PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'Office National d'Études et de Recherches Aérospatiales, au Département Optique et Techniques Associées

Adaptive-optics-corrected image restoration for astronomical and satellite observation: marginal approach by stochastic sampling

Soutenue par
Alix YAN
Le 04 octobre 2023

École doctorale nº127
Astronomie et Astrophysique d'Ile-de-France

Spécialité
Astronomie et Astrophysique

Composition du jury :

Hervé Carfantan IRAP	*Président du jury* *Rapporteur*
Éric Thiébaut CRAL	*Rapporteur*
Anne-Marie Lagrange LESIA	*Examinatrice*
Céline Meillier ICube	*Examinatrice*
Benoît Neichel LAM	*Examinateur*
Laurent Mugnier ONERA	*Directeur de thèse*
Jean-François Giovannelli IMS	*Directeur de thèse*
Cyril Petit ONERA	*Invité*

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

ONERA
THE FRENCH AEROSPACE LAB

Remerciements/Acknowledgements

En arrivant à l'ONERA il y a trois ans, il me semblait que la thèse relevait avant tout d'un apprentissage scientifique et technique. En vérité, ma première expérience professionelle dans la recherche a surtout été marquée par l'aspect humain que l'on a (hélas) souvent tendance à négliger. Ces quelques pages sont donc dédiées aux personnes qui ont rendu possible la rédaction de toutes les autres, qui m'ont accompagnée et soutenue durant ces trois années, et bien au-delà.

Laurent, c'est à toi que j'aimerais adresser en premier ces remerciements. Tu t'es *toujours* rendu disponible pour me voir et travailler avec moi, et je mesure la chance que j'ai eue sur ce point. J'aimerais te remercier pour ton aide précieuse et constante, ton soutien patient, tes conseils bienveillants et tes encouragements magnifaïques : tu as toujours été dans la communication directe et ouverte avec moi, un grand merci pour cela. Enfin, merci pour les madeleines ou les dattes (en fonction des jours), pour les recommandations théâtre et pour ton invitation à déjeuner avec Manuela le jour de ma mi-thèse (dont je me suis vantée auprès de tou.te.s les thésard.e.s), merci beaucoup.

Jean-François, je voudrais te remercier pour ta bonne humeur, ta rigueur, ton Bayésianisme, et ton sens de l'humour en toute circonstance. Toi et Laurent êtes deux chercheurs impressionnants, cependant vous êtes surtout restés très humains. Je pense sincèrement que c'est cela qui m'a convaincue de faire cette thèse avec toi, après avoir suivi ton cours en Master 2 où tu dégageais la même énergie (peut-on dire que j'avais un bon *a priori* ?). Enfin, je te remercie surtout de ne pas m'avoir fouettée quand je venais à Bordeaux te rendre visite, mais au contraire d'avoir grand ouvert la porte très décorée de ton bureau à l'IMS.

Cyril, je sais que tu regrettes de ne pas avoir été aussi disponible que tu l'aurais voulu durant ma thèse, car tu me le dis parfois en t'excusant puis en marchant vite d'un côté ou de l'autre de l'aile B. J'aimerais donc à mon tour te dire que j'ai énormément apprécié les discussions qu'on a eues, qu'elles aient été dans une salle de réunion, à la table du repas de gala des JRIOA à Dijon ou bien contre le radiateur face à mon bureau en regardant l'extérieur ou le matériel dans tes mains. Tu es aussi une personne très humaine et accessible, et je te remercie pour ton aide efficace et nos discussions ouvertes. Tu m'as beaucoup plus aidée que tu ne le penses, merci pour tout.

Romain, un énorme merci pour ton aide et ton soutien, en particulier en début de thèse. C'était très précieux de pouvoir échanger avec toi qui venais de soutenir ta thèse et d'être embauché, donc qui étais le plus proche des codes et des équations que je m'apprêtais à utiliser.

Merci pour ta bienveillance et ta pédagogie, qui font beaucoup de bien et m'ont énormément aidée à prendre confiance en moi. Enfin, merci pour ta gentillesse, pour les pauses thé (avec ou sans perles), pour les jeux en ligne et nos chouettes discussions à Châtillon ou à Marseille, on remet ça quand tu veux !

Je remercie Eric Thiébaut et Hervé Carfantan d'avoir rapporté ma thèse, m'ayant permis d'améliorer significativement le manuscrit. J'aimerais particulièrement insister sur le fait qu'Hervé Carfantan a, de plus, accepté de présider ma soutenance de thèse. Je remercie également Céline Meillier, Benoît Neichel et Anne-Marie Lagrange (dont je regrette l'absence le jour de la soutenance) de m'avoir fait l'honneur d'être également dans mon jury de thèse.

J'aimerais remercier l'école doctorale Astronomie et Astrophysique d'Ile de France, plus particulièrement Thierry Fouchet, Alain Doressoundiram ainsi que Géraldine Gaillant et Jacqueline Plancy, pour leur aide et leur patience.

C'est dans le cadre de ma formation doctorale que j'ai pu suivre le programme de mentorat Femmes et Sciences PSL, qui m'a permis de faire la connaissance de nombreuses chercheuses et de m'ouvrir à d'autres domaines scientifiques. J'aimerais toutes les remercier pour les échanges formels ou non qu'on a eus, plus particulièrement je voudrais remercier Rhita, Sophie (M.), Sophie (T.), Valentine et Elsa du copil, ainsi que les doctorantes du mentorat. Enfin, j'aimerais remercier Alix ma mentore pour nos rendez-vous mensuels : merci pour ton écoute, tes conseils et ta bienveillance. Nos déjeuners / pauses café étaient une bulle d'air pour moi, et te parler m'a permis de réaliser beaucoup de choses sur la recherche et sur moi-même. J'espère qu'on se reverra à l'occasion !

En plus de mes encadrants, je remercie l'équipe Haute Résolution Angulaire (HRA) de m'avoir accueillie durant ces trois ans. J'aimerais remercier plus particulièrement Aurélie (qui a accepté de faire partie de mon comité de suivi de thèse), Joseph et Jean-Marc pour leur gentillesse et leur aide. Nos discussions spontanées devant mon bureau, bien que brèves, ont toujours été un plaisir. Comme autres personnes de l'ONERA, je remercie vivement Marylène pour sa patience, pour son aide précieuse et pour les heures passées au téléphone à propos de réservation d'avion ou de train, ou encore d'envoi postal du manuscrit. De plus, je remercie Chantal qui m'a bien dépannée en crème, pansements ou autres médicaments, et avec qui il est toujours agréable de discuter. Enfin, on ne peut parler de l'équipe HRA sans parler de ses doctorant.e.s ! Je remercie tou.te.s celles et ceux que j'ai eu la chance de croiser durant ma thèse. A Châtillon, je remercie plus particulièrement Chloé mon ancienne co-bureau qui m'a initiée à l'utilisation de l'intranet ONERA, Yann (L.-T.) qui restera à mes yeux toujours un peu doctorant dans l'âme, Antoine & Luca le duo de choc, Hiyam pilier de l'équipe à l'humour légendaire, Valentina pour son style de musique et les délicieuses *arepas* (malgré tes courses et tes lessives à faire), Pablo pour son aide précieuse bien que *nuestra historia de amor se acaba ya y me voy*, Emile pour tous les bons moments passés ensemble que ce soit dans les bars parisiens, les spas dijonais ou les pubs rouennais, Pierre mon co-bureau de l'ambiance très indulgent sur mon chant douteux, Timothée le respo mots croisés (force i), Yann (L.) notre mascotte préférée qui fait les meilleures crêpes et autres douceurs ~~trop alcoolisées~~ bretonnes, et Inès la crack de l'oscillo au grand coeur avec qui nous avons beaucoup de restaurants à faire. Je pense également à Mehdi et Eve, qui nous ont manqué dans les locaux mais que j'étais toujours heureuse de retrouver en dehors. Plus ou moins loin de Châtillon, je remercie Mahawa avec qui j'ai eu la chance d'échanger quelques pas de salsa, Maxime, Arnaud, Nicolas, Hugo, Daniel et Cyril, que j'ai été contente de voir lors de leurs passages à Châtillon/Palaiseau, à Poitiers ou lors des conférences.

I would like to thank the PhD students I have met in Montreal, Nancy or Edinburg for the great times we had together during the conferences, there are precious memories to me. Finally, I want to thank Alexis my twin sister from Marseille, for being such a funny and joyful friend. Thank you for our Zoom calls on AMIRAL (which allowed me to track the evolution of your biceps), for the great times we had in Québec and in Marseille, or for introducing me to the world of stickers. I wish you all the best, in your research career as well as your personal life.

J'aimerais maintenant remercier certaines (ex-)collègues, maintenant amies, qui ont changé mes trois années de thèse. Je veux commencer par l'incontournable Perrine Pépé Peps Tchipeur, coéquipière du duo de putois ainsi que du 29/12. Tu es une chercheuse brillante mais surtout une amie, et je n'aurais pu espérer meilleure co-thésarde pour ces trois années. Merci d'avoir partagé avec moi les galères des calculs (toujours au tableau) et les joies des papiers soumis, les courses dans la forêt et les courses à Tang, merci pour les discussions, les rires, les confidences, les boulettes, les joggs, les soirées et tellement d'autres choses ! Ta positivité m'a boostée et tu étais toujours là dans les moments où j'en avais le plus besoin, merci infiniment. J'ai hâte de fêter dignement la fin de ta thèse, avec un carton de nouilles instantanées j'espère ! *emoji feu*

Léa, la liste de tes qualités est semblable à la liste des endroits qu'on s'est promis de tester, c'est-à-dire qu'elle se remplit toujours plus à chaque fois qu'on se voit. Merci pour tous les moments qu'on a partagés ensemble, les bubble teas instagrammables, John Weng, le bannissement du durian, le Futuroscope, l'appel visio depuis ton salon, toutes ces discussions qui passent trop vite (alors qu'on n'a pas fini de tout se raconter !) et ces fous rires dont je garde précieusement le souvenir. Tu es une personne impressionnante avec un cœur en or, et je suis très heureuse d'être ton amie. Je suis impatiente de poursuivre ce resto-thon parisien avec toi (on ne sait jamais, peut-être que cette liste a une fin), on se voit très bientôt ! *emoji hamster*

Laurie, je crois que la première image qui me vient lorsque je pense à ces trois ans est une image de nous, dans ton bureau, en train de parler autour d'un thé chaud et d'un Kinder country. Il s'agit définitivement d'un des souvenirs les plus doux de ma thèse. Merci pour ta gentillesse, ton empathie, ta générosité, tes conseils, ta patience, ton humour et ta sincérité. Grâce à toi, je me suis immédiatement sentie à l'aise à Châtillon, car il est très facile et en même temps très rassurant de venir se confier à toi. C'était vraiment une chance de t'avoir comme collègue puis comme amie au labo, tu es une très belle personne et je ne te souhaite que le meilleur pour la suite. J'ai hâte de vous revoir, Patapouf et toi ! *emoji koala*

Iva, I would like to start by saying that you are an impressive researcher. I have always admired you for that, ever since we met online during the weekly HRApéros. I feel incredibly lucky and proud that today I can call you both my colleague and my friend. You are caring, funny, honest, trustworthy and very brave. You taught me a lot, but above all it always feels good to be around and talk with you. Thank you for all the great moments we shared together, I am definitely looking forward to our even greater upcoming projects. I do not doubt that they will be filled with tea (or cocktails depending on the time of day), nice food, dance steps and guitar, darts, scribbles on a whiteboard and most importantly, joy. You are an amazing friend, I love you! *Mont d'Or emoji*

Cette thèse est dédiée à mes parents et mon frère (le plus mignon de tous), que j'aime énormément et à qui je dois beaucoup. Je veux les remercier pour tout l'amour et le soutien qu'ils me donnent depuis toujours. Il est évident que je n'aurais pas fait tout cela sans eux. Mathis, Dad, 엄미, 사랑해! ♡

J'aimerais enfin remercier mes ami.e.s. Je ne m'étalerai pas ici (ceux qui savent se reconnaîtront), et espère plutôt pouvoir vous dire combien vous comptez pour moi lorsque nous nous reverrons très prochainement. Vous êtes des perles, merci infiniment d'être là. Un câlin tout particulier à Wario, avec qui j'ai hâte de passer de douces journées en Écosse.

***

Playlist ayant contribué à cette thèse :

- I Will Survive, Gloria Gaynor
- Halo, Beyoncé
- Talkin' Bout a Revolution, Tracy Chapman
- J'aimerais tellement, Jena Lee
- Les gens qui doutent, Anne Sylvestre
- Quelqu'un m'a dit, Carla Bruni
- Balance ton quoi, Angèle
- Love Again, Dua Lipa

Résumé

La restauration d'images corrigées par optique adaptative est particulièrement difficile, du fait de la méconnaissance de la réponse impulsionnelle du système optique (PSF pour point-spread function) en plus des difficultés usuelles. Une approche efficace est de marginaliser l'objet en dehors du problème et d'estimer la PSF et les hyper-paramètres (liés à l'objet et au bruit) seuls avant la déconvolution. Des travaux récents ont appliqué cette déconvolution marginale, combinée à un modèle paramétrique de PSF, à des images astronomiques et de satellites. Cette thèse vise à proposer une extension de cette méthode. En particulier, j'utilise un algorithme Monte-Carlo par chaînes de Markov (MCMC), afin d'inclure des incertitudes sur les paramètres et d'étudier leur corrélation a posteriori. Je présente des résultats détaillés obtenus sur des images astronomiques et de satellites, simulées et expérimentales. Je présente également des premiers éléments sur l'ajout d'une contrainte de support sur l'objet.

Mots-clés : restauration d'image, optique adaptative, déconvolution, turbulence, problèmes inverses

Abstract

Adaptive-optics-corrected image restoration is particularly difficult, as it suffers from the poor knowledge on the point-spread function (PSF). One efficient approach is to marginalize the object out of the problem, and to estimate the PSF and (object and noise) hyper-parameters only before the deconvolution. Recent works have applied this marginal deconvolution, combined to a parametric model for the PSF, to astronomical and satellite images. This thesis aims at extending this previous method, using Markov chain Monte Carlo (MCMC) algorithms. This will enable us to derive uncertainties on the estimates, as well as to study posterior correlation between the parameters. I present detailed results on simulated and experimental, astronomical and satellite data. I also provide elements on the impact of a support constraint on the object.

Keywords: image restoration, adaptive optics, deconvolution, turbulence, inverse problems

Contents

Acronyms

AMIRAL Automatic Myopic Image Restoration ALgorithm

CNRS *Centre National de la Recherche Scientifique*

DFT Discrete Fourier Transform

DM Deformable Mirror

MAP Maximum *A Posteriori*

MCMC Markov Chain Monte Carlo

MISTRAL Myopic Iterative STep-preserving Restoration ALgorithm

MMAE Minimum Mean Absolute Error

MMSE Minimum Mean Square Error

MéO *Métrologie Optique*

OASIS Optimized Astrophysical Simulator for Imaging Systems

ONERA *Office National d'Etudes et de Recherches Aérospatiales*

OTF Optical Transfer Function

P-AMIRAL Parametric Automatic Myopic Image Restoration ALgorithm

PSF Point Spread Function

RMSE Root Mean Square Error

RTC Real-Time Computer

SNR Signal-to-Noise Ratio

VLT Very Large Telescope

WFS WaveFront Sensor

Résumé long en français

L'observation de l'espace depuis le sol possède de nombreuses applications, civiles comme militaires, telles que l'astronomie ou l'observation de satellites. En effet, l'observation des satellites naturels ou artificiels permet de les analyser et de suivre leur évolution depuis la Terre au lieu d'envoyer une sonde dans l'Espace (ce qui est beaucoup plus contraignant et onéreux). Cependant, cette observation depuis le sol est limitée à cause de la turbulence atmosphérique : l'onde plane lumineuse provenant de l'objet est déformée par la turbulence. Cette déformation aléatoire, qui est la manifestation d'aberrations que traverse l'onde, peut être décrite par une carte de phase dans la pupille du télescope. Dans ce travail, la force de la turbulence est rattachée au paramètre de Fried r_0. Il correspond au diamètre d'un télescope qui serait uniquement limité par la diffraction donnant la même résolution qu'un télescope infiniment grand mais limité par la turbulence. Dans les cas que nous allons étudier dans ce travail, le paramètre de Fried est de l'ordre d'une dizaine de centimètre, de nuit, dans le visible.

Une technique permettant de compensation les perturbations liées à la turbulence est l'optique adaptative (OA). Cette dernière peut-être décomposée en trois parties principales : le miroir déformable, le calculateur temps réel et l'analyseur de surface d'onde. L'analyseur de surface d'onde mesure les aberrations résiduelles qui sont envoyées au calculateur temps réel, lequel pilote le miroir déformable pour compenser ces aberrations. Le front d'onde est alors corrigé, mais la phase (résiduelle) dans la pupille du télescope n'est pas totalement constante, se traduisant par un flou résiduel sur l'image corrigée par OA. Ce résidu turbulent est donc celui que nous allons chercher à traiter dans ces travaux. Dans cette thèse, ce résidu va être relié à la qualité de correction par OA, donnée par la variance de phase résiduelle v_ϕ.

Le processus qui, partant de l'objet réel observé, donne l'image effectivement obtenue est décrit par le modèle d'imagerie. Dans notre cas, nous utilisons un modèle d'imagerie assez classique dans la littérature, considérant que l'image $\boldsymbol{i}$ résulte du produit de convolution entre l'objet recherché $\boldsymbol{o}$ et la réponse impulsionnelle du système optique, appelée fonction d'étalement du point (en anglais PSF pour *point spread function*) $\boldsymbol{h}$, auquel s'ajoute du bruit $\boldsymbol{n}$: $\boldsymbol{i} = \boldsymbol{o} * \boldsymbol{h} + \boldsymbol{n}$. Le fait de vouloir retrouver l'objet $\boldsymbol{o}$ à partir des données $\boldsymbol{i}$ revient à résoudre un problème dit inverse, il s'agit plus spécifiquement ici d'un problème de déconvolution. Il s'agit d'un problème difficile, même lorsque la PSF $\boldsymbol{h}$ est totalement connue. Dans notre cas, la PSF n'est que partiellement connue, ce qui rend le problème encore plus difficile : il faut donc estimer l'objet et la PSF.

Pour cela, nous allons utiliser des modèles afin d'utiliser les connaissances physiques sur les grandeurs recherchées pour réduire le nombre d'inconnues. Pour la PSF, les travaux de thèse de R. Fétick ont permis de développer un modèle de PSF corrigée par OA à peu de paramètres [Fétick, 2019a]. Dans ce modèle, la PSF est décrite à partir de la DSP de la phase résiduelle, elle-même décrite en deux parties : une première partie en-dessous de la fréquence de coupure de l'OA où la phase a été partiellement corrigée, et une seconde partie au-dessus de cette fréquence de coupure où la phase est une phase purement turbulente. La première partie est principalement décrite par la variance de phase résiduelle v_ϕ, et la seconde par le diamètre de Fried r_0. La PSF est donc totalement décrite à partir de ces deux paramètres. Ce modèle, qui a été testé et validé sur plusieurs systèmes d'OA (SAXO du VLT/SPHERE, GALACSI du VLT/MUSE, ODISSEE de l'ONERA), est le modèle que nous utiliserons tout au long de ce travail. Pour le bruit, nous considérons essentiellement deux sources de bruit : le bruit de photon et le bruit de lecture, qui peuvent être raisonnablement approximés comme un bruit blanc additif Gaussien, à moyenne nulle. Le bruit est alors décrit par un unique paramètre, sa variance (ou sa précision c'est-à-dire son inverse variance).

Les caractéristiques de la turbulence atmosphérique et donc son impact sur les images sont données plus en détail dans le chapitre 1. On y trouve également la motivation d'utiliser l'OA, une brève description de cette dernière ainsi qu'une description du modèle PSFAO19.

Afin d'estimer l'objet et les paramètres PSF et bruit, nous allons nous placer dans un cadre Bayésien où les informations disponibles sur les quantités recherchées sont incorporées sous la forme de distributions de probabilité : on parle d'*a priori* sur ces quantités. La formule de Bayes permet alors de calculer la probabilité *a posteriori* des quantités d'intérêt (sachant les données), à partir de leurs probabilités *a priori* et de la vraisemblance, c'est-à-dire la probabilité des données sachant ces quantités, donnée par le modèle d'imagerie et le modèle de bruit (Gaussien) dans notre cas. Pour l'objet, nous allons utiliser un *a priori* Gaussien, l'objet est alors décrit par sa moyenne et sa densité spectrale de puissance (DSP). Le modèle utilisé pour la DSP objet est le modèle de Matérn, issu de la littérature.

Une fois la probabilité *a posteriori* écrite, plusieurs possibilités existent afin d'estimer l'objet et la PSF. L'approche historique (qui est également la plus simple) consiste à estimer l'objet et la PSF conjointement, en maximisant la distribution *a posteriori* conjointe c'est-à-dire $p(\boldsymbol{o}, \boldsymbol{h}|\boldsymbol{i})$. Cependant, les travaux antérieurs de L. Blanco [Blanco, 2011] et de A. Levin [Levin, 2009] ont montré que, sous les hypothèses de Gaussianité faites précédemment, la solution obtenue avec cet estimateur conjoint Maximum A Posteriori (MAP) est dégénérée, menant systématiquement à la PSF la plus piquée et donc l'objet le plus mou. En pratique, R. Fétick a également montré le fait que l'estimateur ne donne pas de résultats satisfaisant même en utilisant le modèle paramétrique et une restauration L1-L2 (ou L2 avec contrainte de positivité) [Fétick, 2020b].

Une autre approche dans la littérature consiste à estimer d'abord la PSF avant de restaurer l'objet, en intégrant l'objet hors du problème. De ce fait, le nombre d'inconnues (les paramètres de la PSF) est bien inférieur à la quantité de données disponibles (l'ensemble des pixels de l'image), ce qui donne de bonnes propriétés asymptotiques à cet estimateur dit marginal. Grâce à la marginalisation, il est alors possible d'inclure l'estimation des hyper-paramètres, à savoir certains des paramètres de la DSP objet $\boldsymbol{\theta_o}$ et l'inverse variance du bruit θ_n, dans l'estimation des paramètres de PSF. On maximise donc cette fois $p(\boldsymbol{h}, \boldsymbol{\theta_o}, \theta_n|\boldsymbol{i})$, et une fois la PSF estimée, on restore l'objet en utilisant une méthode de restoration classique en utilisant les paramètres et hyper-paramètres estimés, par-exemple un filtre de Wiener (restauration L2). Cette méthode marginale a été développée et implémentée dans le code AMIRAL (pour Automatic Myopic Image Restoration ALgorithm). Cette méthode, d'abord utilisée pour l'imagerie rétinienne, a ensuite été adaptée à l'observation astronomique en utilisant le modèle PSFAO19 [Fétick, 2019a] dans la version Parametric-AMIRAL.

Le chapitre 2 redonne des éléments sur la déconvolution myope, et décrit les différents estimateurs mentionnés plus haut.

La méthode MAP marginale fonctionne, cependant elle possède plusieurs limitations. Premièrement, elle ne donne pas d'incertitudes sur les quantités estimées. Deuxièmement, comme dit précédemment elle ne permet d'estimer que certains des hyper-paramètres : en effet, l'un des paramètres de la DSP objet doit être fixé en fonction de la classe d'objet observé, on parle de mode quasi non-supervisé. Sans cela c'est-à-dire dans le mode non-supervisé, la méthode ne donne pas de résultat satisfaisant [Fétick, 2020b], et il reste encore à déterminer pourquoi. De plus, il serait souhaitable à terme d'estimer l'ensemble des paramètres. Troisièmement, la méthode MAP marginale ne donne pas de résultats satisfaisants pour des cas de correction par OA moins bonne, ou pour des restorations d'objets plus petits [Lau, 2023].

Les objectifs de ces travaux de thèse sont donc multiples. Le premier but a été de proposer une nouvelle méthode donnant une estimation des quantités d'intérêts ainsi que leurs incertitudes associées. Le second a été de comprendre les limitations de la méthode MAP marginale. Le dernier a été d'améliorer les estimations dans les cas difficiles. Afin de répondre aux deux premiers buts, nous avons exploré plus globalement la distribution *a posteriori* afin de calculer un nouvel estimateur et les incertitudes associées. Quant à la réponse au troisième but, l'approche a été d'ajouter des informations inexploitées jusqu'alors dans la méthode. Nous avons plus spécifiquement travaillé sur l'ajout d'une contrainte de support sur l'objet.

Le chapitre 3 regroupe les différentes contributions faites à l'ancienne méthode MAP marginale au cours de la thèse. On y décrit la modification du modèle de DSP objet que nous proposons afin de découpler davantage ses paramètres tout en décrivant les mêmes cartes de DSP. De plus, on y compare les deux versions de la méthode MAP marginale, à savoir AMIRAL et P-AMIRAL, jusqu'alors utilisées séparément et donc jamais comparées.

Les chapitres 4 à 6 décrivent la nouvelle méthode visant à donner des incertitudes sur les estimées ainsi que les corrélations *a posteriori* des paramètres estimés (chapitre 4), et donnent les résultats obtenus sur des données simulées (chapitre 5) et expérimentales (chapitre 6).

Enfin, le chapitre 7 décrit l'ajout de la contrainte de support sur l'objet, en regardant son impact sur les calculs. Il présente également des résultats à support connu, vrai et approché, pour différents niveaux de bruit.

Introduction

Ground-based high angular resolution imaging has both civilian and military applications, such as astronomy and satellite observation. In particular, the observation of natural or artificial satellites from the ground enables us to analyze them and track their evolution from Earth using large telescopes, instead of sending a spacecraft, which is way more constraining and expensive. These space observations can be done for different purposes: in the case of asteroids, observing them enables astronomers to characterize their chemical composition, which has not evolved much with time due to their relatively small size. This knowledge can then be used, for instance, to better understand the original chemical compounds of planets, thus the first phases of the Solar system formation. Tracking asteroids can also be useful in order to determine their orbit and thus identify the ones that will come close to or even impact Earth. Observing artificial satellites is also useful to follow their different life phases, from post-launch manoeuvres to deorbiting, as well as the potential threats around them. Military applications also consist in the identification and characterization of satellites. Finally, these observations are also important to follow the evolution of space debris, for instance their rotation speed, as they can be a hazard for active spacecrafts and other satellites.

However, these ground-based observations are limited due to atmospheric turbulence, which impacts the incoming light waves by deforming the wavefront, thus impairs the observed data by blurring the object. The observations can be corrected in real time, using an Adaptive Optics (AO) system with the telescope. However, the AO correction is partial and residual blurring remains, impacting mostly high spatial frequencies of the observed object. Therefore, the observation system includes post-processing to restore the high frequencies and retrieve the actual object, taking into account the residual turbulence and the optical limitations of the telescope.

The residual blurring is described by the so-called system point-spread function (PSF). Supposing that the data results from the convolution between the sought object and the PSF, retrieving the object from the data is called a deconvolution problem. Solving this inverse problem is already a difficult task as inverse problems are often ill-posed, but in this case the PSF is not entirely known which complicates even more the problem, given that both the object and the PSF have to be estimated. Such deconvolution problems, where the PSF is unknown or partially known, are called respectively blind or myopic deconvolutions. One possible way to solve inverse problems in general, proposed in the literature, is to do so in the

Bayesian framework, by incorporating knowledge on the unknowns through prior probability distributions. Retrieving the unknowns is then often done using their posterior distribution, taking into account the data, prior information and the imaging model.

As said previously, both the object and the PSF have to be estimated. The historical way to proceed is to estimate them jointly, for instance maximizing the joint probability density function giving the joint Maximum A Posteriori (MAP) estimator [You, 1996; Ayers, 1988; Mugnier, 2004]. However, previous works have shown that it leads to a degenerated solution [Blanco, 2011; Levin, 2009] in the absence of strong constraints. Another efficient way to proceed is to first estimate the PSF by "marginalizing" over the object, *i.e.* by integrating the joint probability density function over all possible objects (with a given prior probability density function), and then to deconvolve the image with the estimated PSF which maximizes the marginal probability density function [Lehmann, 1998; Blanco, 2011]. First used for retinal imaging, this marginal MAP estimator was then adapted to astronomical and satellite observation, and was implemented in a code named AMIRAL, standing for *Automatic Myopic Image Restoration ALgorithm* by Blanco et al. [Blanco, 2011].

In this context, the PSF describes the blurring coming from the partial correction of atmospheric turbulence by AO. Previous works [Fétick, 2019a] defined a sparse physical model for it, taking into account both the turbulence and the AO correction, using a few physical parameters. This model was then incorporated to the marginal MAP estimator, leading to another version of the corresponding myopic image restoration method: Parametric-AMIRAL. This method was implemented and tested on simulated and experimental astronomical data by Fétick et al. [Fétick, 2020b].

P-AMIRAL, the previous version incorporating the AO-corrected PSF model, was not compared to AMIRAL, its former version without the PSF model, and was not applied to satellite images. Moreover, the method only gives access to one element of the posterior distribution which is its maximizer, from which one can infer only local (and not global) information on the posterior distribution. Finally, none of the previous methods takes into account some constraints on the object which could improve the estimation quality.

The present work tackles the problem of AO-corrected image restoration, for astronomical and satellite observations. It aims at extending the previous method P-AMIRAL in order to include more information on the estimated unknowns such as uncertainties and correlation. This is done by sampling the whole marginal posterior distribution using stochastic sampling, which enables us to compute another Bayesian estimator, which minimizes the mean square error (MMSE). In this work, we also aim at exploring the impact of a support constraint on the object in order to make the estimation more robust. This manuscript is divided into seven chapters, the two first giving the general context of the problem, as well as the description of the tools and methods that are used in the last five chapters. The chapters are structured as follows:

- Chapter 1 recalls some characteristics of atmospheric turbulence in order to understand its impact on the data, and the need for Adaptive Optics. General principles on AO systems are also given, as well as their limitations. We tackle the problem for both applications, astronomical and satellite observations, and discuss the differences between the first and the second one. The introduction to turbulence and AO correction finally leads to the description of the AO-corrected PSF model used in this work.

- Chapter 2 recalls some elements on (myopic) deconvolution and describes the estimators used in the literature to solve this inverse problem. We describe these different estimators given our framework and give motivations for a change of the previous marginal MAP estimator, as well as potential options.
- Chapter 3 gives the different contributions that were made to the existing method. Firstly, it describes the modification on the object PSD model that aims at decoupling its parameters. Then, the existing versions of the previous marginal MAP estimator, AMIRAL and P-AMIRAL, are compared by means of simulations, and the conclusion on this comparison is given at the end of the chapter.
- Chapter 4 describes the computation of the marginal MMSE estimator by giving all the elements of the Bayesian framework (likelihood, prior and posterior distributions, hierarchical model). It also describes the Markov Chain Monte Carlo (MCMC) methods we are using in this work.
- Chapter 5 provides the results using this new method, on simulated data. On the astronomical case, we are discussing both the estimated parameters and the derived uncertainties, and test the robustness of the method by running it over several noise realizations. We also compare two different MCMC algorithms in terms of computation speed. Additionally to the uncertainties, the new method provides information on the posterior coupling of parameters which we discuss afterwards. Finally, the method is also tested on (simulated) satellite data.
- Chapter 6 provides the results using the new method, on experimental data, firstly on an asteroid image and secondly on a satellite image. For both contexts, we start by a description of the data and the context, then proceed to a sanity check on the image PSD before giving results on the estimated parameters as well as on the derived uncertainties, and restored images.
- Chapter 7 paves the way for including additional information on the object, through a support constraint. We first give some elements from the literature, before giving details on the different options to include this support constraint, which reveals the need to compute truncated covariance matrices. Different implementations of the truncated matrices computation are given and compared. Finally, we study the impact of different known supports on the estimated parameters on simulated data.

Chapter 1

Observing satellites from the ground: basic concepts and models

Introduction

In this chapter, we give the basic concepts and models we are using in order to tackle the AO-corrected image restoration problem. In Section 1, we describe the impact of atmospheric turbulence on the image. In Section 2, the general principle of AO is given, as well as its limitations. In Section 3, we describe the AO corrected PSF model we are using in this work. Finally, in Section 4, we give the imaging model we will use throughout this work as well as some elements about our application cases and some examples of instruments which are equipped with an AO system.

Contents

1.1 Imaging an extended object through atmospheric turbulence

1.1.1 Physical characteristics of atmospheric turbulence

The atmosphere is a heterogeneous medium where different air volumes are mixed, creating large-scale air motions which transfer kinetic energy to smaller and smaller scale air motions, in an energy cascade. This happens until the energy transfer is dominated by viscous friction in the medium, when the residual energy is dissipated as heat. This energy cascade from large-scale air motions, of characteristic size L_0 (also called the outer scale) around $10\,\mathrm{m} - 100\,\mathrm{m}$, to small-scale air motions, of characteristic size l_0 (the inner scale) around $10^{-3}\,\mathrm{m} - 10^{-2}\,\mathrm{m}$, is described by Kolmogorov in [Kolmogorov, 1941].

The mixing of different air volumes having different temperatures, due to these air motions, leads to irregular changes in the refractive index of the air. This disturbs the propagation of light wave through the atmosphere, which undergo different and random refractions and delays, impacting the electromagnetic wave. This is why, even though the theoretical resolving power of an optical instrument (in our case, a telescope) is limited by the physics of diffraction, in practice this theoretical limit can not be reached because of the optical aberrations, due to the atmospheric turbulence. Their impact is illustrated in Figure 1.1.

The perturbation of the electromagnetic wave during its propagation through turbulence depends on the part of the atmosphere through which the light waves went. To characterize the temperature fluctuations, and so the refractive index fluctuations $\Delta_n(\boldsymbol{r}, h)$, one can define a statistical model for these fluctuations, which can be described by their structure function D_{Δ_n} [Tatarski, 1961]:

$$D_{\Delta_n}(\boldsymbol{\rho}, h) = \langle [\Delta_n(\boldsymbol{r}, h) - \Delta_n(\boldsymbol{r} + \boldsymbol{\rho}, h)]^2 \rangle$$

with $\boldsymbol{\rho}$ the distance between the two points, h the altitude and $\langle . \rangle$ the averaging operator. In the inertial subrange, meaning for structures of size $\rho \in]l_0; L_0[$, this structure function depends on the refractive index structure constant of the turbulence $C_n^2(h)$ which is defined as follows:

$$D_{\Delta_n}(\boldsymbol{\rho}, h) = C_n^2(h) \rho^{2/3}$$

with ρ the modulus of $\boldsymbol{\rho}$. This constant characterizes the strength of atmospheric turbulence at a given altitude h. It has a key role in the description of turbulence and its effects, it depends on a large number of factors (meteorology, geographical localisation, day or night observation,...) and evolves with time.

To quantify the impact of the atmospheric turbulence on image quality, a parameter of interest is the Fried parameter r_0 [Fried, 1966], which is the integration of the refractive index structure constant of the turbulence (characterizing the local turbulence strength), over the line of sight:

$$r_0 = \left[0.423\left(\frac{2\pi}{\lambda}\right)^2 \frac{1}{\cos\gamma} \int_0^\infty C_n^2(h)\mathrm{d}h\right]^{-\frac{3}{5}}$$

where λ is the imaging wavelength, and γ the zenith angle (the angle between the direction of interest and the zenith). The Fried parameter, computed in meters, corresponds to the equivalent diameter that a telescope providing the same angular resolution would have, without optical aberrations meaning in the case of a diffraction-limited observation. This resolution limit is also often given using the seeing parameter $s = \frac{\lambda}{r_0}$.

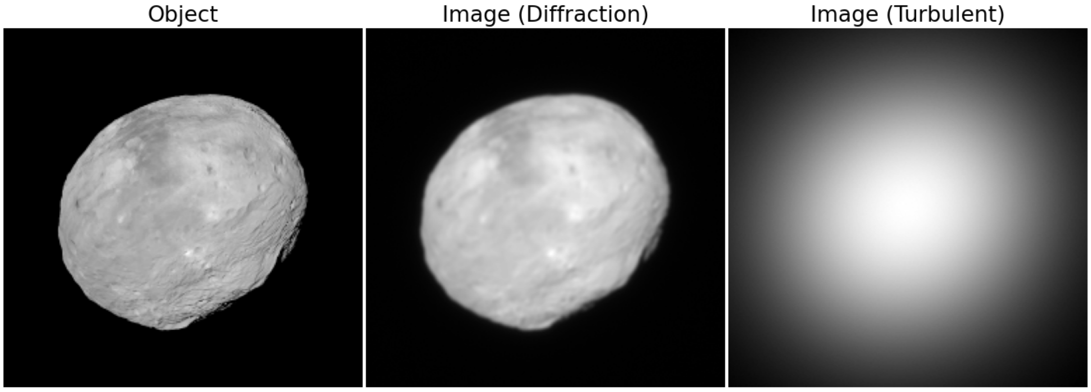

Figure 1.1 – Left: synthetic view of asteroid Vesta, built by OASIS [Jorda, 2010]. Center: simulated diffraction-limited observation of Vesta, using a VLT-like telescope (8 m-diameter telescope). Right: simulated observation of Vesta impaired by atmospheric turbulence, without any correction ($r_0 = 0.15\,\mathrm{m}$ at $\lambda = 550\,\mathrm{nm}$, which is typical value at the VLT's site, at Paranal).

In this work, we suppose that the atmosphere is a stratified medium, as the superposition of homogeneous thin turbulence layers [Roddier, 1981]. Additionally, we suppose that we are within the near field approximation, implying a low perturbation regime where we neglect the scintillation effects, meaning the effects of amplitude-shifting. Within this approximation, the phase of the electromagnetic field's complex amplitude in the telescope aperture (which we will name the turbulent phase) simply writes as the summation of the different phase-shifts coming from the layers through which the wave went. In the inertial subrange, following Kolmogorov's turbulence statistics theoretical model, the turbulent phase power spectral density (PSD) writes:

$$W_\phi(f) = 0.023 r_0^{-5/3} f^{-11/3} \tag{1.1}$$

with f the frequency *i.e.* the inverse of the position in the pupil, in m^{-1}.

1.1.2 Impact of atmospheric turbulence on the image

As shown in Figure 1.1, the atmospheric turbulence impairs the observation by blurring the object. We model its impact on the observed image, combined with the limitations and defects of the optical instruments, by a **point spread function** (PSF) to which the original object is convoluted (as we will write in Section 1.3). This system (telescope and atmosphere) PSF h_{inst}, at imaging wavelength λ, is the squared modulus of the inverse Fourier transform of the electromagnetic field's complex amplitude in the telescope aperture:

$$h_{inst}(\boldsymbol{\xi}) = |\mathrm{FT}^{-1}(P(\lambda \boldsymbol{u}) \exp(j\Phi(\lambda \boldsymbol{u})))|^2(\boldsymbol{\xi}) \tag{1.2}$$

where FT^{-1} denotes the inverse Fourier Transform (modeling here the transformation of the electromagnetic field by the telescope, between the pupil plane and the focal plane), $(\boldsymbol{\xi})$ are angles on the sky in radians, $(\lambda \boldsymbol{u})$ are coordinates in the pupil in meters, $P(.,.)$ is the pupil indicatrix and $\Phi(.,.)$ is the electromagnetic field's complex amplitude. This instantaneous PSF depends on the position we consider on the sky, as well as it depends on time. For a perfect telescope and without turbulence, P is constant within the pupil and Φ is null, the PSF is an Airy disk.

We can consider two different imaging regimes, thus two different types of PSF. The fact of being within a regime or the other depends on the exposure time t_{exp} with respect to the typical variation time of turbulence τ_{turb}, which is around 10 ms, depending on the site [Aristidi, 2020]. If $t_{exp} \ll \tau_{turb}$, the PSF is a **short-exposure** PSF, otherwise if $t_{exp} \gg \tau_{turb}$ it is a **long-exposure** PSF. In the short-exposure case, the PSF h_{inst} is scattered into several speckles, which keep the information on the object (even for high angular frequencies), however due to the short exposure time, the signal-to-noise ratio (SNR) is often low and the PSF is very complex to describe.

On the contrary, a long-exposure PSF h_{opt} corresponds to the integration of short-exposure PSF over time, meaning that the effects of turbulence are averaged:

$$h_{opt} = \langle h_{inst} \rangle_t$$

with $\langle . \rangle_t$ the averaging operator over time (t). The time-averaged speckles form a wider, smoother spot, which corresponds to a loss of resolution on the object. However, the SNR is higher due to the integration time. Roddier [Roddier, 1981] shows that the long-exposure PSF h_{opt} re-writes as the convolution of two PSFs: the first one is called the "static" PSF h_{stat} and corresponds to the static aberrations, the second one is the "atmospheric" PSF h_{atm} corresponding to the impact of atmospheric turbulence:

$$h_{opt} = h_{stat} * h_{atm} \tag{1.3}$$

Moreover, assuming additionally that the phase is stationary and its variance is finite, Roddier relates the atmospheric PSF h_{atm} to the phase PSD W_ϕ through the following expression:

$$\begin{aligned} h_{atm} &= \mathrm{FT}^{-1}(\exp(-B_\phi(0) - B_\phi)) \\ &= \mathrm{FT}^{-1}(\exp(-\sigma_\phi^2 - \mathrm{FT}^{-1}(W_\phi))) \end{aligned} \tag{1.4}$$

with B_ϕ the phase autocorrelation, thus $\sigma_\phi^2 = B_\phi(0)$ the phase variance and W_ϕ the phase PSD. Throughout this manuscript, and as discussed later in Section 1.3, we will only consider long-exposure PSFs in this work.

1.2 The use of Adaptive Optics (AO)

1.2.1 General principle

In order to improve the angular resolution of the optical instruments and to compensate the effects of atmospheric turbulence, several techniques were developed, in particular post-processing techniques [Labeyrie, 1970; Primot, 1990]. However, most of these techniques require to have an exposure time short enough to consider the effects of turbulence as instantaneous (avoiding the turbulence averaging effect of long-exposure images). This implies working on short-exposure data, for which the SNR is significantly lower (even though today the detectors have a smaller and smaller read-on noise) and the PSF is more complex to describe (requiring a great number of parameters to describe the turbulence-induced phase). Thus, using Adaptive Optics (AO) enables us to bypass this problem, given that it corrects the turbulent effects before the images are recorded.

First imagined in the 1950's [Babcock, 1953; Rousset, 1990], AO is a technique which consists in deforming precisely a mirror in order to compensate for light perturbation, meaning correcting the phase in the telescope aperture. This phase in (the plane conjugated to) the aperture is called the wavefront. Therefore, an AO system has three main elements:

— a **WaveFront Sensor** (WFS), which measures the incoming (turbulent) wavefront,

— a **Real-Time Computer** (RTC), which processes the data from the WFS in real-time,

— and finally, the RTC controls a **Deformable Mirror** (DM), which enables the user to impose to the incoming wavefront the desired correction, which should be the complementary of the measured aberrations (if the turbulence-induced aberrations were static).

We illustrate a simplified AO system in Figure 1.2.

1.2.2 Limitations and quality measurement

The wavefront correction using AO is partial: indeed, various processes (noisy measurements, limited number of degrees of freedom for measurements and correction, fast evolution of turbulence,...) undermine the correction level. The quality of the AO correction can be measured according to different criteria. One of them is the Strehl Ratio (SR), which gives the ratio between the actual PSF maximum value and the maximum value of a perfect impulse response of the instrument, without optical aberrations (for instance in the case of a circular pupil, an Airy pattern).

Given that the AO system is correcting the incoming phase in the aperture (wavefront), another way to characterize the quality of correction by AO is the residual phase variance σ_ϕ^2 (in rad^2). This residual phase variance is related to the Strehl Ratio through Maréchal's approximation (in the case of small variance): $SR \approx e^{-\sigma_\phi^2}$. It is also related to the different errors making AO correction partial. The error budget of the AO system sums up the contributions

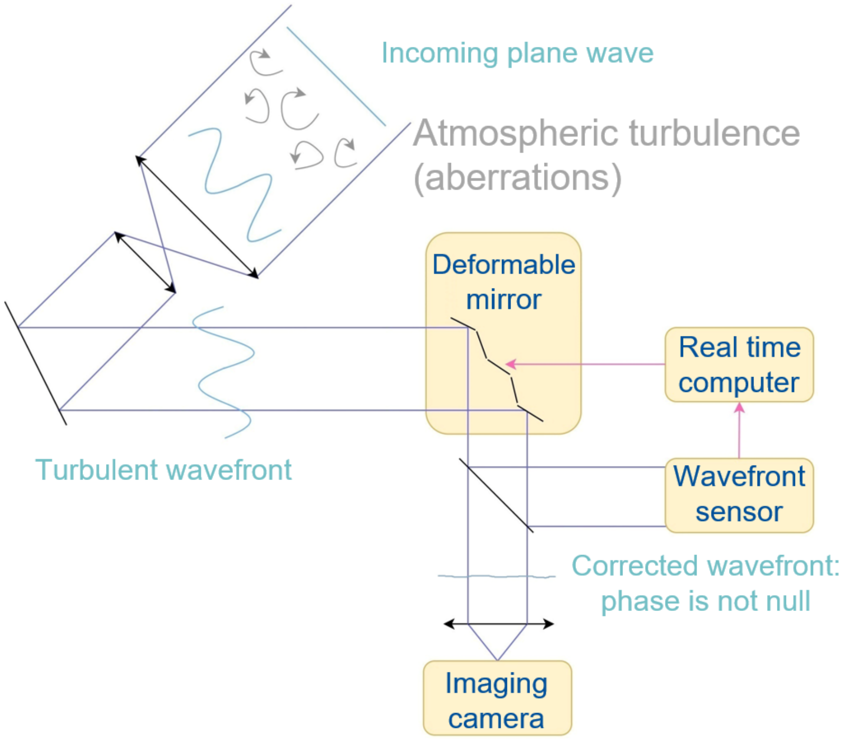

Figure 1.2 – Illustration of a (simplified) adaptive optics system.

of each process, and can be written as follows [Rigaut, 1998]:

$$\sigma_\phi^2 = \sigma_{fit}^2 + \sigma_{temp}^2 + \sigma_{alias}^2 + \sigma_{noise}^2 + \sigma_{aniso}^2$$

with:

- σ_{fit}^2 the fitting error, which is related to the DM. Indeed, the DM is composed of a given number of actuators, determining the maximum spatial frequency that the AO system can correct. The so-called AO cutoff frequency depends on N_{act} the number of actuators on the DM's diameter and on the primary mirror diameter D: $f_{AO} \approx \frac{N_{act}}{2D}$.
- σ_{temp}^2 the temporal error, which includes all delays (exposure time, image capture, RTC computation time, sending of commands, and so on. The system corrects the wavefront but with a slight temporal delay, and the faster the evolution of turbulence, the greater the error.
- σ_{alias}^2 the aliasing error, coming from the fact that the WFS is sampling the phase spatially, which makes possible the aliasing effect leading to a wrong estimation (thus, correction) of frequencies close to the Nyquist frequency of the WFS.
- σ_{noise}^2 the noise propagation error, coming from the propagation of photon noise and readout noise.
- σ_{aniso}^2 the anisoplanetism error, which is due to the spatial decorrelation of turbulence, if the observed object is larger than the isoplanetic patch.

The impact of AO correction of atmospheric turbulence is shown in Figure 1.3.

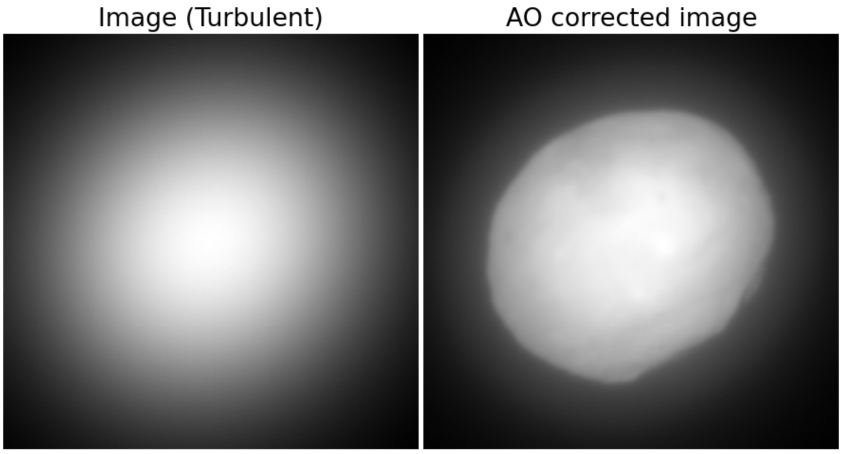

Figure 1.3 – Left: simulated observation of Vesta impaired by atmospheric turbulence, without any correction, given in Figure 1.3 (right). Right: simulated observation of Vesta impaired by atmospheric turbulence, with AO correction. The simulated AO system is a VLT/SPHERE-like system, and $v_\phi = 1.3\,\mathrm{rad}^2$ (which is a typical value for this AO system).

1.2.3 AO-corrected PSF model

Due to the partial correction of the turbulent phase by AO as described in Subsection 1.2.2, the random fluctuations of the corrected wavefront are still impacting the image, in addition to the different measurement noises which are propagated in the AO loop. Thus, the AO-corrected phase is still a statistical process. However, the shape of its PSD, given by Kolmogorov statistics in Equation 1.1, is not the same. Indeed, the low frequencies (below the AO cutoff frequency) were corrected using AO resulting in a lower residual phase PSD at low frequencies. From the description of the PSD of the residual phase, one can then easily have access to the PSF using Equations 1.3 and 1.4.

Conan [Conan, 1994] indeed shows that the description of the PSF as the convolution between a static PSF and an atmospheric PSF is still valid in the case of an AO-corrected PSF.

The static PSF sums the contributions of the telescope and the camera, and depends on the shape of pupil, the central obstruction, the spider and optical aberrations from the telescope and the camera. To describe this static PSF, we take into account the shape and central obstruction, but neither the spider (which rotates and has not much impact) nor the optical aberrations which are unknown: in the data we process in Chapter 6, we consider that the instruments have a high instrumental Strehl ratio *i.e.* the optical aberrations are small.

Describing the residual phase PSD in order to describe the atmospheric PSF is the main idea of the PSFAO19 model, which is developed in [Fétick, 2019a]. The PSFAO19 model has been designed specifically to describe an AO-corrected long-exposure PSF with few physical parameters.

More precisely, and as said previously, the residual phase PSD W_ϕ can be separated in two different spatial frequency zones, depending on the AO cutoff frequency: below this frequency, the AO system partially corrects the effects of turbulence, and above it the turbulent phase PSD follows the theoretical model of turbulence.

$$W_\phi(f) = \begin{cases} A\mathcal{N}_{\alpha,\beta}(1+f^2/\alpha^2)^{-\beta} + C & \text{if } f \leq f_{AO} \\ 0.023 r_0^{-5/3} f^{-11/3} & \text{else.} \end{cases}$$

For the corrected spatial frequencies, a Moffat model is used in order to describe the core of the PSD [Andersen, 2006; Sánchez, 2006]. The main parameter describing the core of the PSD is the amplitude of the Moffat function A, which is very close to the residual phase variance: $v_\phi \approx A + C\mathcal{A}_{AO}$, with $\mathcal{A}_{AO}$ the AO-corrected area in the spatial frequency domain (in the case of a circular AO-corrected area, $\mathcal{A}_{AO} = \pi f_{AO}^2$). C is a constant giving the AO-corrected phase PSD background, which is useful to model the AO-corrected PSD near AO cutoff frequency (where the Moffat function is close to zero). The parameters α (giving the width of the Moffat function) and β (Moffat's power law) do not impact the computation of the residual phase variance, thus they have a less important impact on the PSF. Throughout this work, α, β and C will be considered as constant, known parameters as they do not evolve significantly with the observing conditions [Fétick, 2019a], and their value will be fixed to the values in Table 1.1. Thus, the parameters we are seeking for are r_0 and $v_\phi \approx A + C\mathcal{A}_{AO}$.

Parameter	Value
Moffat width α (m^{-1})	0.05
Moffat power law β	1.5
AO area constant C(rad^{2}m^2)	10^{-10}

Table 1.1 – Moffat fixed parameters

Finally, $\mathcal{N}_{\alpha,\beta}$ is a normalization factor which is used to normalize the integral of the Moffat function over the corrected area:

$$\mathcal{N}_{\alpha,\beta} \triangleq \frac{\beta-1}{\pi\alpha^2}\Big[1 - \Big(1 + \frac{f_{AO}^2}{\alpha^2}\Big)^{1-\beta}\Big]^{-1}$$

which requires that $\beta > 1$.

For the high spatial frequencies, the theoretical Kolmogorov model of turbulence given in Equation 1.1 is used, the main parameter being the Fried parameter r_0 describing the turbulence's strength, taken at the imaging wavelength.

This model, which we will use throughout this work, has been validated by means of simulations, as well as on experimental data, on several AO systems and on different telescopes [Fétick, 2019a; Petit, 2020].

1.3 Our problem: application cases and examples

In this work, we will consider two different application cases: the first case we will describe is an astronomical observation (more precisely, the observation of an asteroid) using the VLT/-SPHERE instrument, and using its AO system SAXO. The second case is the observation of a satellite, on a different telescope and AO system.

1.3.1 Imaging model

In both application cases, we consider an incoherent extended object o_c, meaning that each point of the object emits a light wave spatially incoherent with the others. Note that this object is a continuous distribution but that we are estimating a discrete object. We can write i_c the (continuous) image of this (continuous) object, meaning the result of the object imaging after the turbulence and the optical instruments effects but before the discretization by the sensor and without noise, using the superposition principle and assuming that the object is spatially incoherent:

$$i_c(\vec{r}) = \iint o_c(\vec{\rho}) h_{opt}(\vec{r}, \vec{\rho}) \mathrm{d}\vec{\rho} \tag{1.5}$$

where $h_{opt}(\vec{r}, \vec{\rho})$ is the PSF, namely the impulse response of the optical system, which depends on the considered position in the image $\vec{r}$ and on the position of the object in the sky $\vec{\rho}$.

If we consider that we are within the isoplanetic domain, meaning that the impulse response does not depend on the observed spot, then Equation 1.5 becomes:

$$i_c(\vec{r}) = \iint o_c(\vec{\rho}) h_{opt}(\vec{r} - \vec{\rho}) \mathrm{d}\vec{\rho} \tag{1.6}$$

We recognize here that Equation 1.6 can be written as a convolution product between the object and the point spread function:

$$i_c(\vec{r}) = \left\{ o_c * h_{opt} \right\} (\vec{r}) \tag{1.7}$$

The discrete (noise free) image i_{nl}, of size N pixels, is the integration of the continuous image on each pixel of the detector. One can show that this integration on each pixel is equal to the sampling of the convolution between 1.7 and the shape of a pixel h_d:

$$\boldsymbol{i_{nl}} = Ш_{\Delta} \left\{ o_c * h_{opt} * h_d \right\} \tag{1.8}$$

with Ш the 2D Dirac comb operator (sampling operator), with a discretization step Δ.

We define here $h_c = h_{opt} * h_d$, the point spread function taking into account both the optical aberrations and the detector contribution. One element of the noiseless image $i_{nl}(k, l)$ then writes:

$$i_{nl}(k, l) = \iint h_c(k\Delta - x, l\Delta - y) o_c(x, y) \mathrm{d}x \mathrm{d}y \tag{1.9}$$

If we discretize the object using the Whittaker–Shannon interpolation formula, meaning $o_c(x, y) = \sum_n \sum_m o_{n,m} b_{n,m}(x, y)$ with $o_{n,m}$ the coefficients of the discrete object and $b_{n,m}(x, y)$

the interpolation basis, the noiseless image becomes:

$$\begin{aligned}
i_{nl}(k,l) &= \iint h_c(k\Delta - x, l\Delta - y) \sum_n \sum_m o_{n,m} b_{n,m}(x,y) \mathrm{d}x\mathrm{d}y \\
&= \sum_n \sum_m \iint h_c(k\Delta - x, l\Delta - y) b_{n,m}(x,y) \mathrm{d}x\mathrm{d}y o_{n,m} \\
&= \sum_n \sum_m \left\{h_c * b_{n,m}\right\}(k\Delta, l\Delta) o_{n,m} \\
&= \sum_n \sum_m \left\{h_c * b_{0,0}\right\}(k\Delta - n\Delta, l\Delta - m\Delta) o_{n,m}
\end{aligned}$$

If we define a discrete PSF $\boldsymbol{h}$ so that $h(k,l) \triangleq \left\{h_c * b_{0,0}\right\}(k\Delta, l\Delta)$ and a discrete object $\boldsymbol{o}$ so that $o(n,m) = o_{n,m}$, then:

$$\begin{aligned}
i_{nl}(k,l) &= \sum_n \sum_m h(k-n, l-m) o(n,m) \\
\boldsymbol{i_{nl}} &= \boldsymbol{h} * \boldsymbol{o} \qquad (1.10)
\end{aligned}$$

where $*$ stands here for the discrete 2D convolution operator. If we use the sinc basis $b_{n,m}(x,y) = \frac{1}{\Delta}\mathrm{sinc}(\pi\frac{x-n\Delta}{\Delta}, \pi\frac{y-m\Delta}{\Delta})$ as the interpolation basis, then $b_{0,0}(x,y) = \frac{1}{\Delta}\mathrm{sinc}(\pi\frac{x}{\Delta}, \pi\frac{y}{\Delta})$ and $\tilde{b}_0(f_x, f_y) = \Pi_{1/\Delta}(f_x, f_y)$ with $\Pi_{1/\Delta}$ a $1/\Delta$-long centered door function. Therefore, $h_c * b_{0,0}$ is simply a low-pass filtered version of h_c, and the cut-off frequency of the low-pass filtering is the Nyquist sampling frequency $\frac{1}{2\Delta}$. $\boldsymbol{h}$ is therefore the discretization of the low-pass version of h_c.

The final noisy image suffers from different measurement noises $\boldsymbol{n}$:

$$\boldsymbol{i} = (\boldsymbol{h} * \boldsymbol{o}) \diamond \boldsymbol{n} \qquad (1.11)$$

where $\diamond$ represents a pixel-by-pixel operation.

In our case, we are considering:

- Photon noise, which is associated to the randomness of the arrival of photons on the detector. We consider that it follows a Poisson distribution.
- Read-out noise, which is associated to the electronics during the conversion of the charge in each pixel into an actual signal. We consider that this noise is an additive white Gaussian noise, which is generally the case in the literature.

We approximate noise as additive and independent from the object. In the case of strong flux, the photon noise can be approximated as (non-stationary) Gaussian, which is a hypothesis that has been validated in practice in [Mugnier, 2004]. Then Equation 1.11 becomes:

$$\boldsymbol{i} = \boldsymbol{h} * \boldsymbol{o} + \boldsymbol{n} \qquad (1.12)$$

which is a usual imaging model [Demoment, 1989].

Equation 1.12 can also be written under the following matrix form:

$$\boldsymbol{i} = \boldsymbol{H}\boldsymbol{o} + \boldsymbol{n} \tag{1.13}$$

with $\boldsymbol{H}$ the convolution matrix corresponding to the convolution of the object by $\boldsymbol{h}$: each column of $\boldsymbol{H}$ contains a shifted PSF.

1.3.2 Observing an asteroid on the VLT

The Very Large Telescope (VLT) consists of four telescopes, whose aperture diameter is 8.2 m, and several auxiliary telescopes. It is operated by the European Southern Observatory (ESO), and is located on Cerro Paranal in Chile. The instrument SPHERE, which stands for *Spectro-Polarimetric High contrast Exoplanet REsearch*, is attached to the VLT Unit Telescope 3 (Melipal). It includes an extreme AO system SAXO [Fusco, 2016], which parameters are given in Table 1.2, and its original purpose is the study of exoplanets [Beuzit, 2019]. SPHERE has several sub-systems, one of them is the Zurich IMaging POLarimeter (ZIMPOL) from which some of the images we study here are taken.

Figure 1.4 – Photo of SPHERE (photo credit: Jean-François Sauvage, ONERA).

Parameter	Value
Primary diameter (m)	8.2
Secondary diameter (m)	1.12
Sampling (-)	4.76
AO cutoff frequency (m^{-1})	2.5

Table 1.2 – SPHERE's AO system parameters

The sampling here gives the number of pixels per λ/D. If the sampling is at Shannon-Nyquist frequency, then Sampling $= 2$.

Considering the long exposure time for astronomical images (around one minute), we can assume in this application that we have indeed a long-exposure PSF as discussed in Subsection 1.1.2.

1.3.3 Observing a satellite on MéO

For the ground-based satellite observation, the images we are studying were taken at the Côte d'Azur Observatory (OCA), where there are several telescopes among which the CNRS telescope MéO (standing for *Optical MEtrology*). This telescope, whose aperture diameter is 1.54 m, is fast enough to follow low earth orbit (LEO) satellites, useful for both civilian and military applications. ONERA has designed a prototype AO system for satellite observation, ODISSEE [Petit, 2020], which is installed on this telescope. Its parameters are given in Table 1.3.

Figure 1.5 – Photo of the MéO telescope provided by CNRS.

Parameter	Value
Primary diameter (m)	1.54
Sampling (-)	2.28
AO cutoff frequency (m^{-1})	2.92

Table 1.3 – ODISSEE's parameters

In the case of MéO, the exposure time is shorter (around 100 ms – 1 s, but it can be even shorter) than SPHERE (around a minute). Indeed, the exposure time is a compromise between increasing the image SNR and time scale of satellite evolution (in particular when it is rotating). ODISSEE being a prototype, the AO correction quality is very limited due to fitting, aliasing, as well as temporal error given the fast evolution of the satellite during the observation. However, we will still consider having a long-exposure PSF in this case.

Chapter 2

Image restoration methods

Introduction

In this chapter, we give an overview of the image restoration problem we tackle in this manuscript. In the first section, we recall some elements about the Bayesian framework we are working in to solve our myopic deconvolution problem. In Section 2, we describe the historical joint MAP estimator and its implementation applied to astronomical observation. Section 3 studies the properties of the joint MAP estimator in more depth, in the case of a Gaussian regularization on the object without positivity constraint, showing the degeneracy of the joint MAP criterion. In Section 4, we present the marginal MAP estimator as well as its implementations: AMIRAL, and its parametric version P-AMIRAL. Finally, in Section 5 we give an insight on other possible estimators, with their characteristics.

Contents

2.1 The deconvolution problem

2.1.1 About inverse problems

The observed image is related to the unknown true object by the imaging model given in the previous chapter in Equation 1.12. This imaging equation is said to be forward, meaning that it describes what happened in the causal order (causes leading to consequences). Thus, retrieving the true object given the observed image requires to solve an inverse problem, meaning to invert the forward model to get back to causes from the consequences. Inverse problems can be hard to solve for several reasons: firstly, errors on the imaging model prevent one to have access to the true object and may lead to an error in the estimation. Additionally, perfectly knowing the imaging model does not ensure one to find back the true object. Hadamard stated three conditions to define the well-posedness of an inverse problem [Hadamard, 1902]: the two first conditions are the existence and the unicity of the solution meaning that there is a unique object so that $\boldsymbol{i} = \boldsymbol{H}\boldsymbol{o}$, and the last one is the continuity of the solution meaning that the solution changes continuously with respect to the observations (the inverse of $\boldsymbol{H}$ on $\mathrm{Im}(\boldsymbol{H})$ is continuous).

Inverse problems are often ill-posed, as they often do not respect all three Hadamard conditions. The existence of the solution is not ensured, especially if the imaging model is incorrect. The unicity as well as the continuity of the solution is often lost, due to the structure of $\boldsymbol{H}$ (for instance, if $\boldsymbol{H}$ is non-invertible). Most importantly, the stability/continuity acknowledges the existence of noise, requiring that for a small variation of the image (due to noise), the object only varies a little. This non-continuity problem persists even for inversions in the least square sense [Mugnier, 2008]. In our case, because we work on a discretized problem, the problem is well-posed however it is often unstable, meaning that small errors in the observations may lead to large errors on the solution. In this case, the problem is said to be ill-conditioned. Here, the conditioning term is related to the matrix $\boldsymbol{H}$ which creates the instability.

When the direct equation is the convolution between the unknown and a convolution kernel, eventually to which noise is added, the inverse problem is said to be a deconvolution problem. As its name says, the point spread function (the convolution kernel) describes how altered and blurred is the unknown true object, thus the deconvolution corresponds to a “deblurring” of the observation to retrieve the object.

2.1.2 The Bayesian framework

One possible approach to solve this problem is the Bayesian approach. It consists in using Bayes’ theorem in order to compute the probability of the causes (the unknown object) given the consequences (the observation) [Idier, 2008]:

$$p(\boldsymbol{o}|\boldsymbol{i}) = \frac{p(\boldsymbol{i}|\boldsymbol{o})p(\boldsymbol{o})}{p(\boldsymbol{i})}$$

Bayesian inference writes the probability of the causes given the consequences $p(\boldsymbol{o}|\boldsymbol{i})$, called the posterior probability, as the product of several elements: a prior probability $p(\boldsymbol{o})$ which describes prior knowledge on the unknown, which does not depend on the current observation, and a likelihood term $p(\boldsymbol{i}|\boldsymbol{o})$ which indicates how accurately the observation fits the imaging model and object. The model evidence $p(\boldsymbol{i})$ is not a function of the object $\boldsymbol{o}$.

From the posterior probability, one can compute a Bayesian estimator: the estimated object $\hat{\boldsymbol{o}}$ is defined not as the direct result of the inversion of the forward model, but as an optimum solution with respect to a given criterion / loss function. For example, one possible estimator is the Maximum *A Posteriori*, meaning the object maximizing the posterior probability:

$$\hat{\boldsymbol{o}} = \arg\max_{\boldsymbol{o}} p(\boldsymbol{o}|\boldsymbol{i})$$

Additional details concerning Bayesian estimators will be given in Section 2.5 of this chapter.

2.1.3 Blind (or myopic) deconvolution

In the astronomic or satellite observation cases, the PSF is highly variable due to variable atmospheric turbulence and AO correction. Therefore, in addition to usual difficulties associated to the classic inverse problem of deconvolution, here the PSF is also unknown and must be estimated, and errors on the PSF lead to errors on the unknown object. Deconvolution with an unknown, or poorly known, PSF is called blind, or myopic, deconvolution. In the following sections, we will discuss specific methods in order to address this blind deconvolution problem.

2.2 Joint MAP estimator

2.2.1 Computing the joint MAP criterion

In the case where both object and PSF are unknown, one possible way to estimate them is to do it jointly [Ayers, 1988; You, 1996]. For instance, if we use the Maximum *A Posteriori* (MAP) estimator, it means that we are searching for the object and the PSF that maximize the joint posterior distribution $p(\boldsymbol{o}, \boldsymbol{h}|\boldsymbol{i})$:

$$(\hat{\boldsymbol{o}}, \hat{\boldsymbol{h}}) = \arg\max_{\boldsymbol{o},\boldsymbol{h}} p(\boldsymbol{o}, \boldsymbol{h}|\boldsymbol{i})$$

In our model, we suppose that the object and the PSF are *a priori* independent, thus using Bayes' theorem the joint distribution re-writes:

$$p(\boldsymbol{o}, \boldsymbol{h}|\boldsymbol{i}) = \frac{p(\boldsymbol{i}|\boldsymbol{o}, \boldsymbol{h})p(\boldsymbol{o})p(\boldsymbol{h})}{p(\boldsymbol{i})} \tag{2.1}$$

with $p(\boldsymbol{i}|\boldsymbol{o}, \boldsymbol{h})$ the likelihood, and $p(\boldsymbol{o})$ and $p(\boldsymbol{h})$ prior distributions respectively for the object and for the PSF. $p(\boldsymbol{i})$ is a constant with respect to $\boldsymbol{o}$ and $\boldsymbol{h}$ therefore can be discarded while searching for the MAP solution.

Maximizing the posterior distribution is equivalent to minimizing its anti-logarithm. We can then write J_{jMAP}, the criterion we are minimizing in this problem, as:

$$\begin{aligned} J_{jMAP}(\boldsymbol{o}, \boldsymbol{h}) &= -\ln p(\boldsymbol{o}, \boldsymbol{h}|\boldsymbol{i}) \\ &= -\ln p(\boldsymbol{i}|\boldsymbol{o}, \boldsymbol{h}) - \ln p(\boldsymbol{o}) - \ln p(\boldsymbol{h}) \end{aligned} \tag{2.2}$$

The first term is a data fidelity term, the second and the third ones are regularization terms on respectively the object and the PSF.

2.2.2 Implementing the joint MAP estimator: MISTRAL

One thing that the joint MAP estimator enables is to impose some constraints on the object, such as a positivity constraint. The joint MAP estimator, including this positivity constraint, was implemented in previous algorithm MISTRAL (standing for *Myopic Iterative STep-preserving Restoration ALgorithm*) [Mugnier, 2004]. The minimization of the criterion is done using VMLM-B algorithm [Thiébaut, 2002], including the positivity constraint by means of projected gradient. For the regularization on the PSF, a Gaussian prior distribution for the PSF is defined, with a mean PSF $m_{\boldsymbol{h}}$ and a PSF power spectral density (PSD) $S_{\boldsymbol{h}}$.

In practice, the mean PSF $m_{\boldsymbol{h}}$ and the PSD $S_{\boldsymbol{h}}$ are estimated by an empirical average of PSFs recorded on a star of appropriate magnitude, shortly before and/or after observing the object [Fusco, 2016]. These PSFs should be representative of the possible PSFs and require to be measured on several stars, in order to not introduce an additional error through the regularization. Thus, obtaining representative $m_{\boldsymbol{h}}$ and $S_{\boldsymbol{h}}$ is hardly possible, it is time-consuming and the difference of behaviour of the AO system between the image of a star and the actual PSF is not taken into account: indeed the images on stars taken as PSF realizations are not quite representative as the AO system has a different behavior between a point source and an extended object.

2.3 Specific case: Gaussian object regularization, without positivity

2.3.1 Object prior and likelihood

In this section, we study the joint criterion for a specific regularization on the object. We first need to fully write the object prior as well as the likelihood, in order to re-write the criterion.

In the case of a Gaussian regularization on the object, the object prior is described by its mean $\boldsymbol{m_o}$ and its covariance matrix $\boldsymbol{R_o}$. It writes:

$$p(\boldsymbol{o}) = \det(2\pi\boldsymbol{R_o})^{-1/2}\exp(-\frac{1}{2}(\boldsymbol{o}-\boldsymbol{m_o})^t\boldsymbol{R_o}^{-1}(\boldsymbol{o}-\boldsymbol{m_o})) \tag{2.3}$$

Given the flux levels we are considering in astronomical and satellite imaging, we approximate noise as Gaussian. (As said in Chapter 1, the photon noise Gaussian approximation has been validated in practice in [Mugnier, 2004].) Additionally, noise is taken independent from the object, and is supposed zero-mean, additive and white, which are again reasonable hypothesis given the flux levels in our images. Therefore, noise is defined simply by its covariance matrix $\boldsymbol{R_n}$. Given the imaging equation 1.13, the likelihood $p(\boldsymbol{i}|\boldsymbol{o},\boldsymbol{h})$ writes:

$$p(\boldsymbol{i}|\boldsymbol{o},\boldsymbol{h}) = \det(2\pi\boldsymbol{R_n})^{-1/2}\exp(-\frac{1}{2}(\boldsymbol{i}-\boldsymbol{Ho})^t\boldsymbol{R_n}^{-1}(\boldsymbol{i}-\boldsymbol{Ho})) \tag{2.4}$$

2.3.2 Re-writing the joint MAP criterion

Using the object prior and the likelihood expressions given in Equations 2.3 and 2.4, we can re-write the joint criterion in Equation 2.2 in the specific case we are considering [Blanco, 2011].

$$\begin{aligned} J_{jMAP}(\boldsymbol{o},\boldsymbol{h}) &= -\ln p(\boldsymbol{i}|\boldsymbol{o},\boldsymbol{h}) - \ln p(\boldsymbol{o}) - \ln p(\boldsymbol{h}) \\ &= \frac{1}{2}\ln\det(\boldsymbol{R_n}) + \frac{1}{2}(\boldsymbol{i}-\boldsymbol{Ho})^t\boldsymbol{R_n}^{-1}(\boldsymbol{i}-\boldsymbol{Ho}) \\ &+ \frac{1}{2}\ln\det(\boldsymbol{R_o}) + \frac{1}{2}(\boldsymbol{o}-\boldsymbol{m_o})^t\boldsymbol{R_o}^{-1}(\boldsymbol{o}-\boldsymbol{m_o}) - \ln p(\boldsymbol{h}) \\ &+ \text{constant} \end{aligned}$$

From it, we can compute the minimizer of the criterion. To find the object which minimizes $J_{jMAP}(\boldsymbol{o},\boldsymbol{h})$ for a given $\boldsymbol{h}$, we derive it with respect to the object and nullify it.

$$\frac{\partial J_{jMAP}}{\partial \boldsymbol{o}}(\boldsymbol{o},\boldsymbol{h}) = -\boldsymbol{H}^t\boldsymbol{R_n}^{-1}\boldsymbol{i} + \boldsymbol{H}^t\boldsymbol{R_n}^{-1}\boldsymbol{Ho} + \boldsymbol{R_o}^{-1}\boldsymbol{o} - \boldsymbol{R_o}^{-1}\boldsymbol{m_o}$$

$$\frac{\partial J_{jMAP}}{\partial \boldsymbol{o}}(\boldsymbol{o},\boldsymbol{h}) = 0 \Leftrightarrow \hat{\boldsymbol{o}}(\boldsymbol{h}) = (\boldsymbol{H}^t\boldsymbol{R_n}^{-1}\boldsymbol{H} + \boldsymbol{R_o}^{-1})^{-1}(\boldsymbol{H}^t\boldsymbol{R_n}^{-1}\boldsymbol{i} + \boldsymbol{R_o}^{-1}\boldsymbol{m_o})$$

The matrix $\boldsymbol{H}$ being the convolution matrix by the PSF, it has a Tœplitz structure. Additionally, we consider that the object covariance matrix $\boldsymbol{R_o}$ and noise covariance matrix $\boldsymbol{R_n}$ are also Tœplitz, which means that $(\boldsymbol{o}-\boldsymbol{m_o})$ and noise are stationary, additionally to previous approximations.

We can approximate Tœplitz matrix as circulant matrices, which can be diagonalized in the Fourier basis. If $\boldsymbol{\Lambda_x}$ denotes a diagonal matrix with its diagonal equal to $\boldsymbol{x}$ and $\tilde{\boldsymbol{x}}$ denotes the discrete Fourier transform (DFT) of $\boldsymbol{x}$. Thus, we can write the joint criterion in the Fourier domain, using:

$$\boldsymbol{R_n} = \boldsymbol{F}\boldsymbol{\Lambda}_{\boldsymbol{S_n}}\boldsymbol{F}^{-1}$$

$$\boldsymbol{R_o} = \boldsymbol{F}\boldsymbol{\Lambda}_{\boldsymbol{S_o}}\boldsymbol{F}^{-1}$$

$$\boldsymbol{H} = \boldsymbol{F}\boldsymbol{\Lambda}_{\tilde{\boldsymbol{h}}}\boldsymbol{F}^{-1}$$

with $\boldsymbol{F}$ the Fourier transform matrix, $\boldsymbol{S_n}$ the noise PSD, $\boldsymbol{S_o}$ the object PSD and $\tilde{\boldsymbol{h}}$ the (discretized) OTF, which is the DFT of the PSF.

In the rest of this work, the DFT of $\boldsymbol{x}$ is denoted $\tilde{\boldsymbol{x}}$. For the object and the image, the DFT is normalized so as to have $\boldsymbol{F}$ as a unitary matrix $\boldsymbol{F}^H = \boldsymbol{F}^{-1}$. For the PSF, the DFT is normalized so that $\tilde{h}(0)$ equals the sum of the PSF on the numerical array. Moreover, this value is set to 1 by convention, to express flux conservation.

We did not specify any particuliar regularization for the PSF, we simply write that the term $-\ln p(\boldsymbol{h})$ corresponds to an additional term in the Fourier domain which we will write as $J_h(\tilde{\boldsymbol{h}})$.

We can now re-write the criterion $J_{jMAP}(\boldsymbol{o}, \boldsymbol{h})$ in the Fourier domain:

$$\begin{aligned} J_{jMAP}(\boldsymbol{o}, \boldsymbol{h}) = \frac{N}{2} \ln S_n + \frac{1}{2} \sum_{\boldsymbol{f}} \frac{|\tilde{\boldsymbol{i}}(\boldsymbol{f}) - \tilde{\boldsymbol{h}}(\boldsymbol{f})\tilde{\boldsymbol{o}}(\boldsymbol{f})|^2}{S_n} + \frac{1}{2} \sum_{\boldsymbol{f}} \ln S_o(\boldsymbol{f}) \\ + \frac{1}{2} \sum_{\boldsymbol{f}} \frac{|\tilde{\boldsymbol{o}}(\boldsymbol{f}) - \tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})|^2}{S_o(\boldsymbol{f})} + J_h(\tilde{\boldsymbol{h}}) \end{aligned} \tag{2.5}$$

with N the image size in pixels and $\sum_{\boldsymbol{f}}$ the sum on all pixels in the spatial frequency domain. Moreover, the MAP solution for the object $\hat{\boldsymbol{o}}(\boldsymbol{h})$ writes:

$$\hat{\tilde{\boldsymbol{o}}}(\boldsymbol{h}) = \frac{\tilde{\boldsymbol{h}}^*(\boldsymbol{f})\tilde{\boldsymbol{i}}(\boldsymbol{f}) + \dfrac{S_n}{S_o(\boldsymbol{f})}\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})}{\left|\tilde{\boldsymbol{h}}(\boldsymbol{f})\right|^2 + \dfrac{S_n}{S_o(\boldsymbol{f})}} \tag{2.6}$$

which corresponds to the Wiener filtering when $\boldsymbol{m}_{\boldsymbol{o}} = \boldsymbol{0}$. The MAP solution for the object $\hat{\tilde{\boldsymbol{o}}}(\boldsymbol{h})$ given in Equation 2.6 can be re-injected in the criterion (Equation 2.5) to obtain an expression which does not depend on the object:

$$\begin{aligned} J_{jMAP}(\hat{\boldsymbol{o}}(\boldsymbol{h}), \boldsymbol{h}) &= \frac{N}{2} \ln S_n + \frac{1}{2} \sum_{\boldsymbol{f}} \ln S_o(\boldsymbol{f}) + \frac{1}{2} \sum_{\boldsymbol{f}} \frac{\dfrac{S_n^2}{|S_o(\boldsymbol{f})|^2}\left|\tilde{\boldsymbol{i}}(\boldsymbol{f}) - \tilde{\boldsymbol{h}}(\boldsymbol{f})\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})\right|^2}{S_n\left||\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + \dfrac{S_n}{S_o(\boldsymbol{f})}\right|^2} \\ &+ \frac{|\tilde{\boldsymbol{h}}^*(\boldsymbol{f})(\tilde{\boldsymbol{i}}(\boldsymbol{f}) - |\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f}))|^2}{S_o(\boldsymbol{f})\left||\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + \dfrac{S_n}{S_o(\boldsymbol{f})}\right|^2} + J_h(\tilde{\boldsymbol{h}}) \\ &= \frac{N}{2} \ln S_n + \frac{1}{2} \sum_{\boldsymbol{f}} \ln S_o(\boldsymbol{f}) \\ &+ \frac{1}{2} \sum_{\boldsymbol{f}} \frac{|\tilde{\boldsymbol{i}}(\boldsymbol{f}) - \tilde{\boldsymbol{h}}(\boldsymbol{f})\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})|^2(S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n)^*}{\left|S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n\right|^2} + J_h(\tilde{\boldsymbol{h}}) \\ &= \frac{N}{2} \ln S_n + \frac{1}{2} \sum_{\boldsymbol{f}} \ln S_o(\boldsymbol{f}) + \frac{1}{2} \sum_{\boldsymbol{f}} \frac{|\tilde{\boldsymbol{i}}(\boldsymbol{f}) - \tilde{\boldsymbol{h}}(\boldsymbol{f})\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})|^2}{S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n} + J_h(\tilde{\boldsymbol{h}}) \end{aligned} \tag{2.7}$$

2.3.3 The degeneracy of the joint MAP criterion

In [Blanco, 2011], the authors show in theory the degeneracy of the joint MAP criterion in Equation 2.7, always leading to the sharpest PSF and the smoothest object (meaning, the "less deconvolved" one). It was discussed in the case we describe previously: using a quadratic regularization on the object and without any object constraint such as positivity or support constraint, and previous work in [Fétick, 2020a] also shows in practice that the positivity constraint imposed on the object is not enough to overcome this degeneracy.

