure 3.3. In other words, the on-axis PSF explains the speckles in the image, the off-axis PSF the companion while their SED explains their spectral evolutions.

The diffraction varying with the wavelength, it can be approximated as a scaling effect on the PSF. Following the work of Perrin et al., 2003, Devaney and Thiébaut, 2017 propose a chromatic model of the PSF, based on a series expansion:

$$h(\rho,\lambda) \approx \sum_{k=1}^{K} \gamma(\lambda)^k\, h_k\big(\gamma(\lambda)\,\rho\big) \tag{3.3}$$

with $\gamma(\lambda) = \lambda_{\text{ref}}/\lambda$ a chromatic magnification factor and where h represents either the on-axis PSF $h_\star$ or the off-axis PSF $h_\oplus$. $\{h_k\}_{k\in[\![1,K]\!]}$ is a family of spatial PSF modes taken such that $h_k(\rho) = h_k(\rho, \lambda_{\text{ref}})$ is the k-th spatial mode at an arbitrary reference wavelength.

Assuming that the companion is located outside the coronagraphic mask, the off-axis PSF can be considered as shift-invariant and centered on the companion's angular separation, $\rho_\oplus$. Given the fact that most of the energy of the companion signal is localized closely to its angular separation, it is possible to approximate:

$$h_\oplus(\rho,\lambda) \approx \gamma(\lambda)\, h_\oplus(\gamma(\lambda)\cdot(\rho-\rho_\oplus)) \tag{3.4}$$

with $h_\oplus(\rho) = h_{\oplus 1}(\rho)$ the first mode of the series expansion. With $f_\oplus$, they form the companion component as visible on the right side of Fig. 3.3. To simplify the model of the star component, we consider in the following that one mode of the series expansion model can explain most of the stellar leakages. Considering a possible small misalignment, $\rho_\star$, along the slit between the center of the coronagraphic mask and the star (pointing errors), the on-axis PSF can be expressed:

$$h_\star(\rho,\lambda) \approx \gamma(\lambda)\, h_\star\big(\gamma(\lambda)\cdot(\rho-\rho_\star)\big) \tag{3.5}$$

with $h_\star(\rho) = h_{\star 1}(\rho)$.

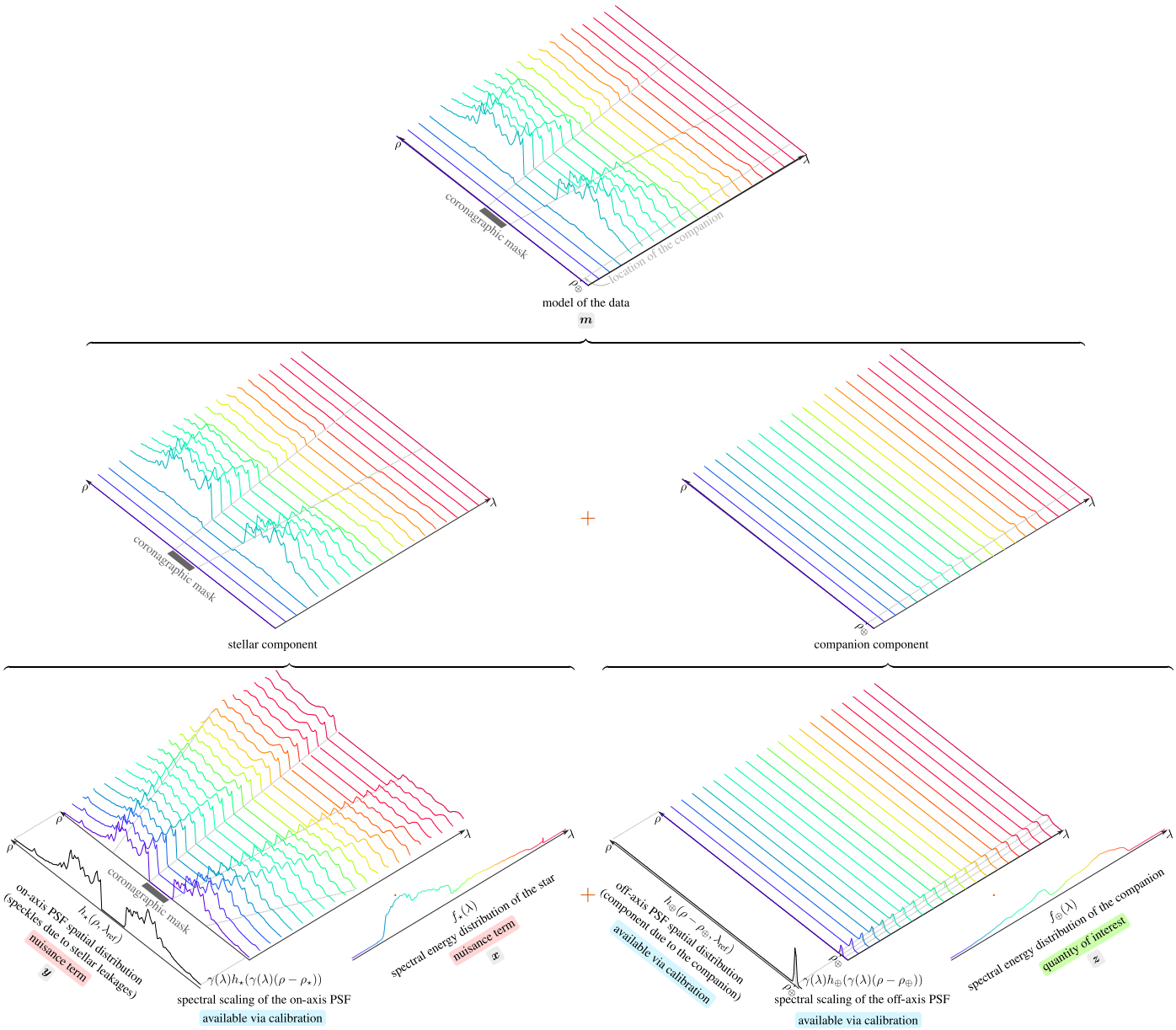

Figure 3.3 – Illustration of the direct model (Eq. (3.2)) as the sum of a stellar component (on the left) and a companion component (on the right).

Thanks to the $\gamma(\lambda)$ factor, the models of Eqs. (3.4) and (3.5) ensure that the off-axis and on-axis PSF are normalized at all wavelengths:

$$\begin{aligned}\int h(\rho,\lambda)\,\mathrm{d}\rho &= \int \gamma(\lambda)\,h(\gamma(\lambda)\,\rho)\,\mathrm{d}\rho \\ &= \int h(s)\,\mathrm{d}s = 1 \end{aligned} \tag{3.6}$$

with $h = h_\star$ or $h_\oplus$ and $s = \gamma(\lambda)\,\rho$ and provided that the PSFs modes are normalized at the reference wavelength λ_{ref}.

3.1.3 Low-rank estimation of the stellar component

As it is preponderant in the data, estimating the stellar component can help extract the companion signal. This Section details the TSVD-based method commonly used to estimate the stellar leakages (Vigan et al., 2008; Mesa, D. et al., 2016).

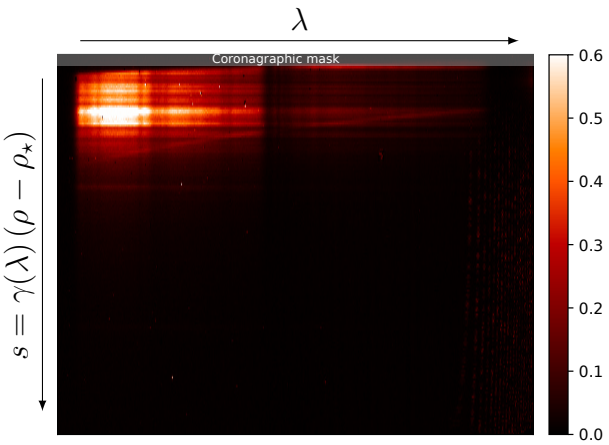

Figure 3.4 – Bottom half of the warped HR 3549 image, corresponding to the side where lies the companion. The companion signal can be seen as a faint curved line.

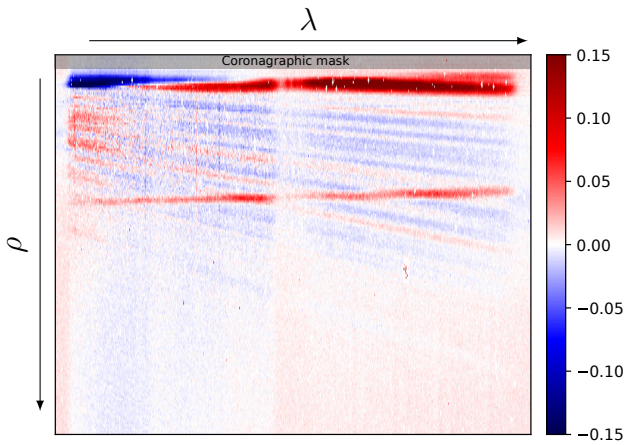

Figure 3.5 – Bottom half of the residual image $\boldsymbol{r}_\oplus$ for the HR 3549 data with stellar leakages estimated by the TSVD method as defined in Eq. 3.12.

By applying a warpping operator to the data, it is possible to form a 2D image $\boldsymbol{d}^{[\mathrm{warp}]} \in \mathbb{R}^{N_1 \times N_2}$ where the first dimension is following the chromatically scaled angular separation of coordinate $s = \gamma(\lambda)\,(\rho - \rho_\star)$ and where the second dimension varies along the wavelength axis (as seen in Fig. 3.4). Each pixel of the resulting image is then modeled by:

$$\begin{aligned} d_{i,j}^{[\mathrm{warp}]} &= m(\rho_\star + s_i^{[\mathrm{warp}]}/\gamma(\lambda_j^{[\mathrm{warp}]}), \lambda_j^{[\mathrm{warp}]}) + b_{i,j}^{[\mathrm{warp}]} \\ &\simeq \sum_{k=1}^{K} \underbrace{\gamma(\lambda_i^{[\mathrm{warp}]})^k\, f_\star(\lambda_i^{[\mathrm{warp}]})}_{\text{spectral component}}\ \underbrace{h_{\star,k}(s_j^{[\mathrm{warp}]})}_{\text{spatial component}} \end{aligned} \tag{3.7}$$

where the approximation of the second line holds considering the noise and the companion component as neglected and taking the expansion series of Eq. (3.3). This is a strong hypothesis but has the advantage of decomposing the warped data into a separable model with a spatial and spectral component following the image grid axes.

From this observation, the commonly used method to estimate the stellar leakages is to use the Singular Value Decomposition (SVD) to describe the warped data:

$$\boldsymbol{d}^{[\mathrm{warp}]} = \sum_{k=1}^{\min(N_1,N_2)} \boldsymbol{u}_k\, \sigma_k\, \boldsymbol{v}_k^\mathsf{T} \tag{3.8}$$

$$\approx \sum_{k=1}^{k_{\max}} \boldsymbol{u}_k\, \sigma_k\, \boldsymbol{v}_k^\mathsf{T} \tag{3.9}$$

where $\boldsymbol{u}_k$ and $\boldsymbol{v}_k$ are the k-th singular vector of the decomposition and σ_k the corresponding singular value. The second line represents the Truncated SVD (TSVD) approximation. Noticing the similitude between Eq. (3.7) and Eq. (3.8), the SVD can be considered as sampling the spectral and spatial components of the warped data respectively via the singular vectors $\boldsymbol{u}_k$ and $\boldsymbol{v}_k$. In other words, it is possible to express the elements of the two singular vectors:

$$u_{k,i} = \beta_k\, \gamma(\lambda_i^{[\mathrm{warp}]})\, f_{\star\,k}(\lambda_i^{[\mathrm{warp}]}) \tag{3.10}$$

$$v_{k,j} = \frac{1}{\beta_k\, \sigma_k}\, h_{\star\,k}(s_j^{[\mathrm{warp}]}) \tag{3.11}$$

where β_k are factors such that the singular vectors $\boldsymbol{u}_k$ are normalized.

From there, the TSVD of $\boldsymbol{d}^{[\text{warp}]}$ should yield a fair approximation of the stellar component in the data, given a suitable maximum number of modes $k_{\max} < \min(N_1, N_2)$. Of course, this approximation holds only for a value of $k_{\max}$ that limits the impact of the companion component and the noise in the resulting model. According to the Eckart–Young–Mirsky theorem (Eckart and Young, 1936; Mirsky, 1960) the TSVD model obtained is the best possible approximation of the warped data of rank $k_{\max}$ in the least squares sense. Commonly used methods extract the companion SED from the residuals formed by the subtraction of the un-warped TSVD model to the data:

$$\boldsymbol{r}_{\oplus} = \boldsymbol{d} - \mathcal{U}\Big(\sum_{k=1}^{k_{\max}} \boldsymbol{u}_k \, \sigma_k \, \boldsymbol{v}_k^{\mathsf{T}}\Big) \tag{3.12}$$

with $\mathcal{U}$ the un-warping operator. Figure 3.5 shows the resulting residuals where the companion component is much more visible and so, easier to extract. Usually, the extraction is done by integrating a small area of the residuals along the companion coordinate.

The TSVD method has however numerous issues in our context. As we need a rectangular warped image that can be interpreted as a matrix by the SVD, some areas of the data need to be discarded, including near the coronagraphic mask and edges in the formed image. Doing so limits the range of acceptable angular distance for the companion, as well as impacts the estimation of the stellar component as the areas discarded can be valuable in the estimation process. As it involves a geometrical transformation of the data, the warping operation done in this method depends greatly on the calibration of the detector. In other words, the relationship between each pixel n and its physical coordinates (ρ_n, λ_n) must be calibrated precisely. It is also important to notice that the TSVD method does not consider the companion component. The resulting model of the stellar component is therefore biased by the companion signal which will result in a negative bias in the residuals $\boldsymbol{r}_{\oplus}$, called *self-subtraction*. Moreover, as it does not consider the noise distribution of the data, the least squares fit performed by the TSVD is sub-optimal. Indeed, as explained in Section 1.1, the distribution of the noise in the data is usually not independent and identically distributed. The interpolation method used to warp the data is also a source of correlation between pixels, thus invalidating the independent hypothesis done by the TSVD method.

3.2 Proposed approach and direct model of the data

3.2.1 Principle of our EXOSPECO method

In this chapter, we propose an alternative approach based on the inverse problems framework. Contrary to state-of-the-art methods[3], the data is not tampered with in an inverse problem approach. Thanks to this framework, the direct model is formulated in the measurement space and consistently discards possible bad pixels in the estimation process. Moreover, it also has the advantage of explicitly taking into account all the components of the data (*e.g.* the companion's signal is not neglected when estimating the speckles component), *i.e.*:

$$\boldsymbol{m} = \boldsymbol{m}_{\star} + \boldsymbol{m}_{\oplus} \tag{3.13}$$

with $\boldsymbol{m}_{\star}$ the star component (speckles) and $\boldsymbol{m}_{\oplus}$ the companion component of the model.

The EXOSPECO algorithm relies on different steps to extract the companion spectrum. Initializing with a star component in our model, we propose to estimate this component by only fitting the part of the data $\boldsymbol{d}$ that is not impacted by the companion. Figure 3.6 presents this first step on the left box. To reduce any self-subtraction bias, an alternated strategy is started with this initial star model $\boldsymbol{m}_{\star[0]}$. At each iteration k (starting with $k = 0$), the companion's component $\boldsymbol{m}_{\oplus[k+1]}$ is estimated

[3] which are more or less based on the TSVD described in Section 3.1.3

from the residuals formed by subtracting the latest star model to the data, *i.e.* $\boldsymbol{r}_{\oplus\,[k]} = \boldsymbol{d} - \boldsymbol{m}_{\star[k]}$ (blue central box in Fig. 3.6). The star component is then updated, yielding $\boldsymbol{m}_{\star[k+1]}$, by fitting the residuals $\boldsymbol{r}_{\star\,[k+1]} = \boldsymbol{d} - \boldsymbol{m}_{\oplus[k+1]}$(orange right box in Fig. 3.6). The two latter steps are repeated and after convergence, we obtain a model of the companion containing its spectrum.

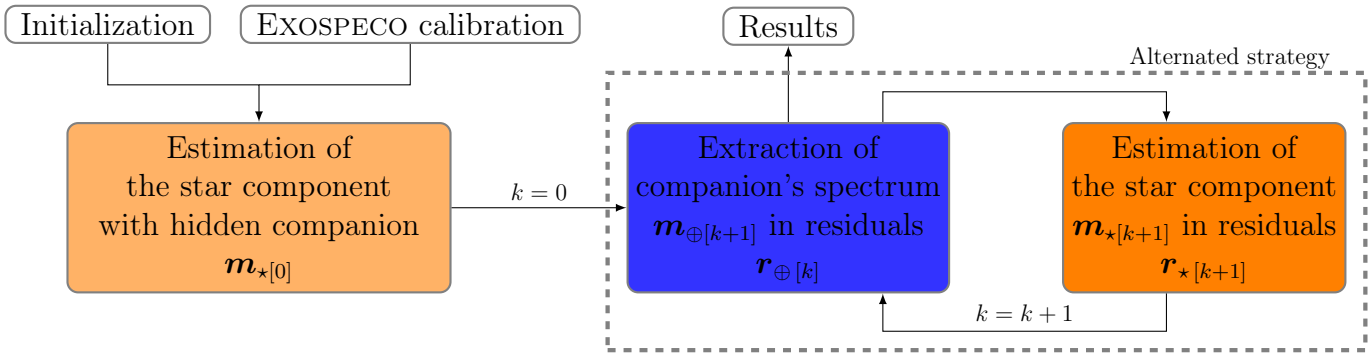

Figure 3.6 – EXOSPECO principle.

3.2.2 Direct model of the data

The light distribution, described in Eq. (3.2), is measured by a detector, meaning that the image plane is spatially sampled. We now describe the construction of the discrete direct model on a finite grid of N elements, or detector's pixels, of coordinates (ρ_n, λ_n).

As the star and companion SEDs, $f_\star$ and $f_\oplus$, and the on-axis PSF, $h_\star$ are unknown, we decide to use linear interpolations on given grids of samples, to model these quantities. This interpolation model will be able to take into account any misalignment of the instrument that could lead to an imperfect sampling of the quantities on the detector (see Section 3.3). For instance, the star SED can be expressed $\forall \lambda \in \mathbb{R}$:

$$f_\star(\lambda) = \sum_{k=1}^{N_{\boldsymbol{x}}} \varphi_\star(\lambda - \lambda_k^\star)\, f_\star(\lambda_k^\star) \tag{3.14}$$

with $\varphi_\star \colon \mathbb{R} \to \mathbb{R}$ an interpolation function and $\boldsymbol{\lambda}^\star \in \mathbb{R}^{N_{\boldsymbol{x}}}$ the interpolation grid of the unknown interpolated star SED, noted $\boldsymbol{x} \in \mathbb{R}^{N_{\boldsymbol{x}}}$. Meaning that for each pixel n of the detector:

$$\begin{aligned} f_\star(\lambda_n) &= \sum_{k=1}^{N_{\boldsymbol{x}}} \overbrace{\varphi_\star(\lambda_n - \lambda_k^\star)}^{F_{\star\,n,k}} \overbrace{f_\star(\lambda_k^\star)}^{x_k} \\ &= \sum_{k=1}^{N_{\boldsymbol{x}}} F_{\star\,n,k}\, x_k = \left(\mathbf{F}_\star\, \boldsymbol{x}\right)_n \end{aligned} \tag{3.15}$$

with $\mathbf{F}_\star \in \mathbb{R}^{N \times N_{\boldsymbol{x}}}$ the spectral linear interpolation operator of the star component. Following the

example of the star SED, we can express the three unknowns for the n-th pixel of the detector:

$$\begin{aligned} f_{\star,n} &= f_\star(\lambda_n) = \sum_{k=1}^{N_{\boldsymbol{x}}} \varphi_\star\big(\lambda_n - \lambda_k^\star\big)\, x_k \\ &= \left(\mathbf{F}_\star\, \boldsymbol{x}\right)_n , \end{aligned} \tag{3.16a}$$

$$\begin{aligned} h_{\star,n} &= h_\star(\gamma_n \cdot (\rho_n - \rho_\star)) = \sum_{k=1}^{N_{\boldsymbol{y}}} \psi_\star\big(\gamma_n \cdot (\rho_n - \rho_\star - \rho_k^\star)\big)\, y_k \\ &= \left(\mathbf{H}_\star\, \boldsymbol{y}\right)_n , \end{aligned} \tag{3.16b}$$

$$\begin{aligned} f_{\oplus,n} &= f_\oplus(\lambda_n) = \sum_{k=1}^{N_{\boldsymbol{z}}} \varphi_\oplus\big(\lambda_n - \lambda_k^\oplus\big)\, z_k \\ &= \left(\mathbf{F}_\oplus\, \boldsymbol{z}\right)_n , \end{aligned} \tag{3.16c}$$

with $\gamma_n = \gamma(\lambda_n)$ and $\psi_\star$ and $\varphi_\oplus$ respectively the spatial interpolation function of the star component and the spectral interpolation function of the companion component and $\boldsymbol{x} \in \mathbb{R}^{N_{\boldsymbol{x}}}$ the star SED, $\boldsymbol{y} \in \mathbb{R}^{N_{\boldsymbol{y}}}$ the on-axes PSF and $\boldsymbol{z} \in \mathbb{R}^{N_{\boldsymbol{z}}}$ the companion SED evenly sampled on their respective interpolation grid $\boldsymbol{\lambda}^\star$, $\boldsymbol{\rho}^\star$ and $\boldsymbol{\lambda}^\oplus$. Note that the spectral sampled grids $\boldsymbol{\lambda}^\star$ and $\boldsymbol{\lambda}^\oplus$ can be chosen equal, but that is not mandatory. In practice, the quantities we want to interpolate can be assumed to be smooth. We choose Catmull and Rom, 1974 cardinal cubic splines as kernels of the interpolation operators.

As for the off-axis PSF (also unknown in our problem), we choose a parametric model. Being shift-invariant and spatially extended on a small portion of the detector, we choose a Gaussian model:

$$h_\oplus\big(\gamma(\lambda)(\rho - \rho_\oplus)\big) = \frac{1}{\sqrt{2\pi}\sigma_\oplus} \exp\left(-\frac{\gamma^2(\lambda)(\rho - \rho_\oplus)^2}{2\sigma_\oplus^2}\right) \tag{3.17}$$

whose single lobe accounts for most of the energy received from the companion. This model has the advantage of being flexible and of only depending on the companion position $\rho_\oplus$ and the PSF standard deviation $\sigma_\oplus$ at the reference wavelength λ_{ref}.

Introducing $\boldsymbol{m} \in \mathbb{R}^N$, $\boldsymbol{\gamma} \in \mathbb{R}^N$ and $\boldsymbol{h}_\oplus \in \mathbb{R}^N$ defined $\forall n \in [\![1, N]\!]$ by:

$$m_n = m(\rho_n, \lambda_n) , \tag{3.18a}$$

$$\gamma_n = \gamma(\lambda_n) = \lambda^{\text{ref}}/\lambda_n , \tag{3.18b}$$

$$h_{\oplus,n} = h_\oplus\big(\gamma_n\,(\rho_n - \rho_\oplus)\big) , \tag{3.18c}$$

the discrete direct model of Eq. (3.2) can be written as:

$$\boxed{\boldsymbol{m}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus) = \boldsymbol{\gamma} \odot \left(\mathbf{H}_\star(\boldsymbol{\nu}_\star)\, \boldsymbol{y}\right) \odot \left(\mathbf{F}_\star\, \boldsymbol{x}\right) + \boldsymbol{\gamma} \odot \boldsymbol{h}_\oplus(\boldsymbol{\nu}_\oplus) \odot \left(\mathbf{F}_\oplus\, \boldsymbol{z}\right)} \tag{3.19}$$

with $\odot$ the Hadamard product, that is the entry-wise multiplication of vectors, $\boldsymbol{\nu}_\oplus = (\rho_\oplus, \sigma_\oplus)$ the parameters of the off-axis PSF, and $\boldsymbol{\nu}_\star$ possible calibration parameters for the stellar leakages like $\rho_\star$ the misalignment between the center of the coronagraphic mask and the location of the star.

To sum up, our model of the data requires estimating the companion SED $\boldsymbol{z}$, the star SED $\boldsymbol{x}$, the on-axis PSF $\boldsymbol{y}$ and the calibration parameters $\boldsymbol{\nu}_\star$ and $\boldsymbol{\nu}_\oplus$.

LSS data are also corrupted by noises, which cannot be neglected. As justified in Section 1.1, these noises contribute to a spurious term that can be approximated by a centered multivariate Gaussian random vector leading to the model of the data:

$$\boldsymbol{d} \sim \mathcal{N}(\boldsymbol{m}, \mathbf{C}) \tag{3.20}$$

with $\boldsymbol{m}$ the expectation of the data, corresponding to the direct model described in Eq. (3.19) and $\mathbf{C}$ the covariance matrix of the noise.

3.3 Geometric calibration: an important step

The model derived in Eq. (3.19) depends on several parameters describing the spatio-spectral distributions. However, to describe precisely the data, the model assumes as known the physical coordinates (ρ_n, λ_n) of each pixel n of the detector. The calibration of the spectral and spatial coordinate dispersion laws is therefore of great importance. In the following, we discuss the estimation of these laws by using the example of the MRS mode of IRDIS on the HR3549 dataset observed on 2015-12-28, knowing that the method is similar in LRS mode.

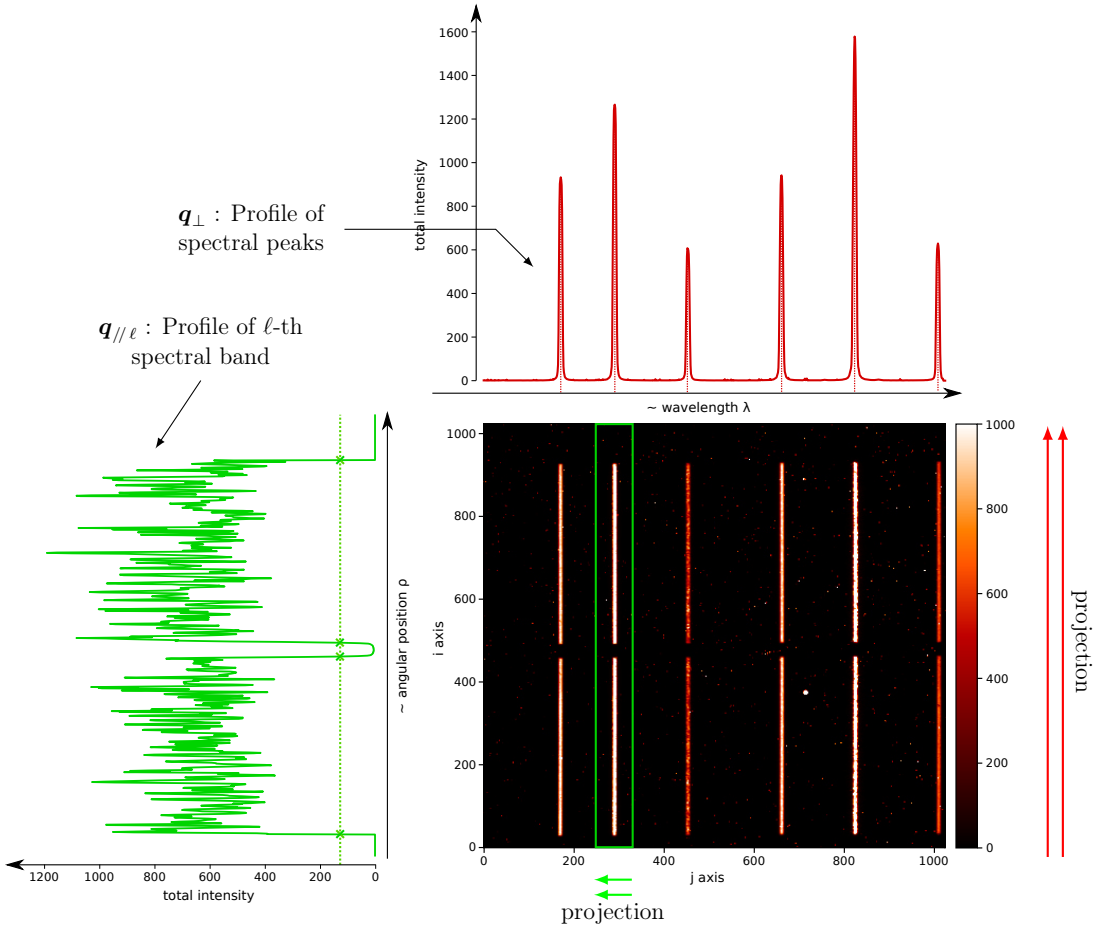

Figure 3.7 – Calibration data and the intermediary results of the geometric calibration process: the profile of the spectral peaks on top and the profile of one spectral band on the left.

3.3.1 Modeling the spectral and spatial dispersions from calibration data

To obtain the spectral and spatial geometric dispersion laws on the image plane, an external calibration procedure is needed. SPHERE/IRDIS calibration provides data where N_λ mono-chromatic laser sources[4] are aligned and illuminate the slit. This gives N_λ bands on the detector, each being divided in two by the coronagraphic mask, as shown in Fig. 3.7. Called in the following $\boldsymbol{d}_{\text{cal}} \in \mathbb{R}^{I\times J}$, $i \in [\![1, I]\!], j \in [\![1, J]\!]$, this calibration image is presumed to have been pre-processed to compensate for bias and non-uniform response of the detector, as discussed in Section 1.1. Bad pixels are supposed identified and represented by a valid pixel mask $\boldsymbol{w}_{\text{msk}}$. Following the detector's Cartesian coordinates (i, j):

$$w_{\text{msk}\ i,j} = \begin{cases} 0 & \text{if } (i,j) \text{ is a bad pixel,} \\ 1 & \text{otherwise.} \end{cases} \tag{3.21}$$

[4] of wavelength 0.9877 μm, 1.1237 μm, 1.3094 μm, 1.5451 μm, 1.73 μm, and 2.015 μm.

We choose to consider two different models of spectral and spatial dispersion laws:

- A standard calibration (sta), considering the spectral and spatial axes as aligned with the axes of the grid of pixels.
- A more complex one (Exospeco), which has more flexibility to account for spatial distortions across the field of view.

These models, giving the physical coordinates of each pixel of the detector, are to be fitted on $\boldsymbol{d}_{\text{cal}}$.

Spectral dispersion law Λ

At pixel $n \sim (i, j)$, it is possible to model the spectral dispersion on the detector as a polynomial of degree P_Λ and of coefficients $\boldsymbol{a}$. The polynomial law allows us to be general while modeling the dispersions with various levels of precision. A first approximation is to consider the wavelength axis to be aligned with the grid of pixels of the detector. This model, used in standard calibration methods, leads to the modeled wavelength value at pixel (i, j):

$$\Lambda_{\text{sta}}(i, j) = \sum_{p=0}^{P_\Lambda} a_p \, j^p \, . \tag{3.22}$$

To take into account local distortions, we propose a more flexible model involving a 2D polynomial:

$$\Lambda_{\text{Exospeco}}(i, j) = \sum_{\substack{p_1, p_2 \\ 0 \leq p_1 + p_2 \leq P_\Lambda}} a_{p_1, p_2} \, i^{p_1} \, j^{p_2} \tag{3.23}$$

Once a set $\mathcal{C}_\ell$ of coordinates (i, j) following the path of each spectral band ℓ has been detected from the calibration data, the coefficients of the polynomial Λ ($= \Lambda_{\text{sta}}$ ou $\Lambda_{\text{Exospeco}}$) can be estimated following a linear least squares fit:

$$\hat{\boldsymbol{a}} = \arg\min_{\boldsymbol{a}} \sum_{\ell=1}^{N_\lambda} \sum_{(i,j) \in \mathcal{C}_\ell} \big(\lambda_\ell - \Lambda(i, j)\big)^2 \tag{3.24}$$

with i and j not necessarily integer coordinates and $\{\lambda_\ell\}_{\ell \in [\![1, N_\lambda]\!]}$ the wavelengths of each spectral band, corresponding to the wavelength of the sources illuminating the slit. This least-squares fit has a closed-form solution for both models as it is quadratic in the parameters, that is the coefficients $\boldsymbol{a}$ of the polynomial.

Spatial dispersion law ϱ

Similarly to our modeling of the spectral dispersion law, it is possible to consider two different models for the spatial dispersion law, based on polynomials of degree P_ϱ and of coefficients $\boldsymbol{s}$. The standard model assuring the alignment between the angular separation axis and the grid of the detector gives the angular separation at the pixel (i, j):

$$\varrho_{\text{sta}}(i, j) = \sum_{p=0}^{P_\varrho} s_p \, i^p \tag{3.25}$$

and the Exospeco model, a 2D-polynomial, which can take into account distortions such as a shear:

$$\varrho_{\text{Exospeco}}(i, j) = \sum_{\substack{p_1, p_2 \\ 0 \leq p_1 + p_2 \leq P_\varrho}} s_{p_1, p_2} \, i^{p_1} \, j^{p_2} \tag{3.26}$$

As can be seen on the calibration image (Fig. 3.7), the estimation of the polynomial law coefficients $\boldsymbol{s}$ can be easily performed by fitting the edges of the coronagraphic mask, of coordinates $(i_\ell^{\text{up}}, j_\ell^{\text{up}}) \in \mathcal{E}_\ell^{\text{up}}$ and $(i_\ell^{\text{down}}, j_\ell^{\text{down}}) \in \mathcal{E}_\ell^{\text{down}}$. Indeed, the edges of the mask are clearly visible in the calibration data where the coronagraphic mask forms a division of size $\Delta\rho$ pixels whatever the wavelength. Given sets of upward and downward edges, we estimate both $\boldsymbol{s}$ and $\Delta\rho$ thanks to the minimization problem:

$$(\hat{\Delta\rho}, \hat{\boldsymbol{s}}) = \underset{\Delta\rho, \boldsymbol{s}}{\arg\min} \left\{ \sum_{\ell=1}^{N_\lambda} \big(\varrho\big(i_\ell^{\text{down}}, j_\ell^{\text{down}}\big) + \Delta\rho/2\big)^2 + \sum_{\ell=1}^{N_\lambda} \big(\varrho(i_\ell^{\text{up}}, j_\ell^{\text{up}}) - \Delta\rho/2\big)^2 \right\} \tag{3.27}$$

which has a closed-form solution, as it is quadratic in $\Delta\rho$ and $\boldsymbol{s}$. Each term represents the coordinates of the edge of the mask (up or down) and their distance to the center.

3.3.2 Detection methods of the spectral bands

From the calibration data $\boldsymbol{d}_{\text{cal}}$, it is possible to deduce the information useful to fit the dispersion laws. More details on the extraction of such information are given in the submitted article Thé et al., 2023, but in a nutshell, the method requires two steps. Firstly, the detection of the paths of each calibrated spectral line in the data $\boldsymbol{d}_{\text{cal}}$. Secondly, the detection of the edges of the coronagraphic mask at each calibrated wavelength.

Detection of the spectral peaks

A weighted projection, approximately along the i axis (accounting for some rotation $\phi_\lambda \simeq 0$ between spectral axes and the pixel grid) is applied to the calibration image $\boldsymbol{d}_{\text{cal}}$ to give a profile $\boldsymbol{q}_\perp$ such as the one shown on top of Fig. 3.7:

$$q_{\perp\, n}(\phi_\lambda) = \frac{\sum_{i,j} \xi_{\perp\, i,j,n}(\phi_\lambda)\, d_{\text{cal},i,j}}{\sum_{i,j} \xi_{\perp\, i,j,n}(\phi_\lambda)}, \tag{3.28a}$$

$$\xi_{\perp\, i,j,n}(\phi_\lambda) = w_{\text{msk},i,j}\, \varphi_{\text{proj}}\big(n - i\, \cos\phi_\lambda - j\, \sin\phi_\lambda\big) \tag{3.28b}$$

where n is the index in the projection and $\varphi_{\text{proj}}\colon \mathbb{R} \to \mathbb{R}$ is introduced to interpolate positions. In practice, we take $\varphi_{\text{proj}}(t) = \max(1 - |t|, 0)$, the linear B-spline, as the interpolating function for projections. Note that, due to the weighting by $\boldsymbol{w}_{\text{msk}}$, defective pixels are discarded in the computed profile. Each peak of the profile $q_{\perp\, n}$ provides an estimate of the average position of the ℓ-th spectral band.

After detecting these peaks, their coordinates are stored in the list $\mathcal{P}(\phi_\lambda)$. To have more precise coordinates of the peaks, an iterative method is done on each line i of the detector to detect the coordinate on the j axis of each peak of the spectral bands at λ_ℓ. To do so, a center of gravity is computed on a small window around $j \in \mathcal{P}(\phi_\lambda)$, for each line. This procedure yields a list $\mathcal{C}_\ell(\phi_\lambda) = \{(i, j)\}$ of coordinates[5] which follows the path of the ℓ-th spectral band.

Detection of the edges of the spectral bands

Knowing the average position $j \in \mathcal{P}(\phi_\lambda)$ of each spectral band ℓ, it is possible to isolate each one of them via a narrow rectangular window $\mathcal{D}_\ell(\phi_\lambda)$ as shown in green in Fig. 3.7:

$$\mathcal{D}_\ell(\phi_\lambda) = \big\{(i, j) \in [\![1, I]\!] \times [\![1, J]\!] \;:\; |\ell - i\, \cos\phi_\lambda - j\, \sin\phi_\lambda| < \delta\big\} \tag{3.29}$$

[5]These coordinates are fractional, being the centers of gravity.

with $\delta \approx 10$ pixels the half-width of the region. Averaging these regions along the j axis, and following the same weighted projection procedure as Eq. 3.28, we obtain a profile $\boldsymbol{q}_{/\!/\,\ell}$ for each line ℓ, as shown on the left of Fig. 3.7:

$$q_{/\!/\,\ell,k}(\phi_\lambda) = \frac{\sum_{(i,j)\in\mathcal{D}_\ell(\phi_\lambda)} \xi_{/\!/\,i,j,k}(\phi_\lambda)\, d_{\mathrm{cal},i,j}}{\sum_{(i,j)\in\mathcal{D}_\ell(\phi_\lambda)} \xi_{/\!/\,i,j,k}(\phi_\lambda)}, \tag{3.30}$$

$$\xi_{/\!/\,i,j,k}(\phi_\lambda) = w_{\mathrm{msk},i,j}\, \varphi_{\mathrm{proj}}(k + i\, \sin\phi_\lambda - j\, \cos\phi_\lambda), \tag{3.31}$$

where k is the index along the projection.

Detecting the edges of the two steps appearing in the profile is quite simple. Each value at index k of the profile is compared with the next one up to a threshold value τ. If $q_{/\!/\,\ell,k} \leqslant \tau \leqslant q_{/\!/\,\ell,k+1}$ and $q_{/\!/\,\ell,k} < q_{/\!/\,\ell,k+1}$, then an ascending edge is detected. If $q_{/\!/\,\ell,k} \geqslant \tau \geqslant q_{/\!/\,\ell,k+1}$ and $q_{/\!/\,\ell,k} > q_{/\!/\,\ell,k+1}$, it is a descending edge. From the four edges detected of the ℓ-th spectral band (crosses in Fig. 3.7), the edges corresponding to the coronagraphic mask are the second and third values. Finding back in $\mathcal{C}_\ell(\phi_\lambda)$ the corresponding index couple (k, j), these two edges are stored in lists $\mathcal{E}^{\mathrm{up}}$ and $\mathcal{E}^{\mathrm{down}}$.

3.3.3 Results of the geometrical calibration

The results of the proposed calibration models are displayed in Fig. 3.8, as blue lines for the spectral law and green lines for the spatial law. In practice, degrees $P_\Lambda = 5$ and $P_\varrho = 1$[6] were chosen. As can be seen on the different zooms shown in the top right of the figure, two-dimensional polynomials are needed to explain local distortions. Choosing this model, we plot on the bottom-right of Fig. 3.8 some iso-wavelength (blue) and iso-angular distance (green) curves, on a zoom-in of the HR3549 dataset, more precisely under the coronagraphic mask. This figure highlights how well our proposed models for the dispersion laws are following the speckles, compared to the standard model. A strong shear effect due to the dispersive elements is visible and taken into account by our model.

Using the two calibration models considered here, we applied the EXOSPECO algorithm (explained in section 3.2.1 and detailed in section 3.4) on the HR3549 dataset previously introduced. Fig. 3.9 shows the residuals, that is the difference between the data and the estimated model of both the star and the companion component. A perfect result would then be to only see the noise, without any residual structures from the stellar leakages or the companion. As can be seen, with the standard calibration model, the resulting residuals present still high levels of artifacts coming from the subtraction of the data by the model of the stellar component. The alternating red and blue features, corresponding to positive and negative values, are the sign of a misalignment between the speckles in the data with our model. Because the standard calibration does not account for the shear effect along the wavelength axis, the extraction of the companion is corrupted by some residuals of the speckles. Contrary to the standard calibration, the EXOSPECO calibration reduces greatly the stellar residuals. The shear effect being taken into account, there are no visible traces of the companion in the residuals which indicates a better extraction of its SED. These results show how critical the geometric calibration step is.

3.4 EXOSPECO EXtraction Of SPEctrum of COmpanion

3.4.1 Defining the objective function

With the geometric calibration giving the coordinates (ρ_n, λ_n) of the n pixel $\forall n \in [\![1, N]\!]$, the direct model of the light distribution in Eq. (3.19), and the model of the data in Eq. (3.20), it is possible to

[6]Note that the small number of points for ρ (edges of the coronagraphic mask) limits the possible degree of the polynomial spatial law.

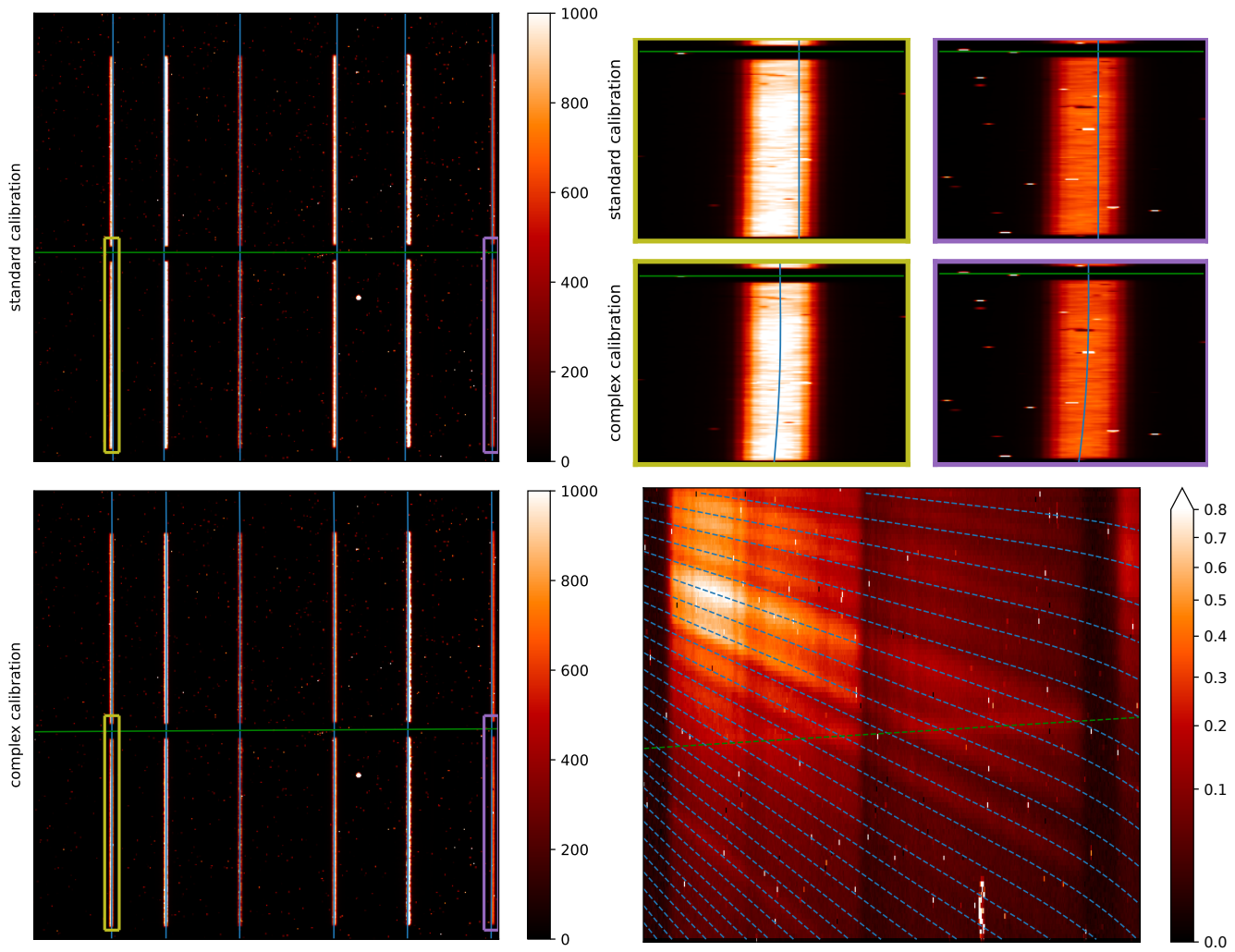

Figure 3.8 – First column: Iso-wavelengths (blue) at the calibrated wavelength and iso-angular separations (green) of the center of the coronagraphic mask, shown on the calibration data of the HR3549 dataset for the standard (top) and proposed model (bottom); Top-right: Zooms of two regions of the figures in the first columns (in yellow and purple) with iso-wavelength and iso-angular separation curves; Bottom-right: Iso-wavelength (blue) and iso-angular distance (green) superposed to the bottom part of HR3549 LSS data under the coronagraphic mask.

follow a maximum *a posteriori* approach (Section 1.3.4) and write the criterion to be minimized in the parameters:

$$\mathcal{C}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{\nu}_\star,\boldsymbol{\nu}_\oplus;\boldsymbol{\mu}) = \overbrace{\left\|\boldsymbol{d}-\boldsymbol{m}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{\nu}_\star,\boldsymbol{\nu}_\oplus)\right\|^2_{\mathbf{W}}}^{\text{data fidelity}} + \underbrace{\mathcal{R}_{\boldsymbol{x}\,\boldsymbol{y}\,\boldsymbol{z}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z};\boldsymbol{\mu})}_{\textit{a priori}} \tag{3.32}$$

with the first term of the criterion measuring the discrepancies between the data and the model. $\mathcal{R}_{\boldsymbol{x}\,\boldsymbol{y}\,\boldsymbol{z}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z};\boldsymbol{\mu})$ is a regularization term imposing priors on $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ and is tuned by hyper-parameters $\boldsymbol{\mu}$. The precision matrix of the data $\mathbf{W}$ is given by a diagonal matrix, *i.e.* $\mathbf{W} = \mathrm{diag}(\boldsymbol{w})$ with:

$$w_n = \begin{cases} 0 & \text{if } n\text{-th pixel is invalid,} \\ 1/\mathrm{Var}(d_n) & \text{otherwise} \end{cases} \tag{3.33}$$

$\mathrm{Var}(d_n)$ denotes the variance of the n-th pixel of the data. Invalid pixels are defined as pixels that do not follow the direct model of Eq. (3.19). This includes the bad pixels detected and flagged in $\boldsymbol{w}_{\mathrm{msk}}$

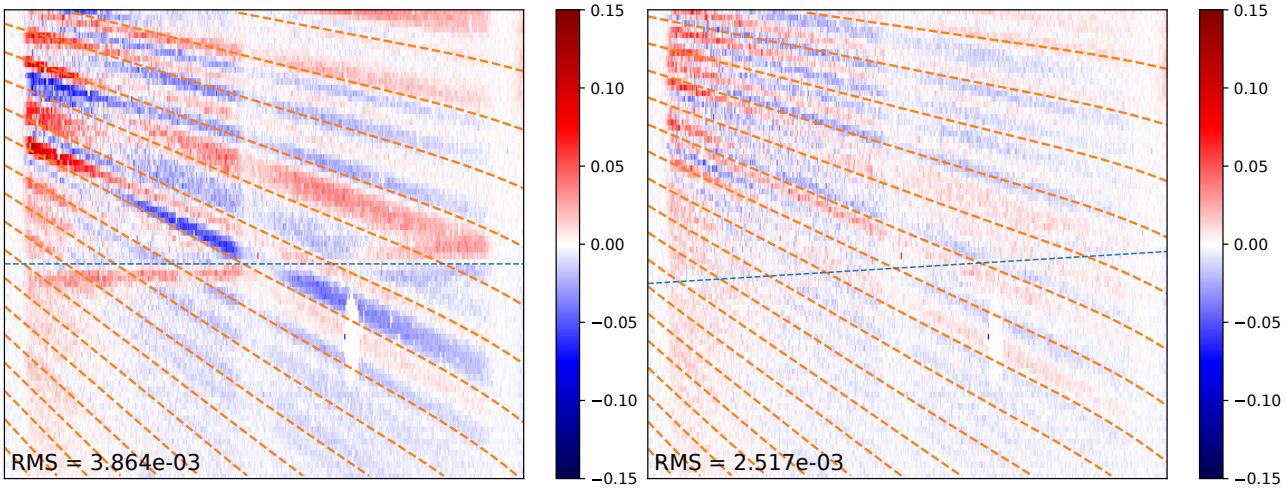

Figure 3.9 – Bottom part of the residuals between the HR3549 scientific data and the direct model with estimated parameters, assuming a *standard* calibration model (left) and the one proposed in EXOSPECO (right). The RMS values of the residuals are given for each model. The iso-levels of ρ/λ which are approximately followed by the dispersed stellar speckles are plotted as orange dashed lines. The position of the companion at the different wavelengths (*i.e.*, at $\rho = \rho_\oplus$) is plotted as a blue dashed line.

(see Eq. (3.21)), pixels outside the field of view and pixels close to the coronagraphic mask that are not well described by the model (see Section 3.5.3).

Estimating the parameters of the direct model amounts to jointly minimizing the criterion (3.32):

$$(\hat{\boldsymbol{x}}(\boldsymbol{\mu}), \hat{\boldsymbol{y}}(\boldsymbol{\mu}), \hat{\boldsymbol{z}}(\boldsymbol{\mu}), \hat{\boldsymbol{\nu}}_\star(\boldsymbol{\mu}), \hat{\boldsymbol{\nu}}_\oplus(\boldsymbol{\mu})) = \underset{\substack{\boldsymbol{x}\geq 0,\, \boldsymbol{y}\geq 0,\, \boldsymbol{z}\geq 0,\\ \boldsymbol{\nu}_\star\in\Omega_\star,\, \boldsymbol{\nu}_\oplus\in\Omega_\oplus}}{\arg\min} \mathcal{C}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus; \boldsymbol{\mu}). \tag{3.34}$$

with $\Omega_\star$ and $\Omega_\oplus$ sets that constrain the possible values of $\boldsymbol{\nu}_\star$ and $\boldsymbol{\nu}_\oplus$, *e.g.* based on physical considerations. The positivity constraints apply to the parameters $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$, as they represent non-negative physical quantities. Note that all the estimators depend on the vector of hyper-parameters $\boldsymbol{\mu}$ which plays an important role in the quality of the estimation.

The star and companion's SEDs, $\boldsymbol{x}$ and $\boldsymbol{z}$ and the on-axis PSF $\boldsymbol{y}$ can be considered mutually independent. Hence the regularization term can be decomposed into the sum of three terms:

$$\mathcal{R}_{\boldsymbol{x}\,\boldsymbol{y}\,\boldsymbol{z}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}; \boldsymbol{\mu}) = \mu_{\boldsymbol{x}}\, \mathcal{R}_{\boldsymbol{x}}(\boldsymbol{x}) + \mu_{\boldsymbol{y}}\, \mathcal{R}_{\boldsymbol{y}}(\boldsymbol{y}) + \mu_{\boldsymbol{z}}\, \mathcal{R}_{\boldsymbol{z}}(\boldsymbol{z}) \tag{3.35}$$

with the vector of hyper-parameters $\boldsymbol{\mu} = (\mu_{\boldsymbol{x}}, \mu_{\boldsymbol{y}}, \mu_{\boldsymbol{z}})$.

We consider *a priori* that the interpolated distributions of $f_\star$, $h_\star$ and $f_\oplus$ are smooth and so are their uniformly sampled versions $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$. We choose Tikhonov's quadratic regularization (Tikhonov and Arsenin, 1977) to favor this smoothness:

$$\mathcal{R}_{\boldsymbol{u}}(\boldsymbol{u}) = \|\mathbf{D}\,\boldsymbol{u}\|_2^2 = \sum_{j=1}^{N_{\boldsymbol{u}}-1} \left((\mathbf{D}\,\boldsymbol{u})_j\right)^2, \tag{3.36}$$

with $N_{\boldsymbol{u}}$ the number of entries in $\boldsymbol{u} = \boldsymbol{x}$, $\boldsymbol{y}$ or $\boldsymbol{z}$ and $\mathbf{D} \in \mathbb{R}^{(N_{\boldsymbol{u}}-1)\times N_{\boldsymbol{u}}}$ a finite difference operator defined by:

$$(\mathbf{D}\,\boldsymbol{u})_j = \begin{cases} u_{j+1} - u_j & \text{if } 1 \leq j < N_{\boldsymbol{u}} - 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.37}$$

Once the criterion is defined, the next step is to find a minimization strategy.

3.4.2 Alternated strategy

Directly solving the problem defined in Eq. (3.34) is difficult. As seen before, in Chapters 1 and 2, the alternated estimation of each component $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{\nu}_\star$, $\boldsymbol{z}$ and $\boldsymbol{\nu}_\oplus$ is a much simpler strategy. Each sub-problem, easier to implement, is then expressed by:

$$\begin{aligned}
\hat{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{r}_\star, \boldsymbol{\nu}_\star; \mu_{\boldsymbol{x}}) &= \underset{\boldsymbol{x}\geq 0}{\arg\min}\, \mathcal{C}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus; \boldsymbol{\mu}) \\
&= \underset{\boldsymbol{x}\geq 0}{\arg\min} \left\{ \|\mathbf{A}_\star\, \boldsymbol{x} - \boldsymbol{r}_\star\|^2_{\mathbf{W}} + \mu_{\boldsymbol{x}}\, \mathcal{R}_{\boldsymbol{x}}(\boldsymbol{x}) \right\}, && \text{(3.38a)} \\
\hat{\boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{r}_\star, \boldsymbol{\nu}_\star; \mu_{\boldsymbol{y}}) &= \underset{\boldsymbol{y}\geq 0}{\arg\min}\, \mathcal{C}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus; \boldsymbol{\mu}) \\
&= \underset{\boldsymbol{y}\geq 0}{\arg\min} \left\{ \|\mathbf{B}_\star\, \boldsymbol{y} - \boldsymbol{r}_\star\|^2_{\mathbf{W}} + \mu_{\boldsymbol{y}}\, \mathcal{R}_{\boldsymbol{y}}(\boldsymbol{y}) \right\}, && \text{(3.38b)} \\
\hat{\boldsymbol{\nu}}_\star(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{r}_\star) &= \underset{\boldsymbol{\nu}_\star \in \Omega_\star}{\arg\min}\, \mathcal{C}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus; \boldsymbol{\mu}) \\
&= \underset{\boldsymbol{\nu}_\star \in \Omega_\star}{\arg\min} \|\boldsymbol{m}_\star(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\nu}_\star) - \boldsymbol{r}_\star\|^2_{\mathbf{W}}, && \text{(3.38c)} \\
\hat{\boldsymbol{z}}(\boldsymbol{r}_\oplus, \boldsymbol{\nu}_\oplus; \mu_{\boldsymbol{z}}) &= \underset{\boldsymbol{z}\geq 0}{\arg\min}\, \mathcal{C}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus; \boldsymbol{\mu}) \\
&= \underset{\boldsymbol{z}\geq 0}{\arg\min} \left\{ \|\mathbf{A}_\oplus\, \boldsymbol{z} - \boldsymbol{r}_\oplus\|^2_{\mathbf{W}} + \mu_{\boldsymbol{z}}\, \mathcal{R}_{\boldsymbol{z}}(\boldsymbol{z}) \right\}, && \text{(3.38d)} \\
\hat{\boldsymbol{\nu}}_\oplus(\boldsymbol{z}, \boldsymbol{r}_\oplus) &= \underset{\boldsymbol{\nu}_\oplus \in \Omega_\oplus}{\arg\min}\, \mathcal{C}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus; \boldsymbol{\mu}) \\
&= \underset{\boldsymbol{\nu}_\oplus \in \Omega_\oplus}{\arg\min} \|\boldsymbol{m}_\oplus(\boldsymbol{z}, \boldsymbol{\nu}_\oplus) - \boldsymbol{r}_\oplus\|^2_{\mathbf{W}}, && \text{(3.38e)}
\end{aligned}$$

with:

$$\begin{aligned}
\forall \boldsymbol{x},\ \mathbf{A}_\star\, \boldsymbol{x} &= \boldsymbol{\gamma} \odot \left(\mathbf{H}_\star(\boldsymbol{\nu}_\star)\, \boldsymbol{y}\right) \odot \left(\mathbf{F}_\star\, \boldsymbol{x}\right), && \text{(3.39a)} \\
\forall \boldsymbol{y},\ \mathbf{B}_\star\, \boldsymbol{y} &= \boldsymbol{\gamma} \odot \left(\mathbf{F}_\star\, \boldsymbol{x}\right) \odot \left(\mathbf{H}_\star(\boldsymbol{\nu}_\star)\, \boldsymbol{y}\right), && \text{(3.39b)} \\
\forall \boldsymbol{z},\ \mathbf{A}_\oplus\, \boldsymbol{z} &= \boldsymbol{\gamma} \odot \left(\boldsymbol{h}_\oplus(\boldsymbol{\nu}_\oplus)\right) \odot \left(\mathbf{F}_\oplus\, \boldsymbol{z}\right), && \text{(3.39c)} \\
\boldsymbol{r}_\star &= \boldsymbol{d} - \boldsymbol{m}_\oplus(\boldsymbol{z}, \boldsymbol{\nu}_\oplus), && \text{(3.39d)} \\
\boldsymbol{r}_\oplus &= \boldsymbol{d} - \boldsymbol{m}_\star(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\nu}_\star), && \text{(3.39e)} \\
\boldsymbol{m}_\star(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\nu}_\star) &= \boldsymbol{\gamma} \odot (\mathbf{H}_\star(\boldsymbol{\nu}_\star)\, \boldsymbol{y}) \odot (\mathbf{F}_\star\, \boldsymbol{x}) = \mathbf{A}_\star\, \boldsymbol{x} = \mathbf{B}_\star\, \boldsymbol{y} && \text{(3.39f)} \\
\boldsymbol{m}_\oplus(\boldsymbol{z}, \boldsymbol{\nu}_\oplus) &= \boldsymbol{\gamma} \odot \boldsymbol{h}_\oplus(\boldsymbol{\nu}_\oplus) \odot (\mathbf{F}_\oplus\, \boldsymbol{z}) = \mathbf{A}_\oplus\, \boldsymbol{z}. && \text{(3.39g)}
\end{aligned}$$

where $\boldsymbol{m}_\star(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\nu}_\star)$ and $\boldsymbol{m}_\oplus(\boldsymbol{z}, \boldsymbol{\nu}_\oplus)$ are the respective contributions of the star and of the companion in the data. The model of the stellar leakages being bi-linear, the criterion to be minimized is non-quadratic in the parameters $\boldsymbol{x}$ and $\boldsymbol{y}$. Using the AMORS algorithm developed in Chapter 2, we propose in the following a strategy that will iteratively estimate the stellar contribution given the companion component, using AMORS, and the companion parameters given the star component, as explained in Fig. 3.6.

Note that the chosen regularizations are convex, and so sub-problems in Eq. (3.38a), (3.38b) and (3.38d), are themselves convex and thus admit a unique solution, provided that $\mathbf{A}_\star^\mathsf{T}\, \mathbf{W}\, \mathbf{A}_\star \neq \mathbf{0}$, $\mathbf{B}_\star^\mathsf{T}\, \mathbf{W}\, \mathbf{B}_\star \neq \mathbf{0}$ and $\mathbf{A}_\oplus^\mathsf{T}\, \mathbf{W}\, \mathbf{A}_\oplus \neq \mathbf{0}$. Without the positivity constraints, estimating $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ would have closed-form solutions. Due to the positivity constraint, we use the quasi-Newton method VMLM-B (Thiébaut, 2002) which implements boundary constraints. However, as the global problem is not jointly convex, the final results of our alternated strategy may be only a local minimum of the criterion. Sub-problems in Eq. (3.38c) and (3.38e) can be constrained by feasible sets $\Omega_\star$ and $\Omega_\oplus$. As these particular problems are non-linear with just a few parameters (*e.g.* $\rho_\star$, $\rho_\oplus$ and $\sigma_\oplus$), we use a derivative-free method such as Brent's method (Brent, 2013), for fitting one parameter, or Newuoa algorithm (Powell, 2006), for fitting multiple parameters (see Section 1.5.3).

3.4.3 Estimation of the stellar component

Considering LSS data without any companion, that is $\boldsymbol{z} = \mathbf{0}$, estimating the stellar parameters amounts to solving Eq. (3.38a), (3.38b) and (3.38c) at each iteration. With a companion, instead of fitting the LSS data $\boldsymbol{d}$, the estimation of the stellar parameters can be done on the residuals $\boldsymbol{r}_\star = \boldsymbol{d} - \boldsymbol{m}_\oplus$.

Looking at the stellar component in more details, it is possible to recognize a bi-linear model in $\boldsymbol{x}$ and $\boldsymbol{y}$, as defined in Section 2.2.1, by Eq. 2.7. Since the regularization chosen in Eq. (3.36) is homogeneous for both $\boldsymbol{x}$ and $\boldsymbol{y}$ of degree $q = 2$ and $r = 2$ respectively, our AMORS algorithm (Alg. 2.1) is an obvious choice to solve the problem. Algorithm 3.1, FITSTAR, takes back the ideas of AMORS, with an additional *autocalibration* step corresponding to the resolution of Eq. (3.38c).

Algorithm 3.1: FITSTAR: Estimation of the stellar component

Input: $\boldsymbol{r}_\star \in \mathbb{R}^N$, $\mathbf{W} \in \mathbb{R}^{N\times N}$, $\boldsymbol{x}_{[0]} \in \mathbb{R}^{N_{\boldsymbol{x}}}$, $\boldsymbol{\nu}_{\star[0]}$, $\mu_{\boldsymbol{x}} > 0$, and $\alpha_{[0]} > 0$.
Output: $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{\nu}}_\star$ a local minimum of $\mathcal{C}$ in $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{\nu}_\star$.

$\mu_{\boldsymbol{y}} \leftarrow 1$
$k \leftarrow 0$
while *not converged* **do**
  while true do
    ▶ Update on-axis PSF
    $\boldsymbol{y}_{[k+1]} \leftarrow \hat{\boldsymbol{y}}\big(\boldsymbol{x}_{[k]}, \boldsymbol{r}_\star, \boldsymbol{\nu}_{\star[k]}; \alpha_{[k]}^{-r}\,\mu_{\boldsymbol{y}}\big)$ ◁ Eq. (3.38b)
    $\alpha_{[k+1/2]} \leftarrow \hat{\alpha}\big(\boldsymbol{x}_{[k]}, \boldsymbol{y}_{[k+1]}; \mu_{\boldsymbol{x}}, \mu_{\boldsymbol{y}}\big)$ ◁ Eq. 2.15
    if $k > 0$ **or** $\alpha_{[k+1/2]} \approx \alpha_{[k]}$ **then break**
    $\alpha_{[k]} \leftarrow \alpha_{[k+1/2]}$
  ▶ Update star SED
  $\boldsymbol{x}_{[k+1]} \leftarrow \hat{\boldsymbol{x}}\big(\boldsymbol{y}_{[k+1]}, \boldsymbol{r}_\star, \boldsymbol{\nu}_{\star[k]}; \alpha_{[k+1/2]}^{q}\,\mu_{\boldsymbol{x}}\big)$ ◁ Eq. (3.38a)
  $\alpha_{[k+1]} \leftarrow \hat{\alpha}\big(\boldsymbol{x}_{[k+1]}, \boldsymbol{y}_{[k+1]}; \mu_{\boldsymbol{x}}, \mu_{\boldsymbol{y}}\big)$ ◁ Eq. 2.15
  ▶ Auto-calibration (optional)
  $\boldsymbol{\nu}_{\star[k+1]} \leftarrow \hat{\boldsymbol{\nu}}_\star\big(\boldsymbol{x}_{[k+1]}, \boldsymbol{y}_{[k+1]}, \boldsymbol{r}_\star\big)$ ◁ Eq. (3.38c)
  $k \leftarrow k + 1$
$\hat{\boldsymbol{x}} \leftarrow \alpha_{[k]}\,\boldsymbol{x}_{[k]}$
$\hat{\boldsymbol{y}} \leftarrow \boldsymbol{y}_{[k]}/\alpha_{[k]}$
$\hat{\boldsymbol{\nu}}_\star \leftarrow \boldsymbol{\nu}_{\star[k]}$

Remark that the initial star's SED $\boldsymbol{x}_{[0]}$ must be nonnegative and its regularization strictly positive, according to Eq. (2.15). Dedicated calibration is usually given with the scientific dataset, obtained by shifting the star from the center of the coronagraphic mask[7]. Extracting the initial star's SED $\boldsymbol{x}_{[0]}$ can then be achieved by applying Alg. 3.2. If no calibration is provided, a simple weighted mean can give a satisfactory starting point:

$$\forall j \in [\![1, N_{\boldsymbol{x}}]\!] : \quad x_{j[0]} = \frac{\sum_{n\in\mathcal{X}_j} w_n\, d_n}{\sum_{n\in\mathcal{X}_j} w_n} \tag{3.40}$$

with w_n the n-th diagonal term of the precision matrix defined in Eq. (3.33) and:

$$\mathcal{X}_j = \Big\{ n \in [\![1, N]\!] \;\Big|\; |\lambda_{\star,j} - \lambda_n| = \min_{j'\in[\![1,N_{\boldsymbol{x}}]\!]} |\lambda_{\star,j'} - \lambda_n| \Big\} \tag{3.41}$$

the set of pixels whose nearest wavelength in the model grid is the j-th one.

[7] This is done by applying a simple tilt on the deformable mirror of the ExAO.

Because we do not know the companion component, it is possible to use Alg. 3.1 with a modified precision matrix, where the pixels of the data impacted the most by the companion contribution are set to zero. The resulting algorithm 3.1 implements the left side and center steps of Fig. 3.6.

3.4.4 Extraction of companion's spectrum in the residuals

Given a stellar model (obtained by FITSTAR (Alg. 3.1) for instance), we extract the companion's SED by fitting the model $\boldsymbol{m}_{\oplus}$ in the residuals $\boldsymbol{r}_{\oplus} = \boldsymbol{d} - \boldsymbol{m}_{\star}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\nu}_{\star})$. Estimating $\boldsymbol{z}$ and $\boldsymbol{\nu}_{\oplus}$ is done by solving Eq. (3.38d) and (3.38e). Alg. 3.2, called FITCOMPANION, applies the alternated strategy to these two problems by iteratively estimating the companion's SED $\boldsymbol{z}$ and refining the off-axis PSF parameters $\boldsymbol{\nu}_{\oplus}$ (right step in Fig. 3.6).

Algorithm 3.2: FITCOMPANION: Estimation of the companion parameters

Input: residual data $\boldsymbol{r}_{\oplus} = \boldsymbol{d} - \boldsymbol{m}_{\star} \in \mathbb{R}^N$, precision matrix $\mathbf{W}$, initial off-axis PSF parameters $\boldsymbol{\nu}_{\oplus[0]}$, $\mu_{\boldsymbol{z}} > 0$.

Output: $\hat{\boldsymbol{z}}$ and $\hat{\boldsymbol{\nu}}_{\oplus}$, a local minimum of $\mathcal{C}$ in $\boldsymbol{z}$ and $\boldsymbol{\nu}_{\oplus}$ given the residuals $\boldsymbol{r}_{\oplus}$.

$k \leftarrow 0$

while *not converged* **do**

- ▶ Update companion SED
- $\boldsymbol{z}_{[k+1]} \leftarrow \hat{\boldsymbol{z}}\left(\boldsymbol{r}_{\oplus}, \boldsymbol{\nu}_{\oplus[k]}, \mu_{\boldsymbol{z}}\right)$ ◁ Eq. (3.38d)
- ▶ Update off-axis PSF (optional)
- $\boldsymbol{\nu}_{\oplus[k+1]} \leftarrow \hat{\boldsymbol{\nu}}_{\oplus}\left(\boldsymbol{z}_{[k+1]}, \boldsymbol{r}_{\oplus}\right)$ ◁ Eq. (3.38e)
- $k \leftarrow k + 1$

$\hat{\boldsymbol{z}} \leftarrow \boldsymbol{z}_{[k]}$

$\hat{\boldsymbol{\nu}}_{\oplus} \leftarrow \boldsymbol{\nu}_{\oplus[k]}$

Note that our FITCOMPANION algorithm requires as input an initial vector of off-axis PSF parameters $\boldsymbol{\nu}_{\oplus}$. This is usually not a problem as we already have a good knowledge of where the companion is in the slit. Other parameters can be estimated from calibration data as the one used to find the initial star's SED $\boldsymbol{x}_{[0]}$. Moreover, Algorithm 3.2 also has an *auto-calibration* step to refine the PSF parameters directly fitting the data.

3.4.5 The EXOSPECO algorithm

Finally, using both the stellar estimation done by FITSTAR (Alg. 3.1) and the companion estimation done by FITCOMPANION (Alg. 3.2) we can define the main Algorithm, EXOSPECO (Alg. 3.3), following the scheme of Fig. 3.6. With EXOSPECO, we can un-mix the companion component from the star component and thus extract the companion's SED. The algorithm is iteratively updating the stellar model and the companion model, taking into account the estimations from the previous iteration. This strategy has the advantage of considerably reducing the self-subtraction bias that would occur without repeating these steps (see section 3.5.2).

Taking the schematic behavior of EXOSPECO shown in Fig. 3.6, the two algorithms FITSTAR and FITCOMPANION describe the center and rightmost boxes. As the companion's SED $\boldsymbol{z}$ is not known at the beginning, we start the method by hiding the companion, giving the leftmost box of Fig. 3.6. To do so, EXOSPECO is initialized with a masked precision matrix in which elements around the companion's angular separation initially estimated, $\rho_{\oplus[0]}$, are set to zero:

$$\mathbf{W}_{\star} = \mathrm{diag}(\boldsymbol{w}_{\star}), \tag{3.42}$$

Algorithm 3.3: EXOSPECO algorithm

Input: The data $\boldsymbol{d}$ and its precision matrix $\mathbf{W}$, the masked precision matrix $\mathbf{W}_\star$, initial estimates $\boldsymbol{x}_{[0]}$, $\boldsymbol{\nu}_{\star[0]}$, and $\boldsymbol{\nu}_{\oplus[0]}$, and hyper-parameters $\mu_{\boldsymbol{x}} > 0$ and $\mu_{\boldsymbol{z}} > 0$.

Output: $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{z}}$, $\hat{\boldsymbol{\nu}}_\star$, and $\hat{\boldsymbol{\nu}}_\oplus$ a local minimum of $\mathcal{C}$.

$\mu_y = 1$
$\boldsymbol{z}_{[0]} = \mathbf{0}$
$\hat{\alpha} = 1$
$k = 0$
while *not converged* **do**
  if $k = 0$ **then**
    ▶ Hide companion
    $\mathbf{W}' = \mathbf{W}_\star$ ◁ Eq. (3.42)
  else
    ▶ Account for companion
    $\mathbf{W}' = \mathbf{W}$
  ▶ Update star leakage model
  $\boldsymbol{r}_{\star[k]} = \boldsymbol{d} - \boldsymbol{m}_\oplus\big(\boldsymbol{z}_{[k]}, \boldsymbol{\nu}_{\oplus[k]}\big)$ ◁ Eq. (3.39d)
  $\big(\boldsymbol{x}_{[k+1]}, \boldsymbol{y}_{[k+1]}, \boldsymbol{\nu}_{\star[k+1]}\big) = \texttt{FitStar}\big(\boldsymbol{r}_{\star[k]}, \mathbf{W}', \boldsymbol{x}_{[k]}, \boldsymbol{\nu}_{\star[k]}, \mu_{\boldsymbol{x}}, \mu_{\boldsymbol{y}}, \hat{\alpha}\big)$
  ▶ Update companion model
  $\boldsymbol{r}_{\oplus[k+1]} = \boldsymbol{d} - \boldsymbol{m}_\star\big(\boldsymbol{x}_{[k+1]}, \boldsymbol{y}_{[k+1]}, \boldsymbol{\nu}_{\star[k+1]}\big)$ ◁ Eq. (3.39e)
  $\big(\boldsymbol{z}_{[k+1]}, \boldsymbol{\nu}_{\oplus[k+1]}\big) = \texttt{FitCompanion}\big(\boldsymbol{r}_{\oplus[k+1]}, \mathbf{W}, \boldsymbol{\nu}_{\oplus[k]}, \mu_{\boldsymbol{z}}\big)$
  $k \leftarrow k + 1$
$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}, \hat{\boldsymbol{\nu}}_\star, \hat{\boldsymbol{\nu}}_\oplus) \leftarrow \big(\boldsymbol{x}_{[k]}, \boldsymbol{y}_{[k]}, \boldsymbol{z}_{[k]}, \boldsymbol{\nu}_{\star[k]}, \boldsymbol{\nu}_{\oplus[k]}\big)$

with $\forall n \in [\![1, N]\!]$:

$$w_{\star,n} = \begin{cases} 0 & \text{if } \gamma_n \left|\rho_n - \rho_\oplus\right| \leq \tau \\ w_n & \text{otherwise} \end{cases} \tag{3.43}$$

where ρ_n and $\gamma_n = \lambda_{\text{ref}}/\lambda_n$ are the angular separation and magnification factor corresponding to n-th pixel and τ is the angular half-width of the mask at a reference wavelength. In practice, we choose the latter as being 2 or 3 times the initial angular standard deviation $\sigma_{\oplus[0]}$ of the off-axis PSF.

Note that Alg. 3.3 only requires tuning properly two hyper-parameters: the hyper-parameter tuning the companion's SED regularization, $\mu_{\boldsymbol{z}}$, and one of the two star's regularizations, $\mu_{\boldsymbol{x}}$ or $\mu_{\boldsymbol{y}}$. The remaining hyper-parameter can be set to an arbitrary value (= 1 for instance). The auto-calibration steps estimating potential misalignment between star and coronagraphic mask, $\boldsymbol{\nu}_\star$, and off-axis PSF parameters $\boldsymbol{\nu}_\oplus$ are non-linear. Solving the self-calibration steps can therefore lead to a poor local minimum. In practice, the corresponding steps in Alg. 3.1 and 3.2 can be deactivated for the first few iterations of Alg. 3.3. In other words, the self-calibration is a refinement when $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ have sufficiently converged.

The results of EXOSPECO are estimations of the star's parameters $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$, the on-axis PSF parameters $\hat{\boldsymbol{\nu}}_\star$, the off-axis PSF parameters $\hat{\boldsymbol{\nu}}_\oplus$ and the companion's SED $\boldsymbol{z}$.

3.5 Validation of the method with injected spectra in real data

3.5.1 Importance of the tuning of the regularization

In inverse problems, tuning the regularization is an important step as it controls the tradeoff between bias and variance of the estimator. As seen in Section 3.4.3, using AMORS reduces the number of

regularization hyper-parameters to two hyper-parameters: the companion's regularization $\mu_{\boldsymbol{z}}$ and one of the star's regularization hyper-parameters $\mu_{\boldsymbol{x}}$ or $\mu_{\boldsymbol{y}}$, the other one being set to one.

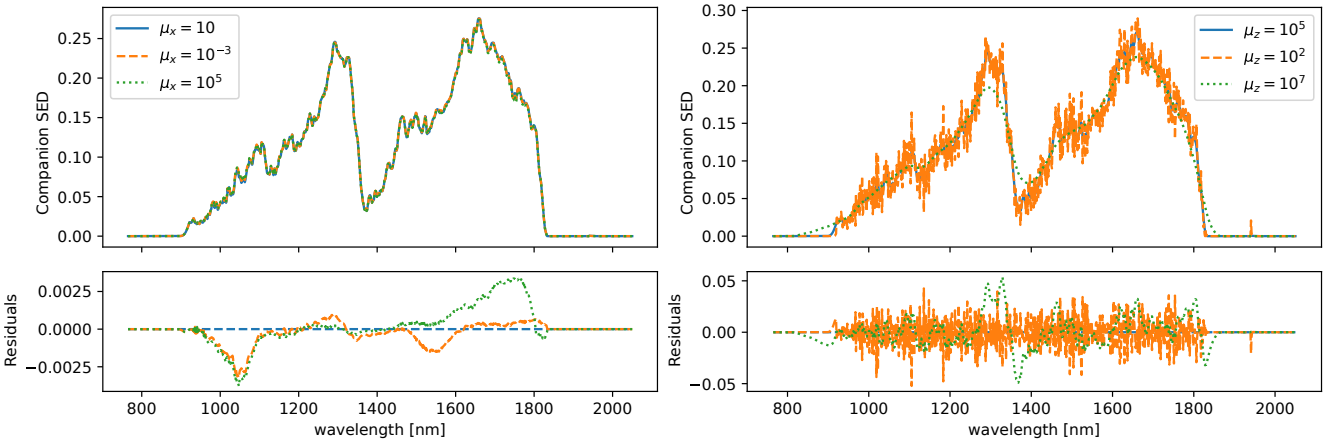

Figure 3.10 – Left panel: profiles of the companion SED $\boldsymbol{z}$, for different levels of the star regularization $\mu_{\boldsymbol{x}}$ at fixed $\mu_{\boldsymbol{z}} = 10^5$. Right panel: profiles of the companion SED $\boldsymbol{z}$ for different levels of the companion regularization $\mu_{\boldsymbol{z}}$ at fixed $\mu_{\boldsymbol{x}} = 10$. The residuals, in the bottom panels, are computed by the difference between the blue profile and the orange and green profiles.

Automatic tuning methods can be applied in a hierarchical way to optimally tune the two hyper-parameters, as explained in Section 1.4. However, practical tests show that the sensitivity of the tuning of the star hyper-parameters is extremely low. Applying ExoSpeco to the same HR3549 dataset as introduced in section 3.3.3, the left side of Fig. 3.10 shows the extracted SEDs for a vast range of star hyper-parameters (from 10^{-3} to 10^5 for $\mu_{\boldsymbol{x}}$ with $\mu_{\boldsymbol{y}} = 1$). As can be seen, the differences between the SEDs are smaller than 1%, showing that the tuning of $\boldsymbol{\mu_x}$ is not of uttermost importance and can be done manually.

Contrary to the star's SED regularization $\mu_{\boldsymbol{x}}$, the companion's SED regularization hyper-parameter $\mu_{\boldsymbol{z}}$ is quite important for the quality of the extraction. Indeed, the right side of Fig. 3.10 shows the companion's SED for multiple values of the hyper-parameters. These tests confirm the importance of the tuning of $\mu_{\boldsymbol{z}}$.

Figure 3.10 demonstrates that our method is mostly driven by the tuning of a single hyper-parameter. ExoSpeco can be considered as an easy-to-use method and can be completely unsupervised by using methods described in Section 1.4 to estimate $\mu_{\boldsymbol{z}}$.

3.5.2 Synthetic injected companion in data and first results

To validate our method, we injected synthetic companion's SED into real data. Acknowledging that it can be considered as an "inverse crime", this validation is however relevant in terms of speckles removal and quality of the un-mixing procedure.

As said before, a calibration of the star's SED is usually undertaken with the measurements of LSS data. It consists of pointing the star with the slit and applying a tilt to the deformable mirror of the ExAO so that the star's SED can be visible on the detector. As it is very bright, a neutral density is inserted in the optical path to avoid any saturation of the detector's pixels. Figure 3.11 presents such calibration data (on the left) for the target HIP65426, observed on 2019-05-20 in MRS mode. Using Alg. 3.2 and dividing the result by a calibration of the neutral density, it is quite straightforward to retrieve the star's SED (right of the figure), called in the following $\boldsymbol{x}_{\text{cal}}$. Note that, thanks to this calibration, it is also possible to fit the off-axis PSF parameters which can be provided as initialization values in Alg. 3.3.

In the following, we choose as a synthetic companion's SED $\boldsymbol{z}^* \in \mathbb{R}^{N_{\boldsymbol{z}^*}}$ a distribution proportional

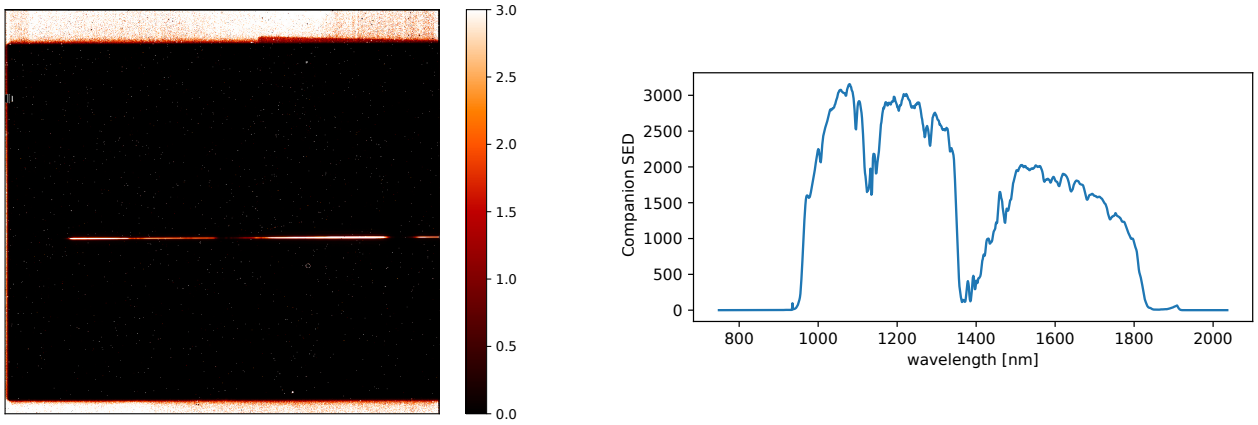

Figure 3.11 – Left: example of calibration data for the star's SED, for the scientific dataset of HIP65426, observed on 2019-05-20 in MRS mode. Right: star's SED of HIP65426, obtained by Alg. 3.2 on the data presented on the left and corrected from the density filter.

to the calibrated star SED $\boldsymbol{x}_{\text{cal}}$:

$$\boldsymbol{z}^* = \chi\, \boldsymbol{x}_{\text{cal}} \tag{3.44}$$

with χ the contrast factor in the astronomical sense. It measures the ratio between the intensity of the star measured without a coronagraphic mask and the companion intensity. Not that we choose this definition conveniently to describe the contrast with only one parameter.

Injecting a synthetic SED ($\chi = 2 \cdot 10^{-4}$, $\rho_{\oplus} = 785$mas) into a HIP65426 dataset (Fig. 3.12 and green curve of Fig. 3.13), we apply ExoSPECO to the resulting data. The regularization hyper-parameter $\mu_{\boldsymbol{z}}$ is tuned to minimize the MSE between the estimated companion's SED and the ground truth $\boldsymbol{z}^*$. Fig. 3.13 shows the companion's SED after 1 iteration (blue dashed line) and after convergence of the algorithm (blue solid line). The figure demonstrates the advantage of iteratively refining the star and companion models. By doing so, our algorithm reduces significantly the self-subtraction bias.

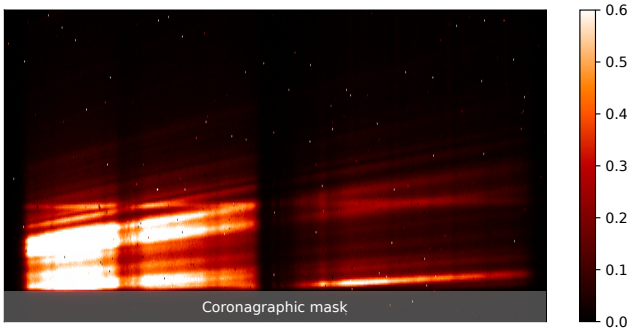

Figure 3.12 – Top part of the HIP65426 data with injected companion's SED of contrast $\chi = 2 \cdot 10^{-4}$ and angular separation $\rho_{\oplus} = 785$mas.

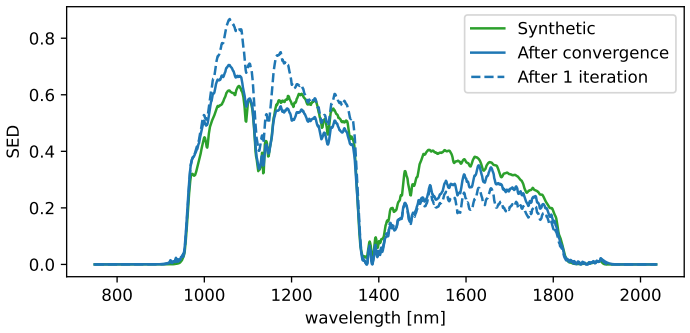

Figure 3.13 – Companion's SED extracted with Alg. 3.3, after one iteration of the algorithm (dashed blue line) and after convergence of the method (solid blue line). The synthetic SED is given for comparison in green.

3.5.3 Extraction quality

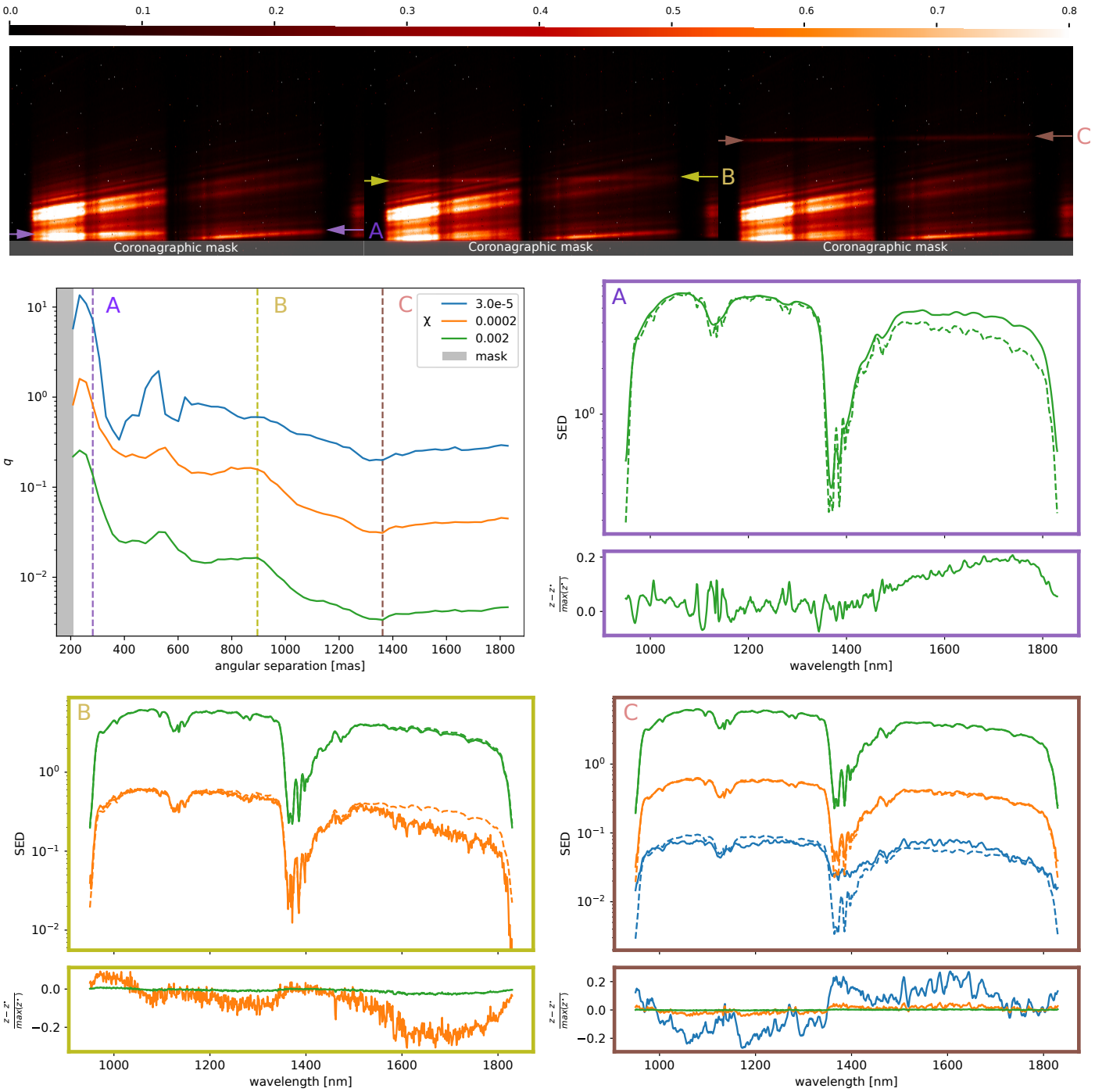

Figure 3.14 – Injections and extraction of synthetic companion's SEDs in the HIP65426 dataset. Top figure: three synthetic data of contrast $\chi = 2 \cdot 10^{-4}$ obtained by injecting a companion's SED at $\rho_\oplus = 273$ mas (A), $\rho_\oplus = 890$ mas (B), and $\rho_\oplus = 1353$ mas (C) (the position is highlighted by the arrows on the figure). Center-left figure: relative absolute error as a function of the angular separation $\rho_\oplus$ of the synthetic SED and for different contrasts (blue: $\chi = 3 \cdot 10^{-5}$, orange: $\chi = 2 \cdot 10^{-4}$, and green: $\chi = 2 \cdot 10^{-2}$). The gray area represents the region invalidated by the coronagraphic mask. The three angular separations of the injections presented in the top figure are marked by the three dashed lines A, B and C. Center-right and bottom figures: extracted SEDs profiles of quality $q > 20\%$ of the injected companion at the angular separation highlighted in the center-left and top figures. The synthetic ground truths are given in dashed lines. Normalized residuals are also given for each SED.

We measure the quality of extractions made by ExoSpeco by the relative mean absolute error

between the obtained companion's SED $\hat{\boldsymbol{z}}$ and the ground truth $\boldsymbol{z}^* \in \mathbb{R}^{N_{\boldsymbol{z}^*}}$:

$$q(\hat{\boldsymbol{z}} \mid \boldsymbol{z}^*) = \frac{\sum_{n=1}^{N_{\boldsymbol{z}^*}} |z_n^* - \hat{z}_n|}{\sum_{n=1}^{N_{\boldsymbol{z}^*}} |z_n^*|} \tag{3.45}$$

We assess the quality of the extraction by plotting the quality criterion q for injections as a function of the angular distances and for different contrasts χ. Figure 3.14 shows the results of this study, highlighting 3 synthetic injections at $\rho_\oplus = 273$ mas (purple A), 890 mas (yellow B) and 1353 mas (brown C) and for 3 contrast values $\chi = 3 \cdot 10^{-5}$ (blue), $2 \cdot 10^{-4}$ (orange) and $2 \cdot 10^{-3}$ (green). Extracted SEDs are given compared to their synthetic ground truth in dashed lines. Normalized residuals are also shown, comparing for different contrasts the errors made by the algorithm. From these results, extractions with contrasts of up to $\chi = 10^{-3}$ with Alg. 3.3 are faithful for a large range of angular separation and even at close range (~ 400 mas) of the coronagraphic mask. For companions at 600 mas or more, the extracted SEDs are still close to what was expected, even for faint companions of $\chi = 10^{-4}$. Farther from the mask, extremely faint companions with a contrast of $\chi = 10^{-5}$ can still be recovered with good fidelity.

As expected, faint companions close to the mask are still very difficult to extract as the quality criterion indicates. Notice in this area the slight underestimation of the SEDs in the H band (1500 to 1800 nm) in the bottom-left and center-right of Fig. 3.14. This can indicate a modeling error, where the star model does not align correctly with the speckles. Another sign of this phenomenon is seen around 550 mas in the center-left figure. This small rising of q is the sign of misalignment between a speckle and our model. From these results, we can see that for $q \sim 2 \cdot 10^{-1}$ the extraction is quite close to the ground truth. Fig. 3.15 shows what contrast value is attainable for a large range of angular separation at this extraction quality. Curves for a more constrained (lower q, dotted in the figure) and more flexible (dashed) quality criterion are plotted. EXOSPECO can obtain faithful extraction for reasonable contrast values with respect to the angular separation. Note that these curves are obtained for the extraction in one LSS frame data (better performances can be expected by processing several frames). It is important to remark that these curves do not convey information about the ability to detect companions, as it is necessary to have detected a companion before using long-slit spectroscopy.

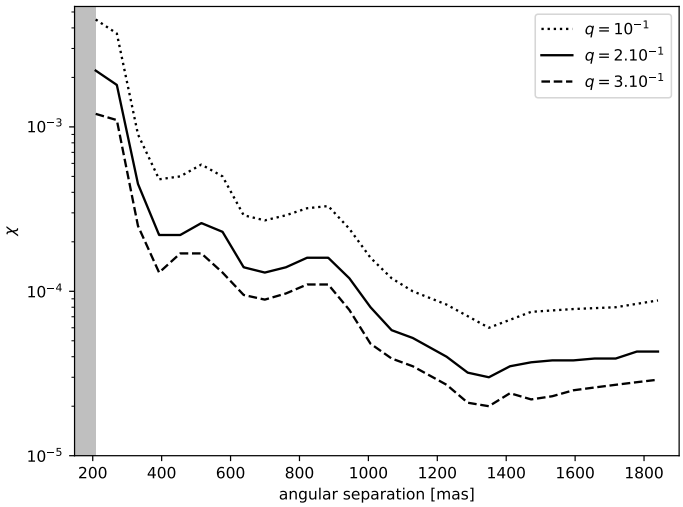

Figure 3.15 – Required contrast χ to achieve a given relative absolute error of $q = 10^{-1}$ (dotted line), $q = 2 \cdot 10^{-1}$ (plain line), and $q = 3 \cdot 10^{-1}$ (dashed line), as a function of the angular separation.

3.5.4 Comparison with a state of the art method

We apply the state-of-the-art TSVD method (explained in Section 3.1.3) on the same data as Section 3.5.2 containing a synthetic companion located at $\rho_{\oplus} = 785$mas and of contrast $\chi = 2 \cdot 10^{-4}$.

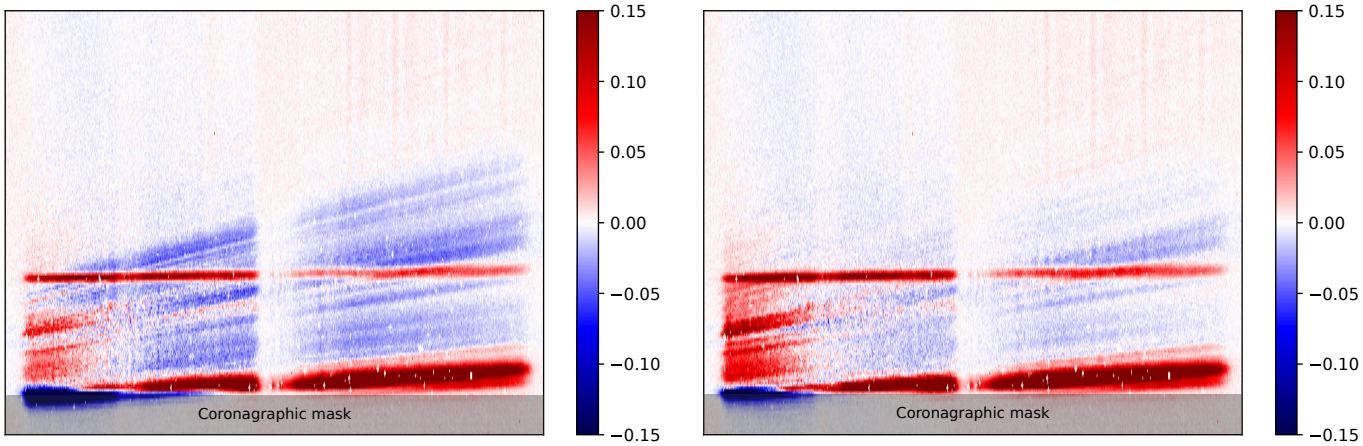

Figure 3.16 – Top area of the residuals between the HIP65426 data and the TSVD stellar model (left figure) ad the EXOSPECO model (right figure). The same synthetic companion as the one in Section 3.5.2 is injected ($\rho_{\oplus} = 785$mas, $\chi = 2 \cdot 10^{-4}$).

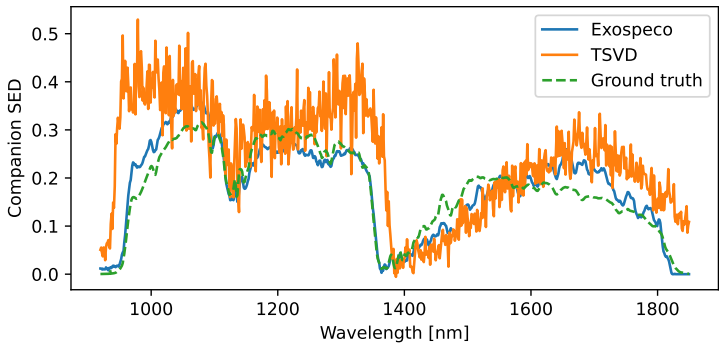

Figure 3.17 – Extracted companion's SED with the TSVD method and our EXOSPECO algorithm for the data with an injected companion at $\rho_{\oplus} = 785$mas and of contrast $\chi = 2 \cdot 10^{-4}$.

Figure 3.16 shows the residuals between the data with synthetic injection and the model of the stellar component done by the TSVD method (left figure) and by our EXOSPECO method (right figure). The TSVD model while reducing the level of residuals is still responsible for a high negative bias on a vast area of the detector. EXOSPECO seems to be a good alternative to such a method, especially when observing the companion component in the residuals. Note the small white areas on each side of the companion signal in the left figure of Fig. 3.16. These features come from the fact that the stellar model learned by TSVD includes some of the companion signal. Comparing the extracted companion by the two methods, Figure 3.17 shows the impact of the self-subtraction and of the extraction method in the TSVD method. In the case of the TSVD method, the extraction of the SED has been done by averaging the residuals around the companion signal on a 7-pixel high window. While being smoother thanks to the regularization, the resulting companion SED obtained by EXOSPECO has closer spectral features from the ground truth with a relative error of $q = 10^{-1}$ than the one obtained by TSVD of relative error $q = 9 \cdot 10^{-1}$.

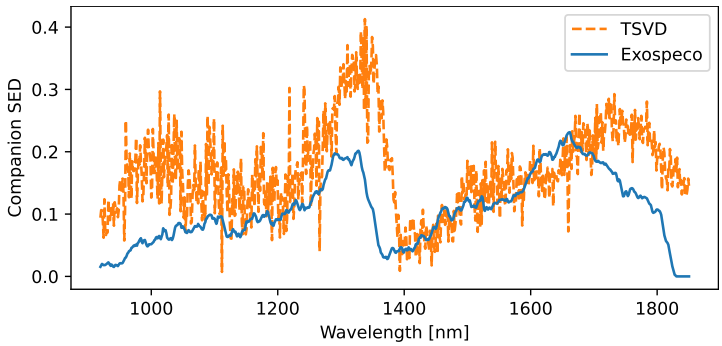

Figure 3.18 – Extracted companion's SED of one frame of the HR3549 data for the TSVD method (orange line) and EXOSPECO (blue line).

3.5.5 Spectrum extraction on real data

We decide to apply the EXOSPECO algorithm to one frame of the HR3549 dataset (Mawet et al., 2012), observed on 2015-12-28 in MRS mode by SPHERE/IRDIS. It presents a companion (brown dwarf) at ~ 830 mas. Figure 3.18 shows the resulting companion's SED using the TSVD method detailed in Section 3.1.3 and the EXOSPECO algorithm. Both methods used the EXOSPECO calibration model composed of 2D-polynomial and shown in Section 3.3.3. Thanks to the regularization, EXOSPECO yields a SED much smoother than with a TSVD method while being optimal in the sense of the maximum *a posteriori*.

3.6 Summary and next steps

This chapter presents a new method, EXOSPECO, to extract the companion's spectrum information from LSS data. It derives from a direct model based on physical considerations of the stellar and companion components. We also present a geometric calibration of the dispersion laws responsible for giving the angular separation and wavelength coordinates of each pixel. We demonstrate how important this step is to exploit at best the data. We also show that by using an alternated estimation strategy, EXOSPECO reduces greatly the self-subtraction bias that affects the state-of-the-art methods. All these properties ensure a good un-mixing between the two components while using simple estimation steps in the iterative framework. Combined with the proposed calibration method, EXOSPECO gives good results when applied to one frame of the data. For a vast range of angular separation, the companion's SED is faithfully retrieved for reasonable contrasts between the star and the companion. The work described in this chapter is part of an article submitted to A&A and the methods can be found as Julia packages (Bezanson et al., 2017) at `https://github.com/SJJThe/ExospecoCalibration` for the calibration and `https://github.com/SJJThe/Exospeco` for EXOSPECO.

However, there are still some imperfections in the residuals where some misalignment between the model and the data can still be seen which may contaminate the extracted SEDs. These phenomena can be explained by modeling errors. Indeed, to build the direct model we made some hypotheses that may be too strong. For instance, the direct model does not take into account any misalignment between the star and the slit in the orthogonal direction of the star/planet axis. Such positioning errors affect the dispersion of the speckles on the detector. Moreover, only the first spatial mode of the series expansion of Eq. (3.3) is taken into account.

Potential improvements are therefore possible, such as taking more spatial modes into account in the on-axis PSF. The inverse problem framework is flexible enough to take into account multiple frames datasets. The model of a T frames dataset $\{\boldsymbol{d}_t\}_{t\in[\![1,T]\!]}$ would then write:

$$\boldsymbol{m}_t(\boldsymbol{x}, \boldsymbol{y}_t, \boldsymbol{z}, \boldsymbol{\nu}_{\star t}, \boldsymbol{\nu}_{\oplus t}) = \boldsymbol{\gamma} \odot (\mathbf{H}_\star(\boldsymbol{\nu}_{\star t})\, \boldsymbol{y}_t) \odot (\mathbf{F}_\star\, \boldsymbol{x}) + \theta_t\, \boldsymbol{\gamma} \odot \boldsymbol{h}_\oplus(\boldsymbol{\nu}_{\oplus t}) \odot (\mathbf{F}_\oplus\, \boldsymbol{z}) \tag{3.46}$$

with $\boldsymbol{y}_t$ depending on the frame to take into account the variations of the speckles throughout the frames, $\boldsymbol{\nu}_{\star t}$ accounting for different misalignments between star and center of the coronagraphic mask and $\boldsymbol{\nu}_{\oplus t} = (\rho_\oplus, \sigma_\oplus, \theta_t)$ the parameters of the off-axis PSF. θ_t is a scaling factor of the companion's SED accounting for changes in the transmission of the off-axis PSF. And rewriting Eq. (3.32), the new criterion:

$$\begin{aligned}\mathcal{C}'\big(\boldsymbol{x}, \{\boldsymbol{y}_t, \theta_t\}_{t\in[\![1,T]\!]}, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus; \boldsymbol{\mu}\big) = &\sum_{t=1}^{T} \|\boldsymbol{d}_t - \boldsymbol{m}_t(\boldsymbol{x}, \boldsymbol{y}_t, \theta_t, \boldsymbol{z}, \boldsymbol{\nu}_\star, \boldsymbol{\nu}_\oplus)\|^2_{\mathbf{W}} \\ &+ \mu_{\boldsymbol{x}}\, \mathcal{R}_{\boldsymbol{x}}(\boldsymbol{x}) + \mu_{\boldsymbol{y}} \sum_{t=1}^{T} \mathcal{R}_{\boldsymbol{y}}(\boldsymbol{y}_t) + \mu_{\boldsymbol{z}}\, \mathcal{R}_{\boldsymbol{z}}(\boldsymbol{z}) \end{aligned} \tag{3.47}$$

can be used with the same strategy ExoSpeco. Using this strategy, it is possible to improve the method by a $\sqrt{T}$ factor (*e.g.* by a factor of ten with a hundred frames).

These improvements can push further the sensitivity of ExoSpeco. Jointly processing more data and accounting for a better model with more modes, our method could extract even extremely faint companions at contrasts of 10^{-5} or 10^{-6}.

Chapter 4

Reconstruction of an extended object from high-contrast images

In Chapter 3 we developed a method for the spectral characterization of sub-stellar companions. Imaging the circumstellar environment is key to understanding the formation of solar systems: the morphology of protoplanetary, transition and debris disks shows the different stages of the formation of exoplanets. The shape of such disks can give us important information on the age and the different steps that take place during the formation of the system. Indeed, while forming, exoplanets aggregate the dust of the disk in their orbit paths, thus creating gaps and spiral arms. Thanks to the dual-band imaging mode of SPHERE/IRDIS (Beuzit et al., 2019; Dohlen et al., 2008), such environments can be imaged. This mode uses the same Extreme Adaptive Optics (ExAO) and coronagraphy of SPHERE than for the long-slit spectroscopy (LSS) mode, but this time to image disks of particles much fainter in intensity than their host star. The data produced by the instrument is formed by pairs of images of the circumstellar environment captured simultaneously on the detector at two different wavelengths. Stellar leakages strongly contaminate the data, in the form of speckles formed by the imperfect correction of phase aberrations in the light path. They are particularly strong close to the coronagraphic mask, as can be seen in the observations shown in Fig. 4.1.

These speckles form a nuisance component in the data. Composed of highly contrasted spatial patterns, they typically are 10^4 brighter than the object of interest, noted $\boldsymbol{x}$ (*e.g.* protoplanetary disks, companions), in the infrared. As discussed in Chapters 1 and 2, the instrument and transfer medium are also responsible for a blurring of $\boldsymbol{x}$ (atmospheric deformations of the wavefront, optical components, imperfect correction of ExAO) and characterized by its off-axis Point Spread Function (PSF) when considering objects outside of the optical axis. Limiting greatly the study of extended objects, we need to use post-processing methods to disentangle the object of interest from other signals present in the data.

As seen throughout this thesis, knowledge of the instrument and the observation method is crucial in the inverse problems framework. The first part of this chapter, Section 4.1, is therefore dedicated to modeling the Angular Differential Imaging (ADI) data produced by IRDIS. Section 4.2 details the most common state-of-the-art methods to recover the object of interest from ADI datasets. The inverse problems approach is then applied in Section 4.3 to derive a reconstruction method and estimate the statistics of the data. Section 4.4 proposes different approximations of the covariance of the data. Finally, Section 4.5 concludes on the validity of these approximations and discusses the reconstruction problem.

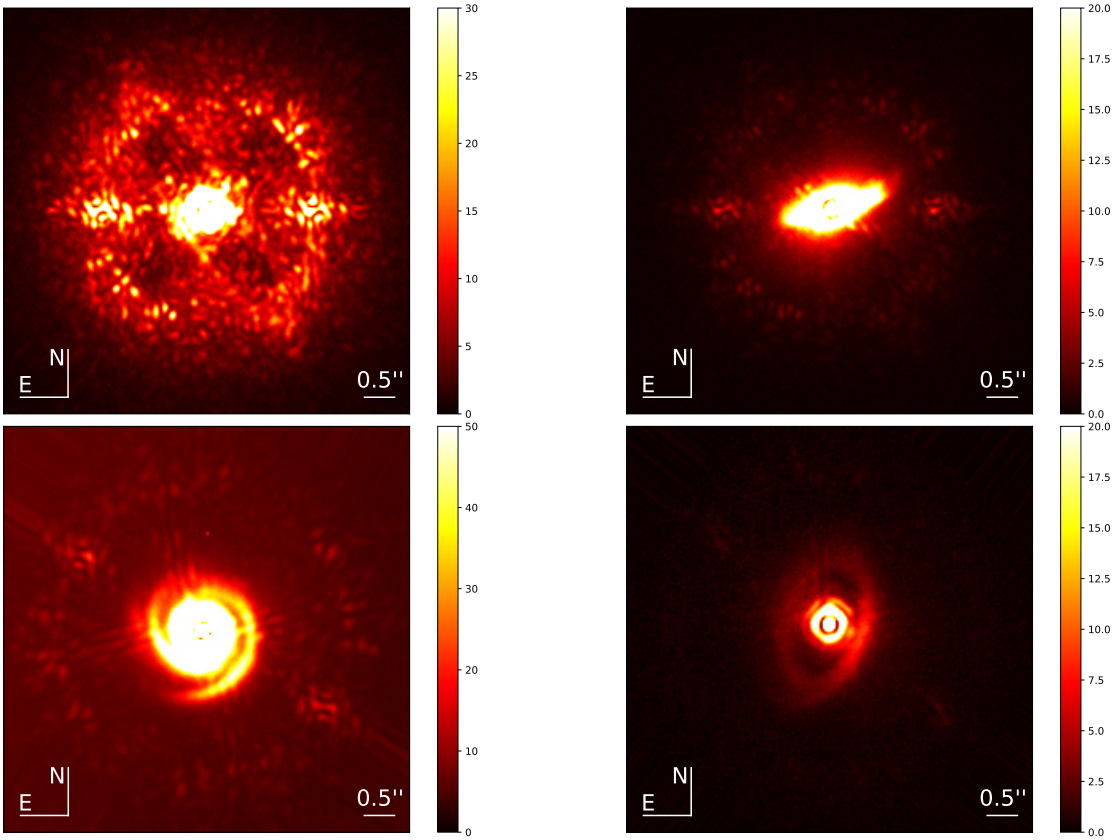

Figure 4.1 – Samples of data acquired with the IRDIS instrument. It presents circumstellar environments of HR 4796 (top-left), RY Lupi (top-right), SAO 206462 (bottom-left) and PDS 70 (bottom-right).

In This Chapter

Prerequisite:

- The formation of the pre-processed data (Section 1.1).
- Maximum *a posteriori* formalism (Sections 1.3.4 and 1.3.5).

Contributions:

- A general algorithm to de-blur, de-mix and de-noise (Flasseur, O. et al., 2021).
- Study of the impact of different approximations of the covariance matrix for modeling speckles.
- Development of an unsupervised procedure to choose the complexity of the covariance model.

4.1 Modeling the observations

4.1.1 Sequence of data

Given the complexity of the speckles in the data, fitting a model of the on-axis PSF can be difficult[1]. In order to recover the object of interest $\boldsymbol{x}$ despite these spurious terms, Angular Differential Imaging (ADI) is commonly used as a method of observation (Vigan et al., 2010). It consists in acquiring multiple images of the region of interest throughout the night, with a de-rotator in pupil-stabilization mode. In doing so, the field of view (night sky) rotates from one acquisition to another due to the Earth's rotation, while the stellar leakages are quasi-static.

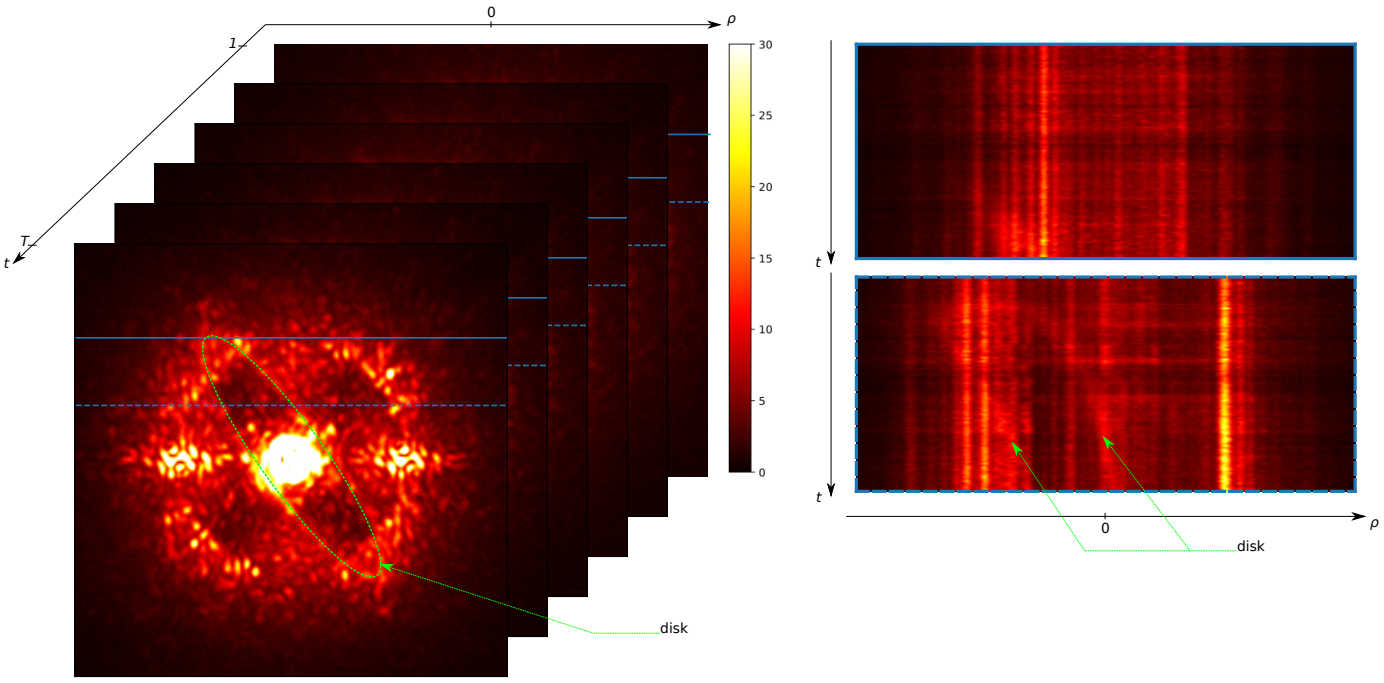

Figure 4.2 – Example of ADI dataset (left) of HR4796 presenting multiple frames and two spatio-temporal slices (right) along the plain and dashed lines of the left figure.

Fig. 4.2 shows a typical ADI dataset (HR4796 observed the 03/02/2015) and spatio-temporal slices of this dataset cut at two different locations, one close (dashed line) to the coronagraphic mask, the other one farther (plain line). This figure illustrates the spatial fluctuations of the nuisance component of the data. The closer to the coronagraphic mask, the stronger the speckles. It is possible to see the apparent rotation of the spatially extended object (a disk here) in the spatio-temporal slices (green arrows in the bottom-right of the figure). Due to this movement throughout the frames, $\boldsymbol{x}$ is not perturbed by the same speckles throughout the time series. Regarding the speckle component, it also means that the object of interest only has a small intermittent contamination impact on them. Un-mixing the object of interest from the speckle component is then possible provided that sufficient rotation of the field of view occurs during the time series. Note that the closer the features of the object are to the coronagraphic mask, the smaller the relative motion between the speckles and the object. In these areas of the data where the speckles are the strongest, the recovery of the object of interest is then particularly challenging as it barely moves.

4.1.2 Preprocessing of the raw data

Detector and instrument calibration must be performed in a first step. Rather than developing a calibration algorithm for the instrument as done for the LSS in Chap. 3, we decided to use the standard

[1] although not impossible as shown by (Devaney and Thiébaut, 2017).

calibration available in the SPHERE-DC pipeline (Pavlov, Feldt, and Henning, 2008; Delorme et al., 2017). Using DARK and FLAT calibration, it follows a procedure similar to the one explained in App. A to correct for the detector's bias (although the correction is not pixel-wise). Bad pixels are automatically identified and interpolated in the data which will unfortunately add correlations between pixels.

In order to take into account the blurring due to the instrument, the off-axis PSF is calibrated by acquiring the image of the star without a coronagraphic mask. To do so, a tilt is applied to the deformable mirror of the ExAO, resulting in the star being shifted away from the coronagraphic mask. By acquiring images of the unmasked star right before and right after the measurements of the ADI dataset, the PSF can be considered as representative of the PSF of the ADI datasets since it is captured under very close atmospherical conditions.

The rotation of the field of view is defined by the parallactic angles measured via the de-rotator in the common path of SPHERE (Montagnier et al., 2007). This information defines the movement of the object of interest on the detector's plane. In practice, the rotation angle is usually not measured at each acquisition. In that case, it is deduced for each frame by means of an interpolation procedure.

Finally, by applying a chessboard modulation on the deformable mirror of the ExAO, four crosswise bright spots are produced on the detector. The geometrical center of the figure formed by these spots yields the position of the optical axis. Thanks to this "waffle" calibration, it is possible to re-center each frame so that the rotation center of the field of view coincides with the geometric center of the frames.

4.1.3 Direct model of the data

Let $\boldsymbol{d}_t \in \mathbb{R}^N$ be the t-th frame of N pixels of the ADI dataset. As discussed in 4.1.1, several nuisance sources are corrupting the data. Separating the different components in $\boldsymbol{d}_t$ and using the hypotheses discussed in Sections 1.1 and 1.2 yields the model:

$$\boldsymbol{d}_t = \mathbf{A}_t\, \boldsymbol{x} + \boldsymbol{m}_\star + \boldsymbol{b}_t \tag{4.1}$$

with $\mathbf{A}_t \in \mathbb{R}^{N\times N}$ the matrix modeling the instrumental response (including the blur and rotation) for the object, $\boldsymbol{x} \in \mathbb{R}^N$ the distribution of light, an image of the object of interest and $\boldsymbol{m}_\star \in \mathbb{R}^N$ the expectation of the quasi-static stellar leakages. $\boldsymbol{b}_t$ represents all the terms that are fluctuating in the data, *i.e.* the fluctuation of the speckles, the photon noise, the read-out noise as well as other sources of uncertainties (rounding errors, model errors, *etc.*) that are still in the data and need to be accounted for. In other words, $\boldsymbol{b}_t$ is a stochastic parameter, centered as it represents only fluctuations in the data. As discussed in the other chapters of this manuscript, we will consider $\boldsymbol{b}_t$ as following a centered multivariate Gaussian distribution, *i.e.* $\boldsymbol{b}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$, with $\mathbf{C}$ its covariance matrix. Note that we make a strong hypothesis here by considering that all components $\boldsymbol{b}_t$ share the same covariance matrix.

As discussed in 4.1.1, the instrument and the observation method imply three major effects on the object of interest $\boldsymbol{x}$: the apparent rotation of the field of view on the detector, the coronagraphic mask's non-uniform transmission and the blurring off-axis PSF of the instrument. $\mathbf{A}_t$ can therefore be decomposed into three linear operators (Flasseur, O. et al., 2021):

$$\mathbf{A}_t = \mathbf{H}\,\boldsymbol{\Gamma}\,\mathbf{Q}_t \tag{4.2}$$

where $\mathbf{Q}_t$ models the rotation of the field of view at the time of acquisition of the frame t. Since each frame has been geometrically re-centered (see 4.1.2), it can be implemented by a simple rotation operator given the angles of rotation of the field of view. $\boldsymbol{\Gamma}$ represents the attenuation by the

coronagraphic mask, stronger near the center optical axis. This operator can be approximated as $\mathbf{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma})$, with $\boldsymbol{\gamma}$ a pixel-wise transmission map. This map is built by a 2D interpolation of a mean profile of the coronagraphic mask transmission, null at the optical center (center of the frame) and equal to 1 farther away. Finally, $\mathbf{H}$ models the instrument off-axis PSF. Under or close to the coronagraphic mask, this PSF is strongly deformed. However, these parts are not well described by our model in Eq. (4.2) and are dominated by speckles. As they are not well contrasted, it is mostly the core of the off-axis PSF that has an effect. Due to the small field of view of the instrument, $\mathbf{H}$ is considered in the following as shift-invariant and modeled by a 2D-convolution (see Section 1.2) applied to the rotated and attenuated image of the object. The frames $\boldsymbol{d}_t$ of the dataset are taken consecutively, meaning that the observation conditions are assumed similar throughout the acquisition time. This motivates the use of a single PSF (measured right before or right after acquiring the ADI dataset).

From all these considerations arise the different parameters of our direct model that are not known after preprocessing. Of course, $\boldsymbol{x}$ is the most important variable of our model to estimate as it represents the objects of interest. However, similarly to Chapter 3, to recover $\boldsymbol{x}$, the stellar leakages in the data $\boldsymbol{m}_\star$, need to be estimated. Noting the significant spatial correlations of the speckles in the data and the fact that the SPHERE-DC pipeline does not produce an estimation of the statistical weights of the data, the covariance $\boldsymbol{C}$ of the data also needs to be estimated.

4.2 State-of-the-art methods

Several methods have been developed to remove the speckles component from the data leaving only the extended objects of interest (Mawet et al., 2012; Pueyo, 2018). The ones used commonly by the community can be separated into four different families:

Methods that rely on a disk model

These methods are based on the estimation of parameters of a model of the disk that constrain its shape and light intensity (Milli et al., 2012; Esposito et al., 2013; Mazoyer et al., 2020). The parameters obtained are the ones that reduce the most the residuals $\boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x}$, with $\boldsymbol{x}$ unknown. Using a parametric model introduces a strong bias if the shape and light distribution of the disk differs from the assumed model or if other objects are also present (*e.g.* companions). In other words, the parametric model does not allow much flexibility compared to a non-parametric estimation of the object of interest.

Methods that rely on an external measurement of the stellar leakages

Here, the methods are using the Reference Differential Imaging technique (RDI) consisting of acquiring another ADI dataset of reference stars where no scientific object has ever been detected (Ruane et al., 2019). This reference dataset is then subtracted from the scientific dataset to reveal the objects of interest. These methods are highly sensitive to variations in observation conditions when acquiring the two datasets which change the speckle component. This reference can however be used to build a linear model of the stellar leakages (*e.g.* a fixed sub-space built by Single-Value Decomposition (SVD) or non-negative decomposition) as explained in Ren et al., 2018. With enough datasets of reference, it is also possible to directly combine each frame, as the LOCI method (Lafreniere et al., 2007), or only the best ones in order to derive a model of the stellar leakages, as in TLOCI (Marois et al., 2014). These methods should however require a vast library of ADI datasets and make the strong hypothesis that the speckles are similar in all these datasets.

Methods that learn a model of the stellar leakages directly from the dataset

These methods deduce, directly from the dataset, a model of the speckles. Like with the methods above, this stellar leakages model is subtracted from the data to reveal the objects of interest.

Classical ADI (cADI) is a commonly used method as it is quite easy to implement. It consists of the subtraction of the temporal median of the dataset to each of its frames. Note that the median is used here to prevent the object of interest to pollute the stellar leakages model that is subtracted. The resulting residual images are then de-rotated to compensate for the rotation of the field of view and temporally averaged. The resulting estimation of the object of interest can be expressed as:

$$\hat{\boldsymbol{x}}^{[\mathrm{cADI}]} = \mathrm{median}_{t'}\left\{\mathbf{Q}_{t'}^{-1}\left(\boldsymbol{d}_{t'} - \mathrm{median}_{t}\{\boldsymbol{d}_t\}_{t\in[\![1,T]\!]}\right)\right\}_{t'\in[\![1,T]\!]} \tag{4.3}$$

with $\mathrm{median}_t\{\boldsymbol{d}_t\}_{t\in[\![1,T]\!]}$ yielding the temporal median of each pixel of the set of frames $\{\boldsymbol{d}_t\}_{[\![1,T]\!]}$ and $\mathbf{Q}_t^{-1}$ the inverse transformation that allows de-rotating to align the object in all frames.

In the same spirit, KLIP methods (Soummer, Pueyo, and Larkin, 2012) are estimating a model of the on-axis PSF by identifying with an SVD the sub-space spanned by the stellar leakages and its complement that contains most of the object of interest.

These methods all suffer from self-subtraction issues (part of the object is lost when subtracting the model of stellar leakages). To counter this phenomenon, Pairet, Cantalloube, and Jacques, 2018 proposed an alternated method that iteratively refines an estimation of the objects of interest and the subspace spanned by the stellar leakages. Based on the TSVD, these methods are however sub-optimal as they do not take into account the statistics of the data, notably the non-uniformity of the variance and the correlated variations of the speckles.

Joint unmixing methods

In contrast to the methods described in the previous paragraph, these methods estimate jointly the two components: the objects and the stellar leakages, by an optimization process following the maximum likelihood or maximum *a posteriori* framework (see Sections 1.3.3 and 1.3.4). These inverse problems methods (Pairet, Cantalloube, and Jacques, 2021; Flasseur, O. et al., 2021) are gaining popularity due to their ability to reduce the self-subtraction effects, thus preserving the photometry of the object while also taking into account instrumental effects such as the PSF blur. In this Chapter, we detail an approach that follows this framework.

4.3 Inverse problems approach

4.3.1 Joint estimation of the two components

We recall that the data processing problem is to disentangle the object of interest $\boldsymbol{x}$ (extended structure like a disk and/or point-like sources like exoplanets) from the spurious signals $\boldsymbol{m}_\star$ and $\boldsymbol{b}_t$. The maximum *a posteriori* (MAP) yields an optimization problem that can be used to derive estimates of the unknowns $\boldsymbol{x}$, $\boldsymbol{m}_\star$ and $\mathbf{C}$ given the series of data-frames $\mathbf{D}$. Following the same methodology as explained in Section 1.3.4, this amounts to solving:

$$\left(\hat{\boldsymbol{x}}, \hat{\boldsymbol{m}}_\star, \hat{\mathbf{C}}\right) = \underset{\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C}}{\arg\max}\ \mathrm{p}(\boldsymbol{x}\,|\,\boldsymbol{\mu})\,\mathrm{p}(\mathbf{D}\,|\,\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C}) \tag{4.4}$$

$$= \underset{\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C}}{\arg\max}\ \mathrm{p}(\boldsymbol{x}\,|\,\boldsymbol{\mu}) \underbrace{\prod_{t=1}^{T}\mathrm{p}(\boldsymbol{d}_t\,|\,\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C})}_{\text{Probability density of } \mathbf{D} \text{ knowing } \boldsymbol{x}} \tag{4.5}$$

where the second line is obtained by considering the mutual independency of the frames of the dataset $\boldsymbol{d}$. $\mathbf{A}_t\,\boldsymbol{x}$ and $\boldsymbol{m}_\star$ being deterministic variables in the direct model of Eq. (4.1) and given our multi-variate Gaussian hypothesis, $\boldsymbol{d}_t$ is also a multi-variate Gaussian. Taking the co-log of the function to maximize, we can obtain the following MAP estimator:

$$\begin{aligned}\big(\hat{\boldsymbol{x}}, \hat{\boldsymbol{m}}_\star, \hat{\mathbf{C}}\big) &= \underset{\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C}}{\arg\min} -\log\Big(\mathrm{p}(\boldsymbol{x}\,|\,\boldsymbol{\mu}) \prod_{t=1}^{T} \mathrm{p}(\boldsymbol{d}_t\,|\,\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C})\Big) \\ &= \underset{\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C}}{\arg\min}\, \mathcal{L}(\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C};\boldsymbol{\mu}) \end{aligned} \tag{4.6}$$

with

$$\mathcal{L}(\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C};\boldsymbol{\mu}) = \overbrace{\sum_{t=1}^{T} \big\|\boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x} - \boldsymbol{m}_\star\big\|^2_{\mathbf{C}^{-1}} + T\,\log\big|\mathbf{C}\big|}^{\text{coming from } -\log\left(\mathrm{p}(\mathbf{D}\,|\,\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C})\right)} + \underbrace{\mathcal{R}(\boldsymbol{x};\boldsymbol{\mu})}_{prior} \tag{4.7}$$

and where $\big|\mathbf{C}\big|$ is the determinant of the matrix $\mathbf{C} \in \mathbb{R}^{N\times N}$ and $\|\boldsymbol{b}_t\|^2_{\mathbf{C}^{-1}} = \boldsymbol{b}_t^\mathsf{T}\,\mathbf{C}^{-1}\,\boldsymbol{b}_t$ is the squared Mahalanobis norm of the vector $\boldsymbol{b}_t = \boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x} - \boldsymbol{m}_\star$ representing the residuals after subtracting the contribution of the objects and the mean stellar leakages. As discussed in Section 1.3.4, the first term of $\mathcal{L}$ measures the discrepancies between our model and each frame of the data. The second term of this function is the regularization term which encodes our *a priori* information on the solution. Its level and shape are controlled by $\boldsymbol{\mu}$, a set of *hyper-parameters* which can contain more than one parameter depending on its specific expression.

Eq. (4.7) defines the problem as both a reconstruction of the component of interest (extended objects and potential point-like sources) and as a statistical characterization of the nuisance component. It also shows that the problem of joint estimation of $\boldsymbol{x}$, $\boldsymbol{m}_\star$ and $\boldsymbol{C}$ is non-convex and so, may require an adequate initialization strategy. A possible and pretty straightforward solution here is to exploit the alternated minimization approach, already explained in the previous chapters, to split the problem into sub-problems that are easier to solve and that can re-use existing algorithms. In practice, we show in the following that, even with this strategy, the estimation problem is not as straightforward as it may seem. From Eq. (4.6) and the considerations resulting from its study, two separate problems appear.

4.3.2 Estimation of the parameters of the nuisance component

The first sub-problem to solve is the characterization of the nuisance component defined by the estimation of the mean star leakages:

$$\hat{\boldsymbol{m}}_\star = \underset{\boldsymbol{m}_\star}{\arg\min}\, \mathcal{L}(\boldsymbol{x},\boldsymbol{m}_\star,\mathbf{C};\boldsymbol{\mu}) \tag{4.8}$$

Writing the gradient of $\mathcal{L}$ with respect of $\boldsymbol{m}_\star$,

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{m}_\star} = \frac{\partial}{\partial \boldsymbol{m}_\star}\Big(\sum_{t=1}^{T} \big\|\boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x} - \boldsymbol{m}_\star\big\|^2_{\mathbf{C}^{-1}}\Big) = -2\mathbf{C}^{-1} \sum_{t=1}^{T} \big(\boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x} - \boldsymbol{m}_\star\big) \tag{4.9}$$

and considering that $\mathbf{C}$ is a semi-definite positive matrix (where invalid pixels have arbitrarily high variances) yields the closed-form solution for this sub-problem:

$$\boxed{\hat{\boldsymbol{m}}_\star(\boldsymbol{x}) = \frac{1}{T} \sum_{t=1}^{T} \big(\boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x}\big)} \tag{4.10}$$

which is the empirical mean of the residual images formed by subtracting the contribution of the object of interest to the data. Note that this estimator does not depend on the statistics of $\boldsymbol{b}_t$, but solely on the object $\boldsymbol{x}$.

The second parameter that defines the nuisance component is the covariance matrix $\mathbf{C}$ of the stochastic component $\boldsymbol{b}_t = \boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x} - \boldsymbol{m}_\star$. Defining $\mathbf{B}$ containing the frames $\boldsymbol{b}_t, t \in [\![1, T]\!]$, it is possible to simplify the problem to:

$$\hat{\mathbf{C}}(\mathbf{B}) = \underset{\mathbf{C}\in\mathbb{S}}{\arg\min} \sum_{t=1}^{T} \left\|\boldsymbol{b}_t\right\|^2_{\mathbf{C}^{-1}} + T\,\log\left|\mathbf{C}\right| \tag{4.11}$$

over a set $\mathbb{S}$ of possible covariance matrices. If there are enough frames and $\mathbb{S}$ is unrestricted, a closed-form solution to this problem exists, that is the sample covariance. However, in practice, we usually need to restrict $\mathbb{S}$ to reduce the number of elements of $\mathbf{C}$ to estimate (see Section 4.4).

Another possible formulation of Eq. (4.11) is to work directly with the precision matrix:

$$\hat{\mathbf{W}}(\mathbf{B}) = \underset{\mathbf{W}\in\mathbb{W}}{\arg\min} \sum_{t=1}^{T} \left\|\boldsymbol{b}_t\right\|^2_{\mathbf{W}} - T\,\log\left|\mathbf{W}\right| \tag{4.12}$$

which will be investigated in Section 4.4.5

4.3.3 Reconstruction of the object of interest

Once the statistics of the nuisance component have been estimated, the problem is then to reconstruct the scientific component $\boldsymbol{x}$ from the data, writing:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\arg\min}\, \mathcal{L}(\boldsymbol{x}, \boldsymbol{m}_\star, \mathbf{C}; \boldsymbol{\mu}) \tag{4.13}$$

To better disentangle the object of interest from the other sources in the data, it is possible to constrain the set of possible solutions to only non-negative ones. Indeed, the object $\boldsymbol{x}$ represents a distribution of light and so cannot be negative. Given this new constraint, Eq. (4.13) can be rewritten under the form:

$$\hat{\boldsymbol{x}}(\boldsymbol{m}_\star, \mathbf{W}; \boldsymbol{\mu}) = \underset{\boldsymbol{x}\geqslant\mathbf{0}}{\arg\min} \sum_{t=1}^{T} \left\|\boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x} - \boldsymbol{m}_\star\right\|^2_{\mathbf{W}} + \mathcal{R}(\boldsymbol{x}; \boldsymbol{\mu}) \tag{4.14}$$

with $\mathbf{W} = \mathbf{C}^{-1}$ the precision matrix. This problem is not unknown in this manuscript as it has been at its center throughout the pages. It is a regularized reconstruction problem from residual images formed by $\boldsymbol{d} - \boldsymbol{m}_\star$.

Here the morphological *a priori* that we want to favor are the sharp edges of smooth extended objects (like disks) on a dark background. Following the commonly used regularization described in Section 1.3.5, the edge-preserving seems well-indicated in this context. Other regularizations are also possible without the problem losing any generality.

This reconstruction step can be done by using the limited memory method VMLM-B (see Section 1.5.2 for details) because of its capacity to include the positivity constraint the fact that it ensures efficient convergence for large-scale problems and can work even for non-quadratic regularization, as long as the criterion $\mathcal{L}$ is differentiable. Indeed, this method only requires the evaluation of the loss function in Eq. (4.14) and its gradient.

4.3.4 Alternated estimation

Separating problem (4.6) in accordance to 4.3.2 and 4.3.3, the estimation process can be done by an alternated strategy as shown in Algorithm 4.1. Doing so, each iteration will estimate the statistics of the residuals formed by the subtraction of the last estimated object to the data, thus progressively removing any possible self-subtraction bias in the residuals. Note that for this algorithm to work, we need to define the method FitPrecision which updates the precision matrix $\mathbf{W}$. This can be done either by solving Eq. (4.11) and inverse the resulting matrix, or directly by solving Eq. (4.12). Solving this problem needs however more details and will be discussed in Section 4.4.

Algorithm 4.1: Reconstruction algorithm from ADI dataset

Input: The T frames $\boldsymbol{d}_t$ and their associated operator $\mathbf{A}_t$, the regularization hyper-parameters $\boldsymbol{\mu}$.
Output: $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{m}}$ and $\hat{\mathbf{C}}$, a local minimum of Eq. (4.6).

$\boldsymbol{x}_{[0]} \leftarrow \mathbf{0}$
$\mathbf{B}_{[0]} \leftarrow \mathbf{0}$
$k \leftarrow 0$
while *not converged* **do**
  ▶ **Update model of the speckles (see 4.3.2)**
  $\boldsymbol{m}_{\star[k+1]} \leftarrow \hat{\boldsymbol{m}}_\star(\boldsymbol{x}_{[k]})$ ◁ Eq. (4.10)
  ▶ **Update covariance of nuisance component (see 4.3.2 and 4.4)**
  $\mathbf{B}_{[k+1]} \leftarrow \left[\boldsymbol{d}_t - \mathbf{A}_t\,\boldsymbol{x}_{[k]} - \boldsymbol{m}_{\star[k+1]}\right]_{t\in[\![1,T]\!]}$
  $\mathbf{W}_{[k+1]} \leftarrow \texttt{FitPrecision}(\mathbf{B}_{[k+1]})$ ◁ discussed in Section 4.4
  ▶ **Reconstruction of object (see 4.3.3)**
  $\boldsymbol{x}_{[k+1]} \leftarrow \hat{\boldsymbol{x}}(\boldsymbol{m}_{[k+1]}, \mathbf{W}_{[k+1]}; \boldsymbol{\mu})$ ◁ Eq. (4.14)
  $k \leftarrow k + 1$
$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{m}}, \hat{\mathbf{C}}) \leftarrow (\boldsymbol{x}_{[k]}, \boldsymbol{m}_{[k]}, \mathbf{C}_{[k]})$

4.4 Modeling the covariance of the nuisance term

To disentangle at best the object of interest from the other spurious signals in the data, we need to estimate precisely the statistics of the stochastic component $\boldsymbol{b}_t$. Indeed, a good approximation of $\mathbf{C}$ (or $\mathbf{W}$) will allow us to better unmix the object that rotates in the data from the fluctuations of $\boldsymbol{b}_t$. Given the size of the covariance matrix, N^2 pixels, and the number $N\,T$ of data entries, fully updating $\mathbf{W}$ in Alg. 4.1 is not doable for reasonable datasets ($T \ll N$). However, by imposing a structure and constraints on the set of possible matrices for $\mathbf{C}$, noted $\mathbb{S}$, it is possible to reduce drastically the number of degrees of freedom.

4.4.1 Evaluation of the performances of the covariance approximation

In this section, we suppose the other parameters which $\boldsymbol{b}_t$ depends on, as fixed variables. We look for an approximation of the covariance matrix that will ensure the whitening of the stochastic component $\boldsymbol{b}_t$, *i.e.* we want to ensure:

$$\boldsymbol{r}_t = \mathbf{V}\,\boldsymbol{b}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \tag{4.15}$$

with $\boldsymbol{r}_t$ the residuals whitened by the Cholesky factorisation matrix $\mathbf{V}$ of the precision matrix, which follows:

$$\begin{aligned}\mathbf{W} = \mathbf{C}^{-1} &= \left(\mathbf{L}\,\mathbf{L}^{\mathsf{T}}\right)^{-1} \\ &= \mathbf{L}^{-\mathsf{T}}\,\mathbf{L}^{-1} = \mathbf{V}^{\mathsf{T}}\,\mathbf{V}\end{aligned} \tag{4.16}$$

with $\mathbf{V} = \mathbf{L}^{-1}$.

The efficiency of the whitening effect of our estimate $\hat{\mathbf{V}}$ of $\mathbf{V}$ will depend greatly on the structural hypothesis and constraints on our set of possible matrices. To evaluate how well our estimation will describe $\boldsymbol{b}_t$, $\forall t \in [\![1, T]\!]$, we propose to use the Leave-One-Out Cross-Validation (LOOCV)(Hastie et al., 2009). It consists on building a set of models $\{\hat{\mathbf{V}}^{[m]}\}_{m\in[\![1,T]\!]}$, where each model $\hat{\mathbf{V}}^{[m]}$ is learned on the set $\{\boldsymbol{b}_t\}_{t\in\{[\![1,T]\!]\setminus\{m\}\}}$ of $T-1$ frames. With this deterministic procedure, each model is applied on a frame that was not used in the learning process:

$$\boldsymbol{r}^{[m]} = \hat{\mathbf{V}}^{[m]}\,\boldsymbol{b}_m \tag{4.17}$$

In other words, the LOOCV measures the capacity of each model $\hat{\mathbf{V}}^{[m]}$ to describe the whole dataset, that is, their capacity to extract generic correlations shared by the whole dataset.

To conduct such a study, we chose to test the estimates of $\mathbf{C}$ on a real dataset that does not contain any known object of interest, HIP 80019, taken the 2015-04-21 by IRDIS in dual-band imaging. We selected the channel of wavelength 2.11μm taken by this observational mode leading to a dataset of $T = 114$ frames of $N = 40804$ pixels. With this dataset, we can consider $\boldsymbol{x} = \mathbf{0}$ in Eq. (4.10) and $\boldsymbol{b}_t = \boldsymbol{d}_t - \hat{\boldsymbol{m}}$. The first row of Figure 4.3 shows different frames of these residuals $\boldsymbol{b}_t$, presenting structures and strong local inhomogeneities. Figure 4.5 presents on its first row the histograms of each frame of the residuals dataset, as Figure 4.6 summarizes in its first entry "centered" their means and standard deviations.

As shown in these figures, throughout this section we will focus on different structural constraints and approximations for the covariance (or precision) matrix. These models from a simple to a complex description of the nuisance component will help us evaluate how important it is to account for the spatial correlations of the speckles.

The LOOCV can also be used to select the level of complexity/sparsity of our approximation, *i.e.* the Effective Number of Degrees of Freedom (ENDoF). Choosing the same criterion that we use to estimate the elements of $\mathbf{V}$, *i.e.* the co-log-likelihood, as a *judge* function[2], the LOOCV study shows how our estimation of the precision matrix explains the dataset without fitting the particular features of the frame it whitens. Fig. 4.7 shows for different approximations and levels of sparsity the co-log-likelihood $\mathcal{L}_{\text{judge}}(\kappa)$ of the set of whitened residuals obtained via Eq. (4.17):

$$\mathcal{L}_{\text{judge}}(\kappa) = \sum_{m=1}^{T}\left[\left\|\boldsymbol{r}^{[m]}\right\|^2 - \log\left|\hat{\mathbf{V}}^{[m]}\right|^2\right] \tag{4.18}$$

with κ characterizing the complexity of the approximation (*e.g.* the number of co-precision elements[3] for each pixel in the precision matrix, the ENDoF). Thanks to this criterion, it is possible to estimate the level of optimal complexity in the sense of $\mathcal{L}_{\text{judge}}$:

$$\hat{\kappa} = \arg\min_{\kappa} \mathcal{L}_{\text{judge}}(\kappa) \tag{4.19}$$

which gives an unsupervised method of selection of the complexity of the approximation.

[2] Although it is possible to choose other criteria.

[3] *i.e.* the precision factors that fall outside or are on the diagonal.

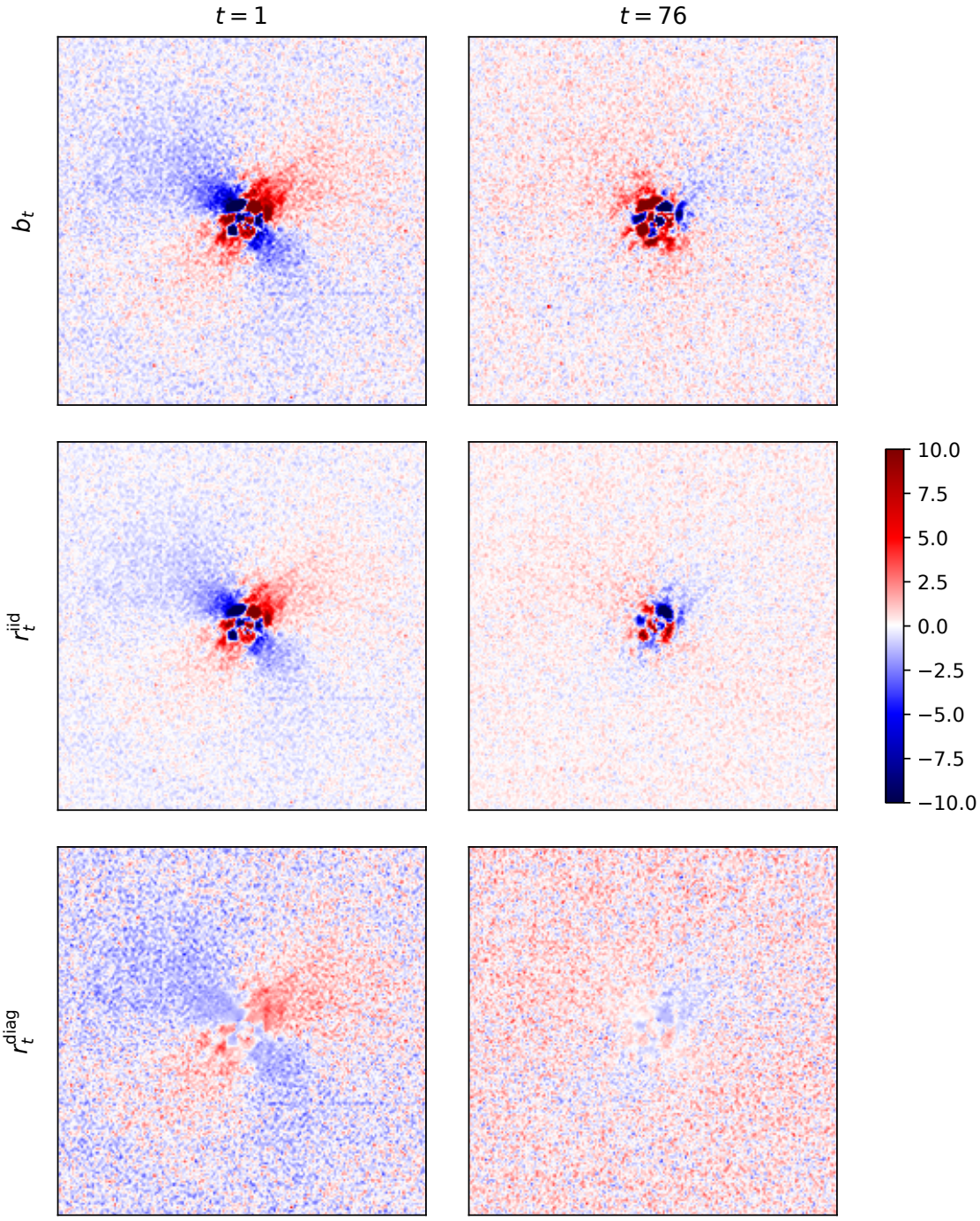

Figure 4.3 – Instances of residuals $\boldsymbol{b}_t$ (top line) and their whitened residuals $\boldsymbol{r}_t$ resulting from an i.i.d. approximation discussed in Section 4.4.2 (center line) and a diagonal covariance introduced in Section 4.4.3 (bottom line) of the covariance matrix.

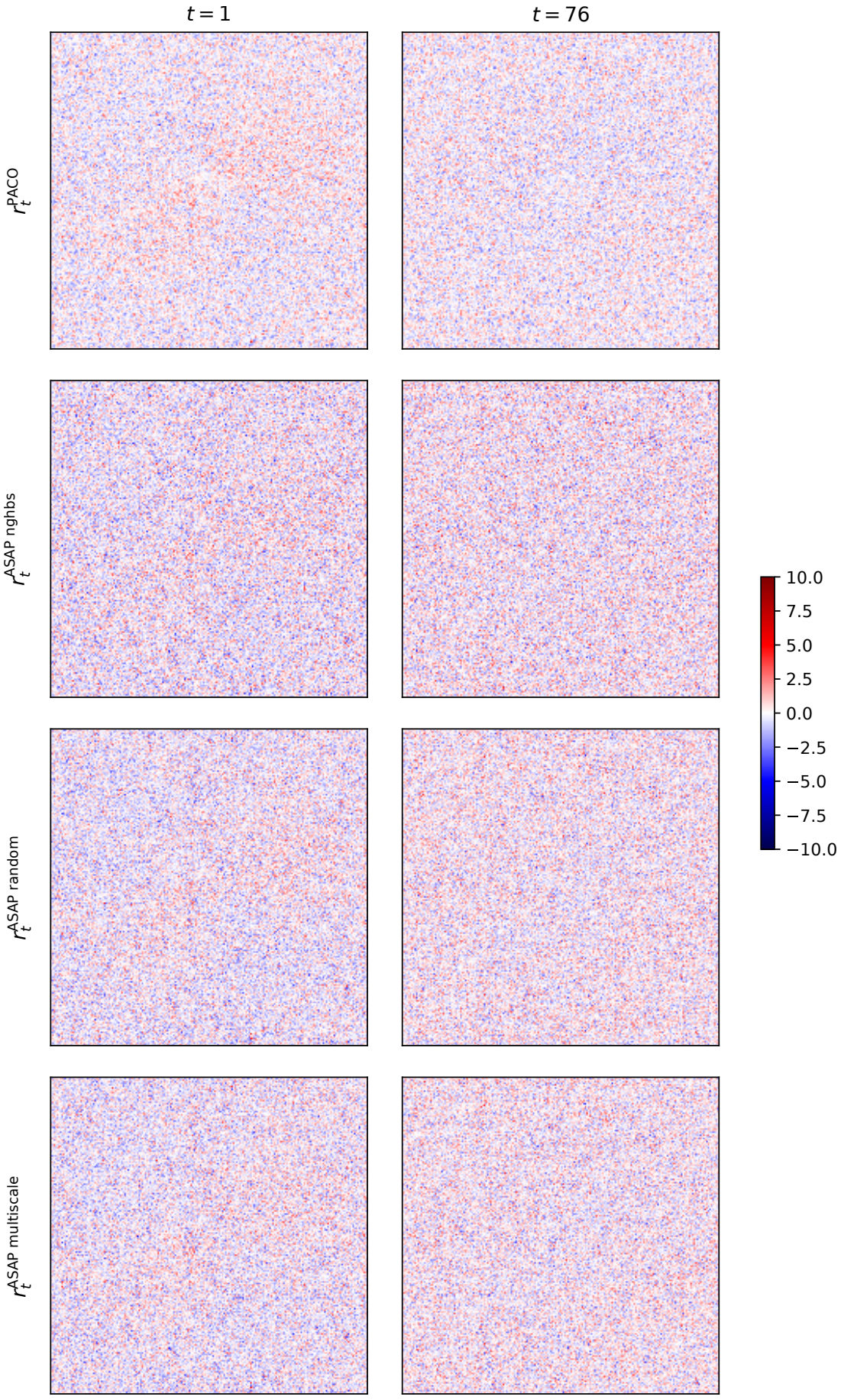

Figure 4.4 – Instances of whitened residuals resulting from different covariance models. The top line presents frames of the residuals obtained with PACO (Section 4.4.4) while the other lines are obtained with the ASAP method and the selection of nearest neighbors without permutation (Section 4.4.6), with a random ordering or a multiscale selection (Section 4.4.7).

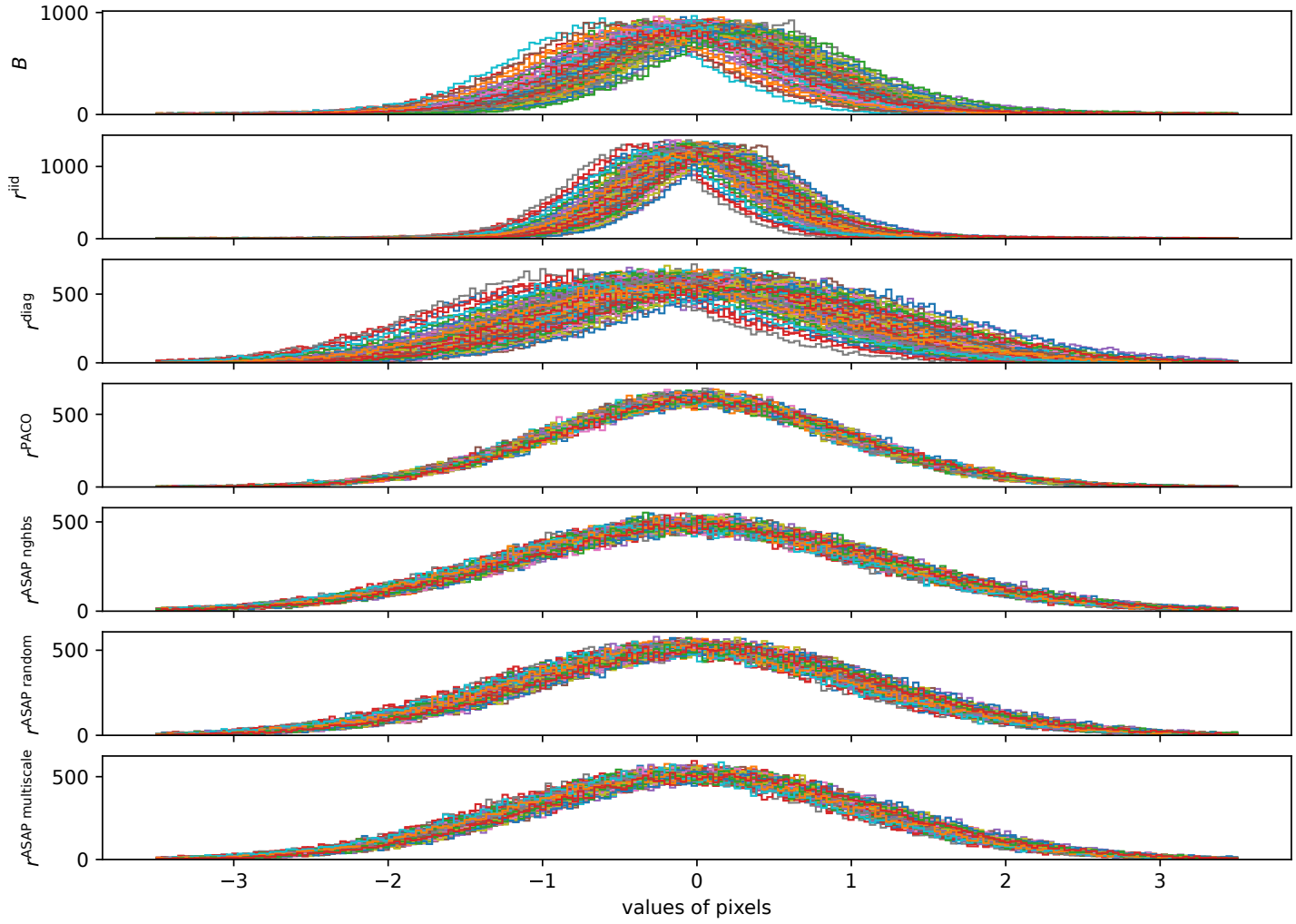

Figure 4.5 – Histograms (upper figure) of the residuals of each frame of the HIP 80019 dataset. The following rows are the histograms of the whitened residuals $\boldsymbol{r}_t$, depending on the chosen covariance model.

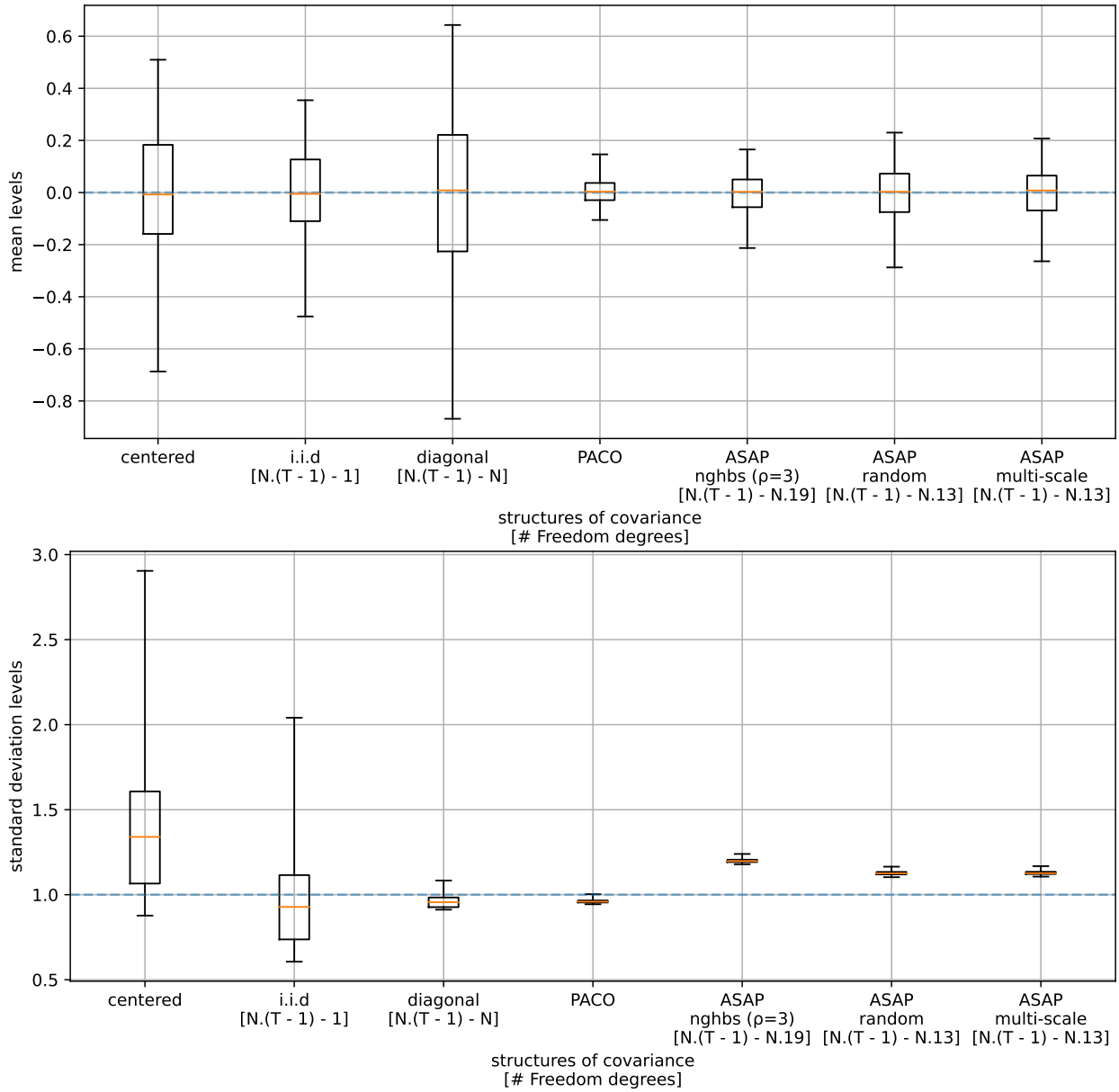

Figure 4.6 – Box plots of the results of the LOOCV study on the HIP 80019 dataset summarizing Fig. 4.5. It shows the means (top) and standard deviations (bottom) of the residuals $\boldsymbol{b}_t$ and its whitened counterparts $\boldsymbol{r}_t$ for different covariance models and $\forall t \in [\![1, T]\!]$. The segments' ends correspond to the minimum and maximum value of the sets while the ends of the boxes to their first and third quartile. The orange lines indicate their median. The mean and standard deviation of the standard normal distribution (respectively 0 and 1) are plotted as blue dashed lines.

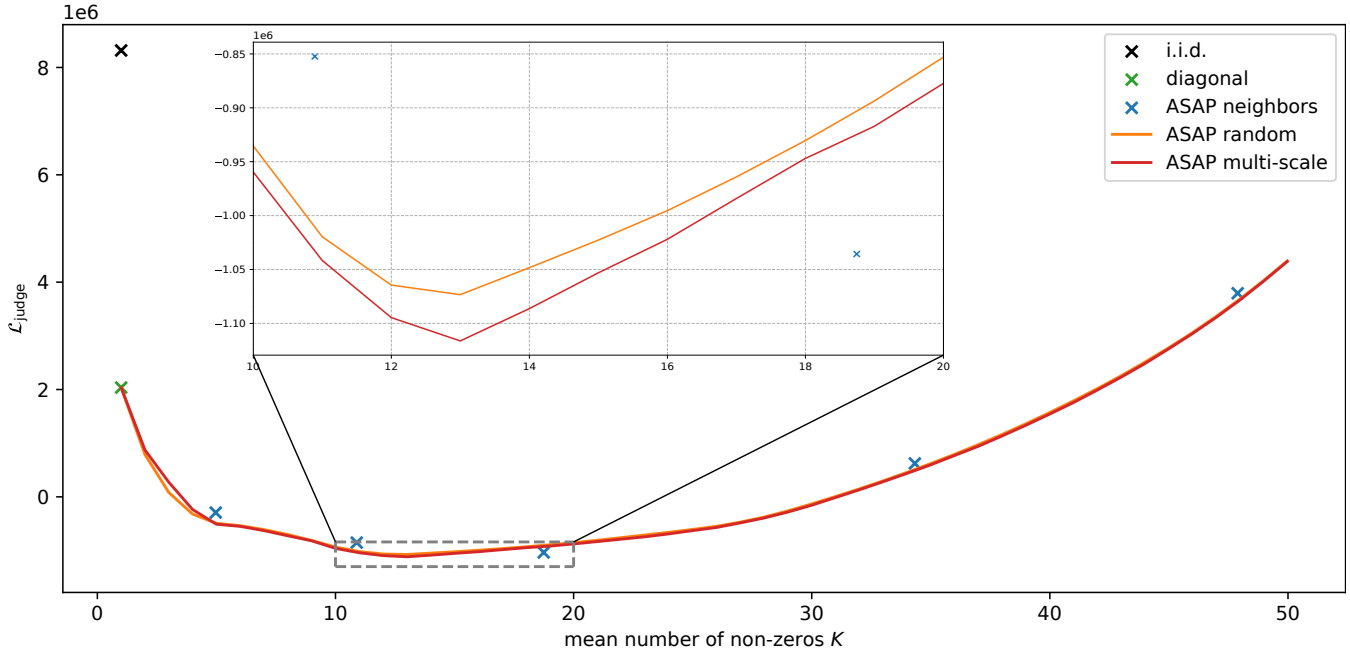

Figure 4.7 – LOOCV likelihood score function $\mathcal{L}_{\text{judge}}$ of Eq. (4.18) computed on the HIP 80019 dataset (N = 40804, T = 114) for multiple amounts of non-zero elements K of co-precision coefficients and different selecting methods. The black cross shows the criterion when considering the stochastic component as i.i.d. (defined in section 4.4.2). The green cross shows the criterion when considering a diagonal structure for the covariance (discussed in Section 4.4.3). The blue samples represent the score when selecting the set $\mathcal{V}_i$ by number of neighbors (section 4.4.6). The random permutation-based selection is shown in orange, while the multi-scale selection method results are shown in red (section 4.4.7).

4.4.2 Constraining the covariance to be proportional to the identity matrix

A very first approximation of the estimator of $\mathbf{C}$ is to consider a diagonal structure, proportional to the identity matrix. Meaning that the stochastic components $\boldsymbol{b}_t$ are supposed independent and identically distributed (i.i.d.) of covariance $\mathbf{C} = \sigma^2 \mathbf{I}$ with $\sigma \geqslant 0$. Eq. (4.11) then becomes:

$$\begin{aligned}\hat{\sigma}^2(\mathbf{B}) &= \underset{\sigma^2 \geqslant 0}{\arg\min} \left\{ \frac{1}{\sigma^2} \sum_{t=1}^{T} \left\| \boldsymbol{b}_t \right\|^2 + TN \log \sigma^2 \right\} \\ &= \frac{1}{TN} \sum_{t=1}^{T} \left\| \boldsymbol{b}_t \right\|^2 \end{aligned} \tag{4.20}$$

In this case, the estimator of the covariance is the sampled variance. Applying the LOOCV to the HIP80019 dataset, the set of whitened residuals $\{\boldsymbol{r}^{\text{iid}\,[m]}\}_{m \in [\![1,T]\!]}$ result from Eq. (4.17):

$$\boldsymbol{r}^{\text{iid}\,[m]} = \frac{\boldsymbol{b}_m}{\hat{\sigma}^{[m]}} \tag{4.21}$$

with

$$\hat{\sigma}^{[m]\,2} = \frac{1}{(T-1)N} \sum_{\substack{t=1 \\ t \neq m}}^{T} \left\| \boldsymbol{b}_t \right\|^2 \tag{4.22}$$

which consists only of a rescaling of the residuals.

Instances of the resulting whitened residuals are presented in the second row of Fig. 4.3 and the histograms of each frame of the whitened residuals are presented in Fig. 4.5. The results of the LOOCV study applied with this model are presented in Fig. 4.6 (second entry, "i.i.d.") and in Fig. 4.7. Obviously, this approximation of the covariance does not explain the distribution of the residuals: a lot of spatial structures are still visible in the residuals in Fig. 4.3, especially close to the coronagraphic mask. The means and standard deviations of the frames in the LOOCV study of Fig. 4.6 are distributed over a vast range of values, indicating the need for a better model.

4.4.3 Constraining the covariance to be diagonal

To reduce the inhomogeneity of the residuals, a more flexible model is needed. We consider here a simple diagonal approximation of the covariance, *i.e.* an approximation in which a different variance is considered at each pixel of the residuals $\boldsymbol{b}_t$. Under this hypothesis, the covariance matrix writes $\mathbf{C} = \text{diag}(\boldsymbol{c})$ with $\text{diag} : \mathbb{R}^N \to \mathbb{R}^{N \times N}$ the operator that forms a diagonal matrix whose diagonal elements are the elements of the vector $\boldsymbol{c} \in \mathbb{R}^N$. Noticing that the problem is separable in each variance c_n of $\boldsymbol{c}$:

$$\begin{aligned}\hat{c}_n &= \underset{c_n \geqslant 0}{\arg\min} \sum_{t=1}^{T} \sum_{k=1}^{N} \frac{1}{c_k} (b_{t,k})^2 + T \sum_{k=1}^{N} \log c_k \\ &= \frac{1}{T} \sum_{t=1}^{T} (b_{t,n})^2 \end{aligned} \tag{4.23}$$

which corresponds to the sample variances of each pixel of the dataset.

Following the same LOOCV, we deduce the new set of residuals

$$\boldsymbol{r}^{\text{diag}\,[m]} = \text{diag}(\hat{\boldsymbol{c}}^{[m]\,-\frac{1}{2}})\, \boldsymbol{b}_m \tag{4.24}$$

with $\hat{\boldsymbol{c}}^{[m]-\frac{1}{2}}$ the element-wise inverse square root of the m-th model variance $\hat{\boldsymbol{c}}^{[m]}$. As before, the results are presented in Figs. 4.3 (third row) and 4.5 (third row). Thanks to this more complex model, the whitened residuals close to the coronagraphic mask are severely reduced compared to those given under an i.i.d. hypothesis. As shown in Fig. 4.6 (third entry, "diagonal") the standard deviations of the whitened frames in the LOOCV study are spread in a smaller range of value and closer to one (except for a few frames where the observation conditions are very poor). The co-log-likelihood criterion is also severely reduced by this approximation, compared to the i.i.d. approximation.

However, some spatial structures from the speckles are still clearly recognizable in the third row of Fig. 4.3, indicating the need for an even more flexible model of $\mathbf{C}$. To take into account these structures, the model of $\mathbf{C}$ must include covariances, which will grasp the spatial correlations coming from the speckles. As said in section 4.3.4, taking all covariances into account is beyond reach. It is necessary to impose some structure on $\mathbf{C}$, for example by considering only sparse matrices to greatly reduce the effective number of degrees of freedom (ENDoF) that have to be estimated.

4.4.4 Modeling the covariances with a patch-based approach

Using the PACO (for PAtch COvariance) methodology (Flasseur, O. et al., 2018) applied to the reconstruction of extended objects, we introduced in Flasseur, O. et al. (2021) a patch-based model of the covariance matrix. It takes into account local correlations, capturing the covariances inside non-overlapping small patches of a few tens of pixels. In other words, each frame of the data is sliced into small K pixel patches ($K \ll N$), the patches, that are used to build local statistics, as seen in Figure 4.8.

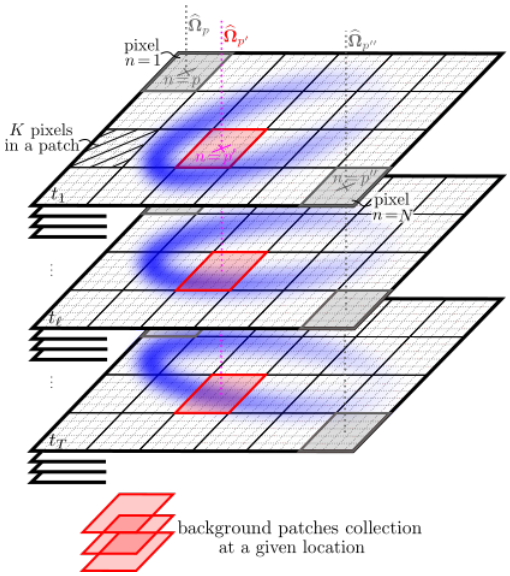

Figure 4.8 – Extraction of patch collections by local learning of the statistics of the nuisance component from ADI sequences. The contribution of off-axis sources is shown in blue. Ω_p designates the statistics $(\hat{\boldsymbol{m}}_\star, \hat{\mathbf{C}}_p)$ of the p-th patch. Source: Flasseur, O. et al. (2021)

Doing so, the resulting covariance matrix has a block diagonal structure characterized by the set of local covariances $\{\hat{\mathbf{C}}_p\}_{p\in\mathbb{P}}$, with $\mathbb{P}$ the set of central indices of the patches. Each small covariance block $\hat{\mathbf{C}}_p$ is learned on the patch extracted at location p from the dataset $\mathbf{P}_p\,\boldsymbol{b}_t, \forall t \in [\![1, T]\!]$, with $\mathbf{P}_p$ the operator extracting the p^{th} patch of K pixels from $\{\boldsymbol{b}_t\}_{t\in[\![1,T]\!]}$. To adapt to the local non-stationarity of the spatial correlations of the speckles, PACO chooses a trade-off between two estimates: an unbiased estimator with large variance and an estimator of reduced variance but which introduces a bias. Using the shrinkage estimator $\breve{\mathbf{C}}_p \in [0, 1]$ (Ledoit and Wolf, 2004), the covariance writes:

$$\hat{\mathbf{C}}_p = (1 - \breve{\rho}_p)\,\breve{\mathbf{C}}_p + \breve{\rho}_p\ \mathrm{diag}(\mathrm{diag}(\breve{\mathbf{C}}_p)) \tag{4.25}$$

with

$$\breve{\mathbf{C}}_p = \frac{1}{T}\sum_{t=1}^{T} \mathbf{P}_p\, \boldsymbol{b}_t\, \boldsymbol{b}_t^{\mathsf{T}}\, \mathbf{P}_p^{\mathsf{T}} \tag{4.26}$$

the full sample covariance estimator of the dataset $\boldsymbol{b}_t \in \mathbb{R}^T$ constrained to the small patch p, and where $\mathrm{diag}(\breve{\mathbf{C}}_p)$ selects the diagonal elements of $\breve{\mathbf{C}}_p$ and so $\mathrm{diag}(\mathrm{diag}(\breve{\mathbf{C}}_p))$ builds a diagonal matrix which diagonal elements are the diagonal elements of $\breve{\mathbf{C}}_p$.

Following the work of Ledoit and Wolf, 2004, Chen et al., 2010, Flasseur, O. et al., 2018 define the best possible shrinkage factor as:

$$\breve{\rho}_p = \arg\min_{\rho_p} \mathrm{E}\left[\|\mathbf{C} - \hat{\mathbf{C}}_p\|_{\mathrm{F}}^2\right] \tag{4.27}$$

with $\mathbf{C}$ the true covariance matrix of the residuals $\boldsymbol{b}_t$. As in practice, $\mathbf{C}$ is unknown, it is replaced in a recursive scheme to replace it with a previous estimate of $\hat{\mathbf{C}}_p$. This procedure converges to a fixed point:

$$\breve{\rho}_p = \frac{\mathrm{tr}(\breve{\mathbf{C}}_p^2) + \mathrm{tr}^2(\breve{\mathbf{C}}_p) - 2\sum_{k=1}^{K}[\breve{\mathbf{C}}_p]_{k,k}^2}{(T+1)\left(\mathrm{tr}(\breve{\mathbf{C}}_p^2) - \sum_{k=1}^{K}[\breve{\mathbf{C}}_p]_{k,k}^2\right)} \tag{4.28}$$

which is used to tune automatically the shrinkage. Note that this parameter adapts itself to the local statistics and so is very flexible to spatially correlated data such as the one studied here. To best evaluate the local covariance, the patch size must be large enough to contain the core of the PSF of the instrument but not too large so that the number of covariance values to estimate is less than the number of frames T. In practice, PACO chooses a patch twice the size of the PSF full width at half maximum, typically leading to $K \simeq 50$ pixels (7×7 patches).

Like for the two previous covariance models, we chose to study the PACO model with the LOOCV. Taking the inverse Cholesky factorization as explained in Eq. (4.15) yields the $\boldsymbol{r}^{\mathrm{PACO}\,[m]}$ residuals shown in Figs. 4.4 (first column), 4.5 (fourth row) and 4.6 (fourth entry). Taking into account the spatial correlations leads to whitened residuals with almost no visible residual structures. Note that the ENDoF for the PACO method is difficult to estimate as it depends directly on the shrinkage. The patched covariance matrix $\hat{\mathbf{C}}_p$ can go from K ENDoF when $\breve{\rho}_p = 1$ to $K\,(K+1)/2$ when $\breve{\rho}_p = 0$.

4.4.5 ASAP- A generic sparse structure of the precision matrix W

However, as the structures in the residuals suggest, the spatial correlations can be spread on farther distances than a patch. We are then looking for an alternative method than PACO, that can be more flexible in its ENDoF to reach correlations at farther distances. An approximation of the covariance can be obtained by working on the estimation of the precision matrix $\mathbf{W} = \mathbf{C}^{-1}$. This approach allows us to work directly with $\mathbf{W}$ in the reconstruction problem of Eq. (4.14), without having to inverse our estimation of the covariance.

We developed in Thiébaut et al., 2022 the ASAP approximation to address this limitation. ASAP is an estimation method that uses a general formalism to learn correlations in the maximum likelihood sense. It is working directly on a sparse factor of the Cholesky factorization $\mathbf{V}$, a lower triangular matrix of positive diagonal entries. Considering the decomposition $\mathbf{W} = \mathbf{V}^{\mathsf{T}}\,\mathbf{V}$, it is possible to re-write the estimation problem of Eq.(4.12) as the minimization of a new loss function:

$$\hat{\mathbf{V}} = \arg\min_{\mathbf{V}\in\mathbb{V}} \left\{ \mathcal{L}_{\mathrm{ASAP}}(\mathbf{B},\mathbf{V}) = \mathrm{tr}\left(\mathbf{V}\,\mathbf{S}\,\mathbf{V}^{\mathsf{T}}\right) - \log\left|\mathbf{V}^{\mathsf{T}}\,\mathbf{V}\right| \right\} \tag{4.29}$$

with $\mathbb{V}$ a set of lower triangular matrices that can ensure a specific sparse structure, $\mathbf{B}$ the dataset of stochastic components $\boldsymbol{b}_t, \forall t \in [\![1, T]\!]$, and $\mathbf{S}$ the sample covariance matrix defined by Eq. (4.31).

Proof: Recalling Eq. (4.12):

$$\hat{\mathbf{W}}(\mathbf{B}) = \underset{\mathbf{W}\in\mathbb{W}}{\arg\min}\Big(\sum_{t=1}^{T}\|\boldsymbol{b}_t\|^2_{\mathbf{W}} - T\,\log|\mathbf{W}|\Big) \qquad \text{(recall of Eq. (4.12))}$$

it is possible to rewrite the function to minimize as:

$$\begin{aligned}
\sum_{t=1}^{T}\|\boldsymbol{b}_t\|^2_{\mathbf{W}} - T\,\log|\mathbf{W}| &= \sum_{t=1}^{T}\|\boldsymbol{b}_t\|^2_{\mathbf{V}^\mathsf{T}\mathbf{V}} - T\,\log|\mathbf{V}^\mathsf{T}\mathbf{V}| \\
&= \sum_{t=1}^{T}\|\mathbf{V}\,\boldsymbol{b}_t\|^2 - T\,\log|\mathbf{V}^\mathsf{T}\mathbf{V}| \\
&= \sum_{t=1}^{T}\mathrm{tr}\Big(\mathbf{V}\,\boldsymbol{b}_t\,\boldsymbol{b}_t^\mathsf{T}\,\mathbf{V}^\mathsf{T}\Big) - T\,\log|\mathbf{V}^\mathsf{T}\mathbf{V}| \\
&= \mathrm{tr}\Big(\mathbf{V}\sum_{t=1}^{T}(\boldsymbol{b}_t\,\boldsymbol{b}_t^\mathsf{T})\,\mathbf{V}^\mathsf{T}\Big) - T\,\log|\mathbf{V}^\mathsf{T}\mathbf{V}| \\
&= T\,\Big(\mathrm{tr}\Big(\mathbf{V}\,\mathbf{S}\,\mathbf{V}^\mathsf{T}\Big) - \log|\mathbf{V}^\mathsf{T}\mathbf{V}|\Big) \qquad (4.30)
\end{aligned}$$

with

$$\mathbf{S} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{b}_t\,\boldsymbol{b}_t^\mathsf{T} \qquad (4.31)$$

the sample covariance matrix. ■

Eq. (4.29) is equivalent to minimizing the Kullback-Leibler divergence between a referenced centered Gaussian distribution of covariance matrix $\mathbf{S}$ and another centered Gaussian distribution of covariance matrix $(\mathbf{V}^\mathsf{T}\mathbf{V})^{-1}$ (Thiébaut and Tallon, 2010; Schäfer, Katzfuss, and Owhadi, 2019). $\mathbf{V}$ being a lower triangular matrix, Eq. (4.29) has a closed form solution but, given the very large number N of pixels, estimating the entire precision matrix ($N\,(N+1)/2$ elements) is not possible considering only $T\,N \ll N\,(N+1)/2$ elements of the data. The solution, as in the PACO method, is to define a sparse approximation of $\mathbf{W}$ which will evaluate the co-precision factors at only given pixels. By defining for the i-th row of $\mathbf{V}$, the set $\mathcal{V}_i$ of column indices (including i) of the K_i structural non-zero elements, Eq.(4.29) becomes:

$$\begin{aligned}
\mathcal{L}_{\mathrm{ASAP}}(\mathbf{B},\mathbf{V}) &= \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} V_{i,j}\,S_{j,k}\,V_{k,i} - \sum_{i=1}^{N}\log(V_{i,i}^2) \\
&= \sum_{i=1}^{N}\Big(\boldsymbol{V}_{i,\mathcal{V}_i}\,\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}\,\boldsymbol{V}_{i,\mathcal{V}_i}^\mathsf{T} - \log(V_{i,i}^2)\Big) \qquad (4.32)
\end{aligned}$$

with $\boldsymbol{V}_{i,\mathcal{V}_i} \in \mathbb{R}^{1\times K_i}$ a row vector containing the non-zero elements of the i^{th} row of $\mathbf{V}$ and $\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}$ the restriction to the $\mathcal{V}_i$ rows and columns of the sample covariance matrix $\mathbf{S}$. As Eq. (4.32) suggests, solving the problem (4.29) can be done independently for each i, leading to separable estimation of each row of $\mathbf{V}$.

The gradient of Eq. (4.32) writes:

$$\frac{\partial\mathcal{L}_{\mathrm{ASAP}}}{\partial\boldsymbol{V}_{i,\mathcal{V}_i}} = 2\,\boldsymbol{V}_{i,\mathcal{V}_i}\,\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i} - \frac{2}{V_{i,i}}\,\boldsymbol{I}_{i,\mathcal{V}_i} \qquad (4.33)$$

with $\boldsymbol{I}_{i,\mathcal{V}_i}$ a row vector containing the restriction of the identity matrix to the i^{th} row and the non-zero column elements defined by $\mathcal{V}_i$. Its elements are all zero except for the last one corresponding to the diagonal element (which is 1). Taking the Hessian of Eq. (4.32) with respect to $\boldsymbol{V}_{i,\mathcal{V}_i}$:

$$\frac{\partial^2 \mathcal{L}_{\text{ASAP}}}{\partial \boldsymbol{V}_{i,\mathcal{V}_i}^2} = 2\,\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i} + \frac{2}{V_{i,i}^2}\,\boldsymbol{I}_{i,\mathcal{V}_i}^{\mathsf{T}}\,\boldsymbol{I}_{i,\mathcal{V}_i} \tag{4.34}$$

we can deduce that if $V_{i,i} \neq 0$ and the length of $\mathcal{V}_i$, K_i is small, the result of (4.34) is a positive-semi-definite matrix. Indeed, as $\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}$ is the restriction of a positive-semi-definite matrix, it is positive-semi-definite if its size does not reach the number of frames T, and $\boldsymbol{I}_{i,\mathcal{V}_i}^{\mathsf{T}}\,\boldsymbol{I}_{i,\mathcal{V}_i}$ is by construction a rank-one nonnegative and symmetric matrix. If $\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}$ is invertible, the function $\mathcal{L}_{\text{ASAP}}$ to be minimized is strictly convex with respect to $\boldsymbol{V}_{i,\mathcal{V}_i}$. The corresponding minimum is expressed as:

$$\hat{\boldsymbol{V}}_{i,\mathcal{V}_i} = \frac{1}{\rho_i}\,\boldsymbol{I}_{i,\mathcal{V}_i}\,\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}^{-1} \tag{4.35}$$

with $\rho_i = \hat{V}_{i,i}$ the i^{th} diagonal element of $\hat{\mathbf{V}}$. Noting that $\hat{V}_{i,i} = \hat{\boldsymbol{V}}_{i,\mathcal{V}_i}\,\boldsymbol{I}_{i,\mathcal{V}_i}^{\mathsf{T}}$,

$$\begin{aligned}\rho_i &= \frac{1}{\rho_i}\,\boldsymbol{I}_{i,\mathcal{V}_i}\,\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}^{-1}\,\boldsymbol{I}_{i,\mathcal{V}_i}^{\mathsf{T}} \\ &= \frac{1}{\rho_i}\,\|\boldsymbol{I}_{i,\mathcal{V}_i}^{\mathsf{T}}\|_{\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}^{-1}}^2\end{aligned} \tag{4.36}$$

which yields

$$\rho_i = \pm\|\boldsymbol{I}_{i,\mathcal{V}_i}^{\mathsf{T}}\|_{\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}^{-1}} \tag{4.37}$$

and so:

$$\boxed{\hat{\boldsymbol{V}}_{i,\mathcal{V}_i} = \frac{\pm\boldsymbol{I}_{i,\mathcal{V}_i}\,\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}^{-1}}{\|\boldsymbol{I}_{i,\mathcal{V}_i}^{\mathsf{T}}\|_{\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}^{-1}}}} \tag{4.38}$$

No matter the sign of ρ_i, the same approximation $\mathbf{V}^{\mathsf{T}}\,\mathbf{V}$ is produced. The computational complexity of this closed-form expression is in $\mathcal{O}(K_i^3)$ as it depends on the number K_i of non-zero elements chosen by the set $\mathcal{V}_i$ in the sparse structure. But given the fact that this is a closed-form solution and that it is separable with respect to i, the computational cost of solving Eq. (4.38) can be shared in a parallel computational strategy.

Observing Eq. (4.32), it is possible to make the changing of parameters defined in Eq. (4.16) so that it depends on the covariance Cholesky factorization instead of the precision Cholesky factorization:

$$\mathcal{L}_{\text{ASAP cov}}(\boldsymbol{b},\mathbf{L}) = \sum_{i=1}^{N}\Big(\boldsymbol{L}_{i,\mathcal{V}_i}^{-1}\,\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}\,\boldsymbol{L}_{i,\mathcal{V}_i}^{-\mathsf{T}} + \log(L_{i,i}^2)\Big) \tag{4.39}$$

Obtaining the gradient of $\mathcal{L}_{\text{ASAP cov}}$ with respect to $\mathbf{L}$ is not easy. In other words, estimating the Cholesky factorization of the covariance does not yield a closed-form solution. Contrary to working with the precision matrix, it then requires an iterative optimization method to minimize the criterion.

4.4.6 Selecting close-by neighbors as non-zero in the sparse approximation of the precision matrix

To have a good approximation of the precision matrix as well as a sparse structure, the design of the patterns of non-zero in the sparse matrix (the set $\mathcal{V}_i$) must be done carefully.

A first selection strategy, similar to PACO, can be to consider the neighbors of each pixel within a distance of a few pixels. Defining a threshold distance ρ from a given pixel i, this strategy selects the set of pixels that are located before the i-th index, in lexicographic order, and not farther away than ρ from that pixel. In other words, the set of non-zero elements can be defined as:

$$\mathcal{V}_i = \left\{ j \in [\![1, i]\!] \,|\, \mathcal{D}(i,j) \leqslant \rho \right\} \tag{4.40}$$

with $\mathcal{D}(i,j)$ denoting the Euclidean distance between the two pixels of indices i and j. This definition describes a half disk set of *parents* pixels that are accounted for in the estimation of Eq.(4.38), as shown in Fig. 4.9. Note that this selection strategy is quite similar to PACO in the case of extended objects reconstruction, *i.e.* REXPACO (Flasseur, O. et al., 2021), with the difference that ASAP uses a shifting window to compute the elements of the precision.

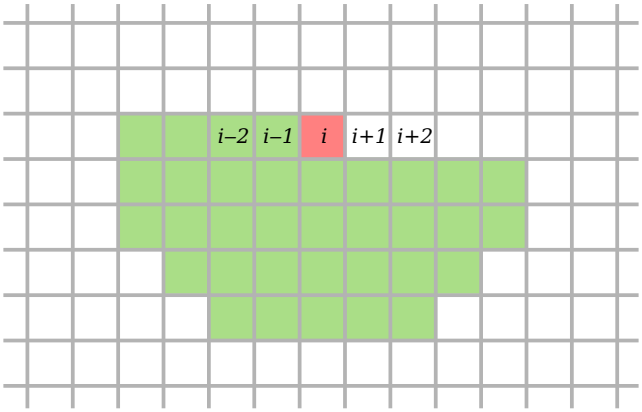

Figure 4.9 – Schematic of the selected parents (in green) of a pixel at index i (in red), in its neighborhood. Source: Thiébaut et al. (2022)

Using the LOOCV study framework, it is possible to evaluate the evolution of the whitening process on the HIP 80019 tested dataset. The whitened residuals are obtained by:

$$\boldsymbol{r}^{\text{ASAP nghbs}\,[m]} = \hat{\mathbf{V}}^{[m]}\, \boldsymbol{b}_m \tag{4.41}$$

with the rows of $\hat{\mathbf{V}}^{[m]}$ structured as Eq. (4.40) dictates and its non-zero elements learned by Eq. (4.38).

Fig. 4.7 presents in blue crosses the co-log-likelihood of Eq. (4.18) for different values of the radius ρ. It indicates the evolution of the describing performances of the ASAP model when taking more, or less, neighbors into account in the estimation of the spatial correlations. Replacing the generic term of complexity κ of Eq. (4.19) by the radius ρ, we obtain an unsupervised method of selection of the number of parents to take into account in the computation of the Cholesky factor $\mathbf{V}$. This allows the method not to rely on a shrinkage approach to improve the sparsity of the covariance matrix. Also, the restricted sampled covariance $\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}$ may be rank deficient if too many parents are selected to participate in Eq. (4.38), leading to a computational error in its inversion. Our method also limits these computational problems.

According to this criterion, a distance of $\rho = 3$ yields the best approximation of the precision matrix. Lower radius values are not enough to grasp the correlations, while if too many parents are selected, the model tends to over-fit the data and lose generality. The results of the whitening process with $\rho = 3$ are presented in Figs. 4.4 (second column) and 4.5 (fifth row) and the comparison of the LOOCV study is shown in Fig. 4.6 (fifth entry, "ASAP nghbs"). Like for PACO, compared to the structures proportional to identity and diagonal, the ASAP learning method allied with this neighborly structure is strongly reducing the spatial features of the speckles. There is however a drawback in the standard deviation of the whitened residuals with a value slightly higher than expected. Still, the co-log-likelihood is improved by this method compared to the others indicating a better description power of ASAP.

4.4.7 Accounting for long range spatial correlations

To take into account correlations at farther distances, one could increase the radius ρ of Eq. (4.40). However, increasing too much ρ can lead to a non-invertible $\mathbf{S}_{\mathcal{V}_i,\mathcal{V}_i}$. As we want to keep the formalism brought by ASAP, the next idea is to *pick* parents at different distances from the pixel. This can be achievable thanks to a permutation.

The set $\mathcal{V}_i$ of non-zero elements taken into account in the computation of the correlations of the i-th pixel does not need to depend on the lexicographic order of the pixels. We define in the following the permutation operator $\mathbf{P}$ builds via the list of ordered indices $\boldsymbol{p}$ which modify the ordering of the pixels in each frame of $\mathbf{B}$:

$$\forall t \in [\![1, T]\!], \forall n \in [\![1, N]\!], [\mathbf{P}\, \boldsymbol{b}_t]_n = b_{t,p_n} \tag{4.42}$$

where b_{t,p_n} is the p_n pixel of $\boldsymbol{b}_t$ (see Fig. 4.10). $\mathbf{P}$ is invertible and its inverse is its transpose matrix, *i.e.* $\mathbf{P}^{-1} = \mathbf{P}^\mathsf{T}$. Given that $\boldsymbol{b}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$, the permuted residual frames also follow a multivariate Gaussian distribution:

$$\mathbf{P}\, \boldsymbol{b}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{P}\, \mathbf{C} \mathbf{P}^\mathsf{T}) \tag{4.43}$$

which gives the permuted precision matrix $\widetilde{\mathbf{W}} = \mathbf{P}^\mathsf{T}\, \mathbf{C}^{-1}\, \mathbf{P}$ and so the permuted Cholesky factor to estimate with ASAP, $\widetilde{\mathbf{V}}$ such that $\widetilde{\mathbf{W}} = \widetilde{\mathbf{V}}^\mathsf{T}\, \widetilde{\mathbf{V}}$. As can be seen, a permutation does not affect the ASAP estimation method but allows us to select parents of pixels without following the lexicographic order and so account for farther spatial correlations.

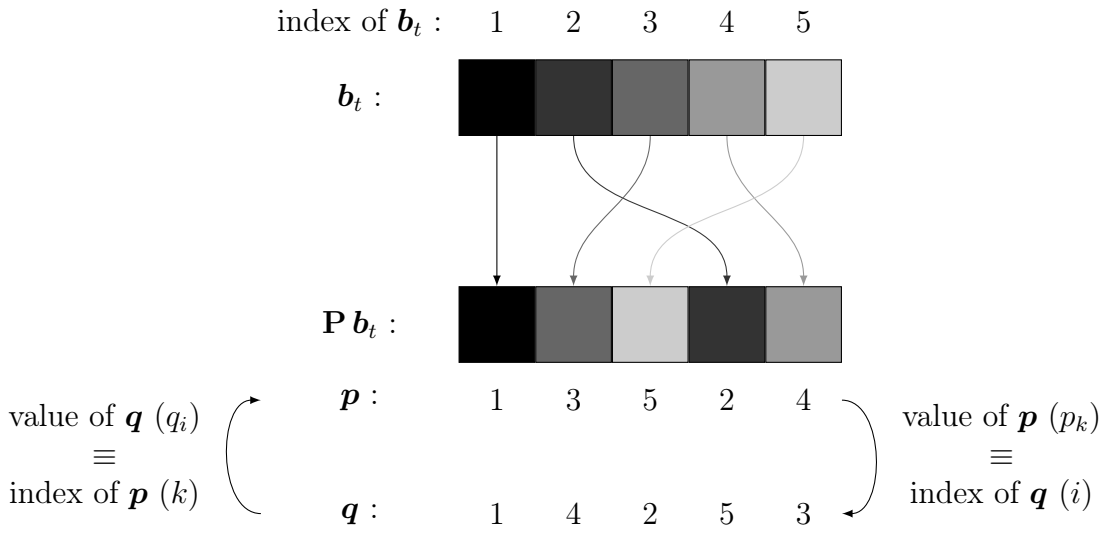

Figure 4.10 – Exemple of permutation by $\boldsymbol{p}$ with $\boldsymbol{q}$ the rank of each of the permuted pixels in $\boldsymbol{p}$.

With this permutation, it is possible to define the same kind of selection process as explained in Section 4.4.6. Like with Eq. (4.40), we want to define a set of parents that are limited by the maximum length of the set K. To do so, it is necessary to define the ranking $\boldsymbol{q}$ of the pixels in the permutation (see Fig. 4.10) such that:

$$\forall k \in [\![1, N]\!],\ q_{p_k} = k \tag{4.44}$$

That is the value of the i-th element of the ranking vector, q_i, represents the index of the i-th pixel of $\boldsymbol{b}_t$ in the permutation $\boldsymbol{p}$:

$$\forall i \in [\![1, N]\!],\ p_{q_i} = i \tag{4.45}$$

Introducing the ranking matrix $\mathbf{J} \in \mathbb{R}^{K \times N}$ such that each of its column i represents the nearest parents of the pixel i taking into account the permutation, we describe in Thiébaut et al. (2022) an

algorithm to build such a matrix which ensures that:

$$J_{\ell,i} = \begin{cases} -1 & \text{if the } \ell\text{-th parent has not been chosen yet} \\ j & \text{otherwise and with } q_j \leqslant q_i \end{cases} \tag{4.46}$$

while respecting the order defined by:

$$\forall \ell \in [\![1, K_i]\!],\ \mathcal{D}(i, J_{\ell,i}) \geqslant \mathcal{D}(i, J_{\ell+1,i}) \tag{4.47}$$

with $K_i \leqslant K \ll N$, K_i being the length of selected parents of the i-th pixel, K the maximum number of parents acceptable and N the number of pixels of one frame of the stochastic component $\mathbf{B}$. Doing so, $(J_{1,i}, J_{2,i}, \ldots, J_{K_i,i})$ is the ordered list of the nearest parents of i, with $J_{K_i,i} = i$. The set of parents chosen for Eq. (4.38) is then defined by:

$$\mathcal{V}_i = \left\{ J_{[\![1,K_i]\!],i} \right\} \tag{4.48}$$

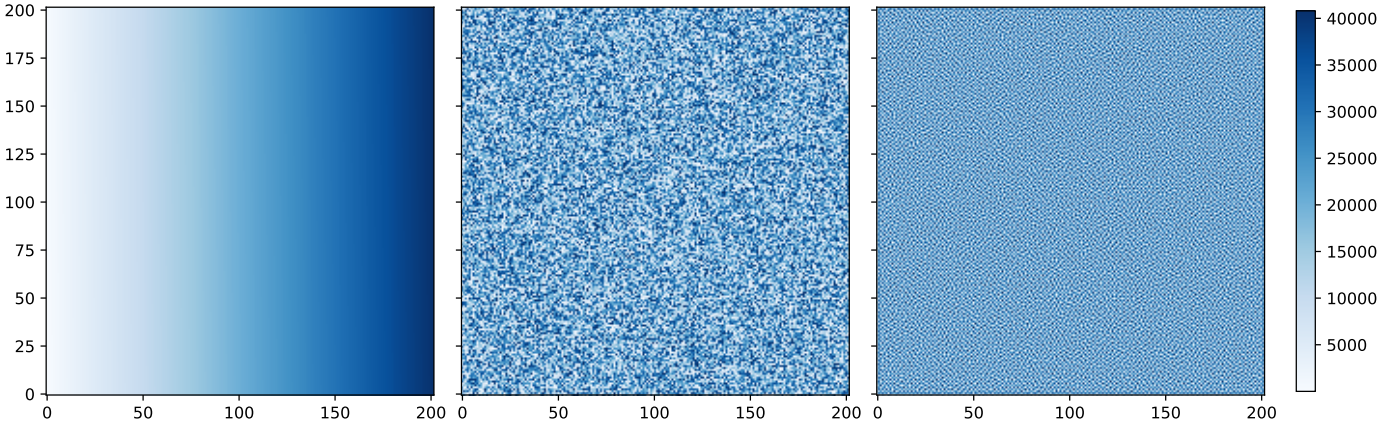

Figure 4.11 – Display the rank of each pixel as shades of blue from the smaller (brighter) to the higher (darker). Three cases are shown: the usual lexicographic order used in Section 4.4.6 (left figure), the order dictated by a random permutation (center figure) and the multi-scale ordering (right figure).

Accounting spatial correlations at random range

From the definitions of Eqs. (4.42) and (4.48), one can choose a random permutation operator. Figure 4.11 shows such permutation in the center figure, representing the ranking of each pixel as shades of blue.

Using the same LOOCV study and score function defined in Eqs. (4.17) and (4.18), Fig.4.7 shows the resulting co-log-likelihood for a randomly ordered permutation (orange line). The criterion shows a better minimization of the co-log-likelihood for fewer parents selected ($K_i = 13$ instead of 19 for the neighborhood selection).

Figure 4.4 presents the resulting whitened residuals via the LOOCV strategy as presented in Eq. (4.41) for $K_i = 13$. The histograms of the frames and resulting means and standard deviations are presented in Figs. 4.5 and 4.6. As can be seen, similar results as with the neighborhood selection can be achieved but with a lower number of parents per pixel. The standard deviation of the subsequent whitened frames is closer to the expected value.

Thanks to the permutation, we are able to reach spatial correlations at higher distances while keeping the same ENDoF. However, as seen in Fig. 4.11, a random permutation creates clusters of parents which prevent certain spatial correlations to be taken into account while favoring others.

Accounting spatial correlations at multi-scale range

To prevent any clustering in the choice of parents, we need to define a structure that will favor selecting parents that are isolated from already selected parents. Defining a structure for the selection of parents derived from FRiM (Thiébaut and Tallon, 2010), we defined in Thiébaut et al., 2022 a sampling of the data at a multi-scale. It delivers an ordering of the pixel that preserves the uniformity of the density of ordered pixels at any stage of the construction of the ranking matrix $\mathbf{J}$. The idea is to monitor the size of the regions between unselected parents and favor the choice of parents that are located in empty areas. This description allows a uniform selection of parents at each scale of the data, thus avoiding disproportionate regions of the data not taken into account in the computation of the correlations.

To estimate the volume of the areas that contains pixels that are not yet ordered, we define the potential around the i-th pixel as:

$$\nu_i = \sum_{\substack{j \in \mathcal{V}_i \\ j \neq i}} \mathcal{F}(\mathcal{D}(i,j)) \tag{4.49}$$

which leads to a set of parents of a pixel defined as the pixels that have the least total potential. $\mathcal{F}$ is a potential function that decreases as the distance between i and j increases. Thiébaut et al., 2022 chooses in practice:

$$\mathcal{F}(d) = \frac{1}{d} \tag{4.50}$$

based on an analogy with particle physics.

Figure 4.11 shows such a structure on the right, compared to the random and lexicographic selections of the parents. Where a random permutation was creating clusters of unordered pixels, this method of ordering at multiscale defines a more uniform pattern. As a result, no region of the data at multi-scale ranges is favored in accounting for spatial correlations.

Applying the same LOOCV methodology, Fig. 4.7 shows the $\mathcal{L}_{\text{judge}}$ criterion with this sparsity pattern. The multi-scale selection of parents reduces even more the co-log-likelihood for the same mean number of parents compared to a random permutation. Figs. 4.6, 4.4 and 4.5 show the whitened residuals $\boldsymbol{r}^{\text{ASAP multiscale}\,[m]}$ of the tested dataset. Once again, accounting for spatial correlations yields really good results, with no discernable spatial structures in the residuals and standard deviations close to one.

All in all, accounting for spatial correlations at multiscale ranges improves greatly the co-log likelihood compared to the diagonal structure of Section 4.4.3. We expect therefore that this model will facilitate more the un-mixing of the object of interest and the speckles in the data when starting the recovery step of Alg. 4.1.

4.5 Early results and next steps

We have shown in this chapter the impact of the approximation of the covariance matrix. Going from a simple to a more complex and complete model of the covariance (or precision matrix), we demonstrated in a high-contrast imaging context the advantages of taking into account long-range spatial correlations in such datasets. I.i.d. approximations of the stochastic components $\boldsymbol{b}_t$, although common in TSVD methods to remove the stellar leakages, are not up to the task and leave strong spatial features in the residuals. Although diagonal approximations do improve the whitening of the residuals, accounting for correlations at all distances strongly reduces any spatial structures, thus better describing the data. Multiple methods to take into account such correlations have been studied in this chapter with the aim to validate the ASAP model developed by Thiébaut et al. (2022).

A new LOOCV method to choose the sparsity and validate the approximation used by ASAP on an ADI dataset without any known object has been detailed. Our results show how the different

approximations reduce the co-log-likelihood criterion by better whitening the residuals. Accounting for correlations strongly reduces such criterion thus improving greatly the whitening process. The sparse approximation used by ASAP is reliable and allows taking into account correlations at close and long ranges. As a result, ASAP can reach spatial correlations at multi-scale ranges without needing a large number of frames to add statistical information.

The aim of this more precise description of the noise in the data is directly linked to how well we can recover extended scientific objects such as protoplanetary disks from these ADI datasets. Indeed, with a better statistical model of the noise in the dataset $\boldsymbol{d}$, we can develop methods that will better disentangle objects of interest from stellar leakages. Figure 4.12 shows the results of the REXPACO algorithm (object of interest and blurred object for comparison purposes) which uses the approximation of the covariance discussed in Section 4.4.4. Compared to State-of-the-art methods, the reconstruction of REXPACO does reduce greatly artifacts due to the stellar leakages. The light distributions reconstructed are similar to predictions done with physical models and observations with other instruments.

Following these encouraging results, the ASAP method can easily be plugged in Alg. 4.1 to estimate the precision matrix. Results on simulated and real ADI datasets still need to be obtained to properly validate the reconstruction method of Section 4.3.4. More long-term improvements include an automatic tuning of the hyper-parameters of the problem, taking into account wavelength correlations delivered by the dual-band imaging mode of IRDIS (two wavelengths) and instruments such as the Integral Field Spectroscopy (IFS) of SPHERE.

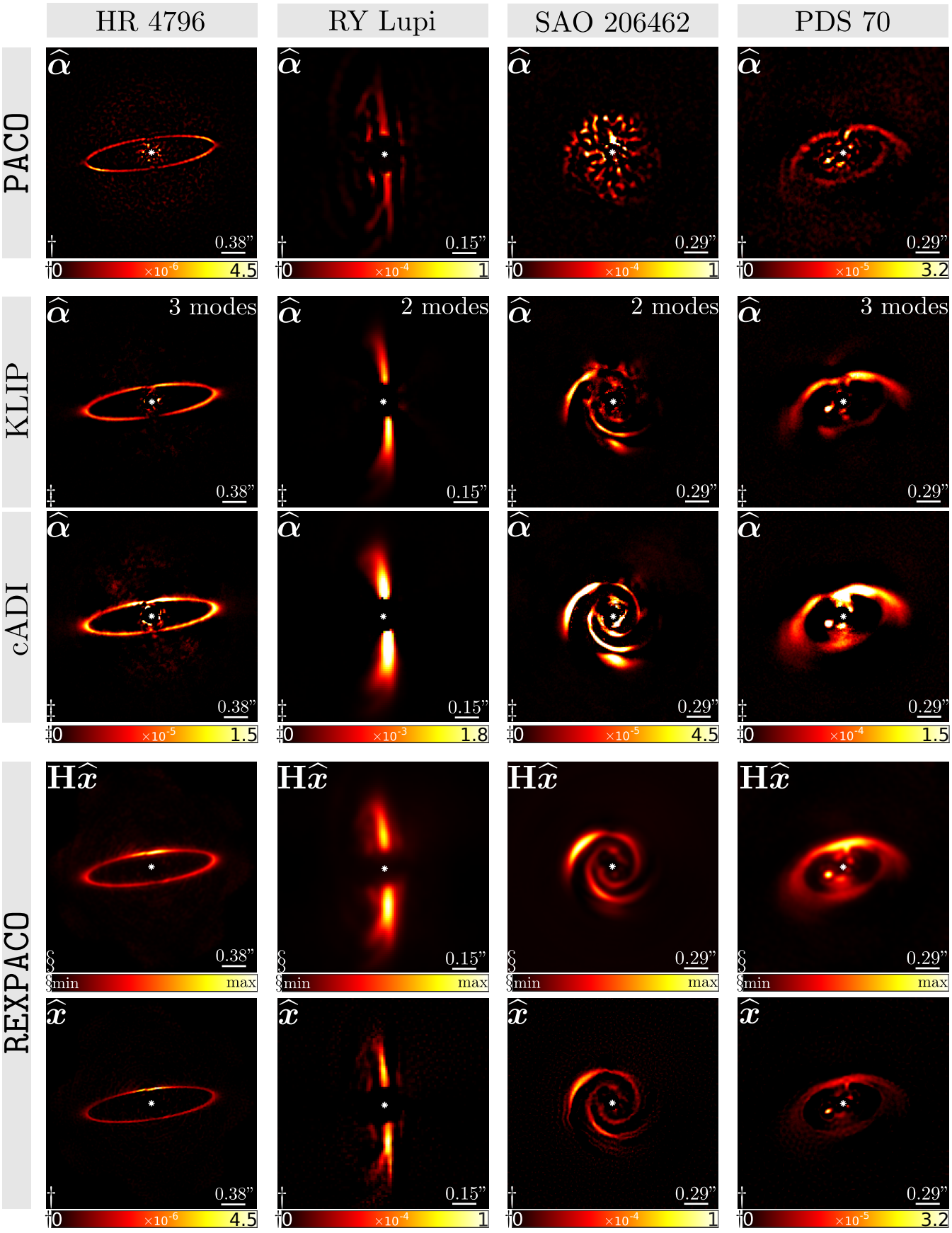

Figure 4.12 – Images of the flux distribution reconstructed by REXPACO comparatively to the PACO, PCA and cADI algorithms. For REXPACO, both the de-blurred reconstruction $\boldsymbol{x}$ and its re-blurred version $\mathbf{H}\,\boldsymbol{x}$ are given for comparison with the other considered methods that do not produce de-blurred images of the flux distribution. Data sets: HR 4796A, RY Lupi, SAO 206462 and PDS 70, as shown as raw data (*i.e.* with stellar leakages) in Fig. 4.1.

Conclusions

High-contrast imaging of circumstellar environments provides essential information to study exoplanets and their formation. Using extreme adaptive optics, instruments such as SPHERE have been developed to reach the necessary resolution to image such environments in the visible and near-infrared. Because the objects of interest are much fainter than their host star (up to 10^{-6} lower in intensity), such instruments use coronagraphy to mask at best the stellar light in the image. Although they are not perfect, these systems usually reduce the stellar contribution in the acquired images by a factor of ~ 10. Using image processing methods pushed further these performances by reducing the so-called stellar leakages by a factor of $\sim 10^2 - 10^3$. Thanks to these instruments allied with processing methods, directly observable exoplanets (usually young gas giants) and disks (protoplanetary, transition, debris) have been detected and studied (Langlois, M. et al., 2021) at angular separations of $\sim 1''$ and even up to $0.3'' - 0.4''$.

In this thesis, I proposed several data processing methods with the aim to improve upon the state-of-the-art. These methods have been developed to extract information such as spectra of companions or shapes of dust disks. More precisely, I focused on three of the main problems that limit high-contrast imaging nowadays: noises corrupting the data, instrumental response blurring the image and stellar leakages much brighter than sub-stellar objects constraining the observations. These problems require the development of post-processing methods that can disentangle objects of interest from other sources present in noisy data. These methods need to be easily usable by the community, validated on real (or realistically simulated) data and flexible to be applied to various datasets or different instruments. Using the inverse problems framework allows great flexibility and easiness in adapting to multiple contexts while leading to an optimal extraction of the information from the data in a statistical sense.

Throughout Chapter 1 and more particularly Chapter 2 was discussed the blurring effect of the instrument. This blur can be estimated jointly with the object of interest but at the cost of solving a large-scale non-convex optimization problem involving bi-linear models, the introduction of additional hyper-parameters and a certain sensitivity to the initialization. I proposed an original approach that uses the optimal scaling of the problem (in the MAP sense) to our advantage. Leading to the AMORS algorithm, this approach allows tuning only one hyper-parameter to adjust the levels of both regularizations. AMORS was developed so that it does not depend on the initial scaling of the problem, which also improves greatly the convergence speed. AMORS can be used as a *black box* algorithm plugged into an unsupervised strategy to tune the remaining hyper-parameters. Beyond the blind deconvolution context, AMORS can be used in any context featuring a bi-linear model such as blind source separation, un-mixing of spatio-spectral components (as illustrated in Chapter 3), low-rank approximation. This work has been presented in two international conferences (Thé et al., 2021; Thé et al., 2022a).

In the context of spectral characterization of exoplanets, we discussed in Chapter 3 the problem of extracting the spectrum of a sub-stellar companion in high-contrast long-slit spectroscopy data. This method of observation gives access to information on its chemical composition. It is however difficult to

use because of stellar leakages that strongly limit the regions where companion signals can be recovered. State-of-the-art methods are usually based on low-rank models built using the TSVD to describe the stellar leakages. Such methods are however sub-optimal as they do not take into account the statistics of the data. The model of the stellar component is usually a transformation of the data (warping, interpolation) obtained by considering the companion signal as negligible and noise as independent and identically distributed. These simplifying assumptions are however not usually correct and lead to the self-subtraction phenomenon: a part of the companion is removed when subtracting the stellar model. This has the effect of inducing a negative bias in the companion component. EXOSPECO was developed precisely to overcome these limitations by jointly estimating the stellar and companion components from the data. To model at best the data, we also developed a custom calibration procedure and demonstrated that accounting for instrumental modeling and calibration leads to improved spectra estimations. Validations on real data with injections of synthetic companions show a reduction of the self-subtraction effects. Our method is easy to use thanks to AMORS and achieves satisfactory spectrum estimation for companions located at angular separation greater or equal to $1''$, with relative errors of less than 20% when processing a single frame, for companions that are up to $5 \cdot 10^{-5}$ weaker than their host-star. The same quality of reconstruction can be reached by EXOSPECO for contrasts up to $\sim 10^{-4}$ for smaller angular separations in the range $0.4''$ to $1''$. The EXOSPECO algorithm has been shared on Github(`https://github.com/SJJThe/ExospecoCalibration`, `https://github.com/SJJThe/Exospeco`) presented in international conferences (Thé et al., 2022b; Thé et al., 2022c) and described in an article submitted to a peer-reviewed journal (Thé et al., 2023).

To better understand the formation of exoplanets, information on the shape of the disks of dust around their host stars are necessary. Yet, observing extended objects in such environments by direct imaging is difficult for the same reasons as described for long-slit spectroscopy: stellar leakages are heavily corrupting the data. Methods of observation such as angular differential imaging have been developed with the aim of better separating these stellar leakages from the object of interest. Statistically modeling the stellar leakages can be complicated and usually requires the use of simplified approximation of their covariances. Like the other reconstruction problems discussed in this thesis, statistical modeling of the data is particularly important. I discussed in Chapter 4 the impact of different statistical models of the noise in ADI datasets. Taking into account correlations improves greatly our capacity to explain the spatial features of the stellar leakages. Following the example of PACO (Flasseur, O. et al., 2018), ASAP aims at capturing spatial correlations with more flexible structures of the sparsity patterns of the covariance matrix. It uses a sparse Cholesky decomposition of the precision matrix, selecting a limited number of elements at multiple scales in the data. Using a leave-one-out cross-validation methodology, I discussed the accuracy of such an approximation and developed an unsupervised method that objectively sets the level of sparsity of the approximation. ASAP gives excellent results in terms of whitening: it removes most of the spatial structures in the residuals.

With these encouraging results, ongoing and possible advances of these methods are focused on two aspects.

From a methodological point of view, there are still a lot of possible experimental tests to validate AMORS on more complex examples. An article for a peer-reviewed journal is in preparation and will focus on confronting AMORS to different initializations, more complex shapes of the PSF and other tests to validate its advantages. Like for REXPACO (Flasseur, O. et al., 2021), ASAP must be systematically tested in the reconstruction context. Extractions of injected synthetic disks in data presenting no known objects are currently being studied to assess the quality of the reconstruction given this model of the precision matrix. An article in a peer-reviewed journal is in preparation, presenting the work detailed in Chapter 4, results on simulations and results on already known targets. The hyper-parameters of both methods are currently tuned thanks to a SURE-based hierarchical strategy which can be time-consuming. To speed up the computational time, I want to develop improvements of the methods detailed in this thesis using unsupervised *on-the-fly* updates of the

hyper-parameters while keeping the quality of the results obtained.

As seen in the results of EXOSPECO (Chapter 3), the modeling of the stellar leakages can still be improved to reduce even more the residuals in long-slit spectroscopy data. To characterize fainter companions, of contrast 10^{-5}, at angular separations of $0.3'' - 0.4''$, the first non-trivial step is to use more spatial modes in the series expansions of the current stellar model. As the star is not perfectly centered in the slit, a more challenging improvement would be to develop a more complex model of the stellar leakages, able to describe their behavior around the mask. Extending the EXOSPECO algorithm to jointly process multi-frame data should also improve the signal-to-noise ratio of the data and so allow to characterize companions at lower contrast values.

Of course, these improvements cannot be separated from advances in the instruments. We need better coronagraphy and ExAO to reduce even further the stellar leakages in the data. Indeed this component is responsible for the photon noise in the acquired data which represents the absolute limit of what the system is capable of discerning. In order to observe Earth-like exoplanets, we need to push the performances of our instruments and processing methods to reach angular separations of $0.1'' - 0.3''$ and contrast values lower than 10^{-6}. With the upcoming SPHERE+ upgrade, it should be possible to observe circumstellar environments at closer angular separations. This will open the path to the next generation of instruments installed on the Extremely Large Telescope which will improve greatly the possible angular resolution. Ground-based telescopes need, more than ever, processing methods that exploit the capacity of these instruments at their maximum. The upcoming flow of data from satellites such as the JWST opens the way to fascinating space observations of these circumstellar environments. However even if they are not limited by the atmosphere, the instruments on board are still limited by the diffraction, giving residuals in the data. These instruments also need efficient post-processing methods to extract correctly the information of interest from their data.

Bibliography

Airy, G. B. (Jan. 1835). "On the Diffraction of an Object-glass with Circular Aperture". In: *Transactions of the Cambridge Philosophical Society* 5, p. 283 (cit. on p. 9).

Ammanouil, R, André Ferrari, D Mary, C Ferrari, and Francesca Loi (Nov. 2019). "A parallel and automatically tuned algorithm for multispectral image deconvolution". In: *Monthly Notices of the Royal Astronomical Society* 490, pp. 37–49. DOI: 10.1093/mnras/stz2193 (cit. on p. 41).

Ayers, G. R. and J. C. Dainty (July 1988). "Iterative blind deconvolution method and its applications". en. In: *Optics Letters* 13.7, p. 547. ISSN: 0146-9592, 1539-4794. DOI: 10.1364/OL.13.000547. URL: https://www.osapublishing.org/abstract.cfm?URI=ol-13-7-547 (visited on 07/15/2020) (cit. on p. 30).

Benichoux, Alexis, Emmanuel Vincent, and Remi Gribonval (May 2013). "A fundamental pitfall in blind deconvolution with sparse and shift-invariant priors". en. In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. Vancouver, BC, Canada: IEEE, pp. 6108–6112. ISBN: 978-1-4799-0356-6. DOI: 10.1109/ICASSP.2013.6638838. URL: http://ieeexplore.ieee.org/document/6638838/ (visited on 07/15/2020) (cit. on p. 30).

Berdeu, Anthony, Soulez, Ferréol, Denis, Loïc, Langlois, Maud, and Thiébaut, Éric (2020). "PIC: a data reduction algorithm for integral field spectrographs - Application to the SPHERE instrument". In: *A&A* 635, A90. DOI: 10.1051/0004-6361/201936890. URL: https://doi.org/10.1051/0004-6361/201936890 (cit. on p. 7).

Beuzit, J. L. et al. (Nov. 2019). "SPHERE: the exoplanet imager for the Very Large Telescope". In: *A&A* 631, A155, A155. DOI: 10.1051/0004-6361/201935251. arXiv: 1902.04080 [astro-ph.IM] (cit. on pp. 2, 69).

Bezanson, Jeff, Alan Edelman, Stefan Karpinski, and Viral B. Shah (2017). "Julia: A Fresh Approach to Numerical Computing". In: *SIAM Review* 59.1, pp. 65–98. DOI: 10.1137/141000671. URL: https://julialang.org/ (cit. on pp. 8, 66).

Blanco, L. and L. M. Mugnier (2011). "Marginal blind deconvolution of adaptive optics retinal images". In: *Opt. Express* 19.23, pp. 23227–23239. DOI: 10.1364/OE.19.023227. URL: http://opg.optica.org/oe/abstract.cfm?URI=oe-19-23-23227 (cit. on p. 31).

Brent, R.P. (2013). *Algorithms for Minimization Without Derivatives*. Dover Books on Mathematics. Dover Publications. ISBN: 9780486143682. URL: https://books.google.fr/books?id=AITCAgAAQBAJ (cit. on pp. 25, 57).

Broyden, C. G. (Mar. 1970). "The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations". In: *IMA Journal of Applied Mathematics* 6.1, pp. 76–90. ISSN: 0272-4960. DOI: 10.1093/imamat/6.1.76. eprint: https://academic.oup.com/imamat/article-pdf/6/1/76/2233756/6-1-76.pdf. URL: https://doi.org/10.1093/imamat/6.1.76 (cit. on p. 25).

Carbillet, Marcel et al. (May 2011). "Apodized Lyot coronagraph for SPHERE/VLT. I. Detailed numerical study". In: *Experimental Astronomy* 30.1, pp. 39–58. DOI: 10.1007/s10686-011-9219-4 (cit. on p. 2).

Catmull, Edwin and Raphael Rom (1974). "A CLASS OF LOCAL INTERPOLATING SPLINES". In: *Computer Aided Geometric Design*. Ed. by ROBERT E. BARNHILL and RICHARD F. RIESENFELD. Academic Press, pp. 317–326. ISBN: 978-0-12-079050-0. DOI: https://doi.org/10.1016/B978-0-12-079050-0.50020-5. URL: https://www.sciencedirect.com/science/article/pii/B9780120790500500205 (cit. on p. 50).

Chan, T.F. and Chiu-Kwong Wong (1998). "Total variation blind deconvolution". In: *IEEE Transactions on Image Processing* 7.3, pp. 370–375. DOI: 10.1109/83.661187 (cit. on p. 30).

Chan, Tony F. and C.K. Wong (Sept. 2000). "Convergence of the alternating minimization algorithm for blind deconvolution". en. In: *Linear Algebra and its Applications* 316.1-3, pp. 259–285. ISSN: 00243795. DOI: 10.1016/S0024-3795(00)00141-5. URL: https://linkinghub.elsevier.com/retrieve/pii/S0024379500001415 (visited on 07/15/2020) (cit. on p. 31).

Charbonneau, David, Timothy Brown, David Latham, and Michel Mayor (Jan. 2000). "Detection of Planetary Transits Across a Sun-like Star". In: *The Astrophysical journal* 529, pp. L45–L48. DOI: 10.1086/312457 (cit. on p. 1).

Chen, Shaobing, D. Donoho, Iain Johnstone, and Michael Saunders (Mar. 1996). "Basis Pursuit". In: (cit. on p. 16).

Chen, Yilun, Ami Wiesel, Yonina C Eldar, and Alfred O Hero (2010). "Shrinkage algorithms for MMSE covariance estimation". In: *IEEE Transactions on Signal Processing* 58.10, pp. 5016–5029 (cit. on p. 86).

Conan, Jean-Marc, Laurent M. Mugnier, Thierry Fusco, Vincent Michau, and Gérard Rousset (1998). "Myopic deconvolution of adaptive optics images by use of object and point-spread function power spectra". In: *Appl. Opt.* 37.21, pp. 4614–4622. DOI: 10.1364/AO.37.004614. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-37-21-4614 (cit. on p. 14).

Craven, Peter and Grace Wahba (1978). "Smoothing Noisy Data with Spline Functions". In: *Numer. Math.* 31.4, 377–403. ISSN: 0029-599X. DOI: 10.1007/BF01404567. URL: https://doi.org/10.1007/BF01404567 (cit. on p. 22).

de Boer, J. et al. (2020). "Polarimetric imaging mode of VLT/SPHERE/IRDIS - I. Description, data reduction, and observing strategy". In: *A&A* 633, A63. DOI: 10.1051/0004-6361/201834989. URL: https://doi.org/10.1051/0004-6361/201834989 (cit. on p. 2).

Delorme, P. et al. (Dec. 2017). "The SPHERE Data Center: a reference for high contrast imaging processing". In: *SF2A-2017: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*. Ed. by C. Reylé et al., p. Di. arXiv: 1712.06948 [astro-ph.IM]. URL: https://ui.adsabs.harvard.edu/abs/2017sf2a.conf..347D (cit. on p. 72).

Denneulin, L., Langlois, M., Thiébaut, É., and Pustelnik, N. (2021). "RHAPSODIE: Reconstruction of High-contrAst Polarized SOurces and Deconvolution for cIrcumstellar Environments". In: *A&A* 653, A138. DOI: 10.1051/0004-6361/202039618. URL: https://doi.org/10.1051/0004-6361/202039618 (cit. on p. 18).

Devaney, Nicholas and Éric Thiébaut (Dec. 2017). "PeX 1. Multispectral expansion of residual speckles for planet detection". en. In: *Monthly Notices of the Royal Astronomical Society* 472.3, pp. 3734–3748. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stx2218. URL: http://academic.oup.com/mnras/article/472/3/3734/4098502/PeX-1-Multispectral-expansion-of-residual-speckles (visited on 07/15/2020) (cit. on pp. 45, 71).

Dohlen, Kjetil et al. (July 2008). "The infra-red dual imaging and spectrograph for SPHERE: design and performance". In: *Ground-based and Airborne Instrumentation for Astronomy II*. Ed. by Ian S. McLean and Mark M. Casali. Vol. 7014. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 70143L. DOI: 10.1117/12.789786 (cit. on pp. 2, 44, 69, 108, 109).

Eckart, Carl and Gale Young (1936). "The approximation of one matrix by another of lower rank". In: *Psychometrika* 1.3, pp. 211–218 (cit. on p. 48).

Eldar, Y.C. (Feb. 2009). "Generalized SURE for Exponential Families: Applications to Regularization". en. In: *IEEE Transactions on Signal Processing* 57.2, pp. 471–481. ISSN: 1053-587X, 1941-0476.

DOI: 10.1109/TSP.2008.2008212. URL: http://ieeexplore.ieee.org/document/4663926/ (visited on 07/03/2020) (cit. on p. 20).

Esposito, Thomas M, Michael P Fitzgerald, James R Graham, and Paul Kalas (2013). "Modeling self-subtraction in angular differential imaging: Application to the HD 32297 debris disk". In: *The Astrophysical Journal* 780.1, p. 25 (cit. on p. 73).

Flasseur, O., Denis, L., Thiébaut, É., and Langlois, M. (2018). "Exoplanet detection in angular differential imaging by statistical learning of the nonstationary patch covariances - The PACO algorithm". In: *A&A* 618, A138. DOI: 10.1051/0004-6361/201832745. URL: https://doi.org/10.1051/0004-6361/201832745 (cit. on pp. 85, 86, 96).

Fletcher, R. (Jan. 1970). "A new approach to variable metric algorithms". In: *The Computer Journal* 13.3, pp. 317–322. ISSN: 0010-4620. DOI: 10.1093/comjnl/13.3.317. eprint: https://academic.oup.com/comjnl/article-pdf/13/3/317/988678/130317.pdf. URL: https://doi.org/10.1093/comjnl/13.3.317 (cit. on p. 25).

Girard, A. (1989). "A Fast 'Monte-Carlo Cross-Validation' Procedure for Large Least Squares Problems with Noisy Data". In: *Numer. Math.* 56.1, 1–23. ISSN: 0029-599X. DOI: 10.1007/BF01395775. URL: https://doi.org/10.1007/BF01395775 (cit. on p. 22).

Goldfarb, Donald (1970). "A Family of Variable Metric Updates Derived by Variational Means". In: *Mathematics of Computation* 24.109, pp. 23–26 (cit. on p. 25).

Hansen, P. (Jan. 2001). "The L-Curve and Its Use in the Numerical Treatment of Inverse Problems". In: vol. 4, pp. 119–142 (cit. on p. 20).

Hastie, Trevor, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman (2009). *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. Springer (cit. on p. 78).

Kattnig, Alain and Jérôme Primot (July 1997). "Model of the second-order statistic of the radiance field of natural scenes, adapted to system conceiving". In: DOI: 10.1117/12.280614 (cit. on p. 14).

Kiefer, J. (1953). "Sequential Minimax Search for a Maximum". In: *Proceedings of the American Mathematical Society* 4.3, pp. 502–506. ISSN: 00029939, 10886826. URL: http://www.jstor.org/stable/2032161 (visited on 10/06/2022) (cit. on p. 25).

Lafreniere, David, Christian Marois, Rene Doyon, Daniel Nadeau, and Etienne Artigau (2007). "A New Algorithm for Point-Spread Function Subtraction in High-Contrast Imaging: A Demonstration with Angular Differential Imaging". In: *The Astrophysical Journal* 660.1, pp. 770–780. DOI: 10.1086/513180. URL: https://doi.org/10.1086/513180 (cit. on p. 73).

Langlois, M. et al. (2021). "The SPHERE infrared survey for exoplanets (SHINE) - II. Observations, data reduction and analysis, detection performances, and initial results". In: *A&A* 651, A71. DOI: 10.1051/0004-6361/202039753. URL: https://doi.org/10.1051/0004-6361/202039753 (cit. on pp. 43, 95).

Latham, David W., Tsevi Mazeh, Robert P. Stefanik, Michel Mayor, and Gilbert Burki (1989). "The unseen companion of HD114762: a probable brown dwarf". In: *Nature* 339, pp. 38–40 (cit. on p. 1).

Ledoit, Olivier and Michael Wolf (2004). "A well-conditioned estimator for large-dimensional covariance matrices". In: *Journal of multivariate analysis* 88.2, pp. 365–411 (cit. on pp. 85, 86).

Levin, Anat, Yair Weiss, Fredo Durand, and William T. Freeman (2011). "Efficient marginal likelihood optimization in blind deconvolution". In: *CVPR 2011*, pp. 2657–2664. DOI: 10.1109/CVPR.2011.5995308 (cit. on p. 31).

Marois, Christian et al. (2014). "GPI PSF subtraction with TLOCI: the next evolution in exoplanet/disk high-contrast imaging". In: *Adaptive Optics Systems IV*. Ed. by Enrico Marchetti, Laird M. Close, and Jean-Pierre Véran. Vol. 9148. International Society for Optics and Photonics. SPIE, pp. 287 –299. DOI: 10.1117/12.2055245. URL: https://doi.org/10.1117/12.2055245 (cit. on p. 73).

Mawet, Dimitri et al. (July 2012). "Review of small-angle coronagraphic techniques in the wake of ground-based second-generation adaptive optics systems". In: *Proceedings of SPIE - The International Society for Optical Engineering* 8442. DOI: 10.1117/12.927245 (cit. on pp. 66, 73).

Mazoyer, Johan et al. (2020). "DiskFM: A forward modeling tool for disk analysis with coronagraphic instruments". In: *Ground-based and Airborne Instrumentation for Astronomy VIII*. Vol. 11447. SPIE, pp. 1080–1099 (cit. on p. 73).

Mesa, D. et al. (2016). "Characterizing HR9sing SPHERE". In: *A&A* 593, A119. DOI: 10.1051/0004-6361/201628740. URL: https://doi.org/10.1051/0004-6361/201628740 (cit. on p. 46).

Milli, J et al. (2012). "Impact of angular differential imaging on circumstellar disk images". In: *Astronomy & Astrophysics* 545, A111 (cit. on p. 73).

Mirsky, Léon (1960). "Symmetric gauge functions and unitarily invariant norms". In: *Quarterly J. Math.* 11, pp. 50–59. DOI: doi:10.1093/qmath/11.1.50 (cit. on p. 48).

Molina, R. (1994). "On the Hierarchical Bayesian Approach to Image Restoration: Applications to Astronomical Images". In: *IEEE Trans. Pattern Anal. Mach. Intell.* 16, pp. 1122–1128 (cit. on pp. 31, 41).

Montagnier, Guillaume et al. (2007). "Pupil stabilization for SPHERE's extreme AO and high performance coronagraph system." In: *Opt. Express* 15.23, pp. 15293–15307. DOI: 10.1364/OE.15.015293. URL: http://opg.optica.org/oe/abstract.cfm?URI=oe-15-23-15293 (cit. on p. 72).

Mugnier, Laurent, Thierry Fusco, and Jean-Marc Conan (Nov. 2004). "MISTRAL: A myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images". In: *Journal of the Optical Society of America. A, Optics, image science, and vision* 21, pp. 1841–54. DOI: 10.1364/JOSAA.21.001841 (cit. on pp. 7, 8).

Nocedal, Jorge (1980). "Updating Quasi-Newton Matrices With Limited Storage". In: *Mathematics of Computation* 35, pp. 773–782 (cit. on p. 25).

"Trust-Region Methods" (1999). In: *Numerical Optimization*. Ed. by Jorge Nocedal and Stephen J. Wright. New York, NY: Springer New York, pp. 64–99. ISBN: 978-0-387-22742-9. DOI: 10.1007/0-387-22742-3_4. URL: https://doi.org/10.1007/0-387-22742-3_4 (cit. on p. 24).

Pairet, Benoît, Faustine Cantalloube, and Laurent Jacques (2018). "Reference-less algorithm for circumstellar disks imaging". In: *arXiv preprint arXiv:1812.01333* (cit. on p. 74).

— (2021). "MAYONNAISE: a morphological components analysis pipeline for circumstellar discs and exoplanets imaging in the near-infrared". In: *Monthly Notices of the Royal Astronomical Society* 503.3, pp. 3724–3742 (cit. on p. 74).

Pavlov, A., M. Feldt, and Th. Henning (Aug. 2008). "Data Reduction and Handling for SPHERE". In: *Astronomical Data Analysis Software and Systems XVII*. Ed. by R. W. Argyle, P. S. Bunclark, and J. R. Lewis. Vol. 394. Astronomical Society of the Pacific Conference Series, p. 581 (cit. on pp. 7, 72).

Perrin, Marshall D., Anand Sivaramakrishnan, Russell B. Makidon, Ben R. Oppenheimer, and James R. Graham (2003). "The Structure of High Strehl Ratio Point-Spread Functions". In: *The Astrophysical Journal* 596.1, pp. 702–712. DOI: 10.1086/377689. URL: https://doi.org/10.1086/377689 (cit. on p. 45).

Perrone, Daniele and Paolo Favaro (June 2014). "Total Variation Blind Deconvolution: The Devil Is in the Details". en. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE, pp. 2909–2916. ISBN: 978-1-4799-5118-5. DOI: 10.1109/CVPR.2014.372. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6909768 (visited on 09/21/2020) (cit. on p. 30).

Powell, M. J. D. (2006). "The NEWUOA software for unconstrained optimization without derivatives". In: *Nonconvex Optimization and Its Applications*. Ed. by G. Di Pillo and M. Roma. Springer Science, pp. 255–297. DOI: 10.1007/0-387-30065-1_16 (cit. on pp. 14, 26, 57).

Pueyo, Laurent (2018). "Direct imaging as a detection technique for exoplanets". In: *Handbook of Exoplanets*, p. 10 (cit. on p. 73).

Ramani, S., T. Blu, and M. Unser (Sept. 2008). "Monte-Carlo Sure: A Black-Box Optimization of Regularization Parameters for General Denoising Algorithms". en. In: *IEEE Transactions on*

Image Processing 17.9, pp. 1540–1554. ISSN: 1057-7149. DOI: 10.1109/TIP.2008.2001404. URL: http://ieeexplore.ieee.org/document/4598837/ (visited on 07/05/2020) (cit. on p. 22).

Ren, Bin, Laurent Pueyo, Guangtun Ben Zhu, John Debes, and Gaspard Duchêne (2018). "Non-negative matrix factorization: robust extraction of extended structures". In: *The Astrophysical Journal* 852.2, p. 104 (cit. on p. 73).

Ruane, Garreth et al. (2019). "Reference Star Differential Imaging of Close-in Companions and Circumstellar Disks with the NIRC2 Vortex Coronagraph at the W. M. Keck Observatory". In: *The Astronomical Journal* 157.3, p. 118. DOI: 10.3847/1538-3881/aafee2. URL: https://doi.org/10.3847/1538-3881/aafee2 (cit. on p. 73).

Rudin, Leonid I., Stanley Osher, and Emad Fatemi (Nov. 1992). "Nonlinear total variation based noise removal algorithms". en. In: *Physica D: Nonlinear Phenomena* 60.1-4, pp. 259–268. ISSN: 01672789. DOI: 10.1016/0167-2789(92)90242-F. URL: https://linkinghub.elsevier.com/retrieve/pii/016727899290242F (visited on 07/15/2020) (cit. on p. 17).

Schäfer, Florian, Matthias Katzfuss, and Houman Owhadi (2019). "Sparse Cholesky Factorization by Kullback–Leibler Minimization". In: *SIAM Journal on Scientific Computing* 43.3, A2019–A2046. DOI: https://doi.org/10.1137/20M1336254 (cit. on p. 87).

Shanno, D. F. (1970). "Conditioning of Quasi-Newton Methods for Function Minimization". In: *Mathematics of Computation* 24.111, pp. 647–656. ISSN: 00255718, 10886842. URL: http://www.jstor.org/stable/2004840 (visited on 10/01/2022) (cit. on p. 25).

Soummer, Rémi, Laurent Pueyo, and James Larkin (2012). "Detection and characterization of exoplanets and disks using projections on Karhunen–Loève eigenimages". In: *The Astrophysical Journal Letters* 755.2, p. L28 (cit. on p. 74).

Stein, Charles M (1981). "Estimation of the mean of a multivariate normal distribution". In: *The annals of Statistics*, pp. 1135–1151 (cit. on p. 20).

Strong, David and Tony Chan (2003). "Edge-preserving and scale-dependent properties of total variation regularization". In: *Inverse Problems* 19.6, S165. DOI: 10.1088/0266-5611/19/6/059. URL: https://dx.doi.org/10.1088/0266-5611/19/6/059 (cit. on p. 18).

Thiébaut, Eric (2002). "Optimization issues in blind deconvolution algorithms". In: *Astronomical Data Analysis II*. Ed. by Jean-Luc Starck and Fionn D. Murtagh. Vol. 4847. International Society for Optics and Photonics. SPIE, pp. 174 –183. DOI: 10.1117/12.461151. URL: https://doi.org/10.1117/12.461151 (cit. on pp. 25, 34, 36, 57).

Thiébaut, Éric and Michel Tallon (2010). "Fast minimum variance wavefront reconstruction for extremely large telescopes". In: 27.5, pp. 1046–1059. DOI: 10.1364/JOSAA.27.001046 (cit. on pp. 87, 92).

Tibshirani, Robert (1996). "Regression Shrinkage and Selection Via the Lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1, pp. 267–288. DOI: https://doi.org/10.1111/j.2517-6161.1996.tb02080.x. eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1996.tb02080.x. URL: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1996.tb02080.x (cit. on p. 16).

Tikhonov, Andrey N. and Vasiliy Y. Arsenin (1977). *Solutions of ill-posed problems*. Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics. Washington, D.C.: John Wiley & Sons, New York: V. H. Winston & Sons, pp. xiii+258 (cit. on pp. 17, 56).

Vigan, A., M. Langlois, C. Moutou, and K. Dohlen (Oct. 2008). "Exoplanet characterization with long slit spectroscopy". In: *A&A* 489.3, pp. 1345–1354. DOI: 10.1051/0004-6361:200810090. arXiv: 0808.3817 [astro-ph] (cit. on pp. 44, 46).

Vigan, Arthur et al. (2010). "Photometric characterization of exoplanets using angular and spectral differential imaging". In: *Monthly Notices of the Royal Astronomical Society* 407.1, pp. 71–82 (cit. on p. 71).

Wahba, Grace et al. (1985). "A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem". In: *The Annals of Statistics* 13.4, pp. 1378–1402 (cit. on p. 20).

Wiener, Norbert (1949). *Extrapolation, interpolation, and smoothing of stationary time series, with engineering applications.* eng. Cambridge: Technology Press of the Massachusetts Institute of Technology (cit. on p. 12).

Wolfe, Philip (1969). "Convergence Conditions for Ascent Methods". In: *SIAM Review* 11.2, pp. 226–235. DOI: 10.1137/1011036. eprint: https://doi.org/10.1137/1011036. URL: https://doi.org/10.1137/1011036 (cit. on p. 24).

You, Yu-Li and M. Kaveh (1996). "Anisotropic blind image restoration". In: *Proceedings of 3rd IEEE International Conference on Image Processing*. Vol. 2, 461–464 vol.2. DOI: 10.1109/ICIP.1996.560885 (cit. on p. 30).

Yu-Li You and M. Kaveh (Mar. 1996). "A regularization approach to joint blur identification and image restoration". en. In: *IEEE Transactions on Image Processing* 5.3, pp. 416–428. ISSN: 10577149. DOI: 10.1109/83.491316. URL: http://ieeexplore.ieee.org/document/491316/ (visited on 09/21/2020) (cit. on pp. 30, 31).

Zhang, Kaihao et al. (Sept. 2022). "Deep Image Deblurring: A Survey". In: *International Journal of Computer Vision* 130, pp. 1–28. DOI: 10.1007/s11263-022-01633-5 (cit. on p. 31).

Personal bibliography

Flasseur, O., Thé, S., Denis, L., Thiébaut, É., and Langlois, M. (2021). "REXPACO: An algorithm for high contrast reconstruction of the circumstellar environment by angular differential imaging". In: *A&A* 651, A62. DOI: 10.1051/0004-6361/202038957. URL: https://doi.org/10.1051/0004-6361/202038957 (cit. on pp. 70, 72, 74, 85, 89, 96).

Thé, Samuel, Éric Thiébaut, Loïc Denis, and Ferréol Soulez (2021). "Exploiting the scaling indetermination of bi-linear models in inverse problems". In: *2020 28th European Signal Processing Conference (EUSIPCO)*, pp. 2358–2362. DOI: 10.23919/Eusipco47968.2020.9287593 (cit. on pp. 29, 95).

— (2022a). "Unsupervised blind-deconvolution with optimal scaling applied to astronomical data". In: *Adaptive Optics Systems VIII*. Ed. by Laura Schreiber, Dirk Schmidt, and Elise Vernet. Vol. 12185. International Society for Optics and Photonics. SPIE, 121853W. DOI: 10.1117/12.2630245. URL: https://doi.org/10.1117/12.2630245 (cit. on pp. 29, 95).

Thé, Samuel, Éric Thiébaut, Loïc Denis, Ferréol Soulez, and Maud Langlois (2022b). "Long-slit spectroscopy characterization of substellar objects with the EXOSPEC algorithm". In: *LYOT - High-contrast observations and image processing techniques* (cit. on pp. 43, 96).

— (2022c). "Long-slit spectroscopy characterization of substellar objects with the EXOSPEC algorithm". In: *Adaptive Optics Systems VIII*. Ed. by Laura Schreiber, Dirk Schmidt, and Elise Vernet. Vol. 12185. International Society for Optics and Photonics. SPIE, 121853Q. DOI: 10.1117/12.2630171. URL: https://doi.org/10.1117/12.2630171 (cit. on pp. 43, 96).

Thé, Samuel et al. (2023). "Characterization of stellar companion from high-contrast long-slit spectroscopy - The EXtraction Of SPEctrum of COmpanion (EXOSPECO) algorithm". Manuscript submitted for publication (cit. on pp. 43, 53, 96).

Thiébaut, Éric, Michel Tallon, Samuel Thé, and Loïc Denis (2022). "Beyond FRiM, ASAP: a family of sparse approximation for covariance matrices and preconditioners". In: *Adaptive Optics Systems VIII*. Ed. by Laura Schreiber, Dirk Schmidt, and Elise Vernet. Vol. 12185. International Society for Optics and Photonics. SPIE, 121852Q. DOI: 10.1117/12.2630192. URL: https://doi.org/10.1117/12.2630192 (cit. on pp. 86, 89, 90, 92).

Appendix A

Pre-processing of IRDIS raw data

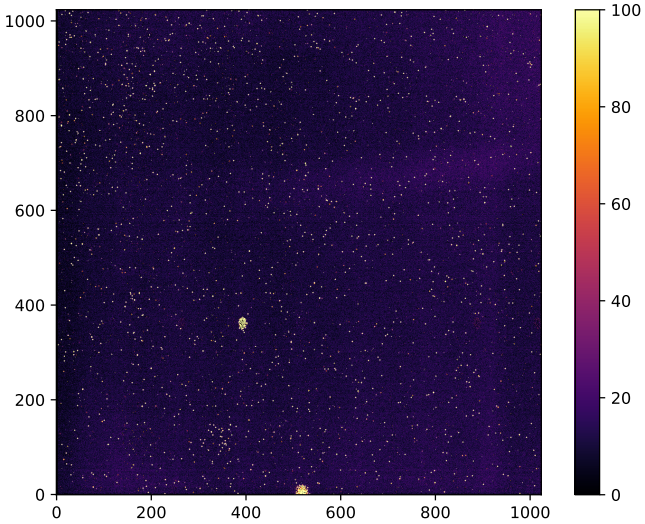

Figure A.1 – Instance of a part of the *background* calibration data of the SPHERE/IRDIS instrument.

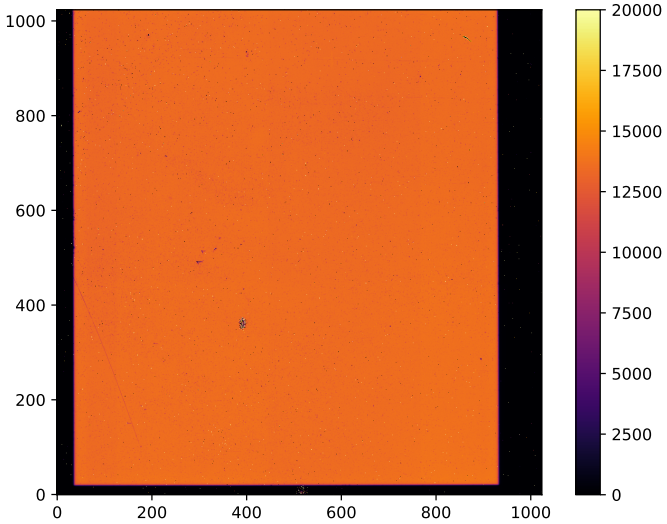

Figure A.2 – Instance of a part of the *flat* calibration data of the SPHERE/IRDIS instrument.

In this appendix, we describe in more details the pre-processing strategy used in Section 1.1 to correct the raw data, thanks to calibration data supposed as taken in the same conditions as the scientific data (*e.g.* detector's temperature, exposition time).

Recalling the model of Eq. (1.1), the output level of the n-th pixel of the detector (the raw data) writes:

$$r_n = \frac{e_n}{g_n} + z_n \tag{A.1}$$

with e_n the number of captured electrons, which follows a Poisson distribution, g_n the ADC's gain and z_n representing the zero-level of the ADC as well as its read-out noise and the rounding errors introduced by the conversion, which follows a Gaussian distribution. Considering that e_n and z_n are independent variables, the statistical model of the raw data writes:

$$\mathrm{E}\big[r_n\big] = \frac{\mathrm{E}\big[e_n\big]}{g_n} + \mathrm{E}\big[z_n\big] \tag{A.2}$$

$$\mathrm{Var}(r_n) = \mathrm{Var}\left(\frac{e_n}{g_n}\right) + \mathrm{Var}(z_n) \tag{A.3}$$

Since that e_n follows a Poisson distribution, its expectation can be separated as an expected number of electrons coming from the object of interest and an expected number of electrons coming from

background sources and dark current:

$$\begin{aligned} \mathrm{E}\big[e_n\big] &= \mathrm{E}\big[e_n^{[\mathrm{obj}]}\big] + \mathrm{E}\big[e_n^{[\mathrm{bg}]}\big] \\ &= q_n\,\phi_n^{[\mathrm{obj}]} + \mathrm{E}\big[e_n^{[\mathrm{bg}]}\big] \end{aligned} \tag{A.4}$$

with q_n accounting for the uneven quantum efficiency and variation in intensity like vignetting, $\phi_n^{[\mathrm{obj}]}$ the flux of light in the detector's plane coming from the object of interest and $\mathrm{E}\big[e_n^{[\mathrm{bg}]}\big]$ the expectation of the number of electrons coming from stray lights and spurious sources such as dark current. Remarking that e_n follows a Poisson distribution, $\mathrm{Var}(e_n) = \mathrm{E}\big[e_n\big]$ and we can rewrite the statistical moments of the raw data of Eq. (A.2) as:

$$\mathrm{E}\big[r_n\big] = \underbrace{\frac{q_n}{g_n}}_{\propto\, 1/a_n}\phi_n^{[\mathrm{obj}]} + \underbrace{\frac{\mathrm{E}\big[e_n^{[\mathrm{bg}]}\big]}{g_n} + \mathrm{E}\big[z_n\big]}_{c_n} \tag{A.5}$$

and

$$\begin{aligned} \mathrm{Var}(r_n) &= \frac{\mathrm{Var}(e_n)}{g_n^2} + \mathrm{Var}(z_n) \\ &= \frac{q_n}{g_n^2}\,\phi_n^{[\mathrm{obj}]} + \frac{\mathrm{E}\big[e_n^{[\mathrm{bg}]}\big]}{g_n^2} + \mathrm{Var}(z_n) \end{aligned} \tag{A.6}$$

From Eq. (A.5) we can see that there is an affine dependence of the expected value of the raw data with respect to the object flux. This suggests applying an affine correction which yields the pre-processed data:

$$d_n = a_n\,(r_n - c_n) \tag{A.7}$$

with a_n and c_n coefficients that need to be estimated.

Following the model of Eq. (A.1), *background* calibrations measure the inherent instrumental background by closing a shutter so that light from external sources does not enter the instrument. Fig. A.1 presents an example of such calibration for the SPHERE/IRDIS instrument (Dohlen et al., 2008). The resulting raw background calibration data can be expressed by:

$$r_n^{[\mathrm{bg}]} = \frac{e_n^{[\mathrm{bg}]}}{g_n} + z_n \tag{A.8}$$

from which follows that taking:

$$\begin{aligned} c_n &= \mathrm{E}\big[r_n^{[\mathrm{bg}]}\big] \\ &= \frac{\mathrm{E}\big[e_n^{[\mathrm{bg}]}\big]}{g_n} + \mathrm{E}\big[z_n\big] \end{aligned} \tag{A.9}$$

will correct the bias in the raw data. In practice, multiple background calibrations are acquired and the expectation can then be approximated by taking the sample mean $\bar{r}_n^{[\mathrm{bg}]}$, leading to the estimate:

$$\hat{c_n} = \frac{\bar{e_n^{[\mathrm{bg}]}}}{g_n} + \bar{z_n} \tag{A.10}$$

Replacing c_n in Eq. (A.7), we can deduce that a_n needs to correct the ratio q_n/g_n.

Using *flat* calibrations (see Fig. A.2), that is exposing the instrument to a constant flux $\phi^{[\mathrm{flat}]}$, the expectation of the resulting raw flat calibration data is:

$$\mathrm{E}\big[r_n^{[\mathrm{flat}]}\big] = \frac{q_n\,\phi^{[\mathrm{flat}]}}{g_n} + c_n \tag{A.11}$$

which gives:

$$\hat{a}_n = \frac{1}{\bar{r}_n^{[\text{flat}]} - \hat{c}_n} \simeq \frac{g_n}{q_n\,\phi^{[\text{flat}]}} \tag{A.12}$$

with $\bar{r}_n^{[\text{flat}]} \approx \mathrm{E}\big[r_n n^{[\text{flat}]}\big]$ the empirical pixel-wise mean of the flat image.

Taking back the Poisson and Gaussian approximation in Eq. (A.1), it is possible to express the expected value of the pre-processed data:

$$\begin{aligned}
\mathrm{E}\big[d_n\big] &= \hat{a}_n \Big(\mathrm{E}\big[r_n\big] - \hat{c}_n\Big) \\
&= \frac{\hat{a}_n}{g_n}\,\mathrm{E}\big[e_n^{[\text{obj}]}\big] \\
&= \frac{\hat{a}_n\,q_n}{g_n}\,\phi_n^{[\text{obj}]} \\
&= \frac{\phi_n^{[\text{obj}]}}{\phi^{[\text{flat}]}}
\end{aligned} \tag{A.13}$$

and its variance, using Eqs. (A.6) and (A.13):

$$\begin{aligned}
\mathrm{Var}(d_n) &= \hat{a}_n^2\,\mathrm{Var}(r_n) \\
&= \frac{\hat{a}_n^2}{g_n^2}\Big(\mathrm{E}\big[e_n^{[\text{obj}]}\big] + \mathrm{E}\big[e_n^{[\text{bg}]}\big]\Big) + \hat{a}_n^2\,\sigma_n^{[\text{ron}]\,2} \\
&= \underbrace{\frac{\hat{a}_n}{g_n}}_{f_n}\,\mathrm{E}\big[d_n\big] + \underbrace{\frac{\hat{a}_n^2\,\mathrm{E}\big[e_n^{[\text{bg}]}\big]}{g_n^2} + \hat{a}_n^2\,\sigma_n^{[\text{ron}]\,2}}_{\sigma_n^2}
\end{aligned} \tag{A.14}$$

with $\sigma_n^{[\text{ron}]\,2} = \mathrm{Var}(z_n)$. The first variance term in Eq. (A.14) comes from the photon noise of the object of interest, while the second term comes from the shot noise of the stray light, and the third term is the read-out noise. Note that with the estimation of a_n described in Eq. (A.12), the pre-processed data is not expressed in ADU but in an arbitrary unit when the flux of the *flat* calibration, $\phi^{[\text{flat}]}$ is unknown.

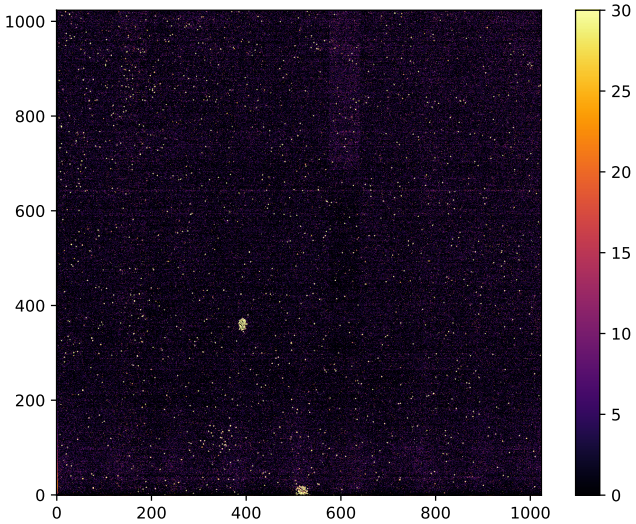

Figure A.3 – Instance of a part of the *dark* calibration data of the SPHERE/IRDIS instrument Dohlen et al., 2008.

In practice, taking the raw *background* calibration data, we can estimate:

$$\hat{\sigma}_n^2 = \hat{a}_n^2\,\widetilde{\mathrm{Var}}\big(r_n^{[\text{bg}]}\big) \tag{A.15}$$

with $\widetilde{\mathrm{Var}}\big(r_n^{[\mathrm{bg}]}\big)$ the sample variance of the *background* calibrations.

As for f_n, it is deduced by Eq. (A.12) and estimating the ADC's gain. To do so, *dark* calibrations are taken which consist of taking an exposition time of zero (see Fig. A.3). In such conditions, the resulting raw dark calibration data is:

$$r_n^{[\mathrm{dark}]} = z_n \tag{A.16}$$

and so:

$$\begin{aligned}\widetilde{\mathrm{Var}}\big(r_n^{[\mathrm{bg}]}\big) - \widetilde{\mathrm{Var}}\big(r_n^{[\mathrm{dark}]}\big) \simeq \frac{\bar{e}_n^{[\mathrm{bg}]}}{g_n^2} &= \frac{\bar{r}_n^{[\mathrm{bg}]} - \bar{r}_n^{[\mathrm{dark}]}}{g_n} \\ \Leftrightarrow \hat{g}_n &= \frac{\widetilde{\mathrm{Var}}\big(r_n^{[\mathrm{bg}]}\big) - \widetilde{\mathrm{Var}}\big(r_n^{[\mathrm{dark}]}\big)}{\bar{r}_n^{[\mathrm{bg}]} - \bar{r}_n^{[\mathrm{dark}]}}\end{aligned} \tag{A.17}$$

Note that in Figures A.1, A.2 and A.3, the response of the detector is not uniform. This justifies the pixel-wise model of Eq. (1.7). It is also possible to observe on these calibrations data the dark or saturated pixels that are called bad pixels in Section 1.1.

In the end, the pre-processed strategy described here gives a statistical model of the pre-processed data:

$$\mathrm{E}\big[d_n\big] = \frac{\phi_n^{[\mathrm{obj}]}}{\phi^{[\mathrm{flat}]}} \tag{A.18}$$

$$\mathrm{Var}(d_n) = \hat{f}_n\,\mathrm{E}\big[d_n\big] + \hat{\sigma}_n^2 \tag{A.19}$$

which describes completely d_n without losing any information compared to the raw data r_n and where the different parameters can be estimated by:

$$\hat{c}_n = \frac{\bar{e}_n^{[\mathrm{bg}]}}{g_n} + \bar{z}_n \tag{A.20}$$

$$\hat{a}_n = \frac{1}{\bar{r}_n^{[\mathrm{flat}]} - \hat{c}_n} \tag{A.21}$$

$$\hat{g}_n = \frac{\widetilde{\mathrm{Var}}\big(r_n^{[\mathrm{bg}]}\big) - \widetilde{\mathrm{Var}}\big(r_n^{[\mathrm{dark}]}\big)}{\bar{r}_n^{[\mathrm{bg}]} - \bar{r}_n^{[\mathrm{dark}]}} \tag{A.22}$$

$$\hat{f}_n = \frac{\hat{a}_n}{\hat{g}_n} \tag{A.23}$$

$$\hat{\sigma}_n^2 = \hat{a}_n^2\,\widetilde{\mathrm{Var}}\big(r_n^{[\mathrm{bg}]}\big) \tag{A.24}$$