



**HAL**  
open science

# Inversion bayésienne myope et non-supervisée pour l'imagerie sur-résolue. Application à l'instrument SPIRE de l'observatoire spatial Herschel.

François Orieux

► **To cite this version:**

François Orieux. Inversion bayésienne myope et non-supervisée pour l'imagerie sur-résolue. Application à l'instrument SPIRE de l'observatoire spatial Herschel.. Traitement du signal et de l'image [eess.SP]. Université Paris Sud - Paris XI, 2009. Français. NNT: . tel-00433540v2

**HAL Id: tel-00433540**

**<https://theses.hal.science/tel-00433540v2>**

Submitted on 20 Nov 2009 (v2), last revised 22 Nov 2009 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° D'ORDRE : 9648



UNIVERSITÉ PARIS-SUD 11  
Faculté des Sciences d'Orsay



**THÈSE DE DOCTORAT**

**SPÉCIALITÉ : PHYSIQUE**

*École Doctorale « Sciences et Technologies de l'Information, des  
Télécommunications et des Systèmes »*

Présentée par :

**François ORIEUX**

Sujet :

**Inversion bayésienne myope et non-supervisée pour l'imagerie  
sur-résolue. Application à l'instrument SPIRE de l'observatoire  
spatial Herschel.**

Soutenue le 16 novembre 2009 devant les membres du jury :

Alain Abergel	Professeur de l'Université Paris-Sud 11	Co-encadrant
Jean-François Giovannelli	Professeur de l'Université Bordeaux 1	Directeur de thèse
Jérôme Idier	Directeur de recherche CNRS	Rapporteur
Guy Le Besnerais	Chercheur ONERA	Examineur
Thomas Rodet	Maître de conférence de l'Université Paris-Sud 11	Co-encadrant, invité
Sylvie Roques	Directeur de recherche CNRS	Rapporteur
Éric Thiébaud	Astronome adjoint	Examineur

---

Inversion bayésienne myope et non-supervisée pour l'imagerie sur-résolue. Application à l'instrument **SPIRE** de l'observatoire spatial Herschel.

François ORIEUX

---

# Remerciements

Il est difficile, me semble-t-il, de savoir et de comprendre ce qu'est une démarche rationnelle et scientifique. A force d'y réfléchir on établit des principes pour mener cette démarche scientifique mais alors c'est leur application rigoureuse qui est difficile. Ce qui par contre semble plus facile c'est de trop se fier à ses idées et d'agir avec certitude pour se rassurer. C'est pourquoi je tiens à remercier l'ensemble des personnes qui m'ont permis d'explorer au mieux ces questions relatives à la rationalité, la démarche scientifique et l'ouverture (très riches dans un contexte où l'on parle d'incertitude et d'inférence!).

Laisser la place à l'apprentissage sous-entend laisser la place à l'erreur et l'imperfection. Cela sous-entend également, lors d'une discussion, de laisser le temps à la personne de réfléchir et d'élaborer sa pensée à voix haute. Je remercie Jean-François Giovannelli pour cette grande qualité que j'ai rarement rencontrée.

J'ai commencé la recherche en stage, encadré par Thomas Rodet. Il a fourni des efforts conséquents pour monter un sujet de thèse et trouver un financement. Je le remercie pour l'ensemble de ce qu'il a fait pour moi pendant mon stage et ma thèse, sur l'encadrement matériel, scientifique et humain quotidien ainsi que tous les à côtés qui sont essentiels.

L'approche proposée dans ce manuscrit emploie des informations qualifiées d'*a priori*. Cependant certains ont des réticences lorsque l'usage de cette nature d'information est explicite. Il arrive même parfois que la mention même du mot *a priori* entraîne des réactions somme toutes étonnantes dans un milieu scientifique composé de gens raisonnables. Alors je remercie Alain Abergel pour son ouverture que je trouve exemplaire. De plus sa critique positive et son expérience de l'analyse quantitative de résultats ont beaucoup contribué à ma vision du métier.

Je dois beaucoup à mes trois encadrants, et je n'ai mentionné qu'une petite partie de ma reconnaissance. D'autres personnes ont également compté et je tiens à remercier en particulier Guy Demoment dont la réflexion, l'étendue des connaissances et l'éthique m'ont beaucoup inspiré dans mes réflexions professionnelles et personnelles.

Je tiens également à remercier mes rapporteurs Jérôme Idier et Sylvie Roques pour leurs remarques et l'examen en détail de mon travail. Je remercie également les examinateurs, Guy Le Besnerais et Éric Thiébaud. Je souhaite remercier l'ensemble des personnes ayant constitué mon environnement scientifique quotidien : Hacheme, Olivier, Aurélien, Nadia, Ali, et plus récemment Nicolas, Aurélia, Dorian, Sha. Une mention spéciale pour Diara.

Enfin merci à ma famille pour m'avoir toujours soutenu. Mon goût pour le savoir, les sciences et la raison ne vient pas de nul part.

Et merci à toi, Camille Maillard. Je suis fier et heureux de mener, à tes côtés, « le combat ordinaire ». Avec toi tout est mieux et je suis partout chez moi. Tu m'es exceptionnelle.

---

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Contexte . . . . .	9
1.1.1	Herschel et SPIRE . . . . .	10
1.1.2	Inversion . . . . .	12
1.1.3	Problématiques . . . . .	15
1.2	Inférence bayésienne . . . . .	16
1.2.1	Loi <i>a posteriori</i> . . . . .	17
1.2.2	Définition d'un estimateur . . . . .	17
1.3	Apport de l'approche bayésienne . . . . .	19
1.3.1	Vraisemblance des données . . . . .	19
1.3.2	Information <i>a priori</i> . . . . .	20
1.3.3	Estimation non-supervisée . . . . .	22
1.3.4	Estimation myope . . . . .	24
1.3.5	Conclusion . . . . .	25
1.4	Algorithmes de calcul . . . . .	25
1.4.1	Optimisation . . . . .	26
1.4.2	Intégration . . . . .	28
1.4.3	Conclusion . . . . .	31
1.5	MADMAP et SANEPIC . . . . .	31
1.5.1	MADMAP . . . . .	31
1.5.2	SANEPIC . . . . .	34
1.6	État de l'art en déconvolution et sur-résolution . . . . .	34
1.6.1	Déconvolution . . . . .	35
1.6.2	Sur-résolution . . . . .	35
1.7	Conclusion . . . . .	36
<b>2</b>	<b>Modèle d'acquisition pour SPIRE</b>	<b>37</b>
2.1	Introduction . . . . .	37
2.2	Modèles physiques . . . . .	38
2.2.1	Ouverture . . . . .	38