A closer look to Equation 2.7 explains why such result could be expected [Levin, 2009]. In the last term, given that the mean object $\boldsymbol{m_o}$ is taken constant, $\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})$ is only non-zero when $f = 0$, whereas $\tilde{h}(0) = 1$ by convention as said in Subsection 2.3.2. Therefore, the numerator does not depend on the PSF. Minimizing the criterion with respect to the PSF only requires maximizing the denominator of the last term:

$$\begin{aligned}\hat{\boldsymbol{h}} &= \arg\min_{\boldsymbol{h}} J_{jMAP}(\hat{\boldsymbol{o}}(\boldsymbol{h}), \boldsymbol{h}) \\ &= \arg\min_{\boldsymbol{h}} \sum_{\boldsymbol{f}} \frac{|\tilde{\boldsymbol{i}}(\boldsymbol{f}) - \tilde{\boldsymbol{h}}(\boldsymbol{f})\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})|^2}{S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n} \\ &= \arg\min_{\boldsymbol{h}} \sum_{\boldsymbol{f}>0} \frac{|\tilde{\boldsymbol{i}}(\boldsymbol{f})|^2}{S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n}\end{aligned}$$

The PSF maximizing the denominator is the one maximizing $|\tilde{\boldsymbol{h}}|^2$. This can be maximized frequency per frequency, in the absence of any structure for the PSF then for each frequency $\tilde{\boldsymbol{h}}$ would be equal to its maximum value meaning 1. For a given PSF structure, and more generally, maximizing $|\tilde{\boldsymbol{h}}|^2$ corresponds to the flattest OTF (in order to maximize its value on all frequencies), thus to the sharpest PSF.

The degeneracy of the joint MAP criterion calls for using another estimator to compute the sought object and PSF.

2.4 Marginal MAP estimator

2.4.1 Computing the marginal likelihood

Another way to estimate the object and the PSF is to first estimate the PSF by computing the so called marginalized posterior probability, meaning integrating the joint posterior density written in Equation 2.1 over the object [Blanco, 2011; Fétick, 2020b]:

$$\begin{aligned}p(\boldsymbol{h}|\boldsymbol{i}) &= \int p(\boldsymbol{o}, \boldsymbol{h}|\boldsymbol{i})\mathrm{d}\boldsymbol{o} \\ &= \frac{1}{p(\boldsymbol{i})}\int p(\boldsymbol{i}, \boldsymbol{o}, \boldsymbol{h})\mathrm{d}\boldsymbol{o}\end{aligned}$$

In practice, we write the marginal posterior distribution following the Bayes' rule as previously, from the marginal likelihood and the prior on the PSF, which is considered independent from the object:

$$p(\boldsymbol{h}|\boldsymbol{i}) = \frac{p(\boldsymbol{h})}{p(\boldsymbol{i})}p(\boldsymbol{i}|\boldsymbol{h})$$

This writing can be extended, not only to the PSF, but to any unknowns other than the object. For the rest of this manuscript, we define vector $\boldsymbol{\theta}$ used hereafter, gathering all unknown parameters other than the object, for instance here the PSF, noise covariance matrix and object mean and covariance matrix. The former equation including all unknown parameters then re-writes:

$$p(\boldsymbol{\theta}|\boldsymbol{i}) = \frac{p(\boldsymbol{\theta})}{p(\boldsymbol{i})} p(\boldsymbol{i}|\boldsymbol{\theta})$$

If we consider the same case as in Section 2.3, noise $\boldsymbol{n}$ is taken Gaussian, white, homogeneous and *a priori* independent from the linear image model $\boldsymbol{Ho}$ with a Gaussian object prior, the image being a linear combination of both is also Gaussian. Therefore, the marginal likelihood writes:

$$p(\boldsymbol{i}|\boldsymbol{\theta}) = (2\pi)^{-N/2} \det(\boldsymbol{R_i})^{-1/2} \exp(-\frac{1}{2}(\boldsymbol{i} - \boldsymbol{i_m})^t \boldsymbol{R_i}^{-1}(\boldsymbol{i} - \boldsymbol{i_m}))$$

with $\boldsymbol{i_m} = \boldsymbol{H m_o}$ et $\boldsymbol{R_i} = \mathbb{E}[\boldsymbol{i}\boldsymbol{i}^t] - \boldsymbol{i_m}\boldsymbol{i_m^t} = \boldsymbol{H R_o H}^t + \boldsymbol{R_n}$. The same way, we define the marginal criterion as the anti-logarithm of the marginal posterior distribution:

$$J_{mMAP}(\boldsymbol{\theta}) = \frac{1}{2} \ln\det(\boldsymbol{R_i}) + \frac{1}{2}(\boldsymbol{i} - \boldsymbol{i_m})^t \boldsymbol{R_i}^{-1}(\boldsymbol{i} - \boldsymbol{i_m}) + \text{constant}$$

2.4.2 Writing the marginal likelihood in the Fourier domain

Given the same circulant approximations as in Subsection 2.3.2, we can write the marginal criterion in the Fourier domain. To develop the calculations, we rewrite separately the two main elements of $J_{mMAP}(\boldsymbol{\theta})$, first $(\boldsymbol{i} - \boldsymbol{i_m})^t \boldsymbol{R_i}^{-1}(\boldsymbol{i} - \boldsymbol{i_m})$ and then $\det(\boldsymbol{R_i})$. To begin:

$$\begin{aligned}
(\boldsymbol{i} - \boldsymbol{i_m})^t \boldsymbol{R_i}^{-1}(\boldsymbol{i} - \boldsymbol{i_m}) &= (\boldsymbol{i} - \boldsymbol{i_m})^t (\boldsymbol{H R_o H}^t + \boldsymbol{R_n})^{-1}(\boldsymbol{i} - \boldsymbol{i_m}) \\
&= (\boldsymbol{i} - \boldsymbol{i_m})^t \boldsymbol{F}^{-1}(\Lambda_{\tilde{\boldsymbol{h}}}\Lambda_{\boldsymbol{S_o}}\Lambda_{\tilde{\boldsymbol{h}}^*} + \Lambda_{\boldsymbol{S_n}})^{-1}\boldsymbol{F}(\boldsymbol{i} - \boldsymbol{i_m}) \\
&= (\tilde{\boldsymbol{i}} - \tilde{\boldsymbol{i}}_{\boldsymbol{m}})^\dagger (\Lambda_{\tilde{\boldsymbol{h}}}\Lambda_{\boldsymbol{S_o}}\Lambda_{\tilde{\boldsymbol{h}}^*} + \Lambda_{\boldsymbol{S_n}})^{-1}(\tilde{\boldsymbol{i}} - \tilde{\boldsymbol{i}}_{\boldsymbol{m}}) \\
&= \sum_{\boldsymbol{f}} \frac{|\tilde{\boldsymbol{i}}(\boldsymbol{f}) - \tilde{\boldsymbol{h}}(\boldsymbol{f})\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})|^2}{S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n}
\end{aligned}$$

Secondly:

$$\begin{aligned}
\det(\boldsymbol{R_i}) &= \det(\boldsymbol{H R_o H}^t + \boldsymbol{R_n}) \\
&= \det(\boldsymbol{F}^{-1}(\Lambda_{\tilde{\boldsymbol{h}}}\Lambda_{\boldsymbol{S_o}}\Lambda_{\tilde{\boldsymbol{h}}^*} + \Lambda_{\boldsymbol{S_n}})\boldsymbol{F}) \\
&= \det(\Lambda_{\tilde{\boldsymbol{h}}}\Lambda_{\boldsymbol{S_o}}\Lambda_{\tilde{\boldsymbol{h}}^*} + \Lambda_{\boldsymbol{S_n}}) \\
&= \prod_{f} \left(S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n\right)
\end{aligned}$$

Finally, the criterion $J_{mMAP}(\boldsymbol{\theta})$ writes:

$$\begin{aligned}
J_{mMAP}(\boldsymbol{\theta}) &= \frac{1}{2} \ln\det(\boldsymbol{R_i}) + \frac{1}{2}(\boldsymbol{i} - \boldsymbol{i_m})^t \boldsymbol{R_i}^{-1}(\boldsymbol{i} - \boldsymbol{i_m}) \\
&= \frac{1}{2}\sum_{\boldsymbol{f}} \ln(S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n) + \frac{1}{2}\sum_{\boldsymbol{f}} \frac{|\tilde{\boldsymbol{i}}(\boldsymbol{f}) - \tilde{\boldsymbol{h}}(\boldsymbol{f})\tilde{\boldsymbol{m}}_{\boldsymbol{o}}(\boldsymbol{f})|^2}{S_o(\boldsymbol{f})|\tilde{\boldsymbol{h}}(\boldsymbol{f})|^2 + S_n}
\end{aligned}$$

We notice that the criterion can be written as:

$$J_{mMAP}(\boldsymbol{\theta}) = \frac{1}{2}\sum_{\boldsymbol{f}} \ln(S_i(\boldsymbol{f})) + \frac{1}{2}\sum_{\boldsymbol{f}} \frac{|\tilde{\boldsymbol{i}}(\boldsymbol{f}) - \tilde{m}_i(\boldsymbol{f})|^2}{S_i(\boldsymbol{f})} \tag{2.8}$$

with image PSD $\boldsymbol{S}_i$ and mean image $\boldsymbol{i}_m$:

$$\begin{aligned} S_i(\boldsymbol{f}) &= S_o(\boldsymbol{f})|\tilde{h}(\boldsymbol{f})|^2 + S_n \\ \tilde{m}_i(\boldsymbol{f}) &= \tilde{h}(\boldsymbol{f})\tilde{m}_o(\boldsymbol{f}) \end{aligned} \tag{2.9}$$

Minimizing the marginal criterion with respect to the image PSD can be seen as finding the image PSD $\boldsymbol{S_i}$ that best fits the empirical PSD $|\tilde{\boldsymbol{i}} - \tilde{\boldsymbol{m}}_{\boldsymbol{i}}|^2$. This can be shown in a simplified scalar case (for a single frequency), where the criterion writes:

$$J_{mMAP}(\boldsymbol{\theta}) = \frac{1}{2}\ln(S_i) + \frac{1}{2}\frac{|\tilde{i} - \tilde{m}_i|^2}{S_i}$$

To find the image PSD S_i minimizing $J_{mMAP}(\boldsymbol{\theta})$, we derive the criterion with respect to $\boldsymbol{S_i}$ and nullify the derivative:

$$\begin{aligned} &\frac{\partial J_{mMAP}}{\partial S_i}(\boldsymbol{\theta}) = \frac{1}{2}\frac{1}{S_i} - \frac{1}{2}\frac{|\tilde{i} - \tilde{m}_i|^2}{S_i^2} = 0 \\ &\implies \frac{1}{S_i} = \frac{|\tilde{i} - \tilde{m}_i|^2}{S_i^2} \\ &\implies S_i = |\tilde{i} - \tilde{m}_i|^2 \end{aligned}$$

The solution for S_i is the empirical PSD $|\tilde{i} - \tilde{m}_i|^2$.

2.4.3 Implementing of the marginal MAP estimator: AMIRAL

Previous work [Blanco, 2011] has developed the marginal MAP estimation code AMIRAL (standing for *Automatic Myopic Image Restoration ALgorithm*). This code was initially used for retinal imaging before it was applied to astronomy. In this case, the estimated PSF is a linear combination of a family of PSFs [Lafrenière, 2007] provided to the method: the algorithm searches for weights to assign to each PSF in the family, in addition the other sought parameters. Defining this family of PSF is quite crucial for the estimation, knowing that the family should ideally be large enough to include the true PSF, but also small enough to have a reasonable number of unknowns with respect to the available data and keep the interesting asymptotic properties of the marginal estimator, which are the same properties as the Maximum Likelihood estimator: it is consistent, meaning that it tends towards the true value when noise tends towards 0 or that the number of data tends towards infinity, and asymptotically efficient.

Another version of the algorithm, P-AMIRAL (standing for Parametric-AMIRAL), estimates the PSF using the PSFAO19 [Fétick, 2019a] model, thus the unknowns are directly the parameters of the PSF model.

For both AMIRAL and P-AMIRAL, the marginal criterion is minimized using the VMLM-B algorithm [Thiébaut, 2002], just like as in the algorithm MISTRAL. However here, the object being marginalized, we can not impose a constraint on it as could be done in the joint estimation.

Remark. *Throughout this manuscript, we will keep using the term "marginal MAP" for the estimator computed by AMIRAL. However, what AMIRAL actually computes is the Maximum Likelihood on the marginal distribution, given that there is no prior on the sought parameters.*

2.5 Short discussion: what about other estimators?

In the previous Section, one could see that the marginal MAP estimator, given its nice asymptotic properties, is a satisfactory option in order to avoid the degeneracy of the joint MAP estimator. However, the MAP estimator gives, as its name implies, only one point in the posterior distribution (its maximizer), and does not enable the user to get more global information on the posterior distribution.

Having the whole posterior distribution would be helpful for different reasons, one being that it would enable the user to compute uncertainties on the estimated parameters. These uncertainties, which are on the estimated parameters at first, can be propagated into uncertainties on the quantities of interest (for instance, the PSF), which are important for the final users (in our case, the astronomers) to determine how trustworthy is the final restored image. Another reason is that studying the posterior distribution would also enable the analysis of the posterior correlation between the estimated parameters, in order to determine possible difficulties in the method linked to strong correlation between parameters: convergence issues, slowness of the method,...

Additionally, having the whole posterior distribution would enable us to compute different estimators, for instance the MMSE (Minimum Mean Square Error) or the MMAE (Minimum Mean Absolute Error) estimator.

Chapter 3

Contributions to AMIRAL

Introduction

In this chapter, we give an insight about the contributions made to the previous method AMIRAL during the PhD. The first section describes the coupling between hyperparameters using the previous object PSD model, and proposes a modification of the model to decouple them. The second section is focused on the comparison of the two existing versions of the previous method: (linear) AMIRAL and Parametric-AMIRAL.

Contents

3.1 Object PSD model modification

As said in Chapter 2 (Subsection 2.3.1), we consider a Gaussian prior for the object, described by its mean and PSD. For the object PSD, previous works considered Matérn's model, which is a circularly-symmetric parametric model [Conan, 1998; Ramani, 2008], writing the object PSD as follows:

$$S_o(f) = \frac{K}{1 + (f/f_0)^p}$$

(with $f = |\boldsymbol{f}|$ the length of the radial frequency.) In this model, K sets the object PSD level in (very) low frequencies, p is the PSD decrease rate at high frequencies, and f_0 is the frequency which divides the model into two regimes.

3.1.1 Coupling between parameters in the previous model

If we have a closer look at Matérn's model, we can notice that in low frequencies, the model mainly depends on K whereas in high frequencies all parameters K, f_0 and p intervene:

$$S_o(f) = \begin{cases} K & \text{if } f \ll f_0 \\ \dfrac{K f_0^p}{f^p} & \text{if } f \gg f_0 \end{cases} \tag{3.1}$$

This coupling leads to some difficulties in the estimation of these hyperparameters in practice. For a mostly unsupervised mode, where hyperparameter p is fixed, the estimated parameter f_0 varies a lot (from 0.4 pix^{-1} to 1.3 pix^{-1}), while it does not have much impact physically: it impacts the regularization at very low frequencies, and does not have much impact in practice. However, this strong variability in the estimation of f_0 impacts the estimation of other parameters, especially K, as K and f_0 vary a lot together. We interpret it as the result of the mixing of the three parameters in high frequencies: indeed, the regularization on the object, coded by the object PSD, is particularly important to restore the high frequencies of the image, as they were particularly attenuated to the point that the signal level is close to the noise's one. However, these frequencies are precisely where all three parameters intervene as seen in Equation 3.1, which would explain the joint evolution of K and f_0, even when p is fixed. This correlation can be seen in Figure 3.1 where the criterion map with respect to K and f_0 is plotted, for given values of PSF and noise parameters: the criterion, as an extended valley, is flat in a direction mixing both parameters.

3.1.2 Proposition of a change of parameters

We propose to slightly change the writing of this model, in order to describe the same object PSDs but to separate more the impact of the parameters. The idea is to have less parameters intervening in the high frequencies, which are particularly sensitive to this regularization. The model we will use in the rest of the manuscript is the following:

$$S_o(f) = \frac{1}{\gamma_o} \frac{1}{k + f^p}$$

Here, the three parameters have similar roles to the ones in Matérn's model: γ_o sets the global PSD level (impacting all frequencies), as it gives the y-intercept of the object PSD slope in logarithmic scale. p is still the PSD decrease rate (slope) at high frequencies. Finally k, similarly to previous f_0, gives the breakpoint between the two regimes of the model. We could write $\gamma_o = \dfrac{1}{K f_0^p}$ and $k = f_0^p$, thus the change of parameters is reversible. With this model, the object PSD in low and high frequencies become:

$$S_o(f) = \begin{cases} \dfrac{1}{\gamma_o k} & \text{if } f \ll f_0 \\ \dfrac{1}{\gamma_o f^p} & \text{if } f \gg f_0 \end{cases}$$

Even though now we have two out of three parameters impacting the low frequencies, they are less sensitive to the regularization thus this expression of the model should still help decoupling the parameters, given that there are now only two parameters which come into play in high frequencies.

3.1.3 Results on criterion maps

In this subsection, we compute maps giving the marginal criterion value ($J_{MAPmarg}(\boldsymbol{\theta})$ given in Equation 2.8) with respect to two out of the three hyperparameters, using both Matérn's model and the modified version. The true PSF (thus, the true r_0 and v_ϕ), as well as the noise precision γ_n are supposed known to compute the criterion. The coolest colors (purple, then blue) correspond to the lowest criterion values, and iso-criterion lines are drawn above the map.

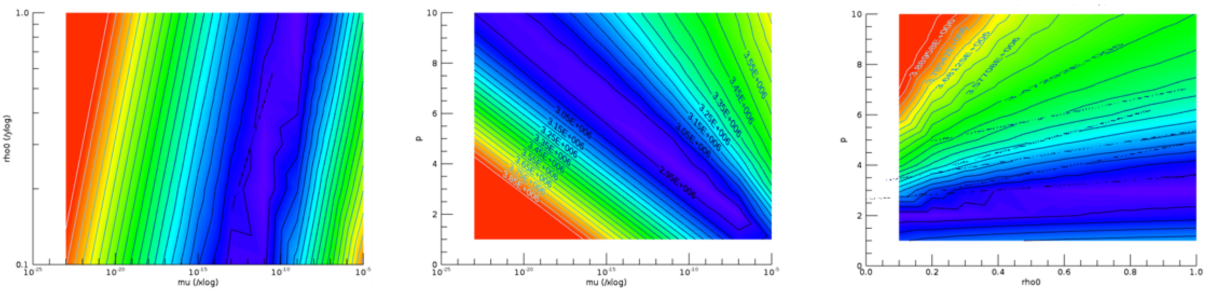

Figure 3.1 – Criterion maps (x, y) with respect to the hyperparameters x and y, using the Matérn's model. (True PSF and noise level are known.) Left: ($1/K$, ρ_0) map, in log-log scale. Center: ($1/K$, p) map, in log-lin scale. Right: (ρ_0, p) map in lin-lin scale.

As can be seen on Figure 3.2, compared to Figure 3.1, the parameter k in the second model (roughly corresponding to f_0 in Matérn's model) has been decoupled from the other parameters using the new model: the iso-criterion lines are either horizontal or vertical, showing that there is no in-determination related to k. Moreover, we can also see on Figure 3.2 that k does have only a minor, barely noticeable impact on the criterion value.

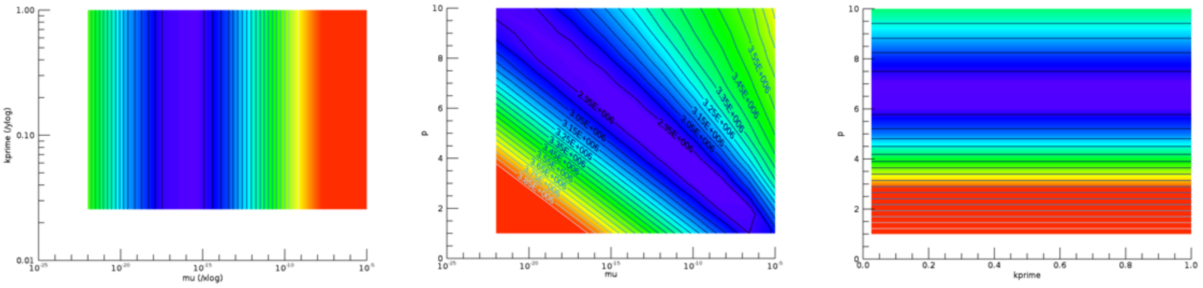

Figure 3.2 – Criterion maps (x, y) with respect to the hyperparameters x and y, using the modified model. (True PSF and noise level are known.) Left: (γ_o, k) map, in log-log scale. Center: (γ_o, p) map, in log-lin scale. Right: (k, p) map in lin-lin scale.

3.1.4 Some elements on convergence

After checking the impact of this new parametrization on criterion maps when the true PSF is known, we now look at the impact of this new version of the object PSD model when we estimate all parameters as in AMIRAL. Both versions are tested on the same generated data where we simulate the observation of asteroid Vesta (same system, same object, same PSF, same noise realization), with the same parameters: for instance the same initialization and the same required convergence threshold. Thus, the criterion values can be fairly compared. In Figure 3.3, we plot the evolution of the criterion with respect to the number of iterations, with both old and new versions. We notice that even though the new PSD parametrization requires more iterations to converge (around 65, against around 45 with the old parametrization), the criterion was better minimize using the new parametrization. Concerning the estimated PSF parameters, the estimations using the old and the new parametrization are very similar, with a

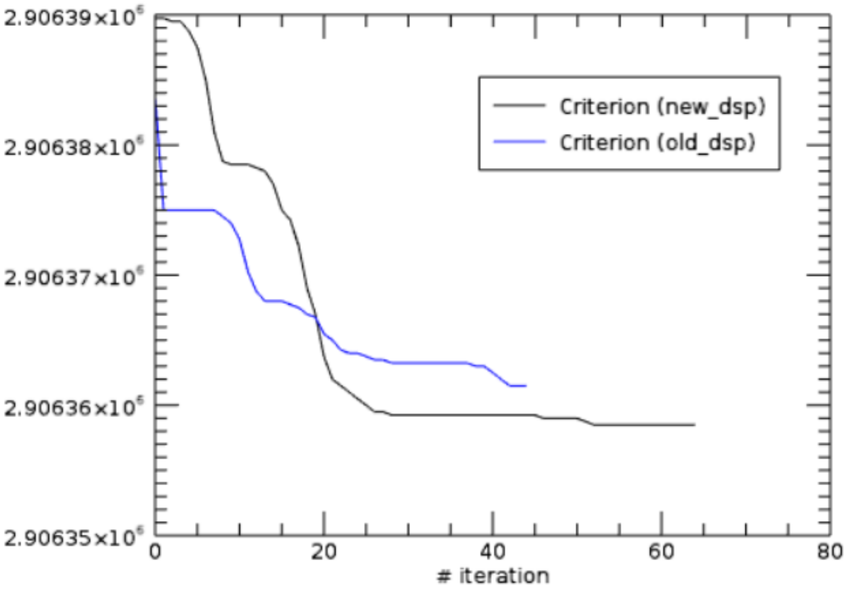

Figure 3.3 – Criterion evolution with respect to the number of iterations, for both versions of the object PSD.

Parameter	Parameters old PSD	Parameters new PSD	True
r_0 (m)	0.15	0.14	0.15
v_ϕ(rad^2)	1.16	1.18	1.30

Table 3.1 – Estimated values for PSF parameters r_0 and v_ϕ using old or new object PSD parametrization, and true values.

difference of 1 cm for r_0 and of 0.02 rad^2, which are what we could typically expect and what we will discuss later in Chapter 5.

3.2 Comparison between AMIRAL and P-AMIRAL

As said in Chapter 2, the marginal MAP estimator has two different implementations. The original version AMIRAL was developed in a retinal imaging context [Blanco, 2011] and estimates a PSF as a linear combination of a family of PSFs. The second version, Parametric-AMIRAL [Fétick, 2020b], estimates the PSF parameters directly using the PSFAO19 model. AMIRAL was successfully applied to satellite imaging [Petit, 2020] before P-AMIRAL was developed, and P-AMIRAL has not been applied to satellite images yet. The purpose of this section is therefore to give some comparison elements between these two versions, on a simulated astronomical observation case.

3.2.1 Simulation conditions

AMIRAL and P-AMIRAL are tested on simulated data, using as the true object the synthetic view of asteroid Vesta, built by the OASIS (standing for Optimized Astrophysical Simulator for Imaging Systems) [Jorda, 2010], on a dark background of size $N = 512 \times 512$ pixels. This view was built based on the observations by the NASA/Dawn spacecraft, which were used by OASIS to simulate the surface of Vesta and build a 3D model.

We simulate its observation from a good astronomical site: The AO system is a “SPHERE-like” AO system, and its parameters are taken identical to those used with the previous method given in [Fétick, 2020b], for comparison purposes. These parameters are summed up in Table 3.2.

The PSF is simulated using the PSFAO19 model [Fétick, 2019a]. Here, true PSF parameters are $r_0 = 0.15$ m and $v_\phi = 1.3$ rad^2 at the imaging wavelength $\lambda = 550$ nm, which corresponds to realistic turbulence and correction conditions. The total flux of the object is set to $F_o = 10^9$ ph (photons), typical from VLT/SPHERE/Zimpol asteroid observations (ESO Large Program ID 199.C-0074). We are approaching the photon noise and the read-out as a Gaussian noise, which variance is equal to the mean value of the object (to approach the Poisson distribution). We consider that this Gaussian approximation is true for the flux levels we are considering, and we are adding as a hypothesis for the simulations that noise is stationary, therefore noise is described by a single parameter, its precision $\gamma_n = N/F_o = 2.62 \times 10^{-4}$ ph^{-2}.

As said previously, AMIRAL estimates a PSF as a linear combination of a PSF family, therefore needing the choice of such a family. Here we choose a family with 9 PSFs, following the PSFAO19 model, with 3 different values for $r_0 = [0.1, 0.15, 0.2]$ m and 3 different values for

$v_\phi = [1.0, 1.3, 2.0]\,\mathrm{rad}^2$, therefore the true PSF with $r_0 = 0.15\,\mathrm{m}$ and $v_\phi = 1.3\,\mathrm{rad}^2$ is in the family. Thus, for these simulations we are testing a favorable case for AMIRAL: the PSF family is relatively small and the true PSF is within the family.

Using P-AMIRAL does not require the choice of a family but implies the choice of minimum and maximum boundaries for the PSF parameters. In this case, we choose $r_0 \in [0.1; 0.2]\,\mathrm{m}$ and $v_\phi \in [1.0; 2.0]\,\mathrm{rad}^2$, thus including the true PSF as well as the PSFs in the family used for AMIRAL.

Parameter	Value
Primary diameter (m)	8.2
Secondary diameter (m)	1.12
Sampling	3
AO cutoff frequency (m^{-1})	1.25
Fried parameter r_0 (m)	0.15
Residual phase variance v_ϕ (rad^2)	1.3
Object total flux (photons)	10^9

Table 3.2 – Simulated parameters

In [Fétick, 2020b], the authors have tested three different modes of estimation using P-AMIRAL: they could decide and change how much information the algorithm has on the object PSD, which is quite crucial for the estimation of the PSF, and finally of the object.

- The first of them was the supervised mode, where the object PSD is fixed following the Matérn's model and its parameters are tuned in order to fit correctly the empirical object periodogram meaning the "true" object PSD. This mode is useful for first simulations, but cannot be used in practice (for instance, on experimental data).
- The second mode is the unsupervised mode, where all object PSD parameters are estimated in addition to the PSF and noise parameters. As shown in [Fétick, 2020b], this mode does not give satisfactory results especially on the residual phase variance v_ϕ. We interpret this result as the consequence of a strong posterior correlation between one of the object PSD parameters p and v_ϕ, which will be discussed in Chapter 5.
- Lastly, they have tested the mostly unsupervised mode where the hyperparameter p, corresponding to the slope of the object PSD in medium-high frequencies, is fixed and the rest is estimated. Indeed, in practice p can be fixed to a reasonable value, according to the class of object that is observed (typically $p \approx 3$ for asteroids).

The authors have shown that in this mostly unsupervised mode, P-AMIRAL gives satisfactory results on the PSF parameters ($r_0 = 0.142\,\mathrm{m}$ and $v_\phi = 1.13\,\mathrm{rad}^2$). Therefore, in the following comparison, we also choose to use the mostly unsupervised mode and fix p to the same value $p = 3$.

3.2.2 Criterion evolution

AMIRAL and P-AMIRAL are run with four different initializations of the sought PSF and hyper-parameters, each of them correspond to a color in the following graphs. We are studying the evolution of different elements with respect to the convergence threshold. This convergence threshold is a relative threshold, computing the evolution of the minimized criterion J with respect to its mean value, between previous and current iterations:

$$\text{threshold} = \left| 2 \frac{J^{(i)} - J^{(i-1)}}{J^{(i)} + J^{(i-1)}} \right|$$

First, we plot the evolution of the criterion with respect to this threshold.

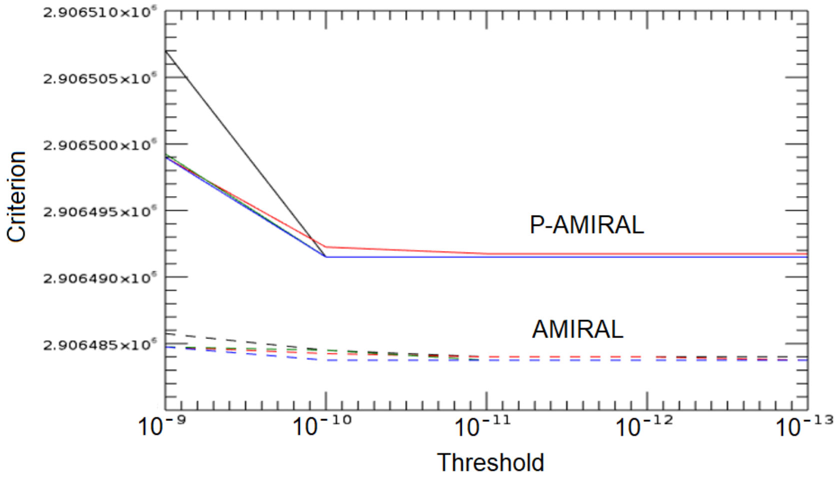

Figure 3.4 – Criterion evolution with respect to threshold, for AMIRAL (in dashed lines) and P-AMIRAL (in solid lines), for different initializations (corresponding to the different colors).

We can notice that P-AMIRAL needs a smaller threshold than AMIRAL to converge (meaning that P-AMIRAL needs a smaller threshold for the criterion not to evolve). For both methods, the criterion does not evolve when the threshold is below 10^{-11}. Therefore, for the following minimizations, we will set the threshold to 10^{-11} for both methods, in order to ensure their convergence.

3.2.3 RMSE on PSF evolution

For the same initializations as previously, we now plot the root mean square error (RMSE) on the OTF, with respect to the threshold, the RMSE being computed as follows: $\epsilon_{RMSE}(\hat{\tilde{h}}) = \sqrt{\langle ||\hat{\tilde{h}} - \tilde{h}||^2 \rangle_f}$, with $\tilde{h}$ the true OTF and $\langle . \rangle_f$ the averaging operator over all spatial frequencies.

As shown in Figure 3.5, P-AMIRAL's estimated PSF is better (the RMSE is smaller), and varies less according to the initialization. This reflects the (slightly better) robustness of P-AMIRAL. Moreover, even though Figure 3.4 suggested that AMIRAL was converging even with a convergence greater smaller than 10^{-11} (due to the fact that the criterion value stopped evolving for thresholds below 10^{-10}), actually it is not exactly the case when looking at the estimated PSF itself.

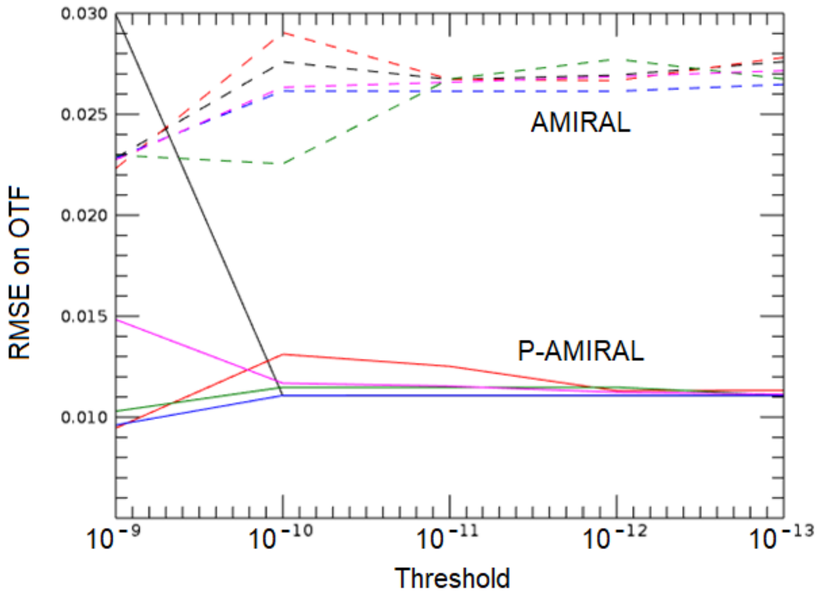

Figure 3.5 – Evolution of the RMSE on the OTF, with respect to threshold, for AMIRAL and P-AMIRAL, for different initializations.

3.2.4 Results on estimated PSF

To have a closer look to the actual estimated shapes of the PSF and the OTF, we plot the results obtained for one of the initializations.

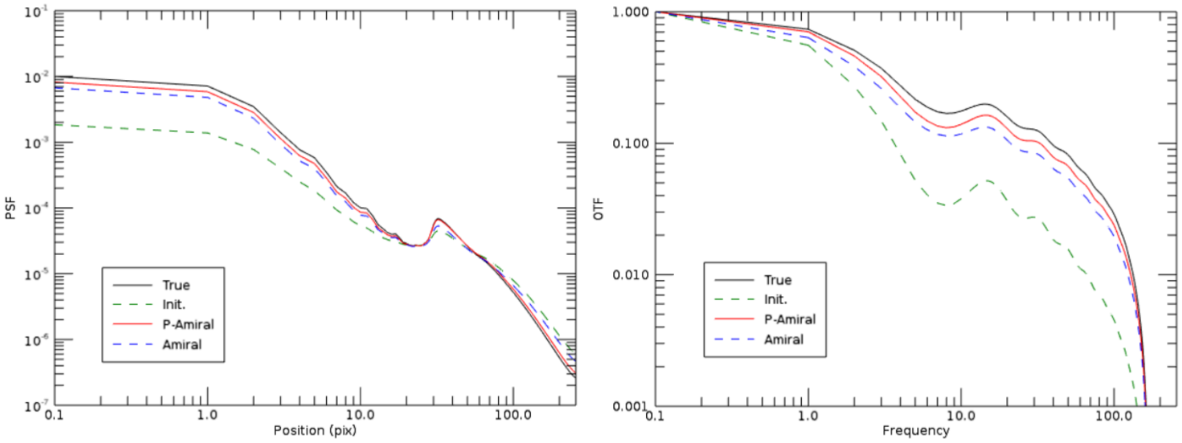

Figure 3.6 – Estimated PSF (left) and OTF (right) for AMIRAL and P-AMIRAL.

As displayed in Figure 3.6 and as seen in the previous subsection, both estimations are very satisfactory. P-AMIRAL estimates the PSF/OTF slightly better than AMIRAL. Indeed, the RMSE on the OTF for AMIRAL is around 2.7%, and the RMSE on the OTF for P-AMIRAL is around 1.1%.

3.2.5 Results on restored images

Finally, we display in Figure 3.7 the restored images obtained using a Wiener filtering, using the estimated PSF obtained with both AMIRAL and P-AMIRAL.

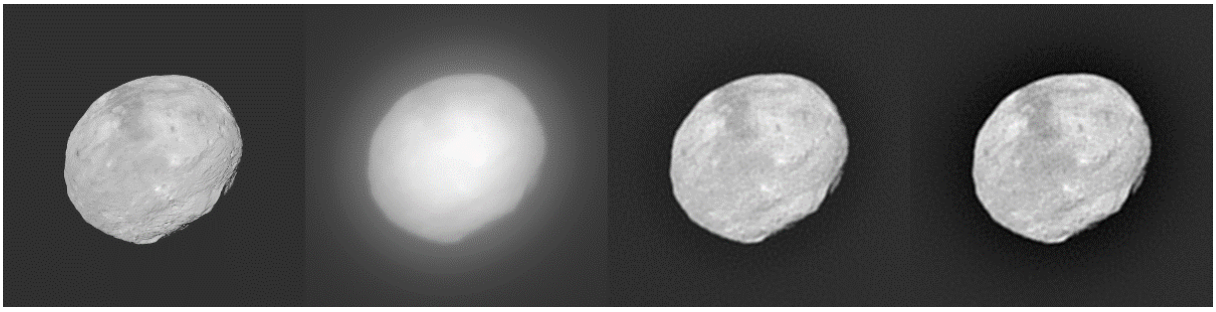

Figure 3.7 – From left to right: true object (synthetic view of Vesta), image (from simulation described in Subsection 3.2.1), P-AMIRAL restored object, AMIRAL restored object.

The results on the restored object are very close for AMIRAL and P-AMIRAL, the details on the object are well restored, coming from the fact that both AMIRAL and P-AMIRAL estimated PSF were very close to the true one. The object restored by AMIRAL is slightly more contrasted, coming from the fact that its estimated OTF is under the one estimated by P-AMIRAL therefore leading to an “over-deconvolution” of the image. Indeed, under-deconvolution means that the object was not “deblurred enough” leading to a residual blur or a loss of details, whereas over-deconvolution means that some frequencies are enhanced more than they should, creating artefacts on the object especially visible on the edges.

3.2.6 Comparison on experimental data

Finally, P-AMIRAL was applied on an experimental image of Vesta [Fétick, 2019b] taken by SPHERE/Zimpol during the *Large Program* (ID 199.C-0074), on June 8, 2018. The goal of this Large Program is to observe a part of the main-belt asteroids which diameter exceeds 100 km, at high angular resolution, with the VLT/SPHERE instrument, throughout their rotation, in order to derive their volume from their 3D shape. Combining this information with the current estimations on their mass enables the characterization of the internal structure of the asteroids, that would deepen the knowledge on their formation and evolution in the Solar system. The observations are done in the visible domain, using the 56.7 nm-width N_R filter of SPHERE/ZIMPOL, centered at 645.9 nm (corresponding to the sampling value in Table 1.2). The AO system parameters are those which were given in Chapter 1, in Table 1.2.

However, using the OASIS software [Jorda, 2010] mentioned in Chapter 5, to simulate the asteroid surface from the data, is harder for ground-based observations because of the atmospheric turbulence blurring the image, even after AO correction. Deconvolving these images is thus necessary in order to get a higher resolution on them, and retrieve useful information for the astronomers: sharp edges (volume), craters on the surface, albedo and rotation.

For this, Vesta seems to be the perfect asteroid candidate as it is the second biggest asteroid of the main-belt, it has a lot of cratering, topography and albedo details, and it is also much observed meaning that a lot of high resolutions data is available, in order to compare the details we restore to other images.

The restored images using a Wiener filtering, which parameters where those estimated by AMIRAL/P-AMIRAL, are shown in Figure 3.9.

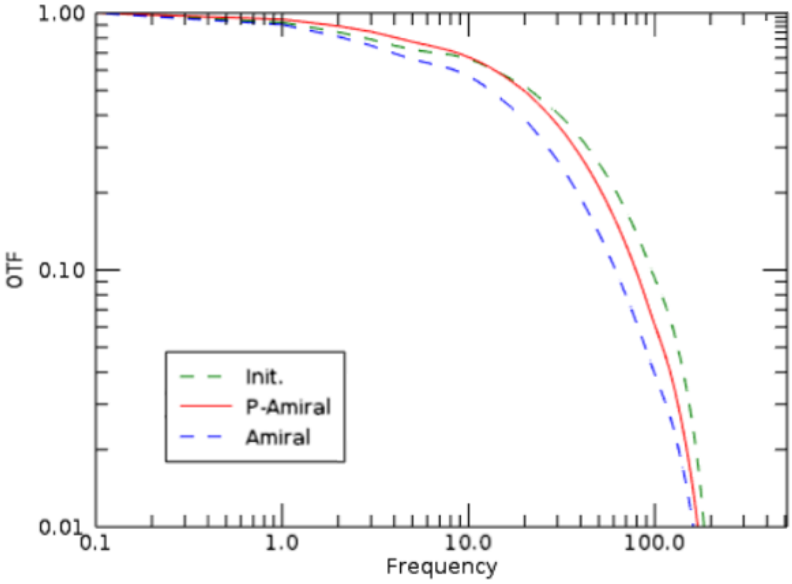

Figure 3.8 – OTF estimated by respectively P-AMIRAL and AMIRAL on experimental data.

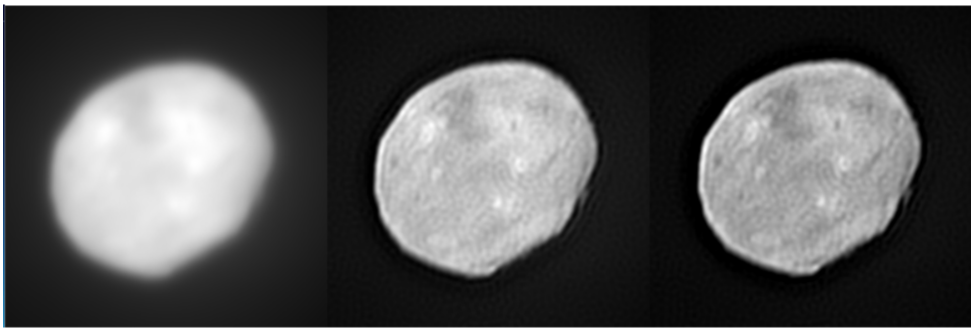

Figure 3.9 – From left to right: experimental image of Vesta, P-AMIRAL restored object, AMIRAL restored object.

As shown in Figure 3.8, P-AMIRAL estimates a slightly "sharper" PSF / higher OTF, which results in very close results on the restored images, slightly more contrasted for AMIRAL. This behavior is indeed the same as what we noticed in the simulated case, and we retrieve several details/craters on Vesta's surface which were lost on the data but which could be seen on the synthetic view shown in Figure 3.7. We notice that there are some artefacts on the surface of the asteroid, and that the edges are particularly bright. These ripples are a known effect of the quadratic regularization (Wiener filtering), and can be attenuated by using a positivity constraint along the quadratic regularization, and also an L1-L2 norm regularization. However, even with L2+ or L1-L2 regularizations [Fétick, 2020b], these bright edges still appear, thus we believe that they are due to the residual error made on the PSF. Indeed, in the simulation case, these bright edges are obtained only if the PSF used for the deconvolution is not the true PSF.

In the same mostly unsupervised mode, with p fixed to 3, the results on PSF parameters are the following: $r_0 = 0.32\,\mathrm{m}$ and $v_\phi = 2.78\,\mathrm{rad}^2$. r_0 might be a little overestimated: indeed, for the ZIMPOL instrument, estimated r_0 values by telemetry data from SPARTA (the real-time computer controlling the extreme AO system SAXO of the SPHERE instrument) or using the MASS/DIMM instruments are respectively within $[15; 32]\,\mathrm{cm}$ and $[9; 23]\,\mathrm{cm}$ [Fétick, 2020a].

3.2.7 Overall discussion

While AMIRAL requires the choice of a family of PSFs, P-AMIRAL assumes a PSF model which the sought PSF must follow. Both methods have their own drawbacks and advantages: on the one hand, because AMIRAL requires to choose the family of PSF accordingly and to look for as much unknowns as the number of PSFs in the family, it can also be more versatile and could work if the PSF does not the PSFAO19 model. On the other hand, using P-AMIRAL implies that the PSFAO19 model is adapted, but there are less unknowns to estimate, in our case the two PSF parameters r_0 and v_ϕ. Thus, though AMIRAL can adapted to other PSF models, for AO-corrected astronomical long-exposure images, the PSFAO19 model gives a parsimonious (meaning, with only a few parameters), physical description of the PSF and is thus suitable for this application. Therefore, in this case, using P-AMIRAL is more adapted as it does not require to choose a PSF family.

On the few tested simulated astronomical observation case, P-AMIRAL gives slightly better results on simulated data, with less dependency with respect to the initialization, than AMIRAL, when the chosen convergence threshold is taken small enough. Both methods give very close results on the simulated and experimental restored images.

Chapter 4

Development of a method to compute the MMSE estimator

Introduction

In this chapter we explicit the method we used to compute the marginal MMSE estimator. In the first section, we describe the models we use in order to introduce the posterior distribution we want to compute. In the second section, we give motivations to use MCMC methods, then we describe the algorithms we have chosen to work with and finally we also give gradient and Fisher information calculations, which can be used for more advanced sampling methods.

Contents

4.1 Computing the posterior distribution

4.1.1 Object prior

As in the specific case described in Chapter 2, as a prior for the object, we consider a Gaussian model described by its mean $\boldsymbol{m_o}$ and its covariance matrix $\boldsymbol{R_o}$. As we suppose that $(\boldsymbol{o} - \boldsymbol{m_o})$ is stationary, $\boldsymbol{R_o}$ is a 2D Tœplitz-block Tœplitz matrix, which we will approximate as circulant, therefore diagonalizable in the Fourier domain. Thus, as in Chapter 2, the object is described hereafter by its mean $\boldsymbol{m_o}$ and its PSD $\boldsymbol{S_o}$ which is the diagonal of the diagonalized $\boldsymbol{R_o}$ matrix.

The mean object $\boldsymbol{m_o}$ is taken constant on all pixels, and because the uncertainty on $\boldsymbol{m_o}$ is small given the number of pixels, it is estimated at the average value of the image.

For the object PSD $\boldsymbol{S_o}$, as described in Chapter 3, we consider a circularly-symmetric model, described by a few parameters:

$$S_o(f) = \frac{1}{\gamma_o} \bar{S}_o(\boldsymbol{f}), \text{ with } \bar{S}_o(\boldsymbol{f}) = 1/(k + f^p) \tag{4.1}$$

In this model, γ_o sets the global PSD level, p is the PSD decrease rate at high frequencies, and k gives the breakpoint between the two regimes of the model.

The object prior distribution writes as follows:

$$p(\boldsymbol{o}|\gamma_o, k) = \Big(\frac{\gamma_o}{2\pi}\Big)^{N/2} \prod_{\boldsymbol{f}} \Big(\bar{S}_o(\boldsymbol{f})^{-1/2} \exp\big[-\frac{\gamma_o}{2} \frac{|\tilde{o}(\boldsymbol{f}) - \tilde{m}_o(\boldsymbol{f})|^2}{\bar{S}_o(\boldsymbol{f})}\big]\Big) \tag{4.2}$$

In previous works [Fétick, 2020b], attempts to estimate hyperparameter p jointly with the other sought parameters has been shown to strongly decrease PSF parameter estimation accuracy. Therefore, we choose to work in a “mostly unsupervised” mode, where p is fixed to a standard value. In the case of astronomical observations of asteroids, a well-fitting empirical value is around $p = 3$, whereas for satellite observation a standard value for p would be around 2.5–2.6.

4.1.2 PSF model

As described in Chapter 1, the specificity of astronomical and satellite observation from ground is the impact of atmospheric turbulence, which is partially corrected by AO in our case. Given the exposure time of the images we are processing here (around 1 s, for a typical variation time of turbulence around 10 ms), we consider having a long-exposure AO-corrected PSF. We will use the PSFAO19 model [Fétick, 2019a], the PSF will then be described by its two main parameters: the Fried parameter r_0 and the residual phase variance v_ϕ. The other parameters of the model, which have much less impact on the PSF, will be fixed to reasonable values, as done in Chapter 3.

4.1.3 Noise prior

Given the flux levels we are considering in astronomical and satellite imaging, as said in Chapter 2, we approximate noise as Gaussian, zero-mean, additive and white, which are again reasonable hypotheses given the flux levels in our images. This noise precision is thus the only parameter we need to define fully the noise statistics, here we denote it by γ_n. The noise PSD is $S_n = 1/\gamma_n$.

4.1.4 Likelihood

Let N be again the image size in pixels. We consider $\boldsymbol{h}$ as an array of the same size as the image and the object, and approximate it as periodic. Given this approximation and the imaging model of Eq. (1.12), and because noise is independent from the object, the joint likelihood $p(\boldsymbol{i}|\boldsymbol{o}, \gamma_n, r_0, v_\phi)$ writes:

$$p(\boldsymbol{i}|\boldsymbol{o}, \gamma_n, r_0, v_\phi) = \Big(\frac{\gamma_n}{2\pi}\Big)^{N/2} \prod_{\boldsymbol{f}} \exp(-\frac{\gamma_n}{2}|\tilde{i}(\boldsymbol{f}) - \tilde{h}(\boldsymbol{f}; r_0, v_\phi)\tilde{o}(\boldsymbol{f})|^2) \tag{4.3}$$

4.1.5 Marginal likelihood

Similarly to what was done in Subsection 2.4.2, given that the image is a linear combination of Gaussian noise and a linear image model with a Gaussian object prior, the image is also Gaussian and the marginal likelihood writes:

$$p(\boldsymbol{i}|\gamma_n, \gamma_o, k, r_0, v_\phi) = (2\pi)^{-N/2} \prod_{\boldsymbol{f}} \Big(S_i(\boldsymbol{f})^{-1/2} \exp\big[-\frac{1}{2}|\tilde{i}(\boldsymbol{f}) - \tilde{m}_i(\boldsymbol{f})|^2 / S_i(\boldsymbol{f})\big]\Big) \tag{4.4}$$

with image PSD $\boldsymbol{S}_i$ and mean image $\boldsymbol{i}_m$:

$$\begin{aligned} S_i(\boldsymbol{f}) &= S_o(\boldsymbol{f})|\tilde{h}(\boldsymbol{f})|^2 + S_n = \frac{1}{\gamma_o}\frac{1}{k + f^p}|\tilde{h}(\boldsymbol{f})|^2 + \frac{1}{\gamma_n} \\ \tilde{m}_i(\boldsymbol{f}) &= \tilde{h}(\boldsymbol{f})\tilde{m}_o(\boldsymbol{f}) \end{aligned} \tag{4.5}$$

4.1.6 Parameter priors

We consider that each parameter γ_n, γ_o, k, r_0 and v_ϕ, which are taken independent, can take any value in a given range. Therefore, in the absence of more information, following the Laplace rule (or principle of insufficient reason), we use uniform priors for each of them [Kass, 1996]. In further works, it could be possible to use telemetry data provided by the AO system, in order to impose more informative priors. The prior interval is taken large enough: from 0.1 to 10 times the usual value of the considered parameter for γ_n, γ_o and k, given knowledge on these parameters. The prior intervals taken for PSF parameters are the following: for r_0 we take $[5\,\text{cm}; 30\,\text{cm}]$ and for v_ϕ we take $[0.5\,\text{rad}^2; 3.0\,\text{rad}^2]$, which correspond to a large range of values taking into account the global knowledge on the AO system and the turbulence.

4.1.7 Hierarchical model

In Figure 4.1 we provide the chosen hierarchical model which sums up the variable interdependency [Bishop, 2007, Chap. 8]. Each upper node (parent) is connected with an edge to a node below (child) and the model says that a child's distribution, given all nodes above, only depends on its parents.

In our model, it means for example that $p(\boldsymbol{i}|\boldsymbol{o}, \gamma_n, \gamma_o, k, r_0, v_\phi) = p(\boldsymbol{i}|\boldsymbol{o}, \gamma_n, r_0, v_\phi)$. Therefore $p(\boldsymbol{i}, \gamma_o, k|\boldsymbol{o}, \gamma_n, r_0, v_\phi) = p(\boldsymbol{i}|\boldsymbol{o}, \gamma_n, r_0, v_\phi) \times p(\gamma_o, k|\boldsymbol{o}, \gamma_n, r_0, v_\phi)$, which means that the image $\boldsymbol{i}$ and object PSD parameters γ_o and k are independent conditionally to the object and the other parameters.

Additionally, the object, the noise variance and the PSF parameters are independent conditionally to object PSD parameters meaning $p(\boldsymbol{o}|\gamma_n, \gamma_o, k, r_0, v_\phi) = p(\boldsymbol{o}|\gamma_o, k)$.

Moreover, as the hierarchical model reads, all $P = 5$ parameters $\boldsymbol{\theta} = \{\gamma_o, k, r_0, v_\phi, \gamma_n\}$ are *a priori* considered independent.

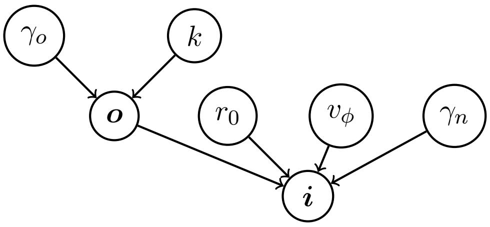

Figure 4.1 – Hierarchical model summing up the inter-dependency between the object, the image and all parameters.

4.1.8 Marginal posterior distribution

The posterior distribution can be easily derived from the parameter priors and marginal likelihood given previously (in Equations 4.1, 4.2, 4.4 and 4.5):

$$\begin{aligned} p(\gamma_n, \gamma_o, k, r_0, v_\phi|\boldsymbol{i}) &= \frac{p(\gamma_n)p(\gamma_o)p(k)p(r_0)p(v_\phi)}{p(\boldsymbol{i})} p(\boldsymbol{i}|\gamma_n, \gamma_o, k, r_0, v_\phi) \\ &= \frac{1}{p(\boldsymbol{i})} U_{\gamma_n}(\gamma_n) U_{\gamma_o}(\gamma_o) U_k(k) U_{r_0}(r_0) U_{v_\phi}(v_\phi) \\ &\times (2\pi)^{-N/2} \prod_{\boldsymbol{f}} \Big(\frac{1}{\gamma_o} \frac{1}{k + f^p} |\tilde{h}(\boldsymbol{f}; r_0, v_\phi)|^2 + \frac{1}{\gamma_n} \Big)^{-1/2} \\ &\times \exp\Big[- \frac{1}{2} \frac{|\tilde{i}(\boldsymbol{f}) - \tilde{h}(\boldsymbol{f}; r_0, v_\phi)\tilde{m}_o(\boldsymbol{f})|^2}{\gamma_o^{-1}(k + f^p)^{-1} |\tilde{h}(\boldsymbol{f}; r_0, v_\phi)|^2 + \gamma_n^{-1}} \Big] \end{aligned} \tag{4.6}$$

Here, for all parameters, there is no conjugate priors (even for γ_n and γ_o) due to the marginalization.

4.2 Markov Chain Monte Carlo (MCMC) methods

4.2.1 Why MCMC?

As previously said in Chapter 2, the MMSE estimator is known to be the mean of the posterior distribution, whereas the MAP estimator is its mode (and the MMAE estimator is known to be the median of the posterior distribution for each sought parameter). Given the complexity of the posterior given in Equation 4.6, there is no known analytical way to calculate it. Indeed, the dependency of the posterior distribution with respect to the sought parameters is complex, and computing the normalization factor $p(\boldsymbol{i})$, which corresponds to the integration of the marginal likelihood and the parameters priors over the parameters: $p(\boldsymbol{i}) = \int_{\boldsymbol{\theta}} p(\boldsymbol{i}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$, is hardly possible analytically.

A way to compute the posterior distribution (and so, the different estimators) is to do it numerically, by drawing samples under the posterior distribution, using a Monte Carlo method.

However, drawing these samples can be difficult, especially for high-dimension problems. The difficulty, using a Monte Carlo method, comes from the fact that these methods draw the random samples independently. Thus, they do not take into account the graphical (hierarchical) model related to the targeted distribution.

A solution is then to use a Markov Chain Monte Carlo (MCMC) method [Gamerman, 2006]. A Markov chain is a process which respects the Markov property, saying that the next state of this process only depends on the current one, and not on all the previous ones. After a while, the process reaches an equilibrium, the distribution we obtain is said to be the stationary distribution. The idea is here to combine the Monte Carlo sampling with a judicious Markov Chain which is built in order to finally settle on the targeted distribution.

4.2.2 Random-Walk Metropolis-Hastings and Gibbs algorithm

The posterior distribution being complex, it is not possible to sample it directly, therefore we use a Metropolis-Hastings algorithm to bypass the problem [Robert, 2004; Villeneuve, 2012; Orieux, 2010; Orieux, 2013]. It consists, for each iteration, in drawing samples under a chosen proposition distribution $q(\boldsymbol{\theta})$ and accepting the samples (else, duplicating the previous value) with a prescribed probability α. For the k-th iteration, α writes:

$$\alpha = \frac{p(\boldsymbol{\theta}^{(prop)}|\boldsymbol{i})}{p(\boldsymbol{\theta}^{(k-1)}|\boldsymbol{i})} \frac{q(\boldsymbol{\theta}^{(k-1)}|\boldsymbol{\theta}^{(prop)})}{q(\boldsymbol{\theta}^{(prop)}|\boldsymbol{\theta}^{(k-1)})} \tag{4.7}$$

Several versions are possible: in particular, we can either draw all the parameters simultaneously (standard Metropolis-Hastings), or separately (Metropolis-Hastings-within-Gibbs). Drawing the parameters together can make the acceptance probability fall (except if we use more advanced, *e.g.* gradient-based algorithms such as MALA or HMC methods), [Robert, 2004; Vacar, 2016; Duane, 1987; Girolami, 2011] whereas drawing parameters individually can slow down the algorithm as it changes parameters one by one and requires more marginal likelihood computations. We will compare both versions, Metropolis-Hastings and Metropolis-Hastings-within-Gibbs algorithms, in this manuscript.

In our case, we use a Random Walk (Metropolis-Hastings) algorithm: the proposed sample for each parameter is drawn under a symmetric (Gaussian) distribution around the current value of the parameter. As the chosen prior distribution is a uniform distribution over an interval, if the proposed parameter is out of its prior interval its acceptance probability is 0, the proposition is rejected thus the previous value is duplicated. The choice of the standard deviation of the Gaussian distribution is based, in our case, on the empirical sensitivity of the PSF or the noise and object PSD to the parameters. This choice only impacts the convergence time of the method, but not its final outcome.

In a standard Gibbs algorithm, each parameter is drawn under its own conditional posterior distribution, which is proportional to the prior of the considered parameter times the marginal likelihood of Eq. (4.4) [Orieux, 2010; Orieux, 2013; Villeneuve, 2012]. The conditional posterior distribution for each parameter writes:

$$p(\theta_n|\boldsymbol{i}, \boldsymbol{\theta}_{m\neq n}) = \frac{p(\theta_n)p(\boldsymbol{i}|\boldsymbol{\theta})}{p(\boldsymbol{i})} \tag{4.8}$$

where θ_n is the considered parameter and $\boldsymbol{\theta}_{m\neq n}$ the four other parameters. In our case, for each iteration, we are using a Metropolis-Hastings algorithm to target $p(\theta_n|\boldsymbol{i}, \boldsymbol{\theta}_{m\neq n})$.

For both versions, asymptotically, the samples are under the marginal posterior distribution for all parameters, and their empirical mean tends towards the expectation of the distribution [Robert, 2004].

The parameters can be initialized to a random value, under the prior distribution of the parameters. For the results given in Chapters 5 and 6, most parameters were initialized to their maximum value to illustrate the convergence. Various initializations were tested, resulting in very close estimations and showing the independence of the solution with respect to the initialization (for a number of iterations large enough). The Random Walk Metropolis-Hastings-within-Gibbs algorithm we use in this work is provided in Algorithm 1.

Algorithm 1 Metropolis-Hastings-within-Gibbs algorithm

```
Define initial $\boldsymbol{\theta}^{(0)}$
for each iteration $k$ do
    for each parameter $\theta_n$ do
        Propose $\theta_n^{prop} \sim \mathcal{N}(\theta_n^{(k-1)}, \sigma_{\theta_n})$
        Acceptance rate $\alpha_n \leftarrow \min\left(1, \frac{p(\theta_n^{prop})p(\boldsymbol{i}|\boldsymbol{\theta}^{(k)}_{m<n}, \theta_n^{prop}, \boldsymbol{\theta}^{(k-1)}_{m>n})}{p(\theta_n^{(k-1)})p(\boldsymbol{i}|\boldsymbol{\theta}^{(k)}_{m<n}, \boldsymbol{\theta}^{(k-1)}_{m\geq n})}\right)$ ▷ Equation 4.7 and 4.8
        Random acceptance $u \sim \mathcal{U}([0;1])$
        if $u < \alpha$ then
            Accept the proposal $\theta_n^{(k)} \leftarrow \theta_n^{prop}$
        else
            Duplicate previous sample $\theta_n^{(k)} \leftarrow \theta_n^{(k-1)}$
        end if
    end for
end for
```

4.2.3 Gradients and Fisher information

More advanced MCMC methods take advantage on the gradient to converge more rapidly towards the solution, such as Metropolis-Adjusted Langevin Algorithms (MALA) [Roberts, 2002; Vacar, 2016]. In these methods, the parameters' values at the current iteration are computed given their values at the previous iteration and a term depending on the gradient and a preconditioning term.

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \tau A \nabla \log p(\boldsymbol{\theta}^{(k)}|i) + \sqrt{2\tau A}\xi^{(k)} \tag{4.9}$$

with τ a fixed step, all $\xi^{(k)}$ independent draws from a multivariate standard normal distribution on $\mathbb{R}^P$ and A the preconditioning matrix, which must be positive-definite.

Several options are available for the preconditioning, one of them is to use the inverse Fisher information, which makes a good approximation of the Hessian [Girolami, 2011] and is positive-definite (unlike the Hessian, which inversion can lead to instabilities as in [Qi, 2002]). The gradient and the Fisher information for each estimated parameter are given in Appendix A.

Chapter 5

Results of the MMSE estimator on simulated data

Introduction

In this chapter we provide detailed results obtained using the MMSE estimator for simulated data, on both astronomical and satellite images. The first section gives the results obtained on an astronomical image, focusing partly on the estimated parameters as well as the associated uncertainties. In Section 2, we test the robustness of the method by trying it on multiple noise realisations. The third section compares two options namely two MCMC algorithms in terms of computational time. Then, in Section 4, we discuss the *a posteriori* coupling between parameters, particularly we stress on the impact of one of the object hyperparameters p. Finally, in Section 5 we present results on a satellite image, similarly to the first section.

Contents

5.1 Results on simulated astronomical images

5.1.1 Simulation conditions

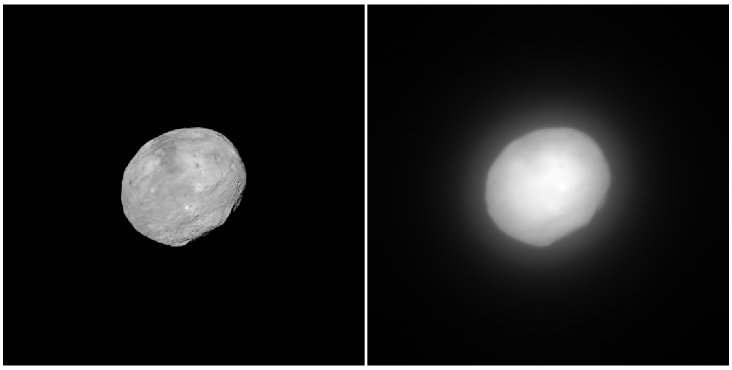

Figure 5.1 – Left: synthetic view of Vesta (true object $\boldsymbol{o}$), of size 512×512. Right: simulated image $\boldsymbol{i}$, with true parameters $r_0 = 0.15\,\mathrm{m}$, $v_\phi = 1.3\,\mathrm{rad}^2$ and $\gamma_n = 2.62\,10^{-4}\,\mathrm{ph}^{-2}$.

The obtained results are shown for the simulated image displayed in Figure 5.1, using as the true object the synthetic view of asteroid Vesta, built by OASIS software [Jorda, 2010], on a dark background of size $N = 512 \times 512$ pixels. The simulated AO system is a “SPHERE-like” AO system, whose parameters are summed up in Chapter 3, in Table 3.2.

The PSF is simulated using the PSFAO19 model [Fétick, 2019a]. Here, true PSF parameters are $r_0 = 0.15\,\mathrm{m}$ and $v_\phi = 1.3\,\mathrm{rad}^2$ at the imaging wavelength $\lambda = 550\,\mathrm{nm}$, which corresponds to realistic turbulence and correction conditions. Noise is taken zero-mean, additive, white and Gaussian with a variance equal to the mean value of the object as a first approximation of the photon noise, just as in Chapter 3. The total flux of the object is set to $F_o = 10^9\,\mathrm{ph}$ (photons), typical from VLT/SPHERE/Zimpol asteroid observations (during ESO’s Large Program, ID 199.C-0074), therefore $\gamma_n = N/F_o = 2.62 \times 10^{-4}\,\mathrm{ph}^{-2}$.

The PSF and PSD parameters are estimated following the proposed marginal method, except the mean object $\boldsymbol{m}_o$, which is taken equal to the average value of the image (because the uncertainty on it is small given the number of pixels), and the object PSD power which is fixed to $p = 3$, which corresponds to a reasonable default value of p for asteroids, as mentioned in Chapter 4.

We are using here a Metropolis-Hastings-within-Gibbs algorithm described in Chapter 4, for which the tuning of the standard deviation of the Gaussian proposition is chosen to be around 0.01 times the allowed range of the prior. Precisely, the used prior interval and tuning for each parameter are given in Table 5.1. This tuning only impacts the acceptance rate for each parameter, which should be carefully looked at to find a balance between small frequent changes and larger but less frequent modifications. Theoretically, a good acceptance rate for a multi-dimensional Gaussian target distribution is around 23% [Gelman, 1997], we will then aim to have acceptance rate around this value.

Parameter	Prior min - max	Step tuning
γ_n (ph^{-2})	$2.62{\times}10^{-5}$ - $2.62{\times}10^{-3}$	$2.62{\times}10^{-6}$
r_0 (m)	0.05 - 0.5	0.001
v_ϕ(rad^2)	0.5 - 3.0	0.01
γ_o (ph^{-2})	$2.62{\times}10^{-15}$ - $2.62{\times}10^{-11}$	$2.62{\times}10^{-14}$
k	0.01 - 10	0.1

Table 5.1 – Prior intervals and tuning of the Gaussian standard deviation for γ_n, r_0, v_ϕ, γ_o and k.

The parameters which have a true value, namely noise and PSF parameters, are initialized to their maximum value ($r_0 = 0.5\,\mathrm{m}$, $v_\phi = 3.0\,\mathrm{rad}^2$, $\gamma_n = 2.62 \times 10^{-3}$ ph^{-2}) in order to clearly notice the convergence of the random samples. For k and γ_o, we initialized them to what we are expecting typically for these parameters: k is initialized to 1, and γ_o to $\gamma_o = N/F_o^2 = 2.62{\times}10^{-13}$. The Gibbs sampler is run for 100 000 iterations, which corresponds to a few hours, in order to check that the chains have indeed converged in practice.

5.1.2 Chains of random samples for the estimated parameters

In Figure 5.2, we plot the samples chains and the corresponding histograms for γ_n, r_0 and v_ϕ. The inspection of Figure 5.2 suggests that chains have a short burn-in period (around 20 000 iterations), followed by a stationary state. As expected from Markov chains, for each parameter the samples are correlated. Moreover, the samples are concentrated in a small interval relatively to their prior interval.

The acceptance rate for most parameters is slightly over 30% (which is a pretty correct rate), except for γ_o and k whose acceptance rate is way higher (respectively 68% and 94%).

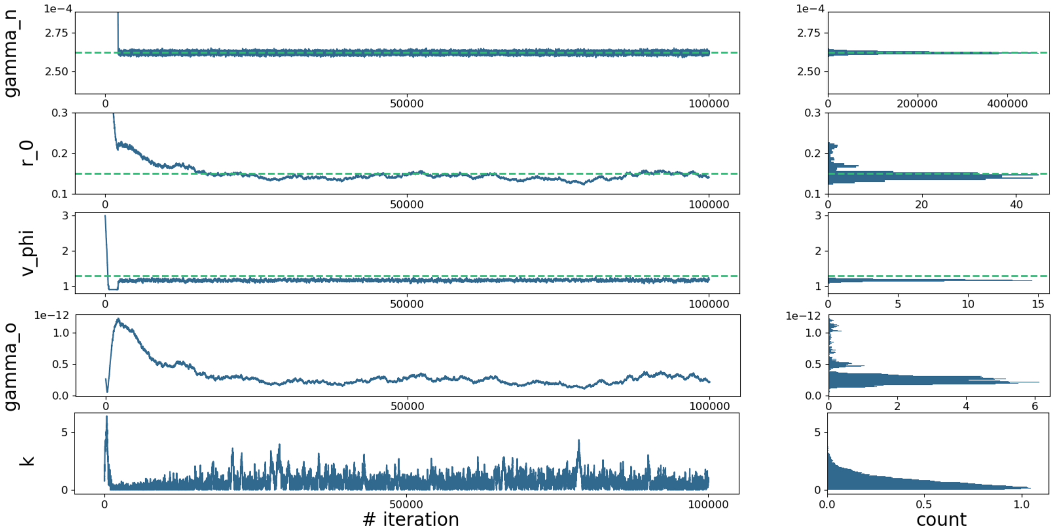

Figure 5.2 – From top to bottom: γ_n, r_0, v_ϕ, γ_o, k. Left: chain of samples for simulated astronomical image. Right: corresponding histogram. True values in dashed line.

Parameter	$m \pm \sigma$	True
γ_n (ph^{-2})	$2.62\times10^{-4} \pm 8.03\times10^{-7}$	2.62×10^{-4}
r_0 (m)	0.142 ± 0.007	0.15
v_ϕ(rad^2)	1.17 ± 0.03	1.30
γ_o (ph^{-2})	$2.37\times10^{-13} \pm 5.41\times10^{-14}$	-
k	0.768 ± 0.594	-

Table 5.2 – Empirical mean value, associated empirical standard deviation of the samples shown in Figure 5.2 and true value, for γ_n, r_0, v_ϕ, γ_o and k for simulated astronomical image, with $p = 3$.

5.1.3 Quality of estimation of the PSF

The empirical mean values m, corresponding to our estimates, and empirical standard deviations σ, corresponding to our predicted uncertainties, for each parameter are displayed in Table 5.2. Firstly, we can note that the error made on the parameters is small: the noise precision is very precisely estimated, with an error smaller than 0.2%, and PSF parameters are also well estimated, with a 5% error on r_0 and a 10% error on v_ϕ. Additionally, the estimated r_0 and v_ϕ are very close to the previous results obtained with P-AMIRAL: for the same conditions [Fétick, 2020b], the estimated PSF parameters were $r_0 = 0.142\,\mathrm{m}$ and $v_\phi = 1.13\,\mathrm{rad}^2$ (compared to $r_0 = 0.142\,\mathrm{m}$ and $v_\phi = 1.17\,\mathrm{rad}^2$ in Table 5.2).

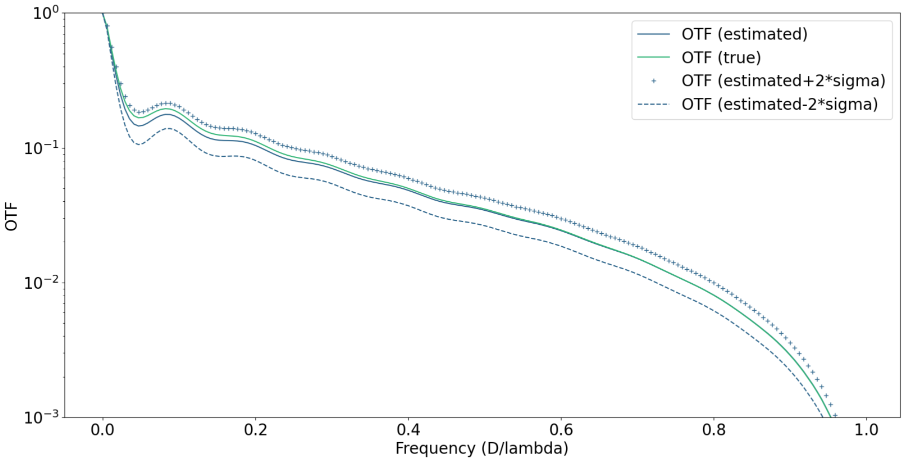

Figure 5.3 – True (in green) and estimated (in blue) OTF for simulated astronomical image, including computed uncertainties (in blue, + and - for upper and lower uncertainty bounds).

We also compare the resulting OTF to the true OTF in Figure 5.3. The slight underestimation of r_0 leads to the lowering of the global OTF level and its impact can mainly be seen at low frequencies. Concerning v_ϕ, its mild underestimation leads to a slower decrease of the OTF and impacts the slope of the latter at medium-high frequencies [Fétick, 2020b]. Thus, we notice that the errors on both parameters partially compensate. We interpret this compensation as the result of the fact that (a) essentially, the method fits the OTF itself (fitting the image PSD), and (b) several couples of parameters lead to close OTFs, in other words that the two parameters are coupled. As a result, the normalized RMSE for the OTF, computed as $\sqrt{\frac{\sum |m_{\tilde{h}} - \tilde{h}|^2}{\sum |m_{\tilde{h}}|^2}}$ with true OTF $\tilde{h}$ and estimated OTF $m_{\tilde{h}}$ is quite small (around 7%).

5.1.4 Results on the derived uncertainties

Concerning the uncertainties derived from our method, we notice in Table 5.2 that the true value for parameter r_0 is in the range $[m_{r_0} \pm 2\sigma_{r_0}]$, and the true v_ϕ is in the interval $[m_{v_\phi} \pm 5\sigma_{v_\phi}]$, therefore the uncertainties on PSF parameters seem under-estimated. We can also compute uncertainties directly on the sought OTF: for each sample (r_0, v_ϕ), we compute the corresponding OTF in order to compute its empirical mean $m_{\tilde{h}}$ and standard deviation $\sigma_{\tilde{h}}$. As shown in Figure 5.3, the true OTF is within the interval $[m_{\tilde{h}} \pm 2\sigma_{\tilde{h}}]$, for all frequencies. Therefore, even though the uncertainties on PSF parameters are somewhat under-estimated, our method gives a very satisfactory uncertainty estimation on the OTF itself.

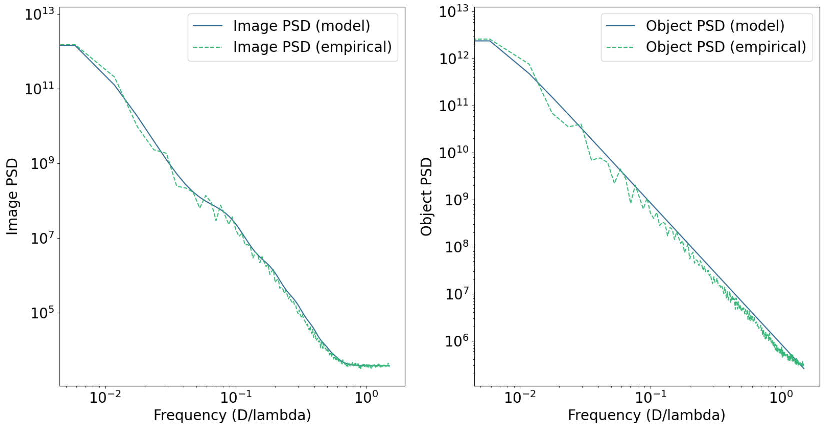

Figure 5.4 – PSDs for simulated astronomical image. Model in solid line, and empirical PSD averaged azimuthally in dashed line. Left: Image PSD. Right: Object PSD.

5.1.5 Results on object and image PSDs

In Figure 5.4 (left), we perform an important sanity check of the method to verify that our model for the image PSD of Eq. (4.5), which combines object PSD, PSF and noise PSD, accurately fits the empirical image PSD averaged azimuthally (*cf.* Eq. (4.4)). Moreover, given the fact that the true object is not the realization of a Gaussian random field following our PSD model, a way to check γ_o and k estimation accuracy is to look at the fitting of our model to the empirical object PSD, averaged azimuthally. As displayed in Figure 5.4 (right), the object PSD model visually fits correctly the empirical object PSD, the slight overestimation of the object PSD being consistent with the slight underestimation of the OTF.

5.1.6 Results on restored image

Finally, Figure 5.5 shows the image in Figure 5.1 restored with the estimated OTF. Many details of the Vesta surface can be seen, that were not visible on the data. Particularly, with our method we retrieve sharp edges of the asteroid from which one can estimate the object volume and sphericity, as well as main crater and albedo features.

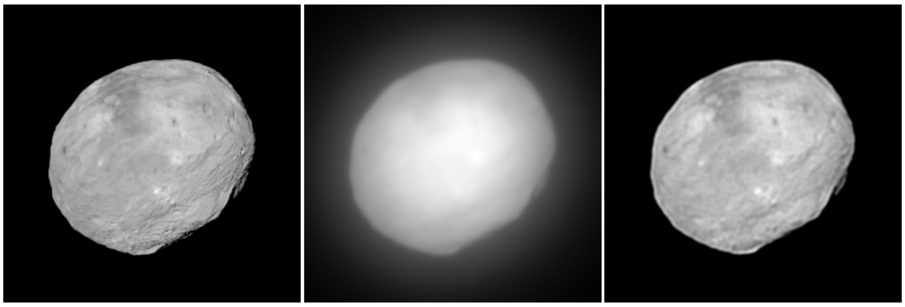

Figure 5.5 – Left and center: true object and image for simulated asteroid observation, 256×256 cropped from Figure 5.1. Right: restored object from the estimated PSF and PSD parameters (using a L2-norm regularization, with positivity constraint), also cropped.

5.1.7 Comparison between MMSE, MMAE and MAP estimators

As mentioned in Chapter 2 and 4, using our method we can compute the MMSE and the MMAE estimator, corresponding to respectively the posterior mean and the posterior median. In practice, we compute the average and the median of our samples after discarding the boiling time. Additionally, we can have an approximation of the MAP by taking the sample maximizing the marginal likelihood among our samples (also discarding the boiling time, when we have not converged towards the posterior distribution). The results using these 3 estimators are showed in Table 5.3. We notice that all 3 results are very close, especially the MMAE and the MMSE results. This is not very surprising, given that the parameters have a pretty narrow and symmetric marginal posterior as displayed in the histograms in Figure 5.2.

Parameter	MMSE	MMAE	MAP	True
γ_n (ph^{-2})	2.62×10^{-4}	2.62×10^{-4}	2.62×10^{-4}	2.62×10^{-4}
r_0 (m)	0.142	0.142	0.140	0.15
v_ϕ(rad^2)	1.17	1.17	1.17	1.30
γ_o (ph^{-2})	2.37×10^{-13}	2.32×10^{-13}	2.16×10^{-13}	-
k	0.768	0.624	0.422	-

Table 5.3 – Empirical mean and median values, as well as value maximizing the likelihood among the samples shown in Figure 5.2 and true value, for γ_n, r_0, v_ϕ, γ_o and k for simulated astronomical image, with $p = 3$.

5.1.8 Results with a more realistic noise

We now simulate the observation of Vesta using the same conditions than in Subsection 5.1.1, except that we now simulate a more realistic noise, using a Poisson distribution to simulate the photon noise, and a stationary Gaussian noise for the read-out noise with a standard deviation of 20 photo-electrons. The results are summed up in Table 5.4. Even if the simulated noise does not exactly match the stationary noise model, all estimated parameters are still very close to the previous estimations. The PSF parameters are still well estimated, with an error around 3% for r_0 and around 8% for v_ϕ. Their associated uncertainties are also still satisfactory for r_0 and, similarly, slightly underestimated for v_ϕ as discussed previously. Thus the violation of the

Parameter	$m \pm \sigma$	True
γ_n (ph^{-2})	$2.57{\times}10^{-4} \pm 7.71{\times}10^{-7}$	-
r_0 (m)	0.145 ± 0.007	0.15
v_ϕ(rad^2)	1.20 ± 0.03	1.30
γ_o (ph^{-2})	$2.49{\times}10^{-13} \pm 5.57{\times}10^{-14}$	-
k	0.717 ± 0.524	-

Table 5.4 – Mean value, associated standard deviation and true value, for γ_n, r_0, v_ϕ, γ_o and k for simulated astronomical image with a more realistic noise (Poisson + Gaussian noise), with $p = 3$ and $m_o = m_i$.

stationary noise hypothesis does not impair the results with our method, given our simulation conditions.

5.2 Tests on several noise realisations

5.2.1 Results on the quality of estimation and derived uncertainties for parameters

To test the robustness of our method to noise, we ran the algorithm for ten different noise realisations, in the simulation conditions described in Section 5.1, with $p = 3$. We compute the bias and standard deviation of the estimated parameters on these ten noise realisations, as well as the maximum error. We also compute the minimum and maximum predicted uncertainty (*i.e.* the standard deviation of the posterior distribution). These values are summarized in Table 5.5.

Parameter	True	Max. error	Empirical bias	Empirical std. dev.	Predicted uncertainty
γ_n (ph^{-2})	$2.62{\times}10^{-4}$	$2.3{\times}10^{-6}$	$9.5{\times}10^{-8}$	$9.0{\times}10^{-7}$	$\in [7.8{\times}10^{-7}, 8.0{\times}10^{-7}]$
r_0 (m)	0.150	0.012	0.006	0.005	$\in [0.006, 0.009]$
v_ϕ(rad^2)	1.30	0.14	0.12	0.01	$\in [0.02, 0.03]$
γ_o (ph^{-2})	-	-	-	$4.1{\times}10^{-14}$	$\in [4.5{\times}10^{-14}, 8.9{\times}10^{-14}]$
k	-	-	-	0.1	$\in [0.4, 0.7]$

Table 5.5 – Summary of results on ten noise realisations: true value, maximum error and bias (if available), standard deviation of estimates for γ_n, r_0, v_ϕ, γ_o and k, and minimum/maximum predicted uncertainty, with $p = 3$.

In these ten cases, we notice very little variations on the estimates: the computed standard deviations (fifth column in Table 5.5) are small with respect to the true values (second column). Moreover, the estimates are satisfactory: first, the errors on the estimated parameters are quite small (third column), particularly on the noise precision (error is less than 1%). For the PSF parameters, the error is always smaller than 11%.

Concerning parameter r_0, the predicted uncertainty is very satisfactory: the true value is always within the interval $[m_{r_0} \pm 2\sigma_{r_0}]$. For parameter v_ϕ, we notice that the error is here dominated by the bias, which is more than ten times greater than the standard deviation (which is not the case for the other parameters). Our interpretation is that this bias is due to the choice of p, as supported by the discussion in Subsection 5.4.3.

5.2.2 Results on derived uncertainties for the OTF

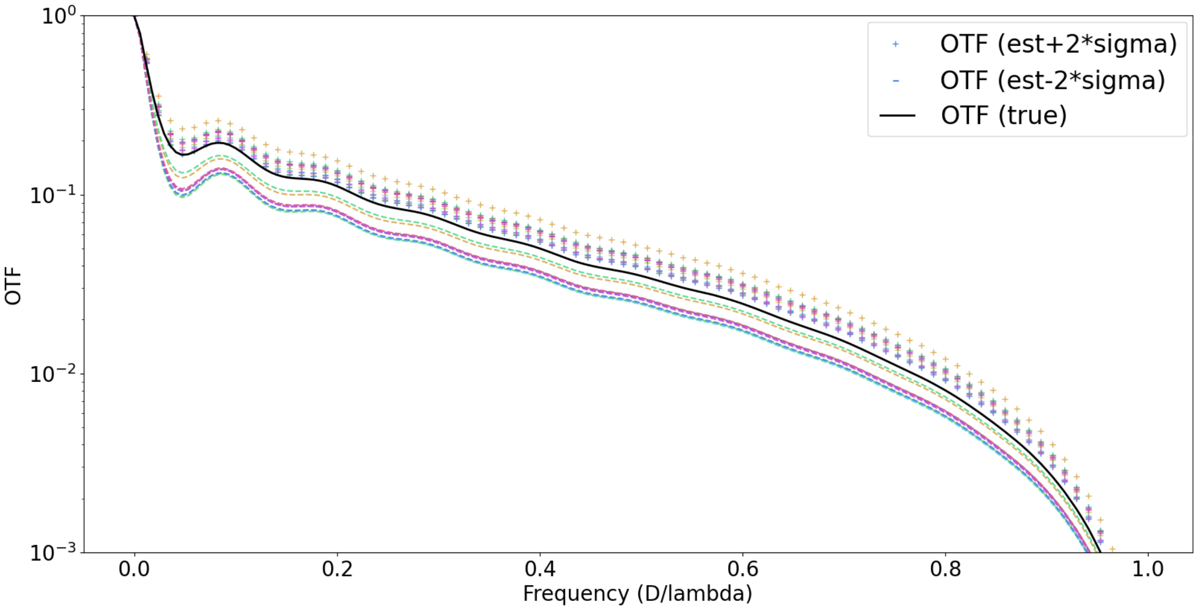

Figure 5.6 – Results on OTF uncertainties for ten realisations of a noisy simulated astronomical image: true OTF (in black) and predicted range $[m_{\tilde{h}} \pm 2\sigma_{\tilde{h}}]$ (+ and - for upper and lower uncertainty bounds, each color corresponds to a noise realisation).

Finally, even though the uncertainties are under-estimated for v_ϕ with the default p, concerning the OTF itself the uncertainties are always well estimated. To compute these uncertainties, for one noise realisation, for each sample of the PSF parameters chains, we computed the corresponding OTF and squared OTF. These OTFs and squared OTFs are averaged for a given frequency length, and this is done for each noise realization. Additionally, with the averaged OTFs $m_{\tilde{h}}$ and squared OTFs $m_{\tilde{h}^2}$, we computed the posterior standard deviation on OTF: $\sigma^2_{\tilde{h}} = m_{\tilde{h}^2} - m^2_{\tilde{h}}$. As shown in Figure 5.6, for *all* ten cases, the true OTF is within the interval $[m_{\tilde{h}} \pm 2\sigma_{\tilde{h}}]$.

We also had a look on the root mean square error on the OTF. With these 10 averaged OTFs $m_{\tilde{h}}$ and the true OTF $\tilde{h}$, we computed the corresponding square error for all noise realizations: $\epsilon = |m_{\tilde{h}} - \tilde{h}|^2$. This error was then averaged on all noise realizations $\langle\epsilon\rangle_n$, as well as the posterior variance on OTF $\langle\sigma^2_{\tilde{h}}\rangle_n$.

In Figure 5.7, we plot the root mean squared error on the OTF $\sqrt{\langle\epsilon\rangle_n}$ for all frequencies. We notice that $\sqrt{\langle\epsilon\rangle_n}$ is smaller than 1.3 times the posterior standard deviation on OTF averaged on noise realisations $\sqrt{\langle\sigma^2_{\tilde{h}}\rangle_n}$, for all frequencies, which gives us confidence that the uncertainties which we derive from the method are reasonable.

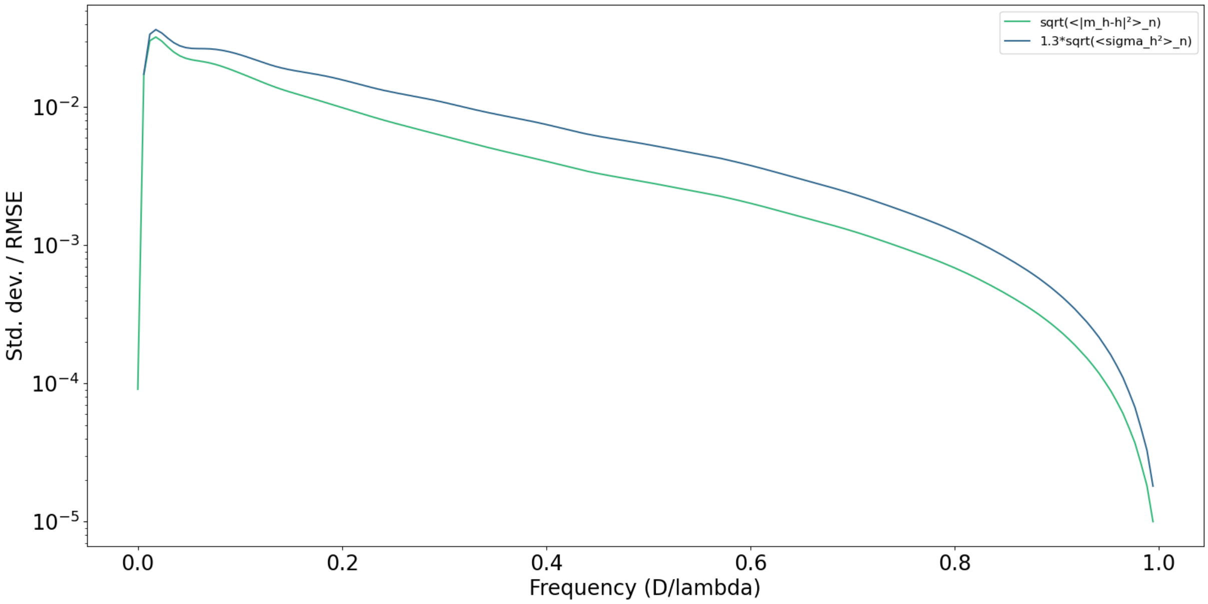

Figure 5.7 – Root mean square error on OTF, averaged on noise realisations, averaged azimuthally (in green) and 1.3 times the estimated variance on OTF on each frequency, averaged on noise realisations, averaged azimuthally (in blue).

5.3 Comparison between Metropolis-Hastings algorithms, within and without Gibbs

5.3.1 Results using the Metropolis-Hastings (without Gibbs) algorithm

Section 5.1 describes results obtained using the Random Walk Metropolis-Hastings-within-Gibbs algorithm. We will here discuss the results for the Random Walk Metropolis-Hastings algorithm. We run this algorithm in the same conditions as mentioned earlier, with the same fixed parameters. The only difference is that the exploration step, namely the standard deviation for the Gaussian proposition, is adjusted in this case to have a similar exploration range as in the Gibbs case. Because we have five parameters to estimate, the standard deviation for each parameter is divided by $\sqrt{5}$. This is done in order to tune the exploration of the posterior distribution to the same “distance” as in the previous case, in order to have an acceptance rate which is close to the optimal one [Gelman, 1997]. As we will see in the following results, this tuning clearly impacts the acceptance rate. The algorithm is also run for 100 000 iterations, to verify convergence.

The computed mean values and standard deviations are reported in Table 5.6. As expected, both are very close to the results obtained with the Metropolis-Hastings-within-Gibbs algorithm given in Table 5.2. Therefore we can state that, for a reasonable tuning of the Gaussian proposition for the Random Walk and a sufficient number of iterations, both algorithm converge towards the same solution.

Parameter	$m \pm \sigma$	True
γ_n (ph^{-2})	$2.61{\times}10^{-4} \pm 7.90{\times}10^{-7}$	$2.62{\times}10^{-4}$
r_0 (m)	0.144 ± 0.007	0.15
v_ϕ(rad^2)	1.18 ± 0.03	1.30
γ_o (ph^{-2})	$2.47{\times}10^{-13} \pm 5.64{\times}10^{-14}$	-
k	0.619 ± 0.411	-

Table 5.6 – Mean value, associated standard deviation and true value, for γ_n, r_0, v_ϕ, γ_o and k for the same conditions as previously, using a Random Walk Metropolis-Hastings algorithm.

5.3.2 Discussion on the computational time for both algorithms

The main difference between both algorithms is then the required time to generate the samples, also related to the number of iterations which are necessary to reach convergence and the acceptance rate for the parameters. Concerning computational time, computing 100 000 iterations requires 1 hour using the Metropolis-Hastings algorithm, 4 hours using the Metropolis-Hastings-within-Gibbs algorithm (the number of hours corresponds to user and system CPU time). The ratio is indeed around 5, because the computational cost is dominated by the computation of the marginal likelihood, which is computed once for a Metropolis-Hasting iteration, and 5 times for a Metropolis-Hastings-within-Gibbs iteration (once per parameter). Additionally, this ratio is less than 5, because the Metropolis-Hastings-within-Gibbs versions carefully computes only the necessary modifications for each parameters and re-uses what remains unchanged (for instance, if a PSF parameter is being modified, the object PSD is not computed redundantly).

In the following Figures 5.8-5.11, we compute the cumulative mean values (meaning, for each iteration the cumulative sum of values, divided by the current number of iterations) $\Sigma_{iter}/iter$ of PSF parameters r_0 and v_ϕ. We discard the first 20 000 iterations, which we consider belonging to the burn-in period (boiling time): thus, the first point on the graph at $iter = 20000$ is the value of the considered parameter at this iteration, the second point is the average value between $iter = 20000$ and $iter = 20001$ and so on. We also plot the mean value m in dashed green line, as well as lower and upper bounds corresponding to a fraction of the standard deviation σ for these parameters (here, half the standard deviation: $[m \pm 0.5\sigma]$). The aim is here to determine the number of iterations that are necessary to reach a certain precision on the estimates, meaning the number of iterations $iter_c$ above which the mean value is within a small (enough) interval: $\Sigma_{iter}/iter \in [m \pm 0.5\sigma]$ for $iter \leq iter_c$.

We notice that the Metropolis-Hastings-within-Gibbs algorithm converges more quickly than the Metropolis-Hastings algorithm, needing less than $iter_c = 30000$ iterations to reach the desired precision level for both parameters, whereas the Metropolis-Hastings algorithm needs more than $iter_c = 50000$ iterations. Finally, in a similar spirit, the acceptance rate for all parameters is slightly under 20% thanks to the change of the standard deviation of the Gaussian proposition (which is divided by $\sqrt{5}$ for all parameters). This acceptance rate is below the one obtained with the Metropolis-Hastings-within-Gibbs algorithm. Without this change of step, the acceptance rate falls to 5%.

As a conclusion, given that the Metropolis-Hastings algorithm needs around 50 000 iterations,

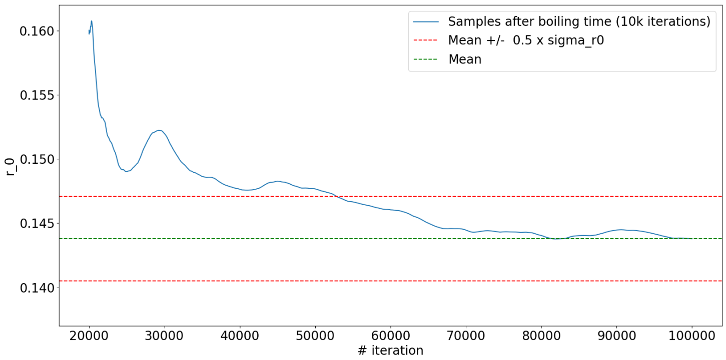

Figure 5.8 – Cumulative mean for r_0 (without boiling time = 20 000 iterations), using Metropolis-Hastings algorithm.

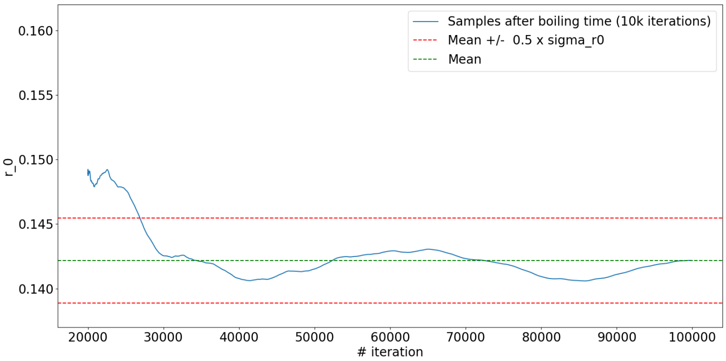

Figure 5.9 – Cumulative mean for r_0 (without boiling time = 20 000 iterations), using Metropolis-Hastings-within-Gibbs algorithm.

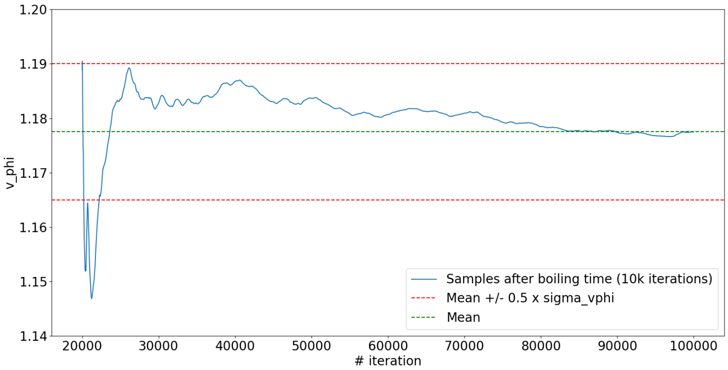

Figure 5.10 – Cumulative mean for v_ϕ (without boiling time = 20 000 iterations), using Metropolis-Hastings algorithm.

each iteration costing around 4 times less than an iteration using the Metropolis-Hastings-within-Gibbs algorithm. It would then be “equivalent” to around 12 000 iterations using the Metropolis-Hastings-within-Gibbs algorithm, whereas using it we effectively need around 30 000 iterations. Therefore, the Metropolis-Hastings algorithm seems more efficient and adapted in our case.

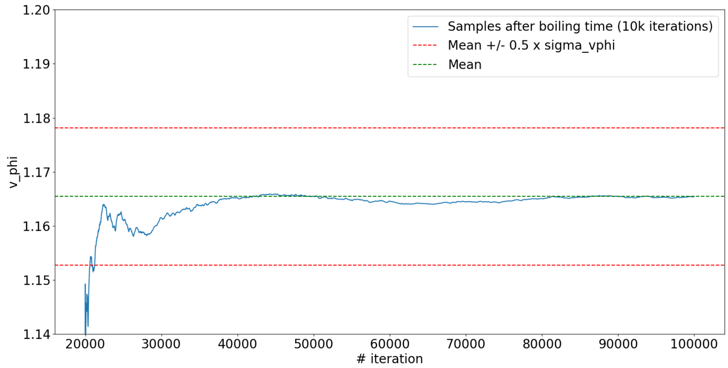

Figure 5.11 – Cumulative mean for v_ϕ (without boiling time = 20 000 iterations), using Metropolis-Hastings-within-Gibbs algorithm.

5.4 Posterior coupling between parameters and impact of hyperparameter p

5.4.1 Marginal posterior scatter plots of the parameters

Sampling the whole posterior distribution, instead of computing a single point of it (for example, the maximum), enables us to study the *a posteriori* coupling of the parameters. In Figures 5.12 and 5.13, we display the scatter graph of the samples, after boiling time, for two different couples of parameters: r_0 and v_ϕ and r_0 and γ_o. Most couples of parameters have a scatter graph similar to Figure 5.12, where the 2D-histogram is rather elliptical and along the axis suggesting that most parameters are not correlated *a posteriori*.

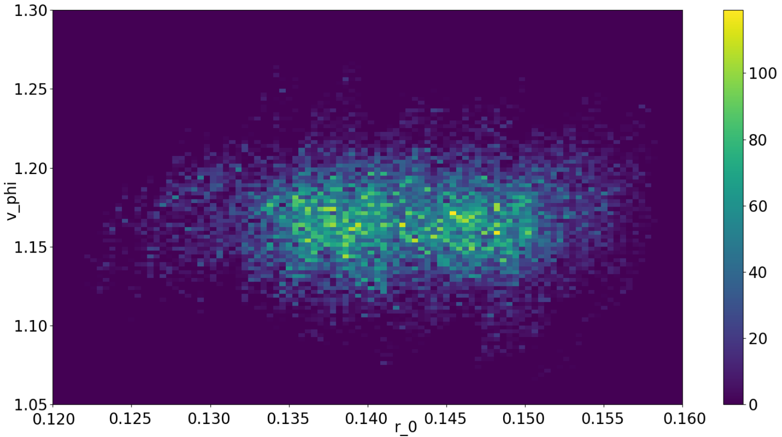

Figure 5.12 – Scatter graph of the samples for r_0 and v_ϕ after boiling time.

The only couple of parameters who does not have an elliptical-like scatter graph, but instead show a strong *a posteriori* correlation, is r_0 and γ_o. We explain this correlation by the fact that as shown in [Fétick, 2020b], r_0 impacts the global level of the OTF wheras γ_o gives the global level of regularization in the object PSD. Therefore, both r_0 and γ_o have a impact on the global level of the image PSD, which is fitted by our method, that explains their strong correlation.

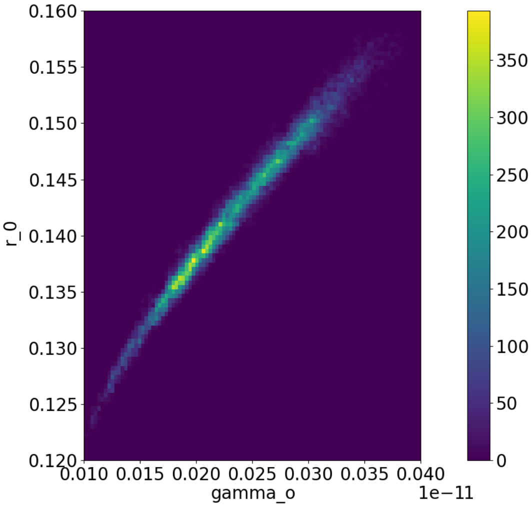

Figure 5.13 – Scatter graph of the samples for r_0 and γ_o after boiling time.

Remark. *The posterior coupling between parameters can also be shown using* ***corner plots****.*

5.4.2 Attempt to estimate p

As said in Chapter 4, attempts in previous works [Fétick, 2020b] to estimate hyperparameter p jointly with the other parameters, in an "unsupervised" mode, has been shown to strongly decrease PSF parameter estimation accuracy. In these works, p was then fixed to a well-fitting empirical value, according to the class of object that was observed (either asteroids, or artificial satellites).

Trying again to estimate p using our new method indeed confirmed this result. We have tested our method in the unsupervised mode, in the exact same conditions than those given in Subection 5.1 (good astronomical site, SPHERE-like AO system, observing Vesta asteroid, with realistic flux and PSF).

We are using here again a Metropolis-Hastings-within-Gibbs algorithm, the tuning of the standard deviation of the Gaussian proposition is taken identical to the previous one given in Table 5.1. The parameters are also initialized the same way as above in Subsection 5.1. The Gibbs sampler is run for 1 000 000 iterations.

For hyperparameter p, it is initialized to $p = 3$, as it is a empirical reasonable value for asteroids. The standard deviation is tuned to 0.1, and the prior interval given was $[2.0 - 4.0]$, which is a pretty large interval for p. The standard deviation tuning corresponds to how sensitive we are on this parameter in magnitude, in practice: as we will see further in Subsection 5.4.3, a change of 0.1 on p has an impact on the results for other parameters, but changing p of 0.01 does not.

In Figure 5.14, we plot the samples chains and the corresponding histograms for all parameters. As one can see, and as said in Subsection 5.4.1, the chains of γ_o and r_0 are strongly correlated, evolving together in a symmetric way, which is consistent with the marginal scatter graph of γ_o and r_0 given in Figure 5.13. Additionally, the chains of p and v_ϕ are also strongly

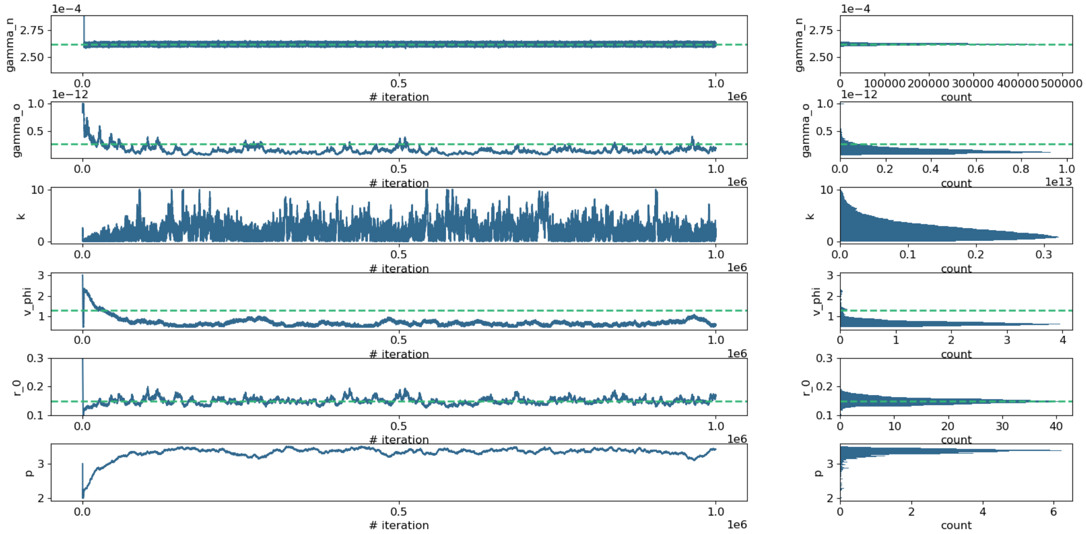

Figure 5.14 – From top to bottom: γ_n, γ_o, k, v_ϕ, r_0, and p. Left: chain of samples for simulated astronomical image. Right: corresponding histogram. True values in dashed line, if available.

correlated, evolving in an anti-symmetric way. The scatter graph of these two parameters is given in Figure 5.15.

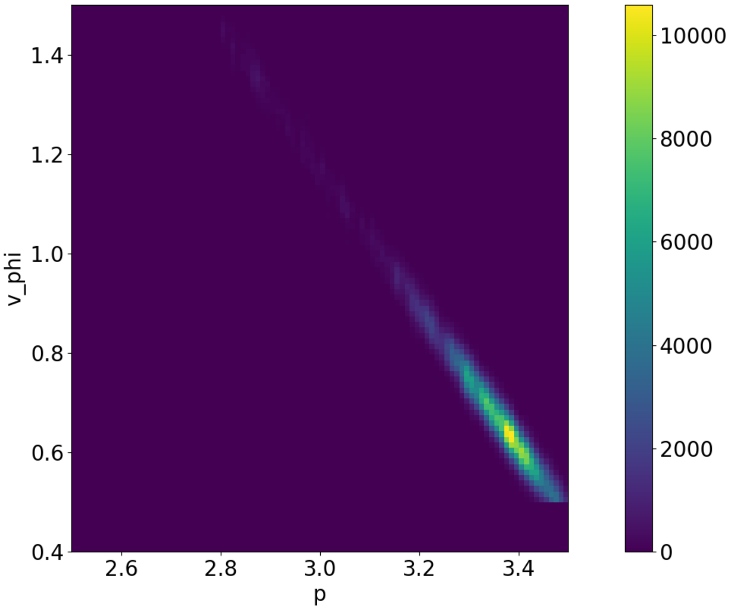

Figure 5.15 – Scatter graph of the samples for v_ϕ and p, after boiling time.

As we can see on the chains as well as on the scatter graph, due to the correlation between the residual phase variance v_ϕ and hyperparameter p, including the estimation of p leads to a strong under-estimation of v_ϕ together with a over-estimation of p (the yellow spot on Figure 5.15).

Similarly to the correlation between r_0 and γ_o that was highlighted in Subsection 5.4.1, we interpret these results as another strong correlation between v_ϕ and the fixed hyperparameter p due to the similar impact they have on the image PSD. Indeed, as shown in [Fétick, 2020b], v_ϕ impacts the slope of the OTF in medium-high frequencies whereas p corresponds to the slope of the object PSD in medium-high frequencies. Therefore, both v_ϕ and p tune the decrease of the image PSD in medium-high frequencies, which can explain their strong posterior correlation.

5.4.3 Changing the tuning of p

Given that the estimation of p is not satisfactory due to the fact that p and v_ϕ have a similar impact on S_i, the goal was then to know how sensitive the tuning of p was. In [Yan, 2022], we have tested our method in the exact same conditions than in Subsection 5.1 for another value for hyperparameter p, tuned slightly differently, towards the "best" value found in [Fétick, 2020b] in the supervised mode $p = 2.91$. Both $p = 2.9$ and $p = 2.91$ were tested with our method, giving the same results.

Parameter	$m \pm \sigma$	True
γ_n (ph^{-2})	$2.62{\times}10^{-4} \pm 7.98{\times}10^{-7}$	$2.62{\times}10^{-4}$
r_0 (m)	0.141 ± 0.006	0.15
$v_\phi(\mathrm{rad}^2)$	1.33 ± 0.02	1.30
γ_o (ph^{-2})	$2.65{\times}10^{-13} \pm 5.39{\times}10^{-14}$	-
k	0.619 ± 0.469	-

Table 5.7 – Mean value, associated standard deviation and true value, for γ_n, r_0, v_ϕ, γ_o and k for simulated astronomical image, with $p = 2.9$.

The posterior mean and standard deviation for each parameter, as well as true values if available, are displayed in Table 5.7. Here again, noise precision γ_n is correctly estimated, to the same precision than in Subsection 5.1 ($\epsilon \leq 0.2\%$). For PSF parameters, we notice that with $p = 2.9$ the true values are with the interval $[m \pm 2\sigma]$, saying that the estimated uncertainties are then satisfactory. Moreover, errors on PSF parameters are now small on both parameters: similarly to the previous case for r_0 ($\epsilon \approx 6\%$, against $\approx 5\%$ for $p = 3.0$) and much smaller for v_ϕ ($\epsilon \approx 2\%$, against $\approx 10\%$ for $p = 3.0$).

Here again, the correlation between v_ϕ and p explains the need of a fine tuning of p in the mostly unsupervised mode, and mostly explains the error (which is mainly a bias) on v_ϕ. However the differences on the restored image, as displayed in Figures 5.16 and 5.17, are quite small, at most around ten times smaller than the global image level.

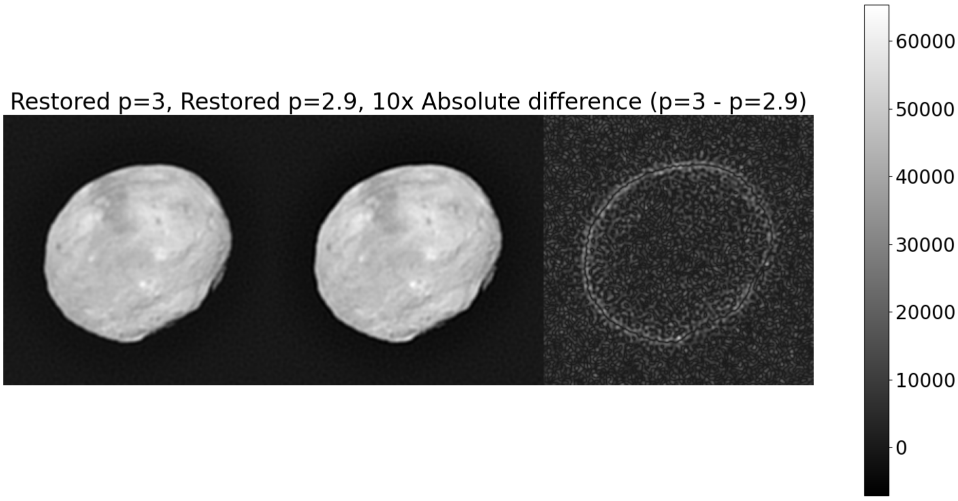

Figure 5.16 – Left: restored image, fixing $p = 3$. Center: restored image, fixing $p = 2.9$. Right: Ten times the absolute difference between the two first images.

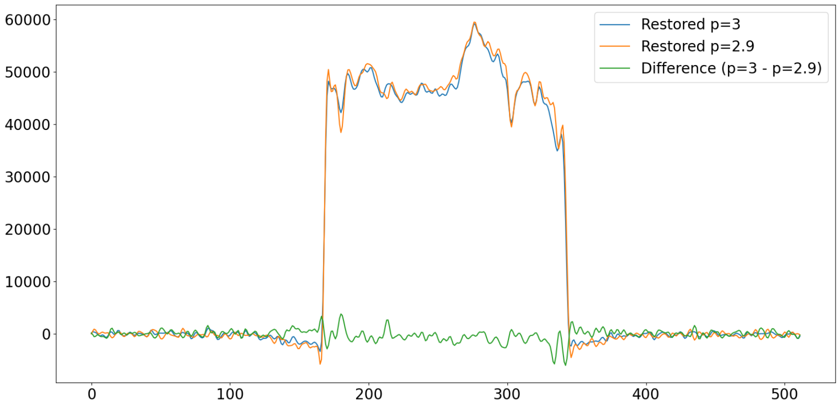

Figure 5.17 – Horizontal sectional plot of the restored images (at $N/2$), fixing $p = 3$ and $p = 2.9$, and their difference.

5.5 Results on simulated satellite data

5.5.1 Simulation conditions

We now show results for a simulated satellite image, using as the true object a synthetic view of the SPOT satellite on a dark background of size 512×512 pixels [Mugnier, 2001]. We simulate its observation using the ODISSEE AO system at OCA [Petit, 2020], and with true PSF parameters $r_0 = 0.10\,\mathrm{m}$ and $v_\phi = 1.85\,\mathrm{rad}^2$, at the imaging wavelength $\lambda = 850\,\mathrm{nm}$, which corresponds to a stronger turbulence, and to a more modest correction than for the astronomical simulation because of a less complex AO system. The noise is taken as zero-mean, additive, white and Gaussian, and its variance is taken equal to the mean value of the object. Here, the mean flux is 10^4 photons per image pixel, corresponding to a somewhat optimistic value. The pixel sampling is close to the Shannon-Nyquist criterion, with slightly more than 2 pixels per λ/D.

The object PSD power p is fixed to an empirical standard value for satellites $p = 2.6$, to fit

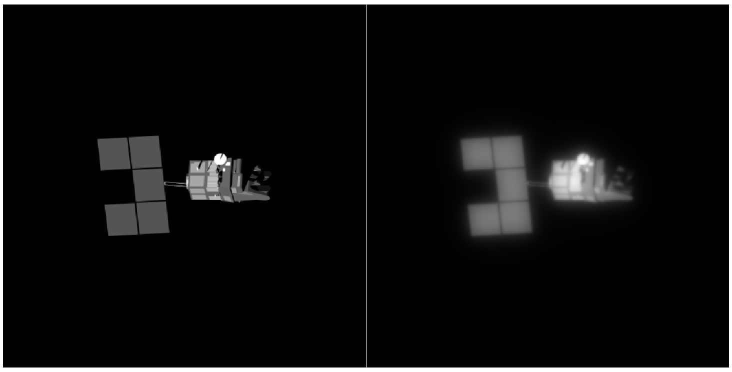

Figure 5.18 – Left: Synthetic view of SPOT (true object), of size 512 × 512. Right: simulated image, with true parameters $r_0 = 0.10\,\mathrm{m}$, $v_\phi = 1.85\,\mathrm{rad}^2$ and $\gamma_n = 1.00\ 10^{-4}\,\mathrm{ph}^{-2}$.

the empirical object PSD. The Gibbs sampler is run for 100 000 iterations. In Figure 5.19, we plot the sample chains and the corresponding histograms for γ_n, r_0 and v_ϕ. Their true values are also represented.

5.5.2 Chains of random samples for the estimated parameters

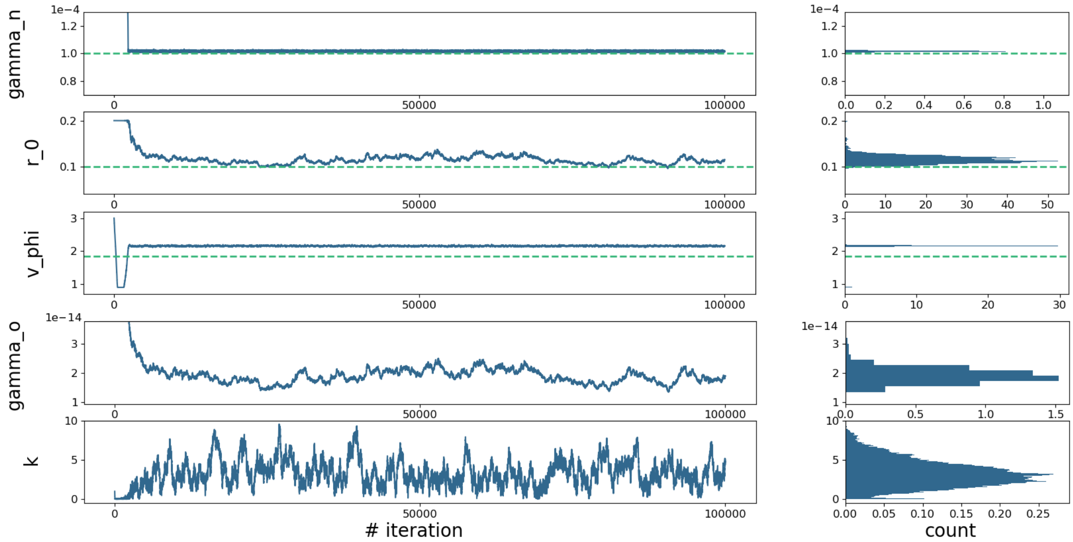

Figure 5.19 – From top to bottom: γ_n, r_0, v_ϕ, γ_o, k. Left: chains of samples for simulated satellite image. Right: corresponding histogram. True values in dashed line.

Similarly to the previous simulations, the empirical mean values m and standard deviations σ of the posterior distribution for each parameter are displayed in Table 5.8. The noise precision γ_n as well as PSF parameters r_0 and v_ϕ are relatively well estimated, with an error of respectively 2%, 14% and 17%. These results are very close to those obtained with P-AMIRAL: for similar conditions, the estimated PSF parameters are $r_0 = 0.112\,\mathrm{m}$ and $v_\phi = 2.16\,\mathrm{rad}^2$.

Parameter	$m \pm \sigma$	True
γ_n (ph^{-2})	1.02 $\times 10^{-4}$ $\pm$ 3.63$\times 10^{-7}$	1.00 $\times 10^{-4}$
r_0 (m)	0.114 $\pm$ 0.008	0.10
v_ϕ(rad^2)	2.16 $\pm$ 0.01	1.85
γ_o (ph^{-2})	1.90 $\times 10^{-14}$ $\pm$ 2.37 $\times 10^{-15}$	-
k	3.14 $\pm$ 1.45	-

Table 5.8 – Mean value, associated standard deviation and true value, for γ_n, r_0, v_ϕ, γ_o and k, for simulated satellite image.

5.5.3 Quality of estimation of the PSF

We notice that the error on PSF parameters is greater for the satellite observation than for the astronomical observation. Our interpretation of these results is that it is due to the spectrum of the satellite object which is less isotropic than Vesta, and therefore does not fit our isotropic power spectral density model as well. This interpretation was checked by complementary simulations: indeed, if we simulate the observation of Vesta instead of SPOT, in the exact same conditions as for satellite observation except that hyperparameter p is re-adjusted to a standard value for asteroids $p = 3$, the errors on PSF parameters are very close to the results obtained in Subsection 5.1, suggesting that the greater error on the PSF comes from the object and not from the simulated AO system or PSF.

We also compare the resulting estimated OTF to the true OTF in Figure 5.20. Here again, as discussed previously, we notice that the errors on both parameters partially compensate, as a result the normalized RMSE for the OTF is quite low (around 8%).

5.5.4 Results on the derived uncertainties

Concerning uncertainties on the PSF parameters, similarly to previous asteroid case, the posterior standard deviation for r_0 is a good uncertainty prediction for this parameter as the true value is within the interval $[m_{r_0} \pm 2\sigma_{r_0}]$. On the contrary, σ_{v_ϕ} is small, giving an under-estimated uncertainty. The reasons for this under-estimation are being investigated, though it should be linked to the more difficult observation conditions simulated here. Additionally, the previous discussion about the correlation between parameters p and v_ϕ in Subsection 5.4.3 is still valid and further work on tuning hyperparameter p for satellite observation should be done.

Concerning the uncertainties on the OTF, we notice again that even though the uncertainties on PSF parameters are under-estimated, the uncertainties on the OTF itself are quite satisfactory as the true OTF is within the interval $[m_{\tilde{h}} \pm 3\sigma_{\tilde{h}}]$.

5.5.5 Results on object and image PSDs

Additionally, the estimations result in a good image PSD fitting, averaged azimuthally, as shown in Figure 5.21. Moreover, as displayed in Figure 5.21, the object PSD model visually fits well the empirical object PSD, averaged azimuthally.

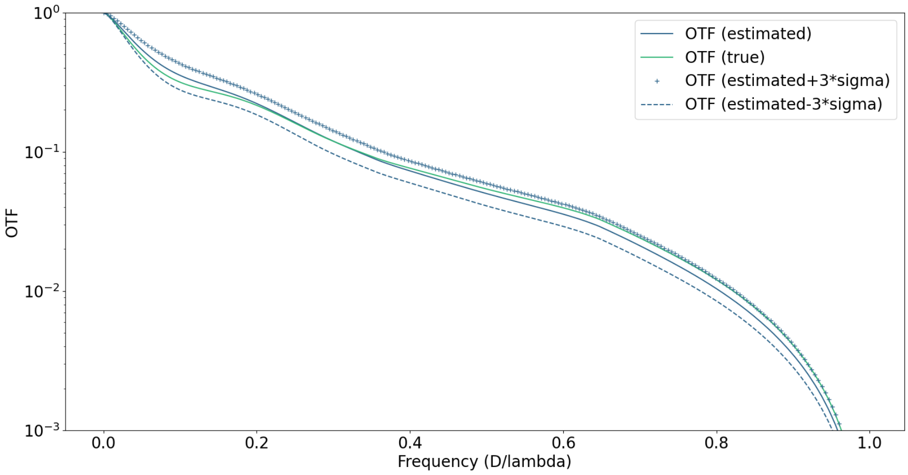

Figure 5.20 – True (in green) and estimated (in blue) OTF for simulated satellite image, including computed uncertainties (in blue, + and - for upper and lower uncertainty bounds).

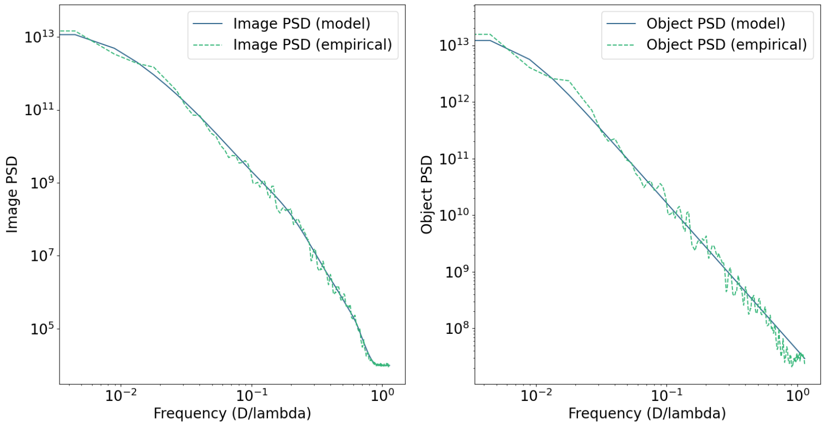

Figure 5.21 – PSDs for simulated satellite image. Model in solid line, and empirical PSD averaged azimuthally in dashed line. Left: image PSD. Right: object PSD.

5.5.6 Results on restored image

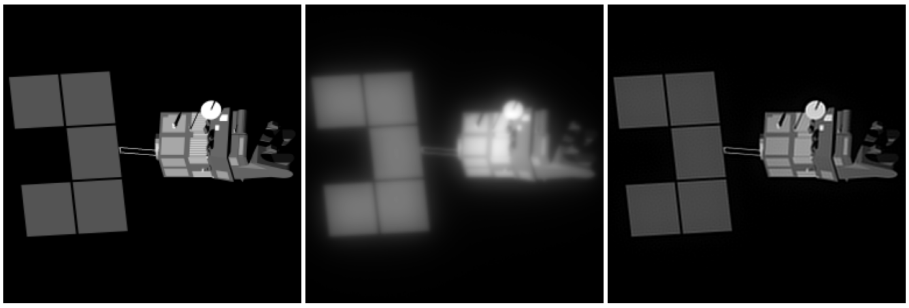

Figure 5.22 – Left and center: True object and image for simulated satellite observation, 256×256 cropped from 512 × 512 (Figure 5.18). Right: restored object from the estimated PSF and PSD parameters (using a L2-norm regularization, with positivity constraint), also cropped.

Finally, Figure 5.22 shows results from the restoration of the image in Figure 5.18 (right) using the estimated OTF. We notice that details of the satellite surface are restored, especially on its main part and its connection with the panel.

Chapter 6

Results of the MMSE estimator on experimental data

Introduction

In this chapter we provide the results obtained using the MMSE estimator for experimental data, for which we remind some elements of the application context, before discussing both the estimated parameters and the restored images. The first section gives these elements on a astronomical image, and the second section on a satellite image.

Contents

6.1 Results on experimental astronomical data

6.1.1 Description of the setup and the data

After testing our method on both astronomical and satellite simulated data, therefore for different turbulence conditions and AO systems, we apply it to experimental images.

Here we process the experimental image of Vesta [Fétick, 2019b] taken by SPHERE/Zimpol during the *Large Program* (ID 199.C-0074), described in Chapter 3.

As estimation method, we are using the MMSE estimator using the Metropolis-Hastings-within-Gibbs algorithm, as described in Chapter 4. As discussed previously, this method is used in a mostly unsupervised mode and hyperparameter p is again fixed to $p = 3$, a reasonable value for asteroids. We run the Gibbs sampler for 500 000 iterations, to verify convergence of the random sample chains.

6.1.2 Results on image PSD, for experimental astronomical image

A good sanity check before looking at the results on the parameters as well as on the restored image is to check that the image PSD model and the empirical image PSD for Vesta fit correctly. We plot both image PSDs (model and empirical ones) in Figure 6.1.

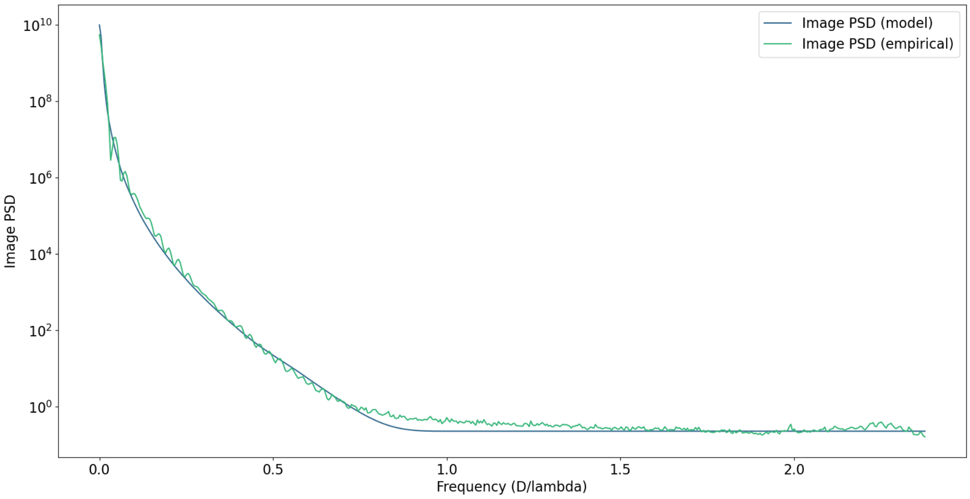

Figure 6.1 – PSD model and Vesta empirical image PSD averaged azimuthally, in dashed line.

As displayed, the image PSD model and the empirical image PSD fit correctly, especially at low and medium frequencies, where signal dominates noise. For high frequencies, where the noise is dominant, we see that the noise floor is not flat (whereas we model the noise as white), and believe it may be due to the data reduction by SPHERE/ZIMPOL's pipeline. This pipeline is known for being bugged, to the point that some researchers have developed an alternative pipeline, which we could not use for the moment.

6.1.3 Results on parameters, for experimental astronomical image

Now that the image PSD fitting was checked, we can look at the parameters' sample chains and corresponding histograms. These are displayed in Figure 6.2. Additionally, results obtained for the five estimated parameters (mean ± standard deviation, as well as acceptance rate τ_{acc}) are summed up in Table 6.1.

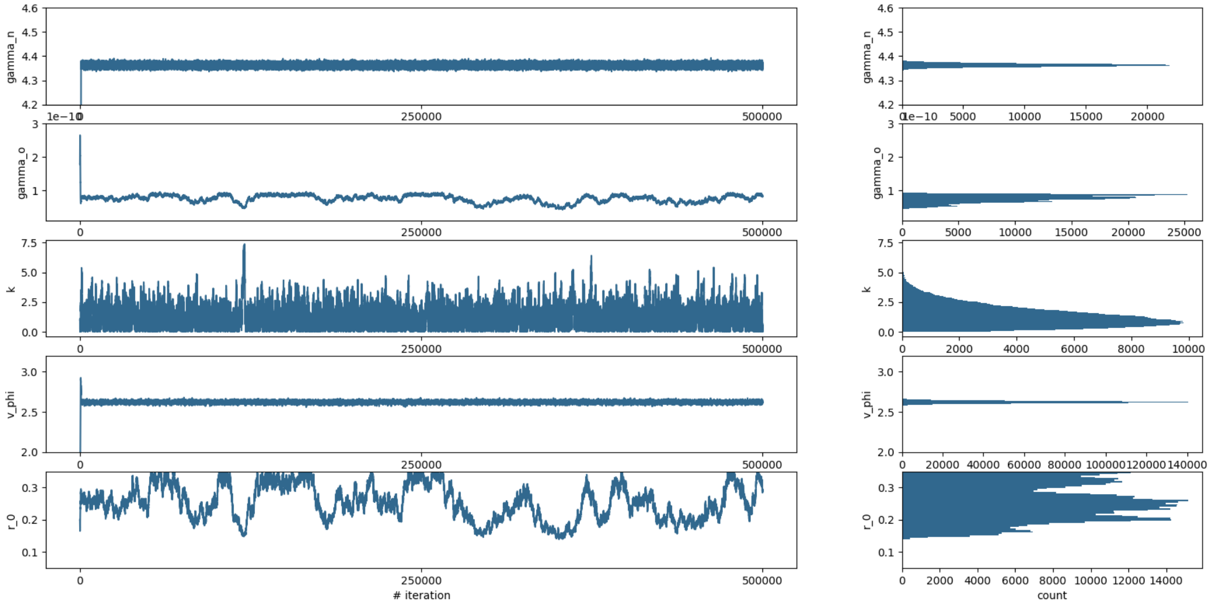

Figure 6.2 – From top to bottom: γ_n, γ_o, k, v_ϕ and r_0. Left: chain of samples for experimental astronomical image. Right: corresponding histogram.

Parameter	$m \pm \sigma$	τ_{acc}
γ_n (ph^{-2})	$4.36 \pm 6.43{\times}10^{-3}$	48%
r_0 (m)	0.254 ± 0.054	83%
v_ϕ(rad^2)	2.62 ± 0.01	43%
γ_o (ph^{-2})	$7.51{\times}10^{-11} \pm 1.07{\times}10^{-11}$	27%
k	1.33 ± 0.87	96%

Table 6.1 – Mean value and associated standard deviation for γ_n, r_0, v_ϕ, γ_o and k for experimental Vesta image, with $p = 3$.

Concerning the chains, we notice that the convergence seems reached quickly for most parameters. However, the uncertainties on the parameters is also larger, especially for r_0 which has a rather wide, multimodal histogram. If we focus on the PSF parameters, we notice that the estimated residual phase variance v_ϕ is close to the value obtained with P-AMIRAL: $v_\phi = 2.62\,\mathrm{rad}^2$ instead of $2.78\,\mathrm{rad}^2$ with P-AMIRAL [Fétick, 2020b] for the same conditions. However, the estimated r_0 is slightly under the one estimated by P-AMIRAL: $r_0 = 0.254\,\mathrm{m}$ instead of $0.32\,\mathrm{m}$ with P-AMIRAL [Fétick, 2020b]. Besides, one can note that the value estimated by P-AMIRAL corresponds to one of the local maxima of the histogram on Figure 6.2. After a check on the known statistics on r_0 for the ZIMPOL instrument [Fétick, 2019a], the newly estimated r_0 seems more likely than the one estimated by P-AMIRAL according to them. Indeed, as said in Chapter 3, for the ZIMPOL instrument, estimated r_0 values by telemetry data from SPARTA or using the MASS/DIMM instruments are respectively within $[15; 32]$ cm and $[9; 23]$ cm [Fétick, 2020a]. Additionally, as one can notice on Figure 6.2 that there is another (smaller) peak in the histogram corresponding to r_0 around $r_0 = 0.32\,\mathrm{m}$, which may thus correspond to a local minimum of the criterion.

If we look now at the hyperparameters, γ_n has a higher value than what we expected for it. Indeed, in Chapter 5, given the considered flux levels we approximated the noise as Gaussian with its variance equal to the mean of the image, which we believe is close to the mean of the object (again due to the small uncertainty on the mean object as said in Chapter 4). Therefore, we were expecting γ_n to be close to $\frac{1}{m_i} = 0.014\,\mathrm{ph}^{-2}$, whereas we have $\gamma_n = 4.36\,\mathrm{ph}^{-2}$.

Additionally, as γ_o gives the global object PSD level, we were expecting it to be close to the sum of the object, taking into account the normalization factor for the Fourier transform with respect to the Parseval's theorem. It means that, in practice here, we were expecting γ_o to be close to $\frac{1}{m_i^2 N} = 1.79 \times 10^{-10}\,\mathrm{ph}^{-2}$, with N the image size in pixels, and we have a relatively close value $\gamma_o = 7.77 \times 10^{-11}\,\mathrm{ph}^{-2}$. Then, the ratio $\mu = \gamma_o/\gamma_n$ was expected to be close to $\frac{1}{m_i N} = 1.31 \times 10^{-8}$, whereas we have $\mu = 2.12 \times 10^{-11}$. However, this value for μ is close to the one obtained with P-AMIRAL: 1.75×10^{-11} [Fétick, 2020b] and these parameters result in a good image PSD fitting as seen in Figure 6.1. Thus we believe that the instrument gain we took into account was incorrect/partial, and these values deserve further investigation for the moment.

Concerning the acceptance rate τ_{acc}, it seems slightly higher than the optimum value [Gelman, 1997] for γ_n, v_ϕ and γ_o, not surprisingly high for k as it was already the case in simulated results, but quite high for r_0 which may require a slightly finer tuning.

Finally, concerning the uncertainties on the PSF parameters we derive from the samples, we notice that the uncertainty on v_ϕ stays to the same level than for simulated images around $0.01\,\mathrm{rad}^2$, whereas the uncertainty on r_0 has strongly increased (from 1 cm on simulated data to 5 cm here), which is probably partly due to the potential local minima mentioned above, whereas fixing p still pretty much sets v_ϕ due to their correlation discussed in Chapter 5.

6.1.4 Results on restored image, for experimental astronomical image

Data and restored object are shown in Figure 6.3. We recognize the same surface features as from the synthetic view of Vesta (built by OASIS after NASA's Dawn spacecraft data) in Figure 5.1. In this experimental case, the bright edge corona starts to appear (on the left side), and the image is slightly granular. This may be due to a slight over-deconvolution *i.e.* to a slight under-estimation of the OTF. This bright edge also appears on the image restored by P-AMIRAL, and is a typical feature we observe on images which are over-deconvolved.

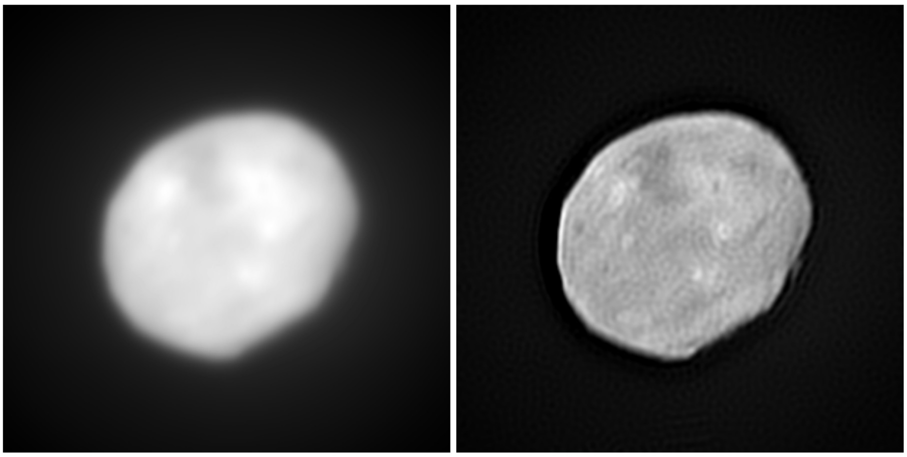

Figure 6.3 – Left: Vesta observed by SPHERE/Zimpol on the European Very Large Telescope (VLT) in Chile [Fétick, 2019b]. Right: restored object with the estimated PSF.

6.2 Results on experimental satellite data

6.2.1 Description of the setup and the data

Finally, we test our method on an experimental image of the ENVISAT satellite. ENVISAT is a large satellite (26 m-long) evolving at around 800 km, it was launched in 2002 and is no more controlled since 2012, which makes it the largest space debris in orbit today.

Figure 6.4 – CAD image of ENVISAT provided by ESA

The image we are processing here was taken at the Observatoire de la Côte d'Azur, using the 1.5 m MéO telescope with ONERA's ODISSEE AO system [Petit, 2020], which parameters are given in Chapter 1, in Table 1.3.

With these parameters, one can compute ENVISAT's size in arcsec and compare it to a resolution element λ/D: ENVISAT is approximately $26\,\mathrm{m}/800\,\mathrm{km} \times 3600 \times 180/\pi \approx 6.7\,\mathrm{rad}$ at the zenith, and $\lambda/D = 850\,\mathrm{nm}/1.5\,\mathrm{m} \times 3600 \times 180/\pi \approx 0.12\,\mathrm{rad}$. ENVISAT is, at the zenith, approximately $6.7/0.12\lambda/D \approx 127\,\mathrm{pix}$.

For this study, we fix hyperparameter p to a reasonable value for satellites ($p = 2.5$), and run here again the Gibbs sampler for 500 000 iterations.

6.2.2 Results on image PSD, for experimental satellite image

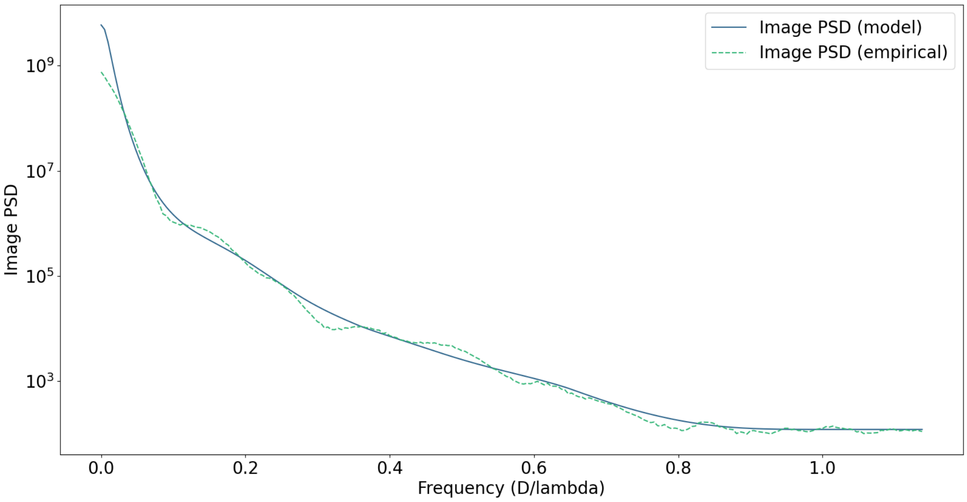

Figure 6.5 – PSD model and Envisat empirical image PSD averaged azimuthally, in dashed line.

As in the experimental astronomical case, we start by comparing the empirical image PSD to our model in Figure 6.5. We have a globally good image PSD fitting of the image PSD model to the empirical Envisat image PSD. The oscillations of the empirical image PSD are likely to come from oscillations of the OTF, which are consistent with the exposure time (≈ 500 ms) which is short with respect to turbulence residuals averaging, and constitutes a deviation to the infinite exposure assumption of our AO-corrected PSF model (but as explained in Chapter 1, this small duration is constrained by the time scale of evolution of satellite attitude and rotation). These oscillations might also come partly from the spectrum of the object itself, showing one limitation of our current method. Indeed, the current object PSD model being azimuthally symmetric, it is not adapted for elliptical or elongated objects (meaning, with an asymmetric PSD).

6.2.3 Results on parameters, for experimental satellite image

Parameter	$m \pm \sigma$	τ_{acc}
γ_n (ph^{-2})	$8.21\times10^{-3} \pm 3.16\times10^{-5}$	22%
r_0 (m)	0.08 ± 0.011	20%
v_ϕ(rad^2)	0.89 ± 0.01	45%
γ_o (ph^{-2})	$2.21\times10^{-11} \pm 8.90\times10^{-12}$	47%
k	9.00 ± 0.94	96%

Table 6.2 – Mean value and associated standard deviation for γ_n, r_0, v_ϕ, γ_o and k for experimental ENVISAT image, with $p = 2.5$.

After the sanity check on the image PSD, we now have a look at the estimated parameters and the derived uncertainties. Our results on the estimated parameters, the uncertainties and the acceptation rates are given in Table 6.2.

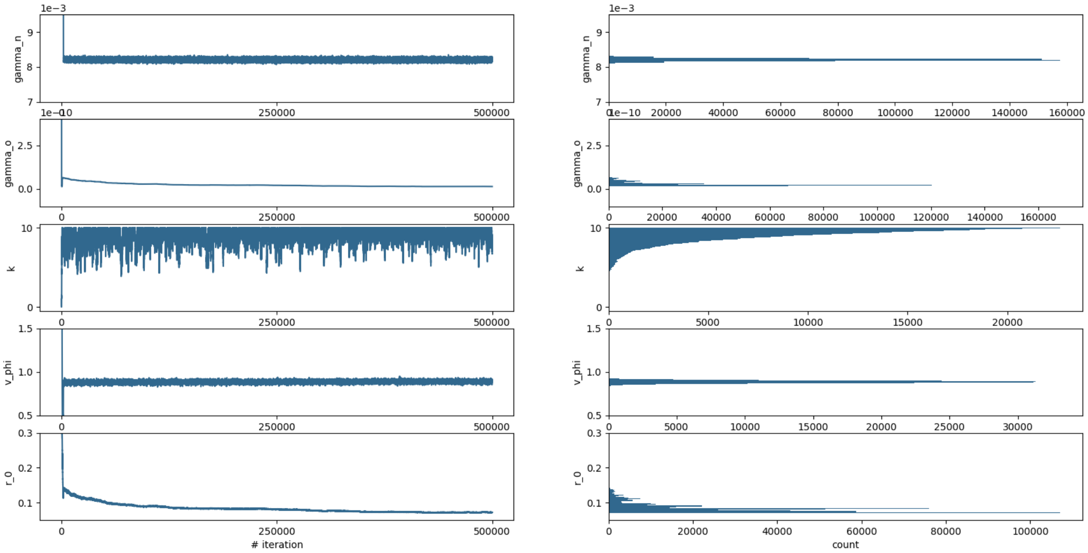

Figure 6.6 – From top to bottom: γ_n, γ_o, k, v_ϕ and r_0. Left: chain of samples for experimental satellite image. Right: corresponding histogram.

Results obtained on PSF parameters are slightly different from the results obtained with P-AMIRAL which we ran for the same conditions: $r_0 = 0.08\,\mathrm{m}$ here instead of $0.06\,\mathrm{m}$ and $v_\phi = 0.88\,\mathrm{rad}^2$ here instead of $0.93\,\mathrm{rad}^2$. Concerning the chains, just as in the experimental astronomical case, we notice that the convergence of the parameters are satisfactory except for r_0 which seems to require way more iterations to converge in the experimental case, compared to simulations in Chapter 5. Finally, determining which method is the closest to the "truth" would require to study furthermore additional data giving information over r_0 and/or v_ϕ such as telemetry data.

The acceptance rate here seems correct for all parameters except k, close to the ones obtained in the simulated case, giving us confidence in the parameters' step tuning. Finally, the derived uncertainties on PSF parameters also seem to be in a reasonable range: uncertainties on v_ϕ are small again due to the p and v_ϕ correlation, and we are expecting a precision close to $1\,\mathrm{cm}$ for r_0.

6.2.4 Results on restored image, for experimental satellite image

Finally, concerning the restored image, as shown in Figure 6.7, we retrieve some elements of the satellite, and as we can check on the CAD model of Envisat in Figure 6.4, the bright spots we obtain on the restored image indeed correspond to instruments and antennas on its surface.

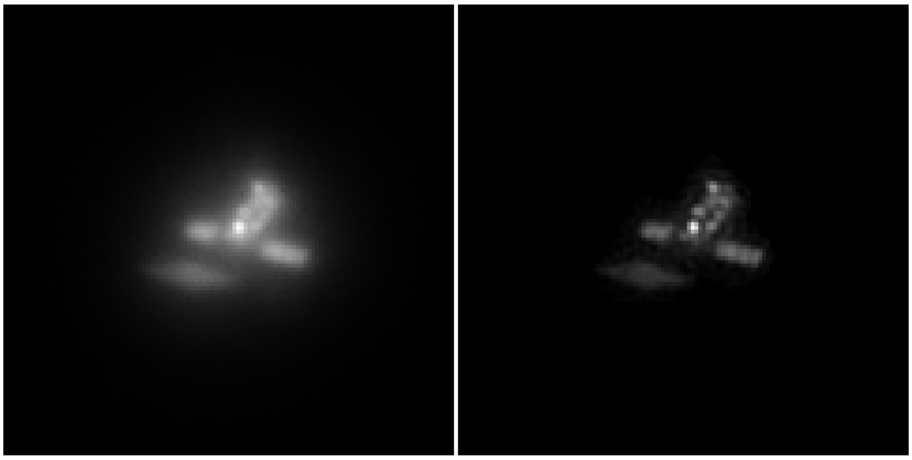

Figure 6.7 – Left: Envisat observed by ODISSEE at the OCA [Petit, 2020]. Right: restored object.

Chapter 7

Adding prior information on the object: support constraint

Introduction

In this chapter, we study the impact of a support constraint as an additional prior information on the object. In the first section, we discuss the different possibilities to model this information. However, any option chosen in Section 1 requires the computation of truncated matrices TAT^t which we develop in Section 2. Finally, Section 3 gives results brought by adding a known support, comparatively to the classic marginal MAP estimation done by P-AMIRAL, on the estimated parameters.

Contents

Adding information on the object, by means of constraints on it, can be particularly relevant and helpful for its estimation. In our case, we could add for instance a positivity constraint (the sought object only having positive-valued pixels) or a support constraint (the object is on a dark background, where pixels are equal to zero). Concerning the latter constraint, this support can be provided elsewhere (from additional information on the object, or based on another object restoration prior to the one we are performing), or can also be estimated along with the other parameters. In this chapter, we focus on the impact of a support constraint when the support is given, and describe several options in order to incorporate the estimation of the support to our framework.

7.1 Modelling the support constraint

7.1.1 Rewriting the imaging model

We still consider the same imaging model given in Equation 1.13, reminded here:

$$\boldsymbol{i} = \boldsymbol{H}\boldsymbol{o} + \boldsymbol{n}$$

We consider that our true object $\boldsymbol{o}$ comes from the truncation of a "full" object which is then put on a dark background. The "full" object which has the same number of pixels N as $\boldsymbol{o}$, is named hereafter $\boldsymbol{o_F}$, and its truncated version which only has M pixels (corresponding to the number of pixels in the support) is named $\bar{\boldsymbol{o}}$. We can also write the rectangular truncation $M \times N$ matrix $\boldsymbol{T}$, naturally giving $\boldsymbol{T^t}$ the zero-padding matrix. The size of this matrix, and more globally the number of pixels in the estimated support M depends on the labels, named ℓ hereafter. The labels correspond to a categorization of the pixels, in our case there are two different categories: either in the support, or out of it. They are either known or estimated with the other parameters, by the method. We can then write that:

$$\boldsymbol{i} = \boldsymbol{H}\boldsymbol{T}_{\boldsymbol{\ell}}^{\boldsymbol{t}}\bar{\boldsymbol{o}} + \boldsymbol{n} \tag{7.1}$$

$$\boldsymbol{i} = \boldsymbol{H}\boldsymbol{T}_{\boldsymbol{\ell}}^{\boldsymbol{t}}\boldsymbol{T}_{\boldsymbol{\ell}}\boldsymbol{o_F} + \boldsymbol{n} \tag{7.2}$$

7.1.2 Marginal approach

One possible approach to take into account the support is to adapt the marginal approach described in Chapter 2 and used throughout the manuscript, but this time using Equation 7.2. The full object $\boldsymbol{o_F}$ is still the Gaussian field considered previously (from which we are "cutting" a piece in the middle, corresponding to the truncated object $\bar{\boldsymbol{o}}$).

Remark. *If Y is an affine transformation of $X \sim \mathcal{N}(\mu, \Sigma)$: $Y = c + BX$, then Y has a multivariate normal distribution with mean expected value $c + B\mu$ and covariance $B\Sigma B^t$.*

Given that the full object is Gaussian, $\boldsymbol{o} = \boldsymbol{H}\boldsymbol{T^t}\boldsymbol{T}\boldsymbol{o_F}$ is also Gaussian. Noise being Gaussian, and taken independent from the object, the support and the PSF, $\boldsymbol{i} = \boldsymbol{H}\boldsymbol{T^t}\boldsymbol{T}\boldsymbol{o_F} + \boldsymbol{n}$ is Gaussian, therefore the marginal likelihood writes:

$$p(\boldsymbol{i}|\gamma_n, r_0, v_\phi, \gamma_o, k; \ell) = (2\pi)^{-N/2} \det(\boldsymbol{R_i})^{-1/2} \exp(-\frac{1}{2}(\boldsymbol{i} - \boldsymbol{m_i})^t \boldsymbol{R_i}^{-1}(\boldsymbol{i} - \boldsymbol{m_i})) \tag{7.3}$$

with $\boldsymbol{m_i} = \boldsymbol{HT^tTm_{o_F}}$ and $\boldsymbol{R_i} = \boldsymbol{HT^tTR_{o_F}T^tTH^t} + \boldsymbol{R_n}$.

One can notice that marginalizing over the full object $\boldsymbol{o_F}$, or the truncated object $\boldsymbol{\bar{o}}$ gives the same marginalized likelihood, so leads to the same approach.

For the hyperparameters and PSF parameters, given that we do not have additional knowledge on them, we choose to take uniform priors for γ_n, γ_o, k, r_0 and v_ϕ.

The posterior distribution in the marginal case for any unknown parameter writes, using the marginal likelihood and the corresponding prior, as:

$$f(\theta|\boldsymbol{i}, \boldsymbol{\theta_{\neq\theta}}; \ell) \propto f(\theta) \det(\boldsymbol{R_i})^{-1/2} \exp(-\frac{1}{2}(\boldsymbol{i} - \boldsymbol{m_i})^t \boldsymbol{R_i}^{-1}(\boldsymbol{i} - \boldsymbol{m_i})) \tag{7.4}$$

Unlike in previous chapters, here the marginal likelihood can not be computed simply using FFTs and requires to manipulate the determinant and the inverse of a $N \times N$ matrix $\boldsymbol{R_i}$, which is something we want to avoid. Thus, we are using the Woodbury matrix identity, to rewrite $\boldsymbol{R_i}^{-1}$ while only inverting smaller $(M \times M)$ matrices:

$$\begin{aligned}\boldsymbol{R_i}^{-1} &= (\boldsymbol{HT^tTR_{o_F}T^tTH^t} + \gamma_n^{-1}\boldsymbol{I_N})^{-1} \\ &= \gamma_n\boldsymbol{I_N} - \gamma_n^2\boldsymbol{HT^t}((\boldsymbol{TR_{o_F}T^t})^{-1} + \gamma_n\boldsymbol{TH^tHT^t})^{-1}\boldsymbol{TH^t}\end{aligned} \tag{7.5}$$

Moreover, we can also rewrite $\det(\boldsymbol{R_i})$ using the matrix determinant lemma, in order to compute only determinants of $M \times M$ matrices:

$$\begin{aligned}\det(\boldsymbol{R_i}) &= \det(\boldsymbol{HT^tTR_{o_F}T^tTH^t} + \gamma_n^{-1}\boldsymbol{I_N}) \\ &= \gamma_n^{-N}\det(\gamma_n\boldsymbol{HT^tTR_{o_F}T^tTH^t} + \boldsymbol{I_N}) \\ &= \gamma_n^{-N}\det(\gamma_n\boldsymbol{TR_{o_F}T^tTH^tHT^t} + \boldsymbol{I_M})\end{aligned} \tag{7.6}$$

7.2 Computing TAT^t matrices: different implementations

As shown in the previous section, we need to compute $M \times M$ matrices resulting from the truncation of a $N \times N$ matrix $\boldsymbol{A}$: $\boldsymbol{TAT^t}$. Indeed, the computation of $\boldsymbol{TAT^t}$ matrices, either $\boldsymbol{TH^tHT^t}$ or $\boldsymbol{TR_{o_F}T^t}$, is particularly crucial as they have to be computed and updated often (either with the PSF parameters update, the object PSD parameters update or the labels update), in order to draw the samples of the parameters. This section describes the implementations of the $\boldsymbol{TAT^t}$-form computation, aiming at gaining time over this crucial element. They are provided in Python in the following section, but can be easily adapted to any other programming language.

7.2.1 Naive versions, in the spatial and the Fourier domain

The naive implementation of such a matrix is to compute it element-wise. This is done by looping twice over the pixels in the support: we compute for each elements p and q in the support, the corresponding element of the matrix $\mathbb{1}_p^t \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^t \mathbb{1}_q$. This can be done in both spatial and Fourier domain.

However, in the spatial domain, it implies to have access to $\boldsymbol{A}$ which is a big matrix $(N \times N)$ which is not always possible. In the case when $\boldsymbol{A}$ can be splitted into two matrices, for instance for $\boldsymbol{A} = \boldsymbol{H}^t\boldsymbol{H}$, computing $\mathbb{1}_p^t \boldsymbol{T}\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{T}^t \mathbb{1}_q$ is feasible: $\mathbb{1}_p^t \boldsymbol{T}\boldsymbol{H}^t$ corresponds then to the PSF, shifted so that it is centered now on the pixel p, therefore $\mathbb{1}_p^t \boldsymbol{T}\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{T}^t \mathbb{1}_q$ computes the (p,q) element of the autocorrelation of the PSF.

```python
import numpy as np

def compute_TATt_spatial(psf,M,sup_coord):
    #computes the autocorrelation element-wise from the PSF
    #INPUTS: PSF, M=number of pixels in the support, sup_coord=(x,y) coordinates of pixels within the support
    #REMARK: M can also be computed from sup_coord as M = len(sup_coord[0]).
    #OUTPUT: TATt

    tatt = np.zeros((M,M))
    for p in range(M):
        xp = sup_coord[0][p]
        yp = sup_coord[1][p]
        psfp = np.roll(np.roll(psf,yp,axis=0),xp,axis=1)
        for q in range(M):
            xq = sup_coord[0][q]
            yq = sup_coord[1][q]
            psfq = np.roll(np.roll(psf,yq,axis=0),xq,axis=1)
            tatt[p,q] = np.sum(psfp*psfq)
    return tatt
```

Listing 7.1 – Python implementation of TATt computation in the spatial domain

To switch from the spatial to the Fourier domain, one can rewrite the computed element as $\mathbb{1}_p^t \boldsymbol{T}\boldsymbol{F}^t\boldsymbol{F}\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{F}^t\boldsymbol{F}\boldsymbol{T}^t \mathbb{1}_q$. This spatial shifting $\mathbb{1}_q$ corresponds to a phase shifting in the Fourier domain $\boldsymbol{F}\boldsymbol{T}^t\mathbb{1}_q$, and $\boldsymbol{F}\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{F}^t$ is the squared norm of the OTF. Computing $\mathbb{1}_p^t \boldsymbol{T}\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{T}^t \mathbb{1}_q$ thus only requires to do a scalar product between the phase shifts and $|\tilde{\boldsymbol{h}}|^2$.

```python
def dirac2D(x,y,N):
    d = np.zeros((N,N))
    d[y,x] = 1
    return d

def myfft2(spatial):
    return np.fft.fft2(spatial)/spatial.shape[0] #normalized to respect Parseval's theorem

def compute_TATt_Fourier(psd,M,sup_coord):
```

```python
    #computes each (p,q) element of TATt = (1p)tTATt(1q)
    #INPUTS: psd=|OTF|^2 or object PSD, M=number of pixels in the support, sup_coord=(x,y) coordinates of pixels within the support
    #OUTPUT: TATt

    Npix = psd.shape[0]
    k,l = np.meshgrid(np.linspace(0,Npix-1,Npix), np.linspace(0,Npix-1,Npix))
    tatt = np.zeros((M,M))
    for p in range(M):
        xp = sup_coord[0][p]
        yp = sup_coord[1][p]
        dp = dirac2D(xp,yp,Npix)
        tf_dp = myfft2(dp)
        for q in range(M):
            xq = sup_coord[0][q]
            yq = sup_coord[1][q]
            dq = dirac2D(xq,yq,Npix)
            tf_dq = myfft2(dq)
            tatt[p,q] = np.real(np.sum(tf_dp.conj()*psd*tf_dq))
    return tatt
```

Listing 7.2 – Python implementation of TATt computation in the Fourier domain

7.2.2 Faster versions using the autocorrelation

As said previously, computing the $\boldsymbol{TH^tHT^t}$ corresponds to computing the autocorrelation of the PSF, only for the elements in the support. Then, instead of computing the autocorrelation element-wise, we can compute it once for all by computing the inverse Fourier transform: of the object PSD in the case where we want $\boldsymbol{TR_{o_F}T^t}$, and of the squared norm of the OTF to compute $\boldsymbol{TH^tHT^t}$. Then, the corresponding $\mathbb{1}_p^t\boldsymbol{TAT^t}\mathbb{1}_q$ element is taken from the pre-computed inverse Fourier transform.

```python
def compute_TATt_autocorr(psd,M,sup_coord):
    #computes TATt using the inverse Fourier transform of the PSD (autocorrelation)
    #INPUTS: psd=|OTF|^2 or object PSD, M=number of pixels in the support, sup_coord=(x,y) coordinates of pixels within the support
    #OUTPUT: TATt

    tatt = np.zeros((M,M))
    autocorr = np.real(np.fft.ifft2(psd))
    for i in range(M):
        for j in range(M):
            xi = sup_coord[0][i]
            xj = sup_coord[0][j]
            yi = sup_coord[1][i]
            yj = sup_coord[1][j]
            tatt[i,j] = autocorr[xi-xj,yi-yj]
    return tatt
```

Listing 7.3 – Python implementation of TATt computation using autocorrelation

To go twice faster, one can also notice that $\boldsymbol{TAT^t}$ is symmetric, only requiring then to built half of the matrix and then easily get the other half (which is true in all methods).

The autocorrelation, computed as the inverse Fourier transform of either the squared OTF or the object PSD, is a vector with the N different elements which are necessary to compute the $N \times N$ matrix $\boldsymbol{A}$. These elements, correctly ordered, correspond to a line (or a column) of $\boldsymbol{A}$. Therefore, it is possible to build directly a line (or a column) of $\boldsymbol{TAT^t}$ by carefully shifting the autocorrelation vector, and only keeping the M elements within the support.

```python
def compute_TATt_autocorrline(psd,M,sup_coord,sup_coord_flat):
    #computes TATt line by line, using the inverse Fourier transform of the PSD (autocorrelation)
    #INPUTS: psd=|OTF|^2 or object PSD, M=number of pixels in the support, sup_coord=(x,y) coordinates of pixels within the support,sup_coord_flat=1D coordinate of pixels within the support (over [0,N-1], N the number of pixels in the image)
    #OUTPUT: TATt

    tatt = np.zeros((M,M))
    autocorr = np.real(np.fft.ifft2(psd))
    for i in range(M):
        # shifting autocorr
        xi = sup_coord[0][i]
        yi = sup_coord[1][i]
        autocorr_roll = np.roll(np.roll(autocorr,xi,axis=0),yi,axis=1)
        # (1D) line reshape
        line = autocorr_roll.ravel()
        # only keeping elements within the support
        tatt[i,:] = line[sup_coord_flat]
    return tatt
```

Listing 7.4 – Python implementation of TATt computation using autocorrelation (line by line)

7.2.3 Comparison of computational time

In order to compare the runtime taken by each version, we run the cProfile profiler for each implementation. We place ourselves in a realistic work frame, and simulate the observation of a small asteroid on a 128 × 128 image (for this, we undersample the 512 × 512 Vesta synthetic object given in Figure 5.1), using the same observation conditions as in previous chapters. We also build the corresponding 128 × 128 PSF. $\boldsymbol{TAT^t}$ is a $M \times M$ matrix, and the number of pixels in the support is $M = 1442$. The Python functions receive as an input either this PSF, or the squared OTF, as well as M and the coordinates of the pixels within the support.

The results are the following: the naive implementation in the Fourier domain given in Listing 7.2 is the slowest, computing the matrix in 1187 s, due to the multiples (costly) Fourier transforms that are done. The naive implementation in the spatial domain in Listing 7.1 takes 131 s, it is much faster than the first one (≈9 times less than the previous method) given that it does not require to do Fourier transforms, however computing the autocorrelation for each pixel is costly. The implementation using the (firstly computed) autocorrelation, but computing the matrix element-wise, in Listing 7.3 takes 2 s, here again it is way faster than the previous one (≈70 times less than with Listing 7.1) due to the fact that the costly computation of the autocorrelation is done once for all. Finally, the implementation computing the matrix line by line in Listing 7.4 enables us to gain even more time, and takes less than 0.1 s (≈20 times less than the previous method). All four methods give the same result for $\boldsymbol{TAT^t}$ (to the machine precision).

7.3 Tests with a known support

7.3.1 Marginal MAP estimator including the support constraint

Now that the issue of $\boldsymbol{TAT^t}$ matrices computation was tackled, we are ready to study the impact of a known support on the estimated parameters. We choose to use the marginal approach described in Subsection 7.1.2 for this study, moreover we choose to use the marginal MAP estimator, meaning that we are aiming at finding the parameters $\boldsymbol{\theta} = \{\gamma_o, k, r_0, v_\phi, \gamma_n\}$ maximizing the marginal likelihood given in Equation 7.3, for given labels ℓ. This means adapting P-AMIRAL in order to take into account the modifications induced by the labels, in the case where they are known.

Given that we are using a marginal MAP estimator, as said in Chapter 2, maximizing the marginal likelihood is equivalent to minimizing its anti-logarithm. Therefore, using Equation 7.3, we can write the minimized criterion as:

$$J_{mMAP_\ell}(\boldsymbol{\theta}) = \frac{1}{2}\ln\det(\boldsymbol{R_i}) + \frac{1}{2}(\boldsymbol{i} - \boldsymbol{m_i})^T \boldsymbol{R_i}^{-1}(\boldsymbol{i} - \boldsymbol{m_i})$$

with mean image $\boldsymbol{m_i}$ and image covariance matrix $\boldsymbol{R_i}$. We then need the inverse of $\boldsymbol{R_i}$ as well as its determinant, computed as follows:

$$\begin{aligned}\boldsymbol{R_i}^{-1} &= (\boldsymbol{HT^tTR_{o_F}T^tTH^t} + \gamma_n^{-1}\boldsymbol{I_N})^{-1}\\ &= \gamma_n\boldsymbol{I_N} - \gamma_n^2\boldsymbol{HT^t}((\boldsymbol{TR_{o_F}T^t})^{-1} + \gamma_n\boldsymbol{TH^tHT^t})^{-1}\boldsymbol{TH^t}\end{aligned}$$

$$\begin{aligned}\text{and }\ \det(\boldsymbol{R_i}) &= \det(\boldsymbol{HT^tTR_{o_F}T^tTH^t} + \gamma_n^{-1}\boldsymbol{I_N})\\ &= \gamma_n^{-N}\det(\gamma_n\boldsymbol{HT^tTR_{o_F}T^tTH^t} + \boldsymbol{I_N})\\ &= \gamma_n^{-N}\det(\gamma_n\boldsymbol{TR_{o_F}T^tTH^tHT^t} + \boldsymbol{I_M})\end{aligned}$$

following the Weinstein–Aronszajn identity. We then have to invert an $M \times M$ matrix twice, instead of inverting a (way bigger) $N \times N$ matrix.

7.3.2 Simulation conditions

Given previous expressions, in order to compute the criterion $J_{mMAP_\ell}(\boldsymbol{\theta})$ we need to manipulate and store $M \times M$ matrices. In order to be able to compute them in practice, we need M, the number of pixels within support, to be small enough. To have only small supports to compute, we decide for these simulations to look at smaller objects: for this, we undersample the synthetic view of Vesta we show in Figure 1.1 (left), which is originally an image of 512×512 pixels, to finally have a 128×128 pixels image. The resulting object is given in Figure 7.1.

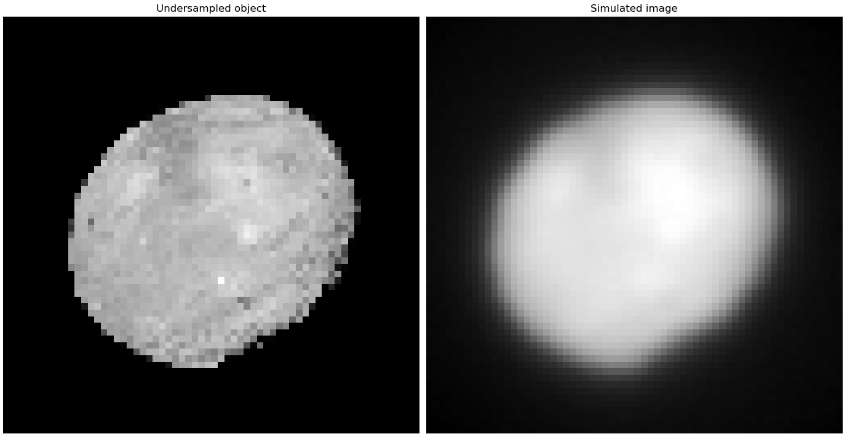

Figure 7.1 – Left: undersampled synthetic view of Vesta (cropped from $128 \times 128 pixels$). Right: simulated AO-corrected observation of the left object using a VLT/SPHERE-like system, with true parameters: $r_0 = 0.15\,\mathrm{m}$, $v_\phi = 1.3\,\mathrm{rad}^2$ and $\gamma_n = 10^{-4}\,\mathrm{ph}^{-2}$.

Apart from the chosen object, the simulation conditions remain the same as in the previous chapters. The simulated AO system is a “SPHERE-like” AO system, whose parameters are summed up in Table 3.2. The PSF is again simulated using the PSFAO19 model [Fétick, 2019a]. As done in previous chapters, we take as true PSF parameters are $r_0 = 0.15\,\mathrm{m}$ and $v_\phi = 1.3\,\mathrm{rad}^2$ at the imaging wavelength $\lambda = 550\,\mathrm{nm}$, and noise is taken zero-mean, additive, white and Gaussian with a variance equal to the mean value of the object. We simulate several noise levels, meaning several mean fluxes of the object $F_o/N = 1/\gamma_n = [10^2, 10^3, 10^4]\,\mathrm{ph}$ (photons), which correspond respectively to slightly pessimistic/low, reasonable and slightly optimistic/high flux levels.

As in previous chapters, all PSF and PSD parameters are estimated following the marginal method except the object PSD power which is fixed to $p = 3$ here also.

The object PSD parameters are initialized to the typical values we are excepting for these parameters: k is initialized to 1, γ_o to $\gamma_o = N/F_o^2$. γ_n is initialized to its true value N/F_o. The PSF parameters are initialized to their extrema values giving the sharpest PSF (leading to the most under-deconvolved image): thus, r_0 is initialized to its maximum value (the weakest turbulence) $1.50\,\mathrm{m}$ and v_ϕ is initialized to its minimum value (the smallest variance meaning the best AO correction) $0.5\,\mathrm{rad}^2$. The aim of such an initialization is to clearly see the impact of adding the support constraint on PSF parameters, putting aside the additional difficulties brought by the hyperparameters.

7.3.3 Results using the true support

To begin with, we compare the results we obtain with and without the support constraint, using as known support the true one, with the highest mean flux $10^4\,\mathrm{ph}$ (photons), to start by a favorable situation where the noise impact is the smallest. The results on the parameters are given in Table 7.1.

Parameter	True	Estimated without support	Estimated with support
γ_n (ph^{-2})	1.00×10^{-4}	1.02×10^{-4}	9.94×10^{-5}
r_0 (m)	0.15	1.50	0.15
v_ϕ(rad^2)	1.30	1.25	1.29
γ_o (ph^{-2})	-	5.16×10^{-12}	1.27×10^{-13}
k	-	0.34	0.21

Table 7.1 – True (if available) and estimated values for γ_n, r_0, v_ϕ, γ_o and k, without and with the support constraint (using the true support), for $F_o/N = 10^4$ ph.

To begin with, we notice that γ_n and v_ϕ are well estimated in both cases, with an error below 4% on v_ϕ and equal to 2% on γ_n without the support constraint (with P-AMIRAL), and an error below 1% on v_ϕ and equal to 6% on γ_n with the support constraint. On the contrary, the estimated r_0 without the support constraint hits the upper bound which is ten times higher than the true value. P-AMIRAL, which was working on 512×512 pixels images, seems to have trouble working on smaller objects, as also discussed in [Lau, 2023]. However, using the support constraint, the estimation of r_0 is then very satisfactory as it exactly gives the true value. Thus, adding the support highly improves the PSF estimation, and enables one to have a satisfactory PSF estimation even for smaller objects.

7.3.4 Results using approached supports (under/over estimated)

However, in practice it is quite unlikely to get the true support, which highly motivates to pursue the study for several supports, underestimated or overestimated. Therefore, in this subsection we run the same marginal MAP method incorporating the support constraint, but using an eroded as well as a dilated version of the true support. Using a dilated support means that some pixels in the background (close to the edge of the object) are said to be in the object, whereas using an eroded support means on the contrary that some pixels on the edge of the object are said to be in the background. Both the erosion and the dilation kernel are of size 7×7 meaning a square of size $7 = 1 + 2 \times \text{samp} = 1 + 2 \times \lambda/D$, with λ/D corresponding to the Full Width at Half Max (FWHM) of the Airy disk, making a good approximation of the uncertainty we have around the support.

As previously, the results on the parameters are given in Table 7.2.

Parameter	True	Dilated support	Eroded support
γ_n (ph^{-2})	1.00×10^{-4}	1.00×10^{-4}	1.00×10^{-5}
r_0 (m)	0.15	0.15	0.16
v_ϕ(rad^2)	1.30	1.30	3.00
γ_o (ph^{-2})	-	3.78×10^{-14}	6.10×10^{-16}
k	-	400	0.0002

Table 7.2 – True (if available) and estimated values for γ_n, r_0, v_ϕ, γ_o and k, with the support constraint, using a dilated and an eroded version of the true support, for $F_o/N = 10^4$ ph.

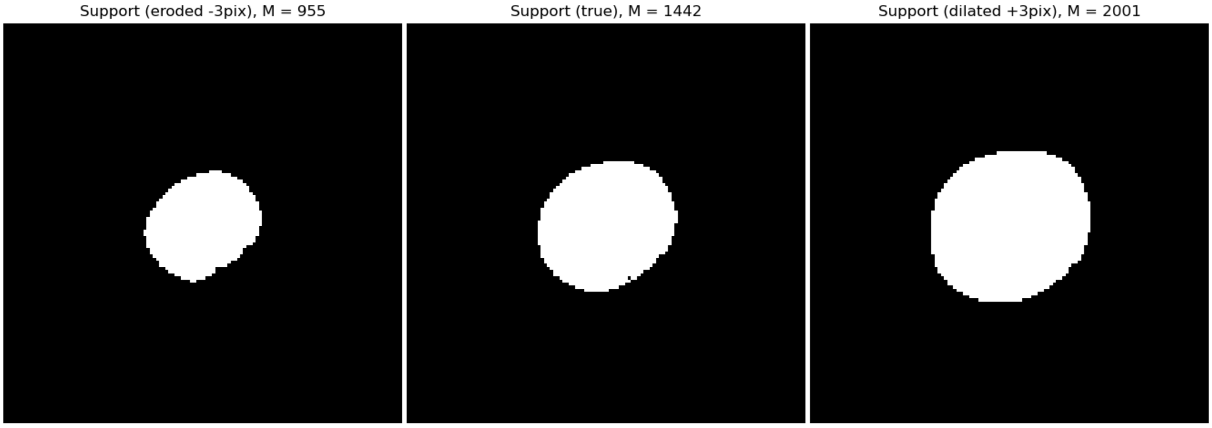

Figure 7.2 – Left: Eroded support using a 7×7 kernel, meaning an erosion of λ/D, $M = 955$. Center: True support, $M = 1442$. Right: Dilated support using a 7×7 kernel, $M = 2001$.

We notice, just as in the previous subsection with the true support, that r_0 estimation is very satisfactory: very close even equal to the true value. Concerning the noise level and the other PSF parameter, they are very well estimated when using the dilated support. However, using the eroded support, the estimated v_ϕ and γ_n are far from the true value: the residual phase variance hits the upper bound and the noise precision is estimated ten times below the true value, whereas they were until now very well estimated even without the support constraint. We explain this as the consequence of assigning the background label to bright pixels in the object. Under our hypotheses, this background "brightness" can then only be modeled by a higher noise level (higher variance, lower precision γ_n), getting closer to the signal level. Thus, in order to compensate for this higher noise level and to still fit the image PSD (which is done in our method, as explained in Chapter 2), the object PSD slope p being fixed, the OTF slope is flattered: thus, the OTF is higher on medium-high frequencies, meaning that the AO correction level is better, leading to a smaller variance v_ϕ.

7.3.5 Tests with different noise levels

In the previous subsection, we simulated the observation of Vesta with a rather high mean flux level $F_o/N = 10^4$ pix. We now do the same simulations but here imposing a smaller mean flux: $F_o/N = 10^3$ pix and $F_o/N = 10^2$ pix, to look at the impact of noise on the results. The results obtained with $F_o/N = 10^3$ pix are very similar to those obtained with $F_o/N = 10^4$ pix, thus we only display here the results on parameters for $F_o/N = 10^2$ pix in Table 7.3.

Here again, we notice that γ_n and v_ϕ are quite well estimated, without and with the constraint support with either the true or the dilated support. However, with this higher noise level, the differences between the marginal MAP estimation with or without the support constraint are also more visible. Indeed, the error on γ_n without the constraint support is around 8% whereas it is closer to 13%-14% with the support constraint (true or dilated support). Concerning the error on v_ϕ, it is around 5%-6% with the support constraint (true or dilated support), whereas it is twice higher (11%) without. For the eroded support, here again the estimated γ_n is way lower than the true value and the estimated v_ϕ hits the upper bound. Finally, concerning r_0,

Parameter	True	Without Sup.	True Sup.	Dilated Sup.	Eroded Sup.
γ_n (ph^{-2})	1.00×10^{-2}	9.92×10^{-3}	9.86×10^{-3}	9.87×10^{-3}	6.08×10^{-3}
r_0 (m)	0.15	1.50	0.15	0.15	0.15
v_ϕ(rad^2)	1.30	1.44	1.23	1.38	3.00
γ_o (ph^{-2})	-	4.25×10^{-8}	3.95×10^{-9}	7.83×10^{-10}	6.10×10^{-12}
k	-	0.5	0.07	0.8	0.006

Table 7.3 – True (if available) and estimated values for γ_n, r_0, v_ϕ, γ_o and k, without and with the support constraint, using the true support, a dilated and an eroded version of it, for $F_o/N = 10^2$ ph.

the estimated value without the support constraint hits the upper bound here again. Using the support constraint, the estimation of r_0 is very satisfactory as it exactly gives the true value, for the true and both approached supports.

Overall, this shows that the support can be known approximately, and still give very satisfactory results on the Fried parameter r_0 which is the one P-AMIRAL could not estimate correctly. This study also shows that slightly overestimating the support (including some background pixels in the true support) leads to better results than underestimating it (removing some pixels from the true support, into the background), due to the fact that the latter leads to an overestimation of noise level and thus of the AO correction level in order to fit the image PSD in global.

7.3.6 Estimating p

As said previously, one motivation for including the support constraint is to differentiate better the object contribution from the PSF contribution. This strong information on the object would enable the estimation of even more parameters. For instance, while we had to work in a mostly unsupervised mode using the previous method, we could think about estimating hyperparameter p using the support constraint.

Using the same simulations with the smallest mean flux $F_o/N = 10^2$ pix, we now look at the results in the fully unsupervised mode displayed in Table 7.4.

Parameter	True	Without Sup.	True Sup.	Dilated Sup.	Eroded Sup.
γ_n (ph^{-2})	1.00×10^{-2}	9.94×10^{-3}	9.91×10^{-3}	9.94×10^{-3}	6.84×10^{-3}
r_0 (m)	0.15	1.50	0.15	0.15	0.16
v_ϕ(rad^2)	1.30	0.5	1.23	1.36	2.94
γ_o (ph^{-2})	-	2.73×10^{-8}	1.27×10^{-9}	8.53×10^{-10}	6.10×10^{-12}
k	-	1.30	3.80	1.33	0.002
p	-	3.59	3.49	2.93	2.00

Table 7.4 – True (if available) and estimated values for γ_n, r_0, v_ϕ, γ_o, k and p, without and with the support constraint, using the true support, a dilated and an eroded version of it, for $F_o/N = 10^2$ ph.

As expected, the results without the constraint support are not satisfactory for both PSF parameters, the r_0 hits the maximum bound as before, and we find the same tendency as in

Chapter 5 for v_ϕ and p in the fully unsupervised mode due to their correlation. Using the support constraint, the estimation of r_0 is very satisfactory as it gives almost exactly the true value, for the true and both approached supports. Moreover, as in the previous case, the estimation of v_ϕ is also very satisfactory using the true and the dilated support. Indeed, if the support is under-estimated, the method tries to compensate with a wider PSF thus a lower AO correction level (higher v_ϕ) and a higher noise variance (lower γ_n) in order to fit the image. Finally, the estimated p using the dilated support is very close to the reasonable $p = 3$ for asteroids, and even closer to the optimal value $p = 2.91$ found in the fully supervised mode [Fétick, 2020b]. We believe that the value of p found using the true support corresponds to the PSD decrease of the object texture meaning the PSD decrease that the full object would have: indeed, inside the object there are no sharp edges thus the PSD decrease is faster (corresponding to a bigger p).

As a conclusion, adding the support constraint highly improves the PSF estimation, and enables one to have a satisfactory PSF estimation even for smaller objects and in the fully unsupervised mode which was not possible with the previous method P-AMIRAL. Similar results have been observed for different noise levels and using different known supports, suggesting that the support constraint impacts greatly the PSF estimation (towards the true PSF) even in the case of stronger noise or approximated support. Again, slightly overestimating the support (including some background pixels in the true support) leads to better results than underestimating it (removing some pixels from the true support, into the background), due to the image fitting done by the method.

7.4 Including the labels' estimation: some elements

Given the improvement brought by the support information shown in the previous result section, one of the next step would then be to estimate the support along with the other parameters. This section gives some elements in order to include the estimation of the labels.

7.4.1 Label priors

In Section 7.1, we re-used the marginal framework developed in previous chapters and adapted it to incorporate the support. One option to estimate the support would then be to add the estimation of the labels to this marginal approach. For the labels, one of the most commonly used model in the probabilistic approaches is the Potts model [Geman, 1984; Ayasso, 2010; Altmann, 2013]: the object is described as a series of homogeneous regions, where the pixels have the same label. As said previously, the N pixels in the object can belong to two different classes, either in or out of the support. The Potts prior writes as follows:

$$f(\ell|\beta_\ell) = C(\beta_\ell)^{-1} \exp(\beta_\ell \nu(\ell)) \tag{7.7}$$

β_ℓ is a "correlation" parameter, it is tuning the typical size of a region. $\nu(\ell)$ counts the number of identical neighbors:

$$\nu(\ell) = \sum_{p\sim q} \delta(\ell_p;\ell_q) \tag{7.8}$$

where δ is the Kronecker function, and $\sim$ defines a neighborhood relation, here we take the four nearest neighbors into account. Finally, $C(\beta_\ell)$ is a normalization constant, which can be pre-computed numerically:

$$C(\beta_\ell) = \sum_{\ell} \exp(\beta_\ell \nu(\ell)) \tag{7.9}$$

7.4.2 Joint approach, with full object

Another possible approach to estimate the support is to return to a joint estimation framework as in Chapter 2, where the object is explicitly estimated with the PSF, but this time using Equation 7.2. The object we consider can either be the full object $\boldsymbol{o_F}$ or the truncated object $\bar{\boldsymbol{o}}$, leading to different expressions for the posterior distributions.

7.4.2.1 Object prior

As said previously, the full object $\boldsymbol{o_F}$ is still the Gaussian field considered previously. Therefore we can use the same Gaussian prior for the full object as in Equation 4.2:

$$\begin{aligned} p(\boldsymbol{o_F}|\gamma_o,k) &= (2\pi)^{-N/2}\det(\boldsymbol{R_{o_F}})^{-1/2}\exp(-\frac{1}{2}(\boldsymbol{o_F}-\boldsymbol{m_{o_F}})^t\boldsymbol{R_{o_F}}^{-1}(\boldsymbol{o_F}-\boldsymbol{m_{o_F}})) \\ &= \Big(\frac{\gamma_o}{2\pi}\Big)^{N/2}\prod_{\boldsymbol{f}}\Big(\bar{S}_o(\boldsymbol{f})^{-1/2}\exp\big[-\frac{\gamma_o}{2}\bar{S}_o(\boldsymbol{f})^{-1}|\tilde{o}_F(\boldsymbol{f})-\tilde{m}_o(\boldsymbol{f})|^2\big]\Big) \end{aligned} \tag{7.10}$$

7.4.2.2 Parameters priors

For the hyperparameters and PSF parameters, several options are possible. If we do not have additional knowledge on them, one possibility is to take uniform priors which are not very informative priors, just as done previously. Another option is to choose wisely the prior distribution in order to have a posterior distribution which is easy to compute at the end. In our case, one can notice that, specifically in the joint approach, using a Gamma law for γ_n and γ_o priors results in having a Gamma law for their posterior distribution, and the posterior Gamma law parameters can be easily written from the parameters of the prior. Such priors which are in the same family as their associated posteriors (given the likelihood) are called conjugated priors [Kass, 1996].

Therefore, we choose to take uniform priors for k, r_0 and v_ϕ, and Gamma priors for γ_n and γ_o : $f(\gamma_n) = G_{\alpha_n^{prior},\beta_n^{prior}}(\gamma_n)$ and $f(\gamma_o) = G_{\alpha_o^{prior},\beta_o^{prior}}(\gamma_o)$. A Gamma prior probability density function writes as follows:

$$f(\gamma) = \frac{\beta}{\Gamma(\alpha)}\gamma^{\alpha-1}\exp(-\beta\gamma)\mathbb{1}_+(\gamma)$$

with Γ the Gamma function, and $\mathbb{1}_+(\gamma)$ the indicator function. The Gamma distribution becomes uninformative when $\alpha \to 0$ and $\beta \to 0$: in this case it becomes a Jeffreys law $f(\gamma) = 1/\gamma$.

7.4.2.3 Likelihood

Knowing the labels, the PSF, and the object, given the noise precision, the likelihood is still Gaussian and writes, as previously in Equation 4.3:

$$\begin{aligned} p(\boldsymbol{i}|\boldsymbol{o_F}, \ell, \gamma_n, r_0, v_\phi) &= \Big(\frac{\gamma_n}{2\pi}\Big)^{N/2} \exp(-\frac{\gamma_n}{2}||\boldsymbol{i} - \boldsymbol{Ho}||^2) \\ &= \Big(\frac{\gamma_n}{2\pi}\Big)^{N/2} \exp(-\frac{\gamma_n}{2}||\boldsymbol{i} - \boldsymbol{H T_\ell^t T_\ell o_F}||^2) \end{aligned} \tag{7.11}$$

We sum up the prior dependencies of the different parameters in the following hierarchical model:

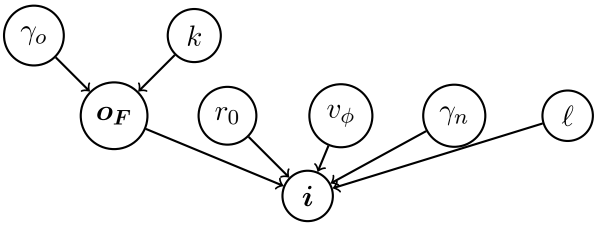

Figure 7.3 – Hierarchical model summing up the inter-dependency between the full object, the image, the labels and all parameters.

7.4.2.4 Conditional posterior distributions

Given the likelihood and the prior distribution, we can write the posterior distributions using the Bayes rule.

For any PSF parameter θ_h, only the prior of the considered parameter and the likelihood depend on it, so its posterior writes:

$$f(\theta_h|\boldsymbol{i}, \boldsymbol{o_F}, \ell, \boldsymbol{\theta}_{\neq\theta_h}) \propto f(\theta_h) \exp(-\frac{\gamma_n}{2}||\boldsymbol{i} - \boldsymbol{H T_\ell^t T_\ell o_F}||^2) \tag{7.12}$$

with $\boldsymbol{\theta}_{\neq x}$ all parameters except the considered parameter x.

For the labels, similarly to the previous computation, the posterior writes:

$$f(\ell|\boldsymbol{i}, \boldsymbol{o_F}, \boldsymbol{\theta}) \propto C(\beta_\ell)^{-1} \exp(\beta_\ell \nu(\ell)) \exp(-\frac{\gamma_n}{2}||\boldsymbol{i} - \boldsymbol{H T_\ell^t T_\ell o_F}||^2) \tag{7.13}$$

For any object parameter θ_o except γ_o (for instance here, k and p if estimated), only the prior of θ_o and the object prior depend on it, so the posterior writes:

$$f(\theta_o|\boldsymbol{i}, \boldsymbol{o_F}, \ell, \boldsymbol{\theta}_{\neq\theta_o}) \propto f(\theta_o) \prod_{\boldsymbol{f}} \Big(\bar{S}_o(\boldsymbol{f})^{-1/2} \exp\big[-\frac{\gamma_o}{2}\bar{S}_o(\boldsymbol{f})^{-1}|\tilde{o}_F(\boldsymbol{f}) - \tilde{m}_o(\boldsymbol{f})|^2\big]\Big) \tag{7.14}$$

For γ_o and γ_n, as said in their prior sections, their posterior is also a Gamma law, and its parameters α and β change as follows:

$$f(\gamma_n|\boldsymbol{i}, \boldsymbol{o_F}, \ell, \boldsymbol{\theta}_{\neq\gamma_n}) \propto \gamma_n^{\alpha_n-1}\exp(-\beta_n\gamma_n)\mathbb{1}_+(\gamma_n) \tag{7.15}$$

with $\alpha_n = N/2 + \alpha_n^{prior}$ and $\beta_n = \dfrac{||\boldsymbol{i} - \boldsymbol{H}\boldsymbol{T_\ell^t}\boldsymbol{T_\ell}\boldsymbol{o_F}||^2}{2} + \beta_n^{prior}$, and

$$f(\gamma_o|\boldsymbol{i}, \boldsymbol{o_F}, \ell, \boldsymbol{\theta}_{\neq\gamma_o}) \propto \gamma_o^{\alpha_o-1}\exp(-\beta_o\gamma_o)\mathbb{1}_+(\gamma_o) \tag{7.16}$$

with $\alpha_o = N/2 + \alpha_o^{prior}$ and $\beta_o = \sum_{\boldsymbol{f}} \dfrac{\bar{S}_o(\boldsymbol{f})^{-1}|\tilde{o}_F(\boldsymbol{f}) - \tilde{m}_o(\boldsymbol{f})|^2}{2} + \beta_o^{prior}$.

Finally, for the full object, its prior and the likelihood being Gaussian, its posterior is also Gaussian and writes as follows:

$$p(\boldsymbol{o_F}|\boldsymbol{i}, \ell, \boldsymbol{\theta}) \propto \det(\boldsymbol{\Sigma_{o_F}})^{-1/2}\exp(-\frac{1}{2}(\boldsymbol{o_F} - \boldsymbol{\mu_{o_F}})^t\boldsymbol{\Sigma_{o_F}^{-1}}(\boldsymbol{o_F} - \boldsymbol{\mu_{o_F}})) \tag{7.17}$$

with $\boldsymbol{\Sigma_{o_F}^{-1}} = \gamma_n\boldsymbol{T_\ell^t}\boldsymbol{T_\ell}\boldsymbol{H^t}\boldsymbol{H}\boldsymbol{T_\ell^t}\boldsymbol{T_\ell} + \boldsymbol{R_{o_F}^{-1}}$ and $\boldsymbol{\mu_{o_F}} = \boldsymbol{\Sigma_{o_F}}(\gamma_n\boldsymbol{T_\ell^t}\boldsymbol{T_\ell}\boldsymbol{H^t}\boldsymbol{i} + \boldsymbol{R_{o_F}^{-1}}\boldsymbol{m_{o_F}})$.

$\boldsymbol{\Sigma_{o_F}^{-1}}$ is a $N \times N$ matrix, so unless the considered image is small enough, computing its inverse is tricky even hardly possible, thus we want to avoid these computations as much as possible. Therefore, we are using the Woodbury matrix identity to write $\boldsymbol{\Sigma_{o_F}}$ in a more convenient way:

$$\begin{aligned}\boldsymbol{\Sigma_{o_F}} &= (\gamma_n\boldsymbol{T_\ell^t}\boldsymbol{T_\ell}\boldsymbol{H^t}\boldsymbol{H}\boldsymbol{T_\ell^t}\boldsymbol{T_\ell} + \boldsymbol{R_{o_F}^{-1}})^{-1}\\ &= \boldsymbol{R_{o_F}} - \gamma_n\boldsymbol{R_{o_F}}\boldsymbol{T_\ell^t}((\boldsymbol{T_\ell}\boldsymbol{H^t}\boldsymbol{H}\boldsymbol{T_\ell^t})^{-1} + \gamma_n\boldsymbol{T_\ell}\boldsymbol{R_{o_F}}\boldsymbol{T_\ell^t})^{-1}\boldsymbol{T_\ell}\boldsymbol{R_{o_F}}\end{aligned}$$

Then, in order to sample this high-dimensional Gaussian object, one possibility is to use a Perturbation-Optimization algorithm as in [Orieux, 2012; Gilavert, 2013].

7.4.3 Joint approach, with truncated object

As said previously, one could also return to the joint estimation framework but considering the truncated object $\bar{\boldsymbol{o}}$. As $\bar{\boldsymbol{o}}$ is a linear transformation of the full object $\boldsymbol{o_F}$ (for which we took a Gaussian prior in Subsection 7.4.2, then the prior for $\bar{\boldsymbol{o}}$ is also Gaussian given the remark in Subsection 7.1.2.

The prior on the truncated object writes:

$$p(\bar{\boldsymbol{o}}|\gamma_o, k, \ell) = (2\pi)^{-M_\ell/2}\det(\boldsymbol{R_{\bar{o}}})^{-1/2}\exp(-\frac{1}{2}(\bar{\boldsymbol{o}} - \boldsymbol{m_{\bar{o}}})^t\boldsymbol{R_{\bar{o}}^{-1}}(\bar{\boldsymbol{o}} - \boldsymbol{m_{\bar{o}}})) \tag{7.18}$$

with $\boldsymbol{R_{\bar{o}}} = \boldsymbol{T_\ell}\boldsymbol{R_{o_F}}\boldsymbol{T_\ell^t}$ and $\boldsymbol{m_{\bar{o}}} = \boldsymbol{T_\ell}\boldsymbol{m_{o_F}}$.

We choose the same priors for labels as well as parameters and hyperparameters, except for γ_o for which we decide to use a uniform prior just like the other object PSD parameters, for simplicity. The likelihood given the truncated object is slightly modified from Equation 7.11:

$$p(\boldsymbol{i}|\bar{\boldsymbol{o}}, \ell, \gamma_n, r_0, v_\phi) = \left(\frac{\gamma_n}{2\pi}\right)^{N/2}\exp(-\frac{\gamma_n}{2}||\boldsymbol{i} - \boldsymbol{H}\boldsymbol{T_\ell^t}\bar{\boldsymbol{o}}||^2) \tag{7.19}$$

The posterior distributions are also slightly modified. For any PSF parameter θ_h, only the prior of the considered parameter and the likelihood depend on it, so its posterior writes:

$$f(\theta_h|\boldsymbol{i}, \bar{\boldsymbol{o}}, \ell, \boldsymbol{\theta}_{\neq\theta_h}) \propto f(\theta_h)\exp(-\frac{\gamma_n}{2}||\boldsymbol{i} - \boldsymbol{H}\boldsymbol{T}_{\boldsymbol{\ell}}^{\boldsymbol{t}}\bar{\boldsymbol{o}}||^2)$$

For any object parameter θ_o (γ_o included), similarly again, only the prior of θ_o and the object prior depend on it, so the posterior writes:

$$f(\theta_o|\boldsymbol{i}, \bar{\boldsymbol{o}}, \ell, \boldsymbol{\theta}_{\neq\theta_h}) \propto f(\theta_o)\det(\boldsymbol{R}_{\bar{\boldsymbol{o}}})^{-1/2}\exp(-\frac{1}{2}(\bar{\boldsymbol{o}} - \boldsymbol{m}_{\bar{\boldsymbol{o}}})^t\boldsymbol{R}_{\bar{\boldsymbol{o}}}^{-1}(\bar{\boldsymbol{o}} - \boldsymbol{m}_{\bar{\boldsymbol{o}}}))$$

For γ_n, as said previously, its posterior is also a Gamma law, and its parameters α and β change as follows:

$$f(\gamma_n|\boldsymbol{i}, \bar{\boldsymbol{o}}, \ell, \boldsymbol{\theta}_{\neq\gamma_n}) \propto \gamma_n^{\alpha_n - 1}\exp(-\beta_n\gamma_n)\mathbb{1}_+(\gamma_n)$$

with $\alpha_n = N/2 + \alpha_n^{prior}$ and $\beta_n = \dfrac{||\boldsymbol{i} - \boldsymbol{H}\boldsymbol{T}_{\boldsymbol{\ell}}^{\boldsymbol{t}}\bar{\boldsymbol{o}}||^2}{2} + \beta_n^{prior}$.

Finally, for the truncated object, its prior and the likelihood being Gaussian, its posterior is also Gaussian and writes as follows:

$$p(\bar{\boldsymbol{o}}|\boldsymbol{i}, \ell, \boldsymbol{\theta}) \propto \det(\boldsymbol{\Sigma}_{\bar{\boldsymbol{o}}})^{-1/2}\exp(-\frac{1}{2}(\bar{\boldsymbol{o}} - \boldsymbol{\mu}_{\bar{\boldsymbol{o}}})^t\boldsymbol{\Sigma}_{\bar{\boldsymbol{o}}}^{-1}(\bar{\boldsymbol{o}} - \boldsymbol{\mu}_{\bar{\boldsymbol{o}}})) \tag{7.20}$$

with $\boldsymbol{\Sigma}_{\bar{\boldsymbol{o}}}^{-1} = \gamma_n\boldsymbol{T}_{\boldsymbol{\ell}}\boldsymbol{H}^{\boldsymbol{t}}\boldsymbol{H}\boldsymbol{T}_{\boldsymbol{\ell}}^{\boldsymbol{t}} + \boldsymbol{R}_{\bar{\boldsymbol{o}}}^{-1}$ and $\boldsymbol{\mu}_{\bar{\boldsymbol{o}}} = \boldsymbol{\Sigma}_{\bar{\boldsymbol{o}}}(\gamma_n\boldsymbol{T}_{\boldsymbol{\ell}}\boldsymbol{H}^{\boldsymbol{t}}\boldsymbol{i} + \boldsymbol{R}_{\bar{\boldsymbol{o}}}^{-1}\boldsymbol{m}_{\bar{\boldsymbol{o}}})$.

The biggest difference with the full object case is the posterior distribution for the labels. In this case, the posterior writes:

$$\begin{aligned} f(\ell|\boldsymbol{i}, \bar{\boldsymbol{o}}, \boldsymbol{\theta}) \propto\ & C(\beta_\ell)^{-1}\exp(\beta_\ell\nu(\ell))\exp(-\frac{\gamma_n}{2}|\boldsymbol{i} - \boldsymbol{H}\boldsymbol{T}_{\boldsymbol{\ell}}^{\boldsymbol{t}}\bar{\boldsymbol{o}}_{\boldsymbol{\ell}}|^2)(2\pi)^{-M_\ell/2} \\ & \times\det(\boldsymbol{T}_{\boldsymbol{\ell}}\boldsymbol{R}_{\boldsymbol{o_F}}\boldsymbol{T}_{\boldsymbol{\ell}}^{\boldsymbol{t}})^{-1/2}\exp(-\frac{1}{2}(\bar{\boldsymbol{o}} - \boldsymbol{T}_{\boldsymbol{\ell}}\boldsymbol{m}_{\boldsymbol{o_F}})^t(\boldsymbol{T}_{\boldsymbol{\ell}}\boldsymbol{R}_{\boldsymbol{o_F}}\boldsymbol{T}_{\boldsymbol{\ell}}^{\boldsymbol{t}})^{-1}(\bar{\boldsymbol{o}} - \boldsymbol{T}_{\boldsymbol{\ell}}\boldsymbol{m}_{\boldsymbol{o_F}})) \end{aligned}$$

As one can notice, the labels intervene way more often, and particularly they change the number of pixels M we are considering (because they change the support).

The hierarchical model in the truncated object case also illustrates this not common dependency.

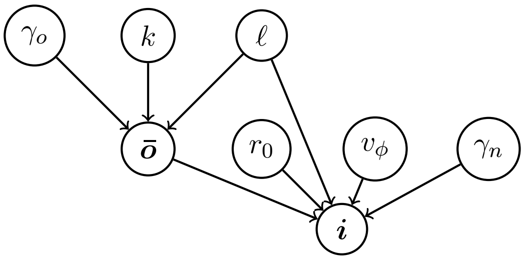

Figure 7.4 – Hierarchical model summing up the inter-dependency between the object, the image and all parameters.

7.4.4 Short discussion

Even though the joint estimator was abandoned in favour of the marginal estimator, due to the joint estimator's degeneracy in the absence of constraints, adding the support information changes the situation. Indeed, introducing this constraint complicates the expression of the posterior distributions using the marginal estimator, given in Equation 7.4. The posterior expressions using the joint estimator, given in Equations 7.12–7.17, especially considering the full object, are way simpler. For the easiest cases, one can even use conjugated priors as said in Subsection 7.4.2.2. However, using the joint estimator also implies to sample the object additionally to all the estimated parameters, which may be done by a Perturbation-Optimization algorithm. Further investigation has to be done in order to be able to choose the most efficient approach between the joint and the marginal ones.

Conclusion

Work summary

The presented work in this thesis was done in the context of adaptive-optics-corrected image restoration, for astronomical and satellite observations. Indeed, restoring AO-corrected images is particularly difficult, as it often suffers from the lack of precise knowledge on the point spread function, in addition to usual difficulties. An efficient approach is to marginalize the object out of the problem and to estimate the PSF and (object and noise) hyperparameters only, before deconvolving the image using these estimates. Recent works [Blanco, 2011; Fétick, 2020b] have implemented such a marginal blind deconvolution method: AMIRAL, based on the MAP estimator, which maximizes the (marginal) posterior distribution. This estimator was also combined with a parametric model of the PSF [Fétick, 2019a], leading to the parametric version of AMIRAL (P-AMIRAL).

However, these two versions (AMIRAL and P-AMIRAL) were used in different contexts and were not compared to each other so far. Additionally, the marginal MAP method only gives access to the maximizer of the marginalized distribution, and does not enable one to infer global information on the posterior distribution. Finally, this previous method does not take into account some constraints on the object, such as a support constraint, which could improve the estimation quality. In this thesis, we extended the previous method P-AMIRAL, in order to derive more information on the estimations.

Firstly, in Chapter 1, we gave a statistical description of the atmospheric turbulence impairing the observations, and defined the Fried parameter r_0 describing the turbulence strength. We also described the principle of AO, partially correcting the turbulent wavefront: the AO correction is limited by several processes introducing (fitting, temporal, aliasing,...) errors and the contributions of these errors are summed up by the residual phase variance v_ϕ, which measures the quality of AO correction. v_ϕ and r_0 are the two main parameters of the AO-corrected PSF model we used throughout this work. Finally, we gave the imaging model as well as the main characteristics of the application cases we study in the rest of the manuscript.

Chapter 2 recalled some elements on image restoration problems. It described the Bayesian framework we work in as well as two different methods in order to solve the myopic deconvolution problem: the historical joint MAP method and the marginal MAP method. To finish, it presented other possible estimators: the MMSE and MMAE estimators, corresponding respectively to the posterior mean and median.

In Chapter 3, we presented our contributions to previous method P-AMIRAL: we proposed a modification of the object PSD model, in order to decouple the parameters in high frequencies: using the modified model, one of the three parameters is decoupled from the two others, and even only has a minor impact on the criterion. We also compared the different versions of the previous method (AMIRAL and P-AMIRAL) on astronomical images, showing that P-AMIRAL gives slightly better results on simulated data, with less dependency with respect to the initialization, than AMIRAL: P-AMIRAL under-estimates less the OTF, meaning that the restored object is less contrasted but has less artefacts. On experimental data, we notice that both versions give similar results, showing the same tendencies with P-AMIRAL estimating a slightly higher OTF than AMIRAL.

A significant part of the work was dedicated to the application of a new marginal blind deconvolution method using MCMC algorithms, extending previous works. Chapter 4 gave the different priors as well as the likelihood derived from the models we used in this thesis, in order to compute the marginal posterior distribution. Its expression being complex, its computation requires the use of numerical methods, thus we used MCMC methods to sample the posterior distribution. In this chapter, we described the two algorithms we used for this work: the Random-Walk Metropolis-Hastings algorithm, and the Gibbs algorithm. Finally, we gave some computational elements to have access to the gradient and the Fisher information, which can be used in faster MCMC methods.

Using the MCMC methods described previously, we sampled the marginal posterior distribution and computed the optimal estimator minimizing the mean square error (MMSE estimator). In Chapter 5, we validated this new estimator on simulated images, for both asteroid and satellite observation contexts. The new method gives accurate estimations of noise and object hyperparameters, as well as satisfactory estimates and derived uncertainties on the OTF, in realistic conditions of simulation and in both contexts, suggesting the broad applicability of the method. For the simulated asteroid images, we computed our estimations for several noise realisations, to check the robustness of our method to noise, both for estimated parameter values and predicted uncertainties. Additionally, we compared the use of a Random-Walk Metropolis-Hastings algorithm within or without a Gibbs algorithm, showing that even though the latter needs less iterations to converge, one iteration being four times more costly than a classical Random-Walk Metropolis-Hastings iteration, it is preferable to use a Random-Walk Metropolis-Hastings algorithm considering the computational time. Finally, we studied the posterior correlation between parameters, which enabled us to better understand the difficulties behind the estimation of hyperparameter p which codes for the decrease of the object PSD. Indeed, as we verified by changing its value, by trying to estimate it and by plotting the correlation between p and one of the PSF parameters, we noticed that the PSF estimation quality is sensitive to the choice of p, which has to be tuned carefully, according to the observed object.

In Chapter 6, we applied the new MMSE estimator to experimental images, in both astronomical and satellite contexts as well: the method is applied to an experimental image of Vesta taken on VLT/SPHERE, and to an experimental image of Envisat observed using ONERA's

AO bench on the 1.54 m-telescope MéO, at OCA. We first checked that for both contexts, the empirical image PSD fits the image PSD model we use in the method. Contrary to the simulated case, the convergence of the estimates is way slower and particularly difficult for r_0, which may require to study additional information (coming for instance from telemetry data) in order to have a more adapted prior distribution for r_0. The restored images however show several details on the surface of the asteroid (craters, albedo) / satellite (bright spots, antennas), which we can recognize from the synthetic images we have of Vesta and Envisat.

Concerning the posterior correlation between parameters discussed in the results on simulated data, we interpreted it as a difficulty to distinguish the contributions of the object and of the PSF to the image, the object PSD and the PSF having correlated parameters which impact the image PSD in the same way. In order to better separate these contributions, we studied the impact of adding a support constraint on the object, which should help getting rid of the indeterminacy. In Chapter 7, we discussed different options to model the support constraint, and gave different computational elements in order to perform the likelihood computation taking into account the labels. To finish, we tested the impact of several (known) supports, the true one and reasonable approximations of the support (over and under-estimated), with different noise levels, on a 128×128 image. These tests show the great improvement brought by the support knowledge, even approximate, on the estimated parameters: more particularly, r_0, which was precisely the one that P-AMIRAL could not estimate on the smaller object, is estimated extremely closely to its true value using the known, true or even approximate, support. Finally, the support knowledge enables us to estimate the hyperparameter p (that had to be given for the previous methods) without degrading the estimation of the PSF parameters.

Future outlook

The current results naturally lead to various questions related to the limitations of the method, thus to possible advances. Some of the prospects, foreseen or ongoing, are given in this section.

To begin with, we currently sample our parameters using either a Metropolis-Hastings-within-Gibbs or a Metropolis-Hastings algorithm, but these algorithms are not the fastest existing in the literature due to the fact that they do not take advantage of the gradient. To accelerate the convergence, a possible development is to use gradient-based methods, such as a Metropolis-adjusted Langevin algorithm [Robert, 2004; Vacar, 2011; Vacar, 2016]. We already provided gradient and Fisher information computation elements in this work, which could be used to switch from the current to faster methods. Additionally, the sampling methods applied in Chapters 4 to 6 can also be used in the support constraint framework developed in Chapter 7. Moreover, for broader results, it would be very meaningful to test our methods on more cases (namely try more data, different simulation conditions and more initializations).

Secondly, in this work, we supposed that the object PSD was circularly symmetric. However, some objects (especially satellites) can be far from isotropic and thus it would be worth considering an anisotropic, 2D PSD model and combining it to the marginal MMSE estimator.

For instance, an elliptical model was developed in [Lau, 2023] and tested on astronomical images. In this model, the hyperparameter p, coding for the object PSD decrease in log scale in medium-high frequencies, has been fixed to 3. In our work, p was fixed to a reasonable value according to the class of the object (either around 3 for asteroids, or around 2.5 satellites). We are currently working on an extension of this model, which would include a parameter similar to the previous p, that could be tuned according to the object class.

Indeed, the PSF estimation is sensitive to the choice of p, likely coming from the difficulty to separate both object and PSF contributions to the image, as discussed in this work. In order to enable the estimation of p, as well as of richer prior models for the object (implying the estimation of even more parameters), adding constraints on the object (namely, support and/or positivity constraints) seems necessary. Thus, one of the first perspectives would then be to adapt and use the our current MCMC method in order to estimate the support, that enables the estimation of p as shown in Chapter 7. Another complementary option would also be to add prior information on the PSF or on its parameters, using for instance data coming from the AO system itself (telemetry data).

Lastly, as starting assumptions in this work, we supposed that the PSF was a circularly symmetric, long-exposure PSF, meaning that the exposure time was bigger than the variation time of turbulence. These assumptions are somehow questionable in the case of the observation of low earth orbit satellites. Indeed, the PSF is likely asymmetric due to the fast evolution of these satellites (possibly fast rotation, short time of appearance) with a strong and directional wind. Additionally, for these observations it is interesting even necessary to reduce the exposure time, in order to have a “static” object in the image series and to exploit temporal redundancy in the series to improve the estimations. The exposure time then becomes closer to or even falls below the typical variation time of turbulence. Therefore, it would be crucial to adopt an asymmetric, short(er)-exposure PSF model for these applications. More globally, it could be key in this case to jointly re-design the system and the post-processing of the images, in order to work with short-exposure images and restore image (temporal) series, taking into account the temporal evolution of parameters, or the redundant information, to improve the estimations.

Appendix A: Gradients and Fisher information computation

In order to compute the gradient and the Fisher information for each estimated parameter, we first derive the criterion J_{mMAP} given in Equation 2.8 with respect to the image PSD:

$$\frac{\partial J_{mMAP}}{\partial S_i(f)} = -\frac{1}{2}(\frac{1}{S_i(f)} - \frac{|\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{(S_i(f))^2})$$

The gradient with respect γ_n, γ_o and k according to the chain rule, is quite easy to write:

$$\begin{aligned}
\frac{\partial J_{mMAP}}{\partial \gamma_n} &= \sum_f \frac{\partial J_{mMAP}}{\partial S_i(f)} \frac{\partial S_i(f)}{\partial \gamma_n} = -\frac{1}{2}\sum_f (\frac{1}{S_i(f)} - \frac{|\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{(S_i(f))^2})\frac{1}{\gamma_n^2} \\
\frac{\partial J_{mMAP}}{\partial \gamma_o} &= -\frac{1}{2}\sum_f (\frac{1}{S_i(f)} - \frac{|\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{(S_i(f))^2})\frac{|\tilde{h}(f)|^2}{\gamma_o^2(k+f^p)} \\
\frac{\partial J_{mMAP}}{\partial k} &= -\frac{1}{2}\sum_f (\frac{1}{S_i(f)} - \frac{|\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{(S_i(f))^2})\frac{|\tilde{h}(f)|^2}{\gamma_o(k+f^p)^2}
\end{aligned}$$

We also use the chain rule to write the gradient for any PSF parameter:

$$\begin{aligned}
\frac{\partial J_{mMAP}}{\partial \theta_{PSF}} &= \frac{\partial}{\partial \theta_{PSF}}[-\frac{1}{2}\sum_f (\ln S_i(f) + \frac{|\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{S_i(f)})] \\
&= -\frac{1}{2}\sum_f (\frac{1}{S_i(f)}\frac{\partial S_i(f)}{\partial \theta_{PSF}} + \frac{1}{S_i(f)}\frac{\partial |\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{\partial \theta_{PSF}} - \frac{|\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{S_i(f)^2}\frac{\partial S_i(f)}{\partial \theta_{PSF}})] \\
&= -\frac{1}{2}\sum_f (\frac{1}{S_i(f)(\gamma_o(k+f^p))}\frac{\partial |\tilde{h}(f)|^2}{\partial \theta_{PSF}} + \frac{1}{S_i(f)}\frac{\partial |\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{\partial \theta_{PSF}} \\
&- \frac{|\tilde{h}(f)\tilde{o}_m(f) - \tilde{\imath}(f)|^2}{(S_i(f))^2(\gamma_o(k+f^p))}\frac{\partial |\tilde{h}(f)|^2}{\partial \theta_{PSF}})
\end{aligned}$$

We now have to compute partial derivatives in the form $\frac{\partial |a\tilde{h}(f) + b|^2}{\partial \theta_{PSF}}$, with $a \in \mathbb{R}$ and $b \in \mathbb{R}$.

$$\begin{aligned}
\frac{\partial |a\tilde{h}(f)+b|^2}{\partial \theta_{PSF}} &= \frac{\partial}{\partial \theta_{PSF}}[\Re(a\tilde{h}(f)+b)^2] + \frac{\partial}{\partial \theta_{PSF}}[\Im(a\tilde{h}(f)+b)^2] \\
&= 2\Re(a\tilde{h}(f)+b)\frac{\partial}{\partial \theta_{PSF}}[\Re(a\tilde{h}(f)+b)] + 2\Im(a\tilde{h}(f)+b)\frac{\partial}{\partial \theta_{PSF}}[\Im(a\tilde{h}(f)+b)] \\
&= 2\Re(a\tilde{h}(f)+b)\times a\frac{\partial \Re(\tilde{h}(f))}{\partial \theta_{PSF}} + 2\Im(a\tilde{h}(f)+b)\times a\frac{\partial \Im(\tilde{h}(f))}{\partial \theta_{PSF}} \\
&= 2\Re\{(\Re(a\tilde{h}(f)+b)+i\Im(a\tilde{h}(f)+b))\times a(\frac{\partial \Re(\tilde{h}(f))}{\partial \theta_{PSF}} - i\frac{\partial \Im(\tilde{h}(f))}{\partial \theta_{PSF}})\} \\
&= 2\Re[(a\tilde{h}(f)+b)\times a\frac{\partial \tilde{h}(f)^*}{\partial \theta_{PSF}}]
\end{aligned}$$

In [Fétick, 2020a], the computation of $\frac{\partial \tilde{h}(f)}{\partial \theta_{PSF}}$ is explicited.

In [Vacar, 2014], the Fisher information is computed in the case where the mean object $\boldsymbol{o}_m = \mathbf{0}$. In this case, given that the OTF (therefore, the PSF parameters) does not appear in the quadratic term anymore $|\tilde{h}(f)\tilde{o}_m(f) - \tilde{i}(f)|^2 = |\tilde{i}(f)|^2$, the first and the second derivative of the criterion with respect to any parameter θ_p writes:

$$\begin{aligned}
\frac{\partial J_{mMAP}}{\partial \theta_p} &= \frac{\partial J_{mMAP}}{\partial S_i}\frac{\partial S_i}{\partial \theta_p} \\
&= -\frac{1}{2}\sum_f (\frac{1}{S_i(f)} - \frac{|\tilde{i}(f)|^2}{S_i(f)^2})\frac{\partial S_i(f)}{\partial \theta_p} \\
\frac{\partial^2 J_{mMAP}}{\partial \theta_p \partial \theta_q} &= -\frac{1}{2}\sum_f \frac{\partial}{\partial \theta_q}\Big[(\frac{1}{S_i(f)} - \frac{|\tilde{i}(f)|^2}{S_i(f)^2})\frac{\partial S_i(f)}{\partial \theta_p}\Big] \\
&= -\frac{1}{2}\sum_f \Big((\frac{1}{S_i(f)} - \frac{|\tilde{i}(f)|^2}{S_i(f)^2})\frac{\partial^2 S_i(f)}{\partial \theta_p \partial \theta_q} + \frac{\partial S_i(f)}{\partial \theta_p}\frac{\partial S_i(f)}{\partial \theta_q}(- \frac{1}{S_i(f)^2} + 2\frac{|\tilde{i}(f)|^2}{S_i(f)^3})\Big)
\end{aligned}$$

In our case, we consider a constant mean image, meaning that $\tilde{o}_m(f) = 0, \forall f \neq 0$. Additionally, we also supposed that $\tilde{h}(0) = 1$ in Chapter 1, meaning that $|\tilde{h}(f)\tilde{o}_m(f) - \tilde{i}(f)|^2$ does not depend on the PSF parameters, therefore these expressions of the first and second derivatives are also valid for any parameter, in the case of a constant $\boldsymbol{o}_m$.

Given that the Fisher information writes $I(\theta_p, \theta_q) = -\mathbb{E}_{\boldsymbol{i}|\boldsymbol{\theta}}\Big[\frac{\partial^2 J_{mMAP}}{\partial \theta_p \partial \theta_q}\Big]$ and that $\mathbb{E}_{\boldsymbol{i}|\boldsymbol{\theta}}[|\tilde{i}(f)|^2] = S_i(f)$, the Fisher information is simply:

$$I(\theta_p, \theta_q) = \frac{1}{2}\sum_f \frac{1}{S_i(f)^2}\frac{\partial S_i(f)}{\partial \theta_p}\frac{\partial S_i(f)}{\partial \theta_q}$$

and can be computed given the derivatives computed previously.

Appendix B: Calendar and tips for final-year PhD students

Dear colleague,

Here is a recap of things you must/should do, in chronological order, during the final year of your PhD. I am writing this recap trying to prevent you from unpleasant, last-minute surprises. However, this recap is ONERA-related, ED127-related, and non-exhaustive given that it is based on my personal experience. I hope deep down that the PhD students who will read this will also modify some items / add more relevant or missing information, maybe even create an online wiki about it, and that the information will be shared throughout the years. Finally, I would like to thank my friends who co-wrote/proofread this appendix!

Remark: The given chronology is a backwards chronology starting from the **defense date**. Following the rules, the defense date is before the end of your contract (given that you are not paid afterwards). One often thinks about a thesis extension at least 6 months before the end of contract = the expected defense date. If you are willing to extend your PhD, I would strongly recommend to start discussing it early, as it is a long and not necessarily successful process, especially at ONERA, even if your funding might be coming from external sources. ONERA has the final word on thesis extensions, no matter the funding.

AT LEAST 7 MONTHS BEFORE: Plan, manuscript language

1) Think about the writing plan with your supervisors, make a detailed plan. It is really important to think about your writing plan in advance and to start "on time". This plan must include buffer time that will be used for: delays (that will occur for sure), reading time for your supervisors, time to take into account your supervisors' comments, simulations/lab work to do or re-do, your vacation (absolutely make sure to plan some! you will need it), your supervisors' vacations (they will not stop themselves from taking it and you better know in beforehand when they are not around to proofread), as well as potential sick leave (we do not choose when we get sick). Then, the rapporteurs need at least 6 weeks to write their report. Finally, the doctoral school asks for 1 month between the reports and the defense, in which also the defense preparation will be done. This means that you should keep about 10 weeks between your manuscript deadline and your defense date. It is much better to overestimate the buffer time than to underestimate it.

2) Think about the manuscript language. For instance, you may consider writing in English, especially if one of your potential reviewers is a non-French speaker, or if you want to ensure

a broader availability of your manuscript. To write your manuscript in English, you need to send a request to the doctoral school before you start writing. Additionally, if you write your manuscript in English, another summary (longer than the abstract) in French is necessary.

4-5 MONTHS BEFORE: Defense committee members

3) Discuss the composition of the defense committee with your supervisors. Be careful of the constraints to meet (as of 2023): there must be between 4 and 8 committee members, there should be at least 50% of “rang A”, at least 50% of external people and at least 25% of women. Afterwards, you need to:

- Contact your doctoral school in order to validate the defense committee,
- Then contact all committee members (with your advisors),
- Once your committee is all set and validated by the doctoral school, make a poll (with all half-days of your defense month), and send it to your committee,
- Set a defense date with your committee and search for a defense room (you can ask your supervisors or the doctoral school for ideas).

3-4 MONTHS BEFORE: ADUM form, finalizing the manuscript

4) Fill in the ADUM form for the defense (typically, for contracts that started before October 1st, do this before the summer break). Official deadline, 3 months before the defense date.

5) Contact the rapporteurs regarding the exact manuscript shipping date. Ask if they need a paper version: if so, get in touch as soon as possible with your supervisors and the secretariat so that the lab can take care of shipping.

Tip: For postal shipping, give yourself at least 10 days (send it via email to imprimerie@onera.fr), then take the paper version to the postal service (before the postal collection at 9:00 am).

1-2 MONTHS PRIOR: Preparing the defense presentation

6) Now that you have handed in your manuscript, talk to your supervisors about any sort of handover (for example code, documentation, ...) that you need to deal with before the end of your work contract.

7) Fill in the CADO form and upload the first version of your manuscript to it.

8) On ADUM, sign the thesis charter (“charte de thèse”) and the conformity certificate (“certificat de conformité”), then send it to the doctoral school. Upload the thesis summaries in English and in French.

9) Start preparing your defense presentation, ask your supervisors how they expect you to prepare it.

1 MONTH PRIOR (after receiving the reports): Organizing the defense

10) On ADUM, check with your supervisors that all the data is correct: defense date and place, correct spelling for the names and first names of your committee, no abbreviations, exact rank, etc. This matters because every document that gets out will come from this data in ADUM, including the diploma.

11) In case some committee members wish to join the defense remotely: ask your doctoral school if it is possible and check the corresponding options on ADUM. The doctoral school will then send you documents on the video conference, that you need to sign.

12) Check that the secretariat is booking train/plane tickets for your committee to come to your defense or at least refunding part of their journey (or ask your supervisors). Your supervisors will likely organize a formal lunch with your committee on the same day like your defense.

13) Apart from the official organization, you will likely have other things to take care of too: Coordinating friends and family coming in for your defense from out of town, the “pot de thèse” (the official reception right after the deliberation of the committee, including food and drinks with the defense committee and your guests), your defense party. Ask far enough in advance whether your lab financially supports the pot de thèse or even organizes it for you. You can also ask your lab/doctoral school (according to where you booked your defense room) either if there is a room for the pot de thèse, available the day of your defense and rather close to your defense room. For all of this, do not hesitate to designate a dedicated buddy or two who can take care of food and drink organization with/for you, as well as help you find a place for your defense party. It will be a stressful time, and it can be very helpful to have some of that taken out of your hands.

2 WEEKS BEFORE: Live stream and final checks

14) If there is a live stream of your defense (and/or if there are committee members remotely), prepare it in advance and check that it is working. Again, do not hesitate to designate a dedicated buddy or two who can take care of the stream during your defense.

15) Send invitations to your defense, including the live stream link.

16) Send to the doctoral school / lab secretariat your defense guests list (according to where you booked your defense room), check that the security has all the required information for your guests to enter.

17) Check that everything is ready for your defense: paper versions of the manuscript and water bottles (at least 1 per committee member and 1 for you), computer hardware (microphone, camera, sound system, laser pointer,...).

AFTER THE DEFENSE

Congratulations doctor! It is very likely that the defense committee has given you some comments on your manuscript that you can improve before submitting its final version on ADUM. You will receive your diploma only:

- after you submit the final version of the manuscript on ADUM,
- after you sign the online release authorization ("autorisation de mise en ligne"),
- after you and your director sign the conformity certificate ("certificat de conformité") of the final version of the manuscript. Official deadline, 3 months after the defense date.

AFTER THE END OF CONTRACT

Oftentimes, you will be attributed the "visiteur privilégié" status at ONERA. This means that you retain access to your ONERA email, ONERA internal systems and your laptop. It can be helpful if you do not have a personal laptop to do the manuscript edits on. However, bear in mind that you will be unemployed and unpaid by ONERA, so my recommendation is to get any technical work (including code handover and documentation) done before. As envisaged by ONERA organization, the "visiteur privilégié" status is destined to allow the PhD students to finalize their manuscripts exclusively.

List of publications and communications

Articles

A. Yan, L. Mugnier, J.-F. Giovannelli, R. J.-L. Fétick, C. Petit. "*Marginalized blind deconvolution of Adaptive Optics corrected images using MCMC methods*". JATIS, November 2023. DOI: 10.1117/1.JATIS.9.4.048004

International Conferences (first author)

A. Yan, L. Mugnier, J.-F. Giovannelli, R. J.-L. Fétick, C. Petit. "*Extending AMIRAL's blind deconvolution of adaptive optics corrected images with Markov chain Monte Carlo methods*". SPIE Astronomical Telescopes and Instrumentation, Montréal, 2022.
DOI: 10.1117/12.2627414

A. Yan, L. Mugnier, J.-F. Giovannelli, R. J.-L. Fétick, C. Petit. "*Marginalized semi-blind restoration of Adaptive-Optics-corrected images using stochastic sampling*". AO4ELT, Avignon, 2023 (To be published).

National Conferences (first author)

A. Yan, R. J.-L. Fétick, L. Mugnier, J.-F. Giovannelli, A. Bonnefois, C. Petit. "*Restauration d'images de satellites corrigées par optique adaptative dans le visible*". JRIOA'8, Dijon, 2021.

A. Yan, L. Mugnier, J.-F. Giovannelli, R. J.-L. Fétick, C. Petit. "*Restauration d'images corrigées par optique adaptative : approche marginale avec échantillonnage bayésien*". JIONC, Paris, 2022.

A. Yan, L. Mugnier, J.-F. Giovannelli, R. J.-L. Fétick, C. Petit. "*Extending marginalized blind deconvolution of AO corrected astronomical images with MCMC methods*". ELBERETH, Paris, 2022.

A. Yan, L. Mugnier, J.-F. Giovannelli, R. J.-L. Fétick, C. Petit. "*Restauration d'images astronomiques corrigées par optique adaptative : méthode marginale étendue par algorithme MCMC*". GRETSI, Nancy, 2022.

A. Yan, L. Mugnier, J.-F. Giovannelli, R. J.-L. Fétick, C. Petit. "*Restauration myope marginale d'images corrigées par optique adaptative avec échantillonnage stochastique*". JIONC, Paris, 2023.

International Workshops

A. Yan, R. J.-L. Fétick, L. Mugnier, J.-F. Giovannelli, A. Bonnefois, C. Petit. "*Image restoration applied to adaptive-optics-corrected satellite observation in the visible*". AO4ASTRO2, 2021.

A. Yan, L. Mugnier, J.-F. Giovannelli, R. J.-L. Fétick, C. Petit. "*Extending marginalized blind deconvolution of AO corrected astronomical images with MCMC methods*". “Interfacing Bayesian statistics, machine learning, applied analysis, and blind and semi-blind imaging inverse problems”, Edinburgh, 2023.

Bibliography

[Altmann, 2013] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret. « Residual Component Analysis of Hyperspectral Images—Application to Joint Nonlinear Unmixing and Nonlinearity Detection ». *Image Processing, IEEE Transactions on* 23 (2013) (cit. on p. 88).

[Andersen, 2006] D. R. Andersen, J. Stoesz, S. Morris, M. Lloyd-Hart, D. Crampton, T. Butterley, et al. « Performance Modeling of a Wide-Field Ground-Layer Adaptive Optics System ». *Publications of the Astronomical Society of the Pacific* 118.849 (2006), pp. 1574–1590 (cit. on p. 15).

[Aristidi, 2020] É. Aristidi, Y. Fanteï-Caujolle, A. Ziad, J. Chabé, C. Renaud, C. Giordano, et al. « Turbulence monitoring at Calern observatory with the generalized differential image motion monitor ». *Adaptive Optics Systems VII.* Ed. by Dirk Schmidt, Laura Schreiber, and Elise Vernet. SPIE, 2020 (cit. on p. 11).

[Ayasso, 2010] H. Ayasso and A. Mohammad-Djafari. « Joint NDT Image Restoration and Segmentation Using Gauss–Markov–Potts Prior Models and Variational Bayesian Computation ». *Image Processing, IEEE Transactions on* 19 (2010), pp. 2265–2277 (cit. on p. 88).

[Ayers, 1988] G. R. Ayers and J. C. Dainty. « Iterative blind deconvolution method and its applications ». *Opt. Lett.* 13.7 (1988), pp. 547–549 (cit. on pp. 6, 22).

[Babcock, 1953] H. W. Babcock. « The Possibility of Compensating Astronomical Seeing ». *PASP* 65.386 (1953), p. 229 (cit. on p. 12).

[Beuzit, 2019] J.-L. Beuzit, A. Vigan, D. Mouillet, K. Dohlen, R. Gratton, A. Boccaletti, et al. « SPHERE: the exoplanet imager for the Very Large Telescope ». *Astronomy & Astrophysics* 631 (2019), A155 (cit. on p. 18).

[Bishop, 2007] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics).* 1st ed. Springer, 2007 (cit. on p. 44).

[Blanco, 2011] L. Blanco and L. M. Mugnier. « Marginal blind deconvolution of adaptive optics retinal images ». *Optics Express* 19 (2011), pp. 23227–23239 (cit. on pp. 3, 6, 24–26, 28, 34, 94).

[Conan, 1994] J.-M. Conan. « Etude de la correction partielle en optique adaptative ». PhD thesis. 1994, 343 P. (Cit. on p. 14).

[Conan, 1998] J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau, and G. Rousset. « Myopic deconvolution of adaptive optics images by use of object and point-spread function power spectra ». *Appl. Opt.* 37 (1998), pp. 4614–4622 (cit. on p. 31).

[Demoment, 1989] G. Demoment. « Image reconstruction and restoration: overview of common estimation structures and problems ». *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37.12 (1989), pp. 2024–2036 (cit. on p. 17).

[Duane, 1987] S. Duane, A.D. Kennedy, B. J. Pendleton, and D. Roweth. « Hybrid Monte Carlo ». *Physics Letters B* 195.2 (1987), pp. 216–222 (cit. on p. 45).

[Fétick, 2020a] R. Fétick. « Traitement d'image en optique adaptative : estimation paramétrique de la réponse impulsionnelle et déconvolution ». Theses. Aix Marseille Université, 2020 (cit. on pp. 25, 40, 72, 99).

[Fétick, 2019a] R. J.-L. Fétick, T. Fusco, B. Neichel, L. M. Mugnier, O. Beltramo-Martin, A. Bonnefois, et al. « Physics-based model of the adaptive-optics-corrected point spread function ». *Astronomy & Astrophysics* 628 (2019), A99 (cit. on pp. 3, 4, 6, 14, 15, 28, 34, 42, 50, 71, 84, 94).

[Fétick, 2020b] R. J.-L. Fétick, L. M. Mugnier, T. Fusco, and B. Neichel. « Blind deconvolution in astronomy with adaptive optics: the parametric marginal approach ». *MNRAS* 496 (2020), pp. 4209–4220 (cit. on pp. 3, 4, 6, 26, 34, 35, 39, 42, 52, 60–63, 71, 72, 88, 94).

[Fétick, 2019b] R. J.L. Fétick, L. Jorda, P. Vernazza, M. Marsset, A. Drouard, Thierry Fusco, et al. « Closing the gap between Earth-based and interplanetary mission observations: Vesta seen by VLT/-SPHERE ». *Astronomy & Astrophysics* 623 (2019), A6 (cit. on pp. 38, 70, 73).

[Fried, 1966] D. L. Fried. « Optical Resolution Through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures ». *J. Opt. Soc. Am.* 56.10 (1966), pp. 1372–1379 (cit. on p. 9).

[Fusco, 2016] T. Fusco, J.-F. Sauvage, D. Mouillet, A. Costille, C. Petit, J.-L. Beuzit, et al. « SAXO, the SPHERE extreme AO system: on-sky final performance and future improvements ». *Adaptive Optics Systems V*. Ed. by Enrico Marchetti, Laird M. Close, and Jean-Pierre Véran. Vol. 9909. International Society for Optics and Photonics. SPIE, 2016, 99090U (cit. on pp. 18, 23).

[Gamerman, 2006] D. Gamerman and H. Lopes. « Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference » (2006) (cit. on p. 45).

[Gelman, 1997] A. Gelman, W. R. Gilks, and G. O. Roberts. « Weak convergence and optimal scaling of random walk Metropolis algorithms ». *The Annals of Applied Probability* 7.1 (1997), pp. 110–120 (cit. on pp. 50, 57, 72).

[Geman, 1984] S. Geman and D. Geman. « Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6.6 (1984), pp. 721–741 (cit. on p. 88).

[Gilavert, 2013] C. Gilavert, S. Moussaoui, and J. Idier. « Rééchantillonnage gaussien en grande dimension pour les problèmes inverses ». *GRETSI 2013*. Brest, France, 2013 (cit. on p. 91).

[Girolami, 2011] M. Girolami and B. Calderhead. « Riemann Manifold Langevin and Hamiltonian Monte Carlo Methods ». *Journal of the Royal Statistical Society Series B: Statistical Methodology* 73.2 (2011), pp. 123–214 (cit. on pp. 45, 47).

[Hadamard, 1902] J. Hadamard. « Sur les problèmes aux dérivés partielles et leur signification physique ». *Princeton University Bulletin* 13 (1902), pp. 49–52 (cit. on p. 21).

[Idier, 2008] J. Idier. *Bayesian Approach to Inverse Problems*. ISTE Ltd and John Wiley & Sons Inc, 2008, p. 384 (cit. on p. 21).

[Jorda, 2010] L. Jorda, S. Spjuth, H. Keller, P. Lamy, and A. Llebaria. « OASIS: a simulator to prepare and interpret remote imaging of solar system bodies ». *Computational Imaging VIII*. Vol. 7533. SPIE. 2010, p. 753311 (cit. on pp. 10, 34, 38, 50).

[Kass, 1996] R. E. Kass and L. Wasserman. « The Selection of Prior Distributions by Formal Rules ». *Journal of the American Statistical Association* 91 (1996), pp. 1343–1370 (cit. on pp. 43, 89).

[Kolmogorov, 1941] A. N. Kolmogorov. « The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers ». *Dokl. Akad. Nauk SSSR*. Vol. 30. 1941, pp. 301–305 (cit. on p. 9).

[Labeyrie, 1970] A. Labeyrie. « Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images ». *SPIE milestone series* 28 (1970), pp. 427–429 (cit. on p. 12).

[Lafrenière, 2007] D. Lafrenière, C. Marois, R. Doyon, D. Nadeau, and É. Artigau. « A New Algorithm for Point-Spread Function Subtraction in High-Contrast Imaging: A Demonstration with Angular Differential Imaging* ». *The Astrophysical Journal* 660.1 (2007), p. 770 (cit. on p. 28).

[Lau, 2023] A. Lau, R. J.-L. Fétick, B. Neichel, O. Beltramo-Martin, and T. Fusco. « Improved prior for adaptive optics point spread function estimation from science images: Application for deconvolution ». *Astronomy & Astrophysics* 673 (2023), A72 (cit. on pp. 4, 85, 97).

[Lehmann, 1998] E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Second. New York, NY, USA: Springer-Verlag, 1998 (cit. on p. 6).

[Levin, 2009] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. « Understanding and evaluating blind deconvolution algorithms ». *IEEE Conference on Computer Vision and Pattern Recognition.* 2009, pp. 1964–1971 (cit. on pp. 3, 6, 26).

[Mugnier, 2008] L. Mugnier. « Des données à la connaissance de l'objet : le problème inverse ». *L'observation en astrophysique.* Ed. by P. Léna, D. Rouan, F. Lebrun, F. Mignard, and D. Pelat. Les Ulis, France: EDP Sciences, 2008. Chap. 9, section 6, pp. 591–613 (cit. on p. 21).

[Mugnier, 2004] L. M. Mugnier, T. Fusco, and J.-M. Conan. « MISTRAL: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images ». *JOSA* 21 (2004), pp. 1841–1854 (cit. on pp. 6, 17, 23).

[Mugnier, 2001] L. M. Mugnier, C. Robert, J.-M. Conan, V. Michau, and S. Salem. « Myopic deconvolution from wave-front sensing ». *J. Opt. Soc. Am. A* 18.4 (2001), pp. 862–872 (cit. on p. 64).

[Orieux, 2012] F. Orieux, O. Feron, and J.-F. Giovannelli. « Sampling High-Dimensional Gaussian Distributions for General Linear Inverse Problems ». *IEEE Signal Processing Letters* 19.5 (2012), pp. 251–254 (cit. on p. 91).

[Orieux, 2010] F. Orieux, J.-F. Giovannelli, and T. Rodet. « Bayesian estimation of regularization and point spread function parameters for Wiener-Hunt deconvolution ». *JOSA* 27 (2010), pp. 1593–1607 (cit. on pp. 45, 46).

[Orieux, 2013] F. Orieux, J.-F. Giovannelli, T. Rodet, and A. Abergel. « Estimating hyperparameters and instrument parameters in regularized inversion. Illustration for Herschel/SPIRE map making ». *Astronomy & Astrophysics* 549 (2013), A83 (cit. on pp. 45, 46).

[Petit, 2020] C. Petit, L. M. Mugnier, A. Bonnefois, J.-M. Conan, T. Fusco, N. Levraud, et al. « LEO satellite imaging with adaptive optics and marginalized blind deconvolution ». *21st AMOS Advanced Maui Optical and Space Surveillance Technologies Conference.* United States, 2020 (cit. on pp. 15, 19, 34, 64, 73, 76).

[Primot, 1990] J. Primot, G. Rousset, and J. C. Fontanella. « Deconvolution from wave-front sensing: a new technique for compensating turbulence-degraded images ». *J. Opt. Soc. Am. A* 7.9 (1990), pp. 1598–1608 (cit. on p. 12).

[Qi, 2002] Y. Qi and T. Minka. « Hessian-based Markov Chain Monte-Carlo Algorithms ». 2002 (cit. on p. 47).

[Ramani, 2008] S. Ramani, D. Van De Ville, T. Blu, and M. Unser. « Nonideal Sampling and Regularization Theory ». *IEEE Trans. Signal Process.* 56.3 (2008), pp. 1055–1070 (cit. on p. 31).

[Rigaut, 1998] F. J. Rigaut, J.-. Veran, and O. Lai. « Analytical model for Shack-Hartmann-based adaptive optics systems ». *Adaptive Optical System Technologies.* Ed. by Domenico Bonaccini and Robert K. Tyson. Vol. 3353. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. 1998, pp. 1038–1048 (cit. on p. 13).

[Robert, 2004] C. P. Robert and G. Casella. *Monte Carlo statistical methods.* Vol. 2. Springer, 2004 (cit. on pp. 45, 46, 96).

[Roberts, 2002] G. O. Roberts and O. Stramer. « Langevin Diffusions and Metropolis-Hastings Algorithms ». *Methodology And Computing In Applied Probability* 4 (2002), pp. 337–357 (cit. on p. 47).

[Roddier, 1981] F. Roddier. « The effects of atmospheric turbulence in optical astronomy ». *Progress in Optics* 19 (1981), pp. 281–376 (cit. on pp. 10, 11).

[Rousset, 1990] G. Rousset, J. Fontanella, Pierre Kern, P. Gigan, and Francois Rigaut. « First diffraction-limited astronomical images with adaptive optics ». *Astronomy and Astrophysics* 230 (1990), pp. L29–L32 (cit. on p. 12).

[Sánchez, 2006] F. Mueller Sánchez, R. I. Davies, F. Eisenhauer, L. J. Tacconi, R. Genzel, and A. Sternberg. « SINFONI adaptive optics integral field spectroscopy of the Circinus Galaxy ». *Astronomy & Astrophysics* 454.2 (2006), pp. 481–492 (cit. on p. 15).

[Tatarski, 1961] V. I. Tatarski, R. A. Silverman, and Nicholas Chako. « Wave Propagation in a Turbulent Medium ». *Physics Today* 14.12 (1961), p. 46 (cit. on p. 9).

[Thiébaut, 2002] É. Thiébaut. « Optimization issues in blind deconvolution algorithms ». *Proceedings of SPIE - The International Society for Optical Engineering* (2002) (cit. on pp. 23, 29).

[Vacar, 2011] C. Vacar, J.-F. Giovannelli, and Y. Berthoumieu. « Langevin and Hessian with Fisher approximation stochastic sampling for parameter estimation of structured covariance ». *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2011, pp. 3964–3967 (cit. on p. 96).

[Vacar, 2014] C. Vacar, J.-F. Giovannelli, and Y. Berthoumieu. « Bayesian Texture and Instrument Parameter Estimation From Blurred and Noisy Images Using MCMC ». *IEEE Signal Processing Letters* 21.6 (2014), pp. 707–711 (cit. on p. 99).

[Vacar, 2016] C. Vacar, J.-F. Giovannelli, and Y. Berthoumieu. « Bayesian Texture Classification From Indirect Observations Using Fast Sampling ». *IEEE Transactions on Signal Processing* 64.1 (2016), pp. 146–159 (cit. on pp. 45, 47, 96).

[Villeneuve, 2012] E. Villeneuve and H. Carfantan. « Hyperspectral data deconvolution for galaxy kinematics with MCMC ». *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. 2012, pp. 2477–2481 (cit. on pp. 45, 46).

[Yan, 2022] A. Yan, L. M. Mugnier, J.-F. Giovannelli, R. Fétick, and C. Petit. « Restauration d'images astronomiques corrigées par optique adaptative : méthode marginale étendue par algorithme MCMC ». *GRETSI 2022*. Nancy, France, 2022 (cit. on p. 63).

[You, 1996] Y.-L. You and M. Kaveh. « A regularization approach to joint blur identification and image restoration ». *IEEE Transactions on Image Processing* 5.3 (1996), pp. 416–428 (cit. on pp. 6, 22).

RÉSUMÉ

La restauration d'images corrigées par optique adaptative est particulièrement difficile, du fait de la méconnaissance de la réponse impulsionnelle du système optique (PSF pour point-spread function) en plus des difficultés usuelles. Une approche efficace est de marginaliser l'objet en dehors du problème et d'estimer la PSF et les hyper-paramètres (liés à l'objet et au bruit) seuls avant la déconvolution. Des travaux récents ont appliqué cette déconvolution marginale, combinée à un modèle paramétrique de PSF, à des images astronomiques et de satellites. Cette thèse vise à proposer une extension de cette méthode. En particulier, j'utilise un algorithme Monte-Carlo par chaînes de Markov (MCMC), afin d'inclure des incertitudes sur les paramètres et d'étudier leur corrélation a posteriori. Je présente des résultats détaillés obtenus sur des images astronomiques et de satellites, simulées et expérimentales. Je présente également des premiers éléments sur l'ajout d'une contrainte de support sur l'objet.

MOTS CLÉS

restauration d'image, optique adaptative, déconvolution, turbulence, problèmes inverses

ABSTRACT

Adaptive-optics-corrected image restoration is particularly difficult, as it suffers from the poor knowledge on the point-spread function (PSF). One efficient approach is to marginalize the object out of the problem, and to estimate the PSF and (object and noise) hyper-parameters only before the deconvolution. Recent works have applied this marginal deconvolution, combined to a parametric model for the PSF, to astronomical and satellite images. This thesis aims at extending this previous method, using Markov chain Monte Carlo (MCMC) algorithms. This will enable us to derive uncertainties on the estimates, as well as to study posterior correlation between the parameters. I present detailed results on simulated and experimental, astronomical and satellite data. I also provide elements on the impact of a support constraint on the object.

KEYWORDS

image restoration, adaptive optics, deconvolution, turbulence, inverse problems