2.2.2	Filtres . . . . .	40
2.2.3	Cornets . . . . .	41
2.2.4	Bolomètres . . . . .	42
2.2.5	Électronique de lecture . . . . .	46
2.2.6	Mode d’observation . . . . .	47
2.2.7	Équation complète du modèle . . . . .	49
2.3	Analyse fréquentielle et possibilités de sur-résolution . . . . .	50
2.3.1	Ciel et pointage . . . . .	51
2.3.2	Convolution et échantillonnage spatial . . . . .	51
2.3.3	Convolution et échantillonnage temporel . . . . .	52
2.3.4	Cas des protocoles <i>Small map</i> et <i>Point source</i> . . . . .	54
2.3.5	Cas du protocole <i>Large map</i> . . . . .	56
2.3.6	Interprétation des spectres dans le cas <i>Large map</i> . . . . .	56
2.4	Calcul explicite du modèle d’acquisition . . . . .	61
2.4.1	Modèle de ciel . . . . .	61
2.4.2	Calculs explicites . . . . .	63
2.4.3	Cas du protocole <i>Large map</i> . . . . .	66
2.4.4	Cas des protocoles <i>Point source</i> et <i>Small map</i> . . . . .	69
2.4.5	Expression complète des modèles . . . . .	69
2.5	Conclusion . . . . .	70
<b>3</b>	<b>Inversion</b>	<b>73</b>
3.1	Modèles probabilistes . . . . .	73
3.1.1	Modélisation de la composante basse fréquence . . . . .	74
3.1.2	Modélisation de la régularité spatiale du ciel . . . . .	76
3.1.3	Estimation jointe et marginalisation de dérive thermique . . . . .	84
3.1.4	Loi <i>a posteriori</i> étendue . . . . .	89
3.2	Estimateur et algorithme de calcul . . . . .	91
3.2.1	Algorithme MCMC proposé . . . . .	92
3.2.2	Échantillonnage de l’objet par optimisation . . . . .	95
3.2.3	Critère et gradients pour l’optimisation . . . . .	97
3.3	Calculs efficaces d’une sortie modèle et du transposé . . . . .	99
3.3.1	Économies de mémoire et de calculs . . . . .	99
3.3.2	Structure invariante et modèle de la littérature en sur-résolution . . . . .	102
3.3.3	Approximation du hessien . . . . .	105
3.4	Conclusion . . . . .	107
<b>4</b>	<b>Étude expérimentale</b>	<b>109</b>
4.1	Conditions de simulation . . . . .	109

4.1.1	Ciels étudiés . . . . .	109
4.1.2	Balayage . . . . .	111
4.1.3	Modèle instrument . . . . .	112
4.1.4	Simulation de données . . . . .	113
4.1.5	Algorithmes . . . . .	114
4.1.6	Outils d'évaluation . . . . .	115
4.2	Moindres carrés régularisés . . . . .	116
4.2.1	Restauration du <i>cirrus</i> . . . . .	116
4.2.2	Variance <i>a posteriori</i> et incertitude . . . . .	118
4.2.3	Variation du niveau de bruit . . . . .	120
4.2.4	Biais et variance de l'estimateur . . . . .	122
4.2.5	Autres types de ciels . . . . .	122
4.3	Estimation non-supervisée . . . . .	124
4.3.1	Estimation du ciel . . . . .	126
4.3.2	Estimation et chaînes des hyper-paramètres . . . . .	130
4.3.3	Conclusion . . . . .	131
4.4	Conclusion . . . . .	131
<b>5</b>	<b>Conclusions et perspectives</b>	<b>135</b>
5.1	Bilan des travaux . . . . .	135
5.1.1	Contributions majeures . . . . .	135
5.1.2	Autres apports . . . . .	136
5.2	Perspectives . . . . .	137
5.2.1	Perspectives algorithmiques et calculatoires . . . . .	137
5.2.2	Conditions d'échantillonnages pour la sur-résolution . . . . .	137
5.2.3	Bruit impulsionnel . . . . .	138
5.2.4	Estimation de paramètres instruments . . . . .	138
5.2.5	Application à d'autres instruments . . . . .	138
5.2.6	Sur-résolution myope et non-supervisée . . . . .	138
5.2.7	Utilisation conjointe des trois bandes . . . . .	139
5.2.8	Extension du modèle pour le ciel . . . . .	139
<b>A</b>	<b>Physique des bolomètres</b>	<b>143</b>
A.1	Principes . . . . .	143
A.2	Bilan de puissance . . . . .	143
A.3	Linéarisation . . . . .	146
A.4	Réponse à un signal constant . . . . .	147
A.5	Réponse à un signal variable . . . . .	148
A.6	Réponse en tension . . . . .	149

<b>B Compléments sur le modèle direct</b>	<b>151</b>
B.1 Grille hexagonale . . . . .	151
B.2 Décomposition sur des gaussiennes et protocole <i>Large map</i> . . . . .	152
B.3 Calculs explicites pour les protocoles <i>Point source</i> et <i>Small map</i> . . . . .	155
B.4 Autres modèles d'acquisition . . . . .	156
<b>C Calculs d'hivers rigoureux</b>	<b>157</b>
C.1 Convolution de deux gaussiennes . . . . .	157
C.2 Convolution d'une gaussienne et d'un sinus cardinal . . . . .	158
C.3 Lemme d'inversion de matrices partitionnées . . . . .	162
C.4 Calculs du filtre de Wiener–Hunt et inverse du hessien . . . . .	162
<b>D Valeur numérique des paramètres du modèle</b>	<b>167</b>
<b>E Algorithmes</b>	<b>169</b>
<b>F Déconvolution myope non-supervisée</b>	<b>177</b>
<b>Références</b>	<b>190</b>

# Chapitre 1

## Introduction

### 1.1 Contexte

L'information au sujet des objets astrophysiques est presque exclusivement accessible à travers le rayonnement électromagnétique qui nous en parvient. Les caractéristiques de ce rayonnement sont ensuite analysées pour en déduire des propriétés physiques des objets observés. Par exemple, la présence de raies d'émission ou d'absorption à certaines longueurs d'onde renseigne sur la composition chimique des objets. Plus particulièrement le rayonnement provenant de l'émission étendue, comme les galaxies ou les nuages de gaz, renseigne sur la densité ou la température de ces objets.

Pour observer ce rayonnement, des observatoires terrestres ou spatiaux sont conçus afin de collecter le flux lumineux. Les objets sont à des distances très grandes, considérées infinies, de l'observateur et la ligne de visée, ou la direction d'observation, est caractérisée par deux angles notés ici  $\alpha$  et  $\beta$ . La distance est perdue et tout le flux provenant d'une ligne de visée est intégré. De plus le rayonnement est caractérisé par une troisième dimension, la longueur d'onde  $\lambda$ .

La problématique consiste ici à estimer le rayonnement inconnu  $\phi(\alpha, \beta, \lambda)$  suivant les trois dimensions. Cette tâche sera plus ou moins difficile en fonction des caractéristiques de l'instrument et de ses capacités à mesurer l'information dans les trois dimensions. Pour le cas des dimensions spatiales la focalisation, par une optique, de la lumière sur un plan focal permet la formation d'une image. Pour être discriminant en longueur d'onde cette fois il faut employer des capteurs sensibles à un certain domaine spectral par exemple en employant des filtres à longueurs d'onde ou des systèmes dispersifs comme les réseaux de diffraction.

Les instruments fournissent une mesure du flux de photons à travers toute une chaîne de collecte et de détection. De nombreux éléments sont employés pour produire un signal dans une grandeur physique facile à manipuler. Les instruments possédant toujours des limitations, l'information provenant du signal astrophysique d'intérêt est dégradée ou déformée. L'étude et la modélisation de l'ensemble de ces phénomènes physiques à l'origine des données pour un ciel fixé est appelé « problème direct ».

La question de l'estimation du ciel inconnu  $\phi$  connaissant le signal acquis (les données  $\mathbf{y}$ ) est appelé « problème inverse ». En langage courant il s'agit connaissant la conséquence  $\mathbf{y}$  d'en déduire la cause  $\phi$ <sup>1</sup>. Sans rentrer dans les détails, alors que le modèle direct est généralement bien-posé mathématiquement (les équations sont « stables ») ce n'est généralement pas le cas des problèmes inverses qualifiés dans ce cas de mal-posés. Lever la dégénérescence, en physique par exemple, est un problème mal posé puisque plusieurs combinaisons des paramètres d'entrées peuvent expliquer

---

<sup>1</sup>Lorsque l'objet inconnu est dans un espace à dimension fini nous utiliserons la notation  $\mathbf{x}$  au lieu de  $\phi$ .

le phénomène observé.

Dans la suite nous décrivons l'instrument SPIRE sur lequel nous avons concentré nos travaux. C'est un instrument unique, qui a demandé pour sa conception des technologies de pointe, et qui bien que permettant d'accéder à des informations inaccessibles jusqu'à maintenant, possède des limitations notamment en résolution spatiale. De nombreux problèmes abordés dans cette thèse s'appliquent également à d'autres instruments tel que PACS d'Herschel.

### 1.1.1 Herschel et SPIRE

Herschel, représenté figure 1.1, est un observatoire spatial de l'Agence Spatiale Européenne pour l'observation astronomique infrarouge et sub-millimétrique. L'atmosphère étant opaque pour ces longueurs d'ondes, il a été lancé le 14 mai 2009 par une fusée Ariane 5 (en même temps que le satellite Planck). Comme tous les observatoires spatiaux, c'est un instrument unique qui a nécessité des technologies parfois uniquement développées dans des laboratoires de recherche. Le programme officiel a débuté en 2000, mais le projet a environ vingt ans, et de nombreux acteurs européens (laboratoires, instituts, *etc.*) ont participé à sa conception. Le coût total du projet est de l'ordre du milliard d'euros.

L'observation de l'infrarouge lointain permet d'accéder à une information sur la formation et la vie des étoiles ainsi que des galaxies. Les propriétés de nuages de poussière et de gaz présents dans le milieu interstellaire sont également accessibles à l'aide de ce rayonnement. La lumière émise par ces objets possédant une étendue spatiale est appelée « émission étendue » pour la distinguer du rayonnement émis par des objets ponctuels tels que les étoiles.

Un élément central d'Herschel est le télescope, décrit partie 2.2.1, qui focalise la lumière et forme une image sur le plan focal. De type Ritchey-Chrétien, son miroir primaire fait 3,5 mètres de diamètre. C'est le plus grand miroir lancé dans l'espace jusqu'à maintenant et sa conception a nécessité une technicité de pointe pour qu'il ait un poids raisonnable. La dimension du miroir est importante car elle détermine à la fois la quantité de flux collecté et la résolution angulaire des données, c'est-à-dire la capacité naturelle à discriminer deux points rapprochés.

La résolution angulaire des données à la sortie du miroir dépend de la longueur d'onde. En effet plus la longueur d'onde est grande, moins la résolution est bonne, comme décrit dans la partie sur l'ouverture 2.2.1. Dans le cas de l'observatoire Herschel les longueurs d'onde observées sont de l'infrarouge lointain au sub-millimétrique. Ce sont de très grandes longueurs d'onde et la perte en résolution est importante. La grande dimension du miroir permet donc en partie de compenser cette perte en résolution.

L'observatoire comprend trois instruments placés sur le plan focal. On trouve :

- HIFI (Heterodyne Instrument for the Far Infrared) qui est un spectromètre haute résolution ;
- PACS (Photodetector Array Camera and Spectrometer) qui est un photomètre à bolomètres et un spectromètre à réseau à photoconducteurs et observe de 61 à 210  $\mu\text{m}$  ;
- SPIRE (Spectral and Photometric Imaging REceiver) est un imageur et un spectromètre par transformée de Fourier, tous deux à base de bolomètres. Il observe des longueurs d'onde de 200 à 600  $\mu\text{m}$ .

Dans ce travail nous nous sommes concentrés uniquement sur les trois photomètres de l'imageur de SPIRE. Les photomètres observent la même étendue spatiale du ciel en  $(\alpha, \beta)$  dans trois bandes spectrales différentes et relativement étroites. Ils ne sont pas destinés à une étude fine du contenu spectral du ciel mais plutôt à une mesure spatiale.

Les trois photomètres ont leur bande spectrale centrée en 250  $\mu\text{m}$ , 360  $\mu\text{m}$  et 520  $\mu\text{m}$  environ. Ces photomètres possèdent des filtres en longueur d'onde (partie 2.2.2) pour sélectionner la bande



**Figure 1.1** – Image synthétique du satellite Herschel. Le bouclier thermique est en arrière plan du miroir primaire et secondaire. Juste en dessous du miroir primaire se trouve le cryostat qui contient tous les instruments. Le bloc inférieur contient l'électronique ainsi que les instruments de mesure, de télémétrie, de traitement de données, *etc.* Il mesure plus de sept mètres de haut et pèse plus de trois tonnes.

spectrale et de cornets (partie 2.2.3) qui sont des intégrateurs spatiaux et concentrent le rayonnement sur les bolomètres (partie 2.2.4). Les bolomètres sont une des particularités de l'instrument. Ce sont des détecteurs de chaleur, très sensibles, assez peu employés jusqu'à maintenant pour les observatoires spatiaux à cause de difficultés techniques. Ces capteurs produisent un signal électrique qui est ensuite numérisé. Il y a bien sûr d'autres éléments sur le trajet optique (miroir dichroïque, lentilles, guide d'onde, *etc.*) qui modifient dans une moindre mesure le rayonnement.

Pour permettre une grande sensibilité et observer des flux très faibles, les bolomètres sont placés dans un bain thermique à 0,3 Kelvin. Cependant plusieurs difficultés sont présentes. Tout d'abord le miroir est trop grand pour qu'il soit lui-même refroidi : l'observatoire émet un rayonnement parasite reçu par le capteur. L'observatoire est protégé du rayonnement du soleil par un bouclier, visible figure 1.1, mais la protection n'est pas totale. Enfin la température du bain n'est pas parfaitement stable. Il y a une dérive en température visible pendant les observations qui perturbent les mesures du flux incident. Pour cela, deux bolomètres dit aveugles sont intégrés au sein de chaque photomètre de SPIRE. Ce sont des bolomètres comme les autres, placés à chaque extrémité, mais qui ne reçoivent aucun flux provenant du ciel. Ils sont cependant sensibles à la dérive thermique qui influence l'ensemble du capteur et permettent donc d'en effectuer une mesure plus directe.

Le pointage du télescope n'est pas fixe dans le temps, mais des balayages ou des pointages successifs décrits partie 2.2.6 sont effectués. Ces protocoles permettent de mesurer des étendues spatiales plus grandes que le champ de vue d'un photomètre. De plus, de la redondance spatiale dans les données est introduite à l'aide de faibles incréments de pointage.

L'ensemble de ces éléments constitue une chaîne d'acquisition complexe. La définition formelle d'un modèle d'acquisition comprenant les éléments décrits ici fait l'objet du chapitre 2 consacré à sa modélisation et à son analyse. L'objectif est de permettre la meilleure interprétation possible de l'information présente dans les données.

Pour déduire une information nouvelle sur le ciel observé à partir des données acquises, nous avons suivi une démarche d'*inversion*. Il s'agit d'inverser les données produites pour estimer l'ensemble des variables inconnues en prenant en compte le processus d'acquisition.

### 1.1.2 Inversion

Les éléments d'un processus d'acquisition, comme celui de SPIRE, sont généralement modélisés et mis ensemble sous la forme d'un système  $\mathcal{H}$  qui d'un ciel défini continûment  $\phi$  (par exemple  $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ ) décrit les données discrètes observées  $\mathbf{y}$

$$\mathbf{y} = \mathcal{H}(\phi).$$

L'estimation de l'inconnue revient ici à résoudre l'équation par rapport à  $\phi$  lorsque  $\mathbf{y}$  est connue c'est-à-dire de déterminer

$$\phi = \mathcal{H}^{-1}(\mathbf{y}).$$

Malheureusement, le problème ainsi défini n'est généralement pas bien-posé au sens de Hadamard [Had01, Idi01a, Two63]. Pour qu'un problème soit bien-posé il faut :

- que la solution existe,
- qu'elle soit unique,
- qu'elle soit continue vis-à-vis des données (*i.e.*, stable : une petite variation dans les données n'introduit pas de grande variation dans la solution).

En dimension infinie une des ces conditions n'est souvent pas remplie. Pour une description plus complète et détaillée nous invitons le lecteur à consulter [Idi01a], chapitre 1 ou sa traduction anglaise [Idi08].

En pratique les problèmes se définissent souvent en dimension finie, par exemple en exprimant la solution définie continûment  $\phi$  à l'aide d'un nombre fini de coefficients de décompositions  $\mathbf{x}$  tel que  $\phi = f(\mathbf{x})$ , voir partie 2.4.1,  $f$  étant l'interpolateur de Shannon par exemple. De plus l'incertitude est généralement modélisée au travers d'un terme de perturbation  $\mathbf{b}$  et d'une loi d'addition

$$\mathbf{y} = \mathcal{H}(\phi) + \mathbf{b} = \mathcal{H}(f(\mathbf{x})) + \mathbf{b}.$$

Ce terme représente le fait qu'il est impossible de retrouver la solution qui explique parfaitement les données au travers du modèle. Dans cette situation la solution  $\hat{\phi}$  du problème est définie comme l'élément optimisant un critère (coût) par rapport aux données selon une certaine distance, par exemple quadratique

$$\hat{\phi} \text{ minimise } \|\mathcal{H}(\phi) - \mathbf{y}\|^2.$$

En dimension finie (discret–discret) il existe toujours au moins un minimiseur [Idi01a] obtenu en annulant la dérivée première du critère, mais cette solution n'est pas nécessairement unique et des problèmes de stabilités sont très souvent rencontrés. Si la solution n'est pas unique (le problème est sous-déterminé par exemple), celle-ci peut être établie en choisissant parmi l'ensemble des solutions celle à norme minimale (selon une distance à définir) appelée *solution généralisée*.

Malgré le fait qu'il est toujours possible de se ramener à une solution unique avec la *solution généralisée*, celle-ci n'est pas nécessairement stable (même si elle est continue vis-à-vis des données). Dans le cas discret–discret, si l'opérateur est linéaire le modèle s'écrit sous la forme

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{b}$$

où  $\mathbf{H}$  est une matrice qui peut être examinée au travers des ses valeurs singulières  $\lambda_i$ . Avec la *solution généralisée*, les valeurs singulières nulles qui autorisent plusieurs solutions sont « exclues » mais il peut y avoir un problème avec les valeurs singulières proches de zéro. Dans ces cas là, l'inverse généralisée, qui fait intervenir l'inverse de ces valeurs singulières  $\lambda_i^{-1}$ , provoque une amplification et une instabilité puisqu'il y a division par des valeurs très petites. C'est un problème de conditionnement.

Pour fixer les choses prenons comme exemple la convolution, présente dans notre modélisation de SPIRE. Sous certaines hypothèses de périodicité, la convolution dans l'espace de Fourier s'écrit

$$\hat{\mathbf{y}} = \mathbf{\Lambda}\hat{\mathbf{x}} + \hat{\mathbf{b}}$$

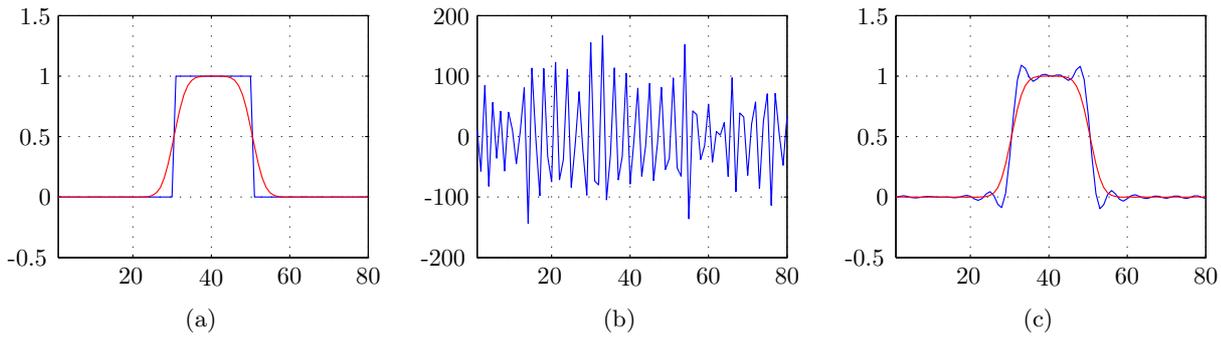
avec la fonction de transfert

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_m \end{pmatrix}$$

une matrice diagonale et  $\circ$  symbolise le signal dans l'espace de Fourier. Si la distance est quadratique, la *solution généralisée*  $\hat{\hat{\mathbf{x}}}$ , obtenue avec le filtre inverse, s'écrit

$$\hat{\hat{\mathbf{x}}} = \mathbf{\Lambda}^{-1}\hat{\mathbf{y}}.$$

Si des valeurs propres  $\lambda_i$  sont petites parce que la fonction de transfert atténue certaines fréquences, le filtre inverse  $\mathbf{\Lambda}^{-1}$  aura tendance à les amplifier par  $\lambda_i^{-1}$  pour compenser. Dans ce cas, les composantes de  $\mathbf{b}$  qui est large bande se trouveront amplifiées et une petite variation dans les données introduisent de grandes variations dans la solution. Le problème est mal conditionné. Pour illustrer ce problème, la figure 1.2 présente un cas où le noyau de convolution est une gaussienne et



**Figure 1.2** – Caractère mal conditionné du filtre inverse. La figure 1.2(a) représente le signal d’origine en bleu et le signal mesuré en rouge après dégradation par filtrage et ajout d’un bruit sous une loi gaussienne blanche *idd*. Le bruit est imperceptible. La figure 1.2(b) correspond à la *solution généralisée* et la figure 1.2(c) la solution des moindres carrés régularisés en bleu avec un *a priori* quadratique sur les différences.

la *solution généralisée* figure 1.2(b) est inacceptable. Ce problème de conditionnement est le plus rencontré en pratique.

La solution aux problèmes mal-posés consiste à introduire de l’information supplémentaire pour limiter l’espace des solutions ou favoriser certaines solutions. Pour cela plusieurs approches existent.

- Une approche, très souvent employée de manière implicite, est le contrôle de dimensions. Elle consiste à réduire l’espace des solutions à un espace de dimension moindre pour rendre le problème bien-posé. L’utilisation d’un modèle paramétrique à quelques paramètres pour exprimer un objet continu est un exemple. L’ajout de contraintes, comme la positivité, en est un autre. Le rejet de solutions qui ne sont pas physiquement vraisemblables est également possible. Ces approches évitent les difficultés en définissant, à l’origine, un problème bien-posé. Elles injectent une information *a priori*, souvent jugée plausible puisqu’elle vient d’une connaissance d’expert ou de ceux intéressés par l’application.
- Certaines approches consistent à éliminer ce qui est considéré comme gênant. La *décomposition tronquée en valeurs singulières* (TSVD) est un exemple [Nas76]. Elle consiste à reconstruire l’inconnue uniquement avec les valeurs singulières  $\lambda_i$  suffisamment importantes. Par construction cette solution est stable numériquement. Cependant se pose le choix de la limite où le seuil est appliqué. De plus il n’est plus possible de restaurer des composantes spectrales considérées trop dégradées. Un autre exemple est le filtrage, en changeant de représentation ou non, pour éviter l’amplification du bruit. Par exemple le signal peut être décomposé sur une base où il aurait une représentation *a priori* parcimonieuse puis un seuil peut être appliqué sur les coefficients jugés non significatifs [KMR97]. Dans ces approches il faut choisir une représentation adaptée et le filtrage élimine une partie de l’information à propos du signal d’intérêt.
- Certaines approches exploitent le comportement de méthodes itératives pour calculer la solution en limitant le nombre d’itérations [Ric72]. Ces approches reposent sur l’observation empirique que certains algorithmes restaurent généralement les composantes spectrales les plus significatives puis au fur et à mesure restaurent les plus faibles. En limitant le nombre d’itérations on évite la restauration des composantes qui provoquent l’instabilité. Ces approches souffrent d’un manque de formalisme qui limite leur compréhension et leur justification. En outre la détermination de l’arrêt des itérations n’est pas claire et elles sont sensibles à l’initialisation y compris dans les cas de critères convexes.
- Enfin une autre classe de méthode, appelée régularisation, fait intervenir explicitement dans le

problème un caractère désiré de la solution. Par exemple la solution est relativement régulière, continue par morceaux ou composée de pics. Cette information est souvent formalisée sous la forme de critère composite

$$\hat{\phi} \text{ minimise } \| \mathcal{H}(\phi) - \mathbf{y} \|^2 + \alpha \mathcal{F}(\phi)$$

où  $\mathcal{F}$  modélise le caractère désiré sur la solution et l’hyper-paramètre  $\alpha$  gère le compromis entre les deux sources d’information. C’est l’approche historique développée par les auteurs tels que Tikhonov [Tik63] ou Twomey [Two63]. Un exemple est illustré figure 1.2(c). Ces approches ont très souvent un équivalent avec les modèles probabilistes bayésiens [Dem89], ces derniers apportant un formalisme et des outils très riches. Elles demandent cependant une réflexion et un échange avec les experts pour déterminer au mieux l’information disponible et le type de caractère désiré dans la solution. Cette caractéristique est un avantage puisqu’elle pose directement dans la définition du problème même la formalisation de l’information injectée dans l’inférence, sans passer par des artifices algorithmiques ou l’élimination de composantes gênantes comme dans les méthodes précédentes.

Nous voyons ici qu’il y a presque toujours un compromis à faire entre l’information provenant des données et une information disponible par ailleurs sur un caractère recherché de la solution. Ce compromis intervient par la définition d’un espace des solutions, par un seuil de troncature, un arrêt des itérations ou encore un paramètre explicite  $\alpha$ .

### 1.1.3 Problématiques

Nous pouvons maintenant établir les problèmes soulevés par l’inversion des données acquises par l’instrument SPIRE.

- L’ouverture, ou le miroir, introduit une dégradation de la résolution angulaire (modélisé par une convolution). Malgré la dimension du miroir, les longueurs d’ondes sont grandes donc l’atténuation des hautes fréquences spatiales également. De même les bolomètres dégradent la résolution temporelle. Ils possèdent un temps de réaction avec un comportement type passe-bas.

Par conséquent le miroir, les cornets et les bolomètres dégradent la résolution spatiale et temporelle. Non seulement le modèle d’acquisition sera complexe, mais le problème de déconvolution soulevé est connu pour son mauvais conditionnement [Hun72, GY95, Gio08].

- Pour permettre une bonne sensibilité, l’angle solide d’ouverture (le diamètre) des cornets est important. La conséquence est un sous-échantillonnage par rapport à la fonction de transfert du miroir. La fréquence d’échantillonnage est deux fois plus petite que la fréquence de coupure du miroir ce qui entraîne un repliement de spectre. Pour résoudre ce problème de repliement, des protocoles d’observation adaptés introduisent de la redondance dans les données avec des incréments de pointages inférieurs à la taille d’un capteur.

En exploitant la redondance spatiale dans les données, introduite par les protocoles d’observation, il est possible de reconstruire une image possédant une résolution supérieure à la résolution naturelle des données. C’est un problème typique de sur ou super-résolution [PPK03, RCLBG06, VSVV07, EF99].

- L’environnement évolue en température et introduit une composante basse fréquence visible dans les données (c’est ce que les astronomes appellent « bruit en  $1/f$  » pour mentionner le caractère basse-fréquence). L’origine principale est la dérive thermique du bain thermique et les bolomètres aveugles fournissent une mesure plus directe de cette composante [SBL<sup>+</sup>08, CCB<sup>+</sup>06].
- Ces problèmes peuvent être abordés avec une approche problème inverse mais se posent deux questions : la nature de l’information *a priori* sur l’objet d’intérêt et les autres incon-

nues ainsi que la question de l'estimation des paramètres de compromis entre les différentes sources d'information. Le formalisme bayésien fournit des réponses solides que nous présenterons [Dem89, Gio08, BMK09].

- Le modèle instrument est un modèle paramétrique bâti sur la physique. Certains paramètres peuvent être mal connus, peuvent évoluer pendant la vie de l'instrument ou le temps d'une observation. Par exemple les caractéristiques des bolomètres dépendent de la température ambiante contrôlée par un cryostat. Dans ce cas l'estimation de ces paramètres peut être importante.
- Enfin des particules à hautes énergies frappent les détecteurs. Ce sont des événements impulsifs qui influent sur le comportement du capteur de façons variable, le plus simple étant un pic en température d'un seul bolomètre, le plus complexe étant une influence électrique sur l'ensemble du capteur. Ce problème n'a pas été abordé dans les travaux de cette thèse.

Le travail présenté ici propose un formalisme, reposant sur une démarche d'inversion bayésienne classique, pour répondre principalement aux principaux points c'est-à-dire le problème de convolution, de repliement, de la dérive en température, le réglage du compromis entre les différentes sources d'information et l'estimation de paramètres instrument. L'apport des travaux de cette thèse est exposé dans les chapitres 2 et 3. Le chapitre 4 présente une étude expérimentale des résultats des chapitres précédents.

Le cadre bayésien permet de mettre en évidence les informations utilisées mais il offre également une grande richesse d'outils permettant d'aborder de nombreux problèmes. Il permet notamment une solution naturelle à la question du compromis à faire entre les différentes sources d'information.

## 1.2 Inférence bayésienne

L'approche bayésienne offre un cadre cohérent, complet et exhaustif pour mener une inférence *logique* lorsque les informations disponibles sont entachées d'incertitude. Pour une présentation détaillée de ses paradigmes nous invitons le lecteur à consulter les livres d'Edwin T. Jaynes [Jay03] et David MacKay [Mac03], tous deux largement reconnus et disponibles en libre téléchargement. Pour une présentation en français, le lecteur pourra se référer à l'excellent support de cours de Master 2 Recherche de Guy Demoment [Dem]. Pour une approche plus statistique les livres de Christian Robert [Rob05, Rob96] sont une référence.

Les travaux de Jaynes (ainsi que Cox ou encore Jeffreys) présentés dans le polycopié de Guy Demoment établissent la théorie des probabilités, et particulièrement l'approche bayésienne, comme une extension de la logique booléenne à des événements incertains. Il s'agit de règles logiques qui permettent d'établir des relations entre événements lorsque l'on dispose seulement d'une information incomplète sur le fait « vrai » ou « faux ». Dans ce cas, la mesure de probabilité  $p(\phi)$ , prise entre 0 pour « faux » et 1 pour « vrai », permet de *mesurer quantitativement* le degré de connaissance d'une hypothèse  $\phi$ . Il s'agit donc d'un outil quantitatif, qui obéit à des règles bien établies, logiques et précises pour mesurer une information mais également la manipuler et l'explorer. L'intérêt de cette théorie est qu'elle est établie à partir d'axiomes de bases, appelés axiomes de Cox, qui définissent ce qui est attendu d'une théorie pour mener une déduction logique quantitative en environnement incertain.

Selon l'approche fréquentiste la probabilité est établie à l'aide du concept de limite de fréquence de réalisation de variable « aléatoire » : une pièce parfaite (qui est une hypothèse) lancée un nombre fictif de fois verra la fréquence relative de réalisation de l'évènement « face » tendre vers 0,5. Dans l'approche bayésienne les probabilités ne sont pas des caractéristiques établies par une limite de fréquences. Le raisonnement se situe au niveau des connaissances : en l'absence d'autre information, il n'y a aucune raison de penser que le lancé donnera « face » plutôt que « pile ». Les événements

sont équiprobables et par conséquent les deux probabilités valent chacune 0,5. Il n'est pas fait appel à un nombre fictif d'événements (ou au concept de chance) mais à une démarche logique et quantitative établissant l'information disponible. Si par exemple la connaissance que la pièce est truquée est disponible, alors les probabilités sont adaptées pour refléter cette connaissance.

### 1.2.1 Loi *a posteriori*

Un des outils de base de la théorie bayésienne est la règle de Bayes

$$p(\phi|\mathbf{y}, I) = \frac{p(\mathbf{y}|\phi, I)p(\phi, I)}{p(\mathbf{y}, I)}$$

qui est une formule d'*inversion* [Rob05]. En effet elle établit l'information  $p(\phi|\mathbf{y}, I)$  sur le ciel à l'aide de l'information fournie par les données  $p(\mathbf{y}|\phi, I)$ . Pour cela la formule fait également apparaître l'information disponible sur le ciel  $p(\phi|I)$  lorsque l'on ne connaît que les hypothèses *a priori*  $I$ . La règle de Bayes peut également être vue comme une formule de mise-à-jour : des informations *a priori*  $p(\phi|I)$  on déduit l'information sur le même objet  $p(\phi|\mathbf{y}, I)$  une fois les données acquises. La partie 1.3 présente plus en détail la vraisemblance des données  $p(\mathbf{y}|\phi, I)$  et la loi *a priori*  $p(\phi|I)$ .

Souvent l'environnement logique  $I$  n'est pas noté puisqu'il est toujours du même côté de la barre de conditionnement. Appelé parfois modèle, l'environnement logique réunit ce qui est considéré *vrai*. Il comprend les lois de la physique, le modèle d'acquisition, les espaces de définition des inconnues et des données, ou encore les lois de probabilités *a priori etc.* Dans les travaux menés dans cette thèse, l'environnement logique définit en particulier le modèle d'acquisition des données. Le modèle le plus classique déjà mentionné est noté

$$\mathbf{y} = \mathcal{H}_\theta(\phi) + \mathbf{b} \tag{1.1}$$

qui considère que les données proviennent d'un signal d'intérêt transformé par un opérateur  $\mathcal{H}$ , paramétré par  $\theta$  dont la sortie est dégradée par un signal inconnu  $\mathbf{b}$  appelé bruit. Ici le bruit est un modèle pour prendre en compte l'incertitude sur le processus d'acquisition et ne correspond pas à une réalité. L'environnement logique est fondamental puisqu'il établit qu'il y a toujours quelque chose qui est supposé connu. Aucun traitement ne peut être fait sans *a priori*, hypothèses ou encore postulat. Dans le cas contraire la seule chose possible est le stockage des données.

La mise en commun de ces différentes sources d'information s'effectue à l'aide de la règle de Bayes pour former la loi *a posteriori*. Cette loi est donc la collection de toute l'information utilisée pour résoudre le problème de l'estimation de l'objet inconnu  $\phi$ .

### 1.2.2 Définition d'un estimateur

La loi *a posteriori* collecte *toute* l'information disponible sur l'objet inconnu, c'est-à-dire le ciel  $\phi$ , lorsque l'on connaît les données  $\mathbf{y}$  et les informations *a priori* disponibles par ailleurs. Cependant, en dehors des cas où l'objet est un scalaire ou un vecteur à deux coordonnées, l'exploration et l'interprétation de la loi est très difficile. C'est particulièrement le cas en traitement du signal ou de l'image où les objets sont définis dans des espaces de grande dimension.

La solution consiste à choisir un point caractéristique, un estimateur, de cette loi. Ce point, résume l'information présente dans la loi mais une partie de l'information est ignorée. Pour déterminer ce point caractéristique, l'approche bayésienne fait appel à la théorie de la décision [Rob05] : on détermine de façon objective une fonction  $C(\mathbf{x}, \mathbf{x}^*)$  qui représente le coût à se tromper lorsque le vrai est  $\mathbf{x}^*$ . L'estimateur est la fonction des données qui minimise le coût moyen par rapport à

la probabilité de tous les possibles

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \min_{\mathbf{x}} \int C(\mathbf{x}(\mathbf{y}), \mathbf{x}^*) p(\mathbf{x}^* | \mathbf{y}) d\mathbf{x}^*.$$

L'estimateur obtenu est alors le meilleur estimateur au sens de cette fonction de coût et de la loi *a posteriori*. Formellement la détermination de cette fonction de coût est importante. En pratique il est souvent fait appel aux mêmes coûts classiques :

- le coût tout ou rien

$$C(\mathbf{x}, \mathbf{x}^*) = \begin{cases} 0 & \text{si et seulement si } \mathbf{x}(\mathbf{y}) = \mathbf{x}^* \\ 1 & \text{sinon} \end{cases}$$

aboutit au choix de l'argument du maximum de la loi soit

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \max_{\mathbf{x}} p(\mathbf{x} | \mathbf{y}).$$

Cet estimateur est appelé le *maximum a posteriori* ou MAP.

- le coût quadratique  $C(\mathbf{x}, \mathbf{x}^*) = \|\mathbf{x} - \mathbf{x}^*\|^2$  qui correspond à l'erreur quadratique moyenne aboutit au choix de la moyenne de la loi

$$\hat{\mathbf{x}}(\mathbf{y}) = \int \mathbf{x} p(\mathbf{x} | \mathbf{y}) d\mathbf{x}.$$

Cet estimateur est appelé la *moyenne a posteriori*, ou EAP, et c'est le meilleur au sens de l'erreur quadratique moyenne.

- le coût absolu  $C(\mathbf{x}, \mathbf{x}^*) = |\mathbf{x} - \mathbf{x}^*|$  aboutit au choix de la médiane de la loi si  $\mathbf{x}$  est un scalaire

$$\hat{\mathbf{x}}(\mathbf{y}) \quad \text{tel que} \quad p(\mathbf{x}^* < \hat{\mathbf{x}}(\mathbf{y})) = \frac{1}{2}$$

appelé la *médiane a posteriori*.

Il est classique de mettre en avant les avantages et les inconvénients des estimateurs de la moyenne et du maximum *a posteriori*. En effet si la loi est bi-modale, c'est-à-dire que la masse de probabilité se concentre principalement autour de deux points alors la moyenne se placera entre les deux, à un endroit où la probabilité peut être très faible. Le cas d'un élément fortement probable, un pic, mais où la grande partie de la masse de probabilité est située autour d'un autre point est un autre cas pathologique concernant le maximum cette fois. Celui-ci se placera sur le pic, alors que la masse de probabilité autour de celui-ci peut être très faible.

Les deux cas illustrent le fait qu'il faut garder à l'esprit qu'il est réducteur de limiter l'interprétation de la loi à un point. On peut se poser la question, du point de vue de la connaissance, de la justification de limiter l'interprétation de la loi à un point lorsqu'elle est bi-modale. Il est possible de justifier par exemple en évoquant le fait qu'un des modes n'est pas réaliste physiquement. Cependant dans ce cas c'est la définition des informations disponibles (ici *a priori*) qu'il faut réviser pour refléter cette connaissance. Et ceci peut être fait potentiellement par la vraisemblance en modifiant le modèle d'acquisition ou les *a priori*.

Le cas où la loi *a posteriori* est globalement uni-modale mais dont le lobe est très large est également intéressant. La loi *a posteriori* modélise l'information et l'incertitude sur l'objet. Si le lobe est très large, plutôt que d'extraire un point caractéristique, il s'agit d'un signe que l'information n'est pas suffisante pour prendre une décision raisonnable. Cette information est donc également portée par la loi *a posteriori* et dans le cas gaussien par exemple, il s'agit de la matrice de covariance *a posteriori*. Dans le cas d'une image par exemple, la diagonale de cette matrice fournit une carte d'incertitude.

### 1.3 Apport de l'approche bayésienne

L'approche bayésienne apporte une solution aux problèmes mal-posés puisqu'elle fait apparaître naturellement les informations nécessaires pour rendre le problème bien déterminé, en partie à travers les *a priori*. Dans de nombreux cas d'ailleurs les méthodes de régularisation présentées dans la partie 1.1.2 possèdent une interprétation bayésienne. Le fait de manipuler des lois de probabilité offre plusieurs avantages. Tout d'abord c'est l'approche qui exploite au mieux les données puisqu'elle permet d'exploiter un modèle réaliste qui serait mal-conditionné. Ensuite elle offre un cadre logique, cohérent et riche pour modéliser une grande variété d'informations. Enfin elle permet de répondre à des questions de façon naturelle comme l'estimation des paramètres de compromis entre les différentes sources d'information (hyper-paramètres), là où les autres approches peinent à trouver des réponses solides.

**Remarque.** *On peut arguer également du fait que contrairement à l'approche fréquentiste qui généralement définit des estimateurs ad hoc et étudie leurs propriétés (de convergence par exemple) après coup, l'approche bayésienne permet, par la théorie de la décision, de définir les propriétés recherchées (comme la minimisation de l'erreur quadratique moyenne) et d'en déduire l'estimateur qui les remplit.*

#### 1.3.1 Vraisemblance des données

Si on ne souhaite injecter aucune information *a priori* sur le bruit  $\mathbf{b}$  alors le problème d'estimation de  $\phi$  à partir de  $\mathbf{y}$  est indéterminé. Il s'agit donc de déterminer une information *a priori* disponible sur  $\mathbf{b}$ . Cette information entachée d'incertitude est modélisée par une loi de probabilité  $p_{\mathbf{b}}(\mathbf{b}|I)$ .

Le couplage entre le modèle d'acquisition et la loi du bruit permet de définir la loi  $p(\mathbf{y}|\phi)$  des données lorsque  $\phi$  est connu. Cette loi est nécessaire pour bâtir la loi *a posteriori*. Pour l'établir on peut remarquer que

$$p(\mathbf{y}|\phi) = p_{\mathbf{b}}(\mathbf{y} - \mathcal{H}(\phi)|\phi).$$

Il s'agit donc de la loi du bruit dont la moyenne est modifiée par la sortie du modèle d'acquisition  $\mathcal{H}(\phi)$

$$\mathbb{E}_{p(\mathbf{y}|\phi)}[\mathbf{y}] = \mathcal{H}(\phi) + \mathbb{E}_{p(\mathbf{b})}[\mathbf{b}].$$

Elle illustre le fait que même en connaissant l'objet d'intérêt, les données sont incertaines à cause du manque d'information sur  $\mathbf{b}$ . Cette loi de probabilité du point de vue de  $\phi$ , lorsque  $\mathbf{y}$  est fixé, est une fonction de  $\phi$  appelée vraisemblance de l'inconnue attachée aux données.

Comme nous l'avons vu, la vraisemblance seule définit souvent des problèmes mal-posés. Les approches qui n'exploitent pas d'informations disponibles par ailleurs peuvent être confrontées à des difficultés avec les estimateurs définis uniquement à l'aide de la vraisemblance. La solution dans ce cas consiste à modifier la vraisemblance pour rendre le problème bien posé, c'est-à-dire à limiter l'interprétation possible de l'information présente dans les données. Face au problème de déconvolution présenté partie 1.1.2, une solution serait ici de considérer que la réponse instrumentale n'atténue pas les fréquences du signal d'origine. L'instabilité est ainsi évitée mais il est alors impossible de restaurer des fréquences qui pourraient l'être.

L'approche bayésienne permet de prendre en compte des informations disponibles par ailleurs et permet alors de tirer pleinement profit de l'information présente dans les données en évitant le caractère mal-posé. Dans le cas de la déconvolution, si le compromis entre les différentes sources d'information est bien réglé, alors le processus d'estimation restaurera l'inconnue à partir des données,

jusqu'à ce que l'information soit trop incertaine. Alors l'*a priori* complète l'information manquante des données.

L'approche bayésienne met en évidence qu'il ne peut y avoir une entière confiance dans les données et qu'il est possible de faire appel à des informations *a priori*. Plutôt que de dissimuler ces informations au travers d'hypothèses implicites, elle les prend en compte explicitement. Il est alors possible d'exploiter au mieux toute l'information utile présente dans les données.

Concernant la loi de  $\mathbf{b}$  on considère souvent en pratique que l'on connaît seulement les deux premiers cumulants, par exemple qu'il est de moyenne nulle et que l'énergie de sa dispersion est finie. Dans ce cas le choix le moins compromettant [Jay03] c'est-à-dire qui introduit le moins d'information, est la loi normale

$$p_{\mathbf{b}}(\mathbf{b}|\mathbf{R}_{\mathbf{b}}) = (2\pi)^{-N/2} |\mathbf{R}_{\mathbf{b}}|^{-1/2} \exp \left[ -\frac{1}{2} \|\mathbf{b}\|_{\mathbf{R}_{\mathbf{b}}^{-1}}^2 \right].$$

C'est la loi du bruit que nous utiliserons en pratique pour notre inversion comme décrit dans le chapitre 3. La covariance  $\mathbf{R}_{\mathbf{b}}$  permet par exemple de modéliser un bruit corrélé dans le temps et l'espace comme le signal parasite affectant les données de SPIRE. Si d'autres informations sont à disposition, comme le fait que des événements impulsionnels de grandes amplitudes sont fortement probables, ou encore que le comptage de photons induit une loi de Poisson, il faut alors changer la loi de  $\mathbf{b}$  pour modéliser cette information.

Pour conclure, la vraisemblance est le terme qui collecte toute l'information provenant des données et est par conséquent fondamentale pour notre inférence. L'exploitation de cette information passe par un modèle d'acquisition  $\mathcal{H}$ , le chapitre 2 étant dédié à sa définition, et la modélisation de l'incertitude par la loi du bruit  $\mathbf{b}$ .

### 1.3.2 Information *a priori*

Tout comme pour le bruit, la loi *a priori* modélise les informations disponibles sur l'objet d'intérêt, avant toute acquisition de données. Ces informations sont généralement très nombreuses et variées. La loi *a priori* permet de modéliser ces informations ainsi que leur caractère incertain. Mentionnons par exemple les contraintes de positivité (il n'existe pas de flux négatif de photons), la régularité (un nuage de gaz n'est pas un tapis de fakir), la présence de plusieurs zones différentes comme la distinction entre un nuage et l'extérieur. La partie suivante décrit plus en détails les modèles probabilistes étudiés dans la littérature.

Les probabilités offrent une grande richesse et flexibilité pour modéliser ces informations et il existe plusieurs guides pour les déterminer et les formaliser.

- Les contraintes de la physique sont les plus naturelles. Citons par exemple les contraintes de positivité [MIB05] ( $p(\phi < 0) = 0$ ) ou les contraintes de support [GC05] (spatiale, temporel, *etc.*).
- Les informations d'experts sont très employées en pratique. Ce ne sont en général pas des informations aussi strictes que les contraintes. Elles définissent une certaine qualité recherchée dans la solution. Dans l'exemple présenté sur la déconvolution figure 1.2, il y a une information de régularité de la solution [Two63], c'est-à-dire qu'elle ne possède pas de grandes variations (d'un pixel à l'autre). On utiliserait alors une loi de probabilité qui modélise le fait que de grandes variations sont peu probables. Une image composée de zones aux caractéristiques différentes est un autre exemple.
- Enfin, il est possible de faire appel à des règles formelles, ou des règles par défaut. Dans ce cas on utilise plutôt des principes. On peut citer le « principe de raison insuffisante » de Laplace qui aboutit généralement à une loi uniforme sur l'ensemble des solutions possibles. Il

Il y a également les principes d'information minimum, ou de maximum d'entropie, motivés par l'envie d'injecter le minimum d'information dans le problème d'inférence [KW96].

Dans tous les cas, il est possible de faire appel à l'ensemble des modèles probabilistes pour modéliser ces informations. Pour cela nous présentons deux grandes classes de modèle très utilisés que sont les champs de Gibbs-Markov et les modèles à variables cachées, les deux pouvant être combinées.

**Les champs de Gibbs-Markov.** Ces modèles, introduits entre autre par les frères Geman en traitement de l'image [GG84], ont été inspirés par des travaux en physique statistique. Le principe est simple : à partir d'une description locale de l'objet à l'aide de caractéristiques markoviennes, on déduit un champ de Gibbs qui décrit un comportement global. Pour une présentation détaillée nous invitons le lecteur à se référer à [Idi01a], chapitre 7. Un champ de Gibbs très classique [Dem89, MMK06, Fér06] et beaucoup utilisé dans les travaux de cette thèse est le champ gaussien corrélé. Dans le cas d'une image, la structure markovienne précise que la probabilité de la valeur  $x_i$  d'un pixel  $i$  dépend de la valeur de ses voisins  $x_j, j \in \mathcal{V}$  au travers du paramètre  $r_{ij}$

$$p(x_i|\mathcal{V}) \propto \exp \left[ -\frac{1}{2} \sum_{j \in \mathcal{V}} r_{ij} |x_i - x_j|^2 \right].$$

Autrement dit, il est plus probable au sens d'une loi gaussienne, ou d'un potentiel quadratique, que  $x_i$  ait une valeur proche de ses voisins qu'une valeur éloignée. Le champ de Gibbs correspondant, c'est-à-dire la loi de *toute* l'image  $\mathbf{x} \in \mathbb{R}^N$ , est

$$p(\mathbf{x}) = (2\pi)^{-N/2} |\mathbf{R}_{\mathbf{x}}|^{-1/2} \exp \left[ -\frac{1}{2} \mathbf{x}^t \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{x} \right]$$

qui est un champ gaussien corrélé dont la covariance  $\mathbf{R}_{\mathbf{x}}$  s'obtient avec les  $r_{ij}$ . Ici le champ de Markov a été défini comme gaussien c'est-à-dire que nous avons utilisé une fonction de potentiel quadratique qui croît très vite vers l'infini. Autrement dit, les grandes différences entre pixels voisins sont très peu probables. D'autres fonctions de potentiel sont possibles. Par exemple la valeur absolue (une loi de Laplace) qui pénalise moins les différences, les potentiels dits  $L_2 - L_1$  quadratique autour de l'origine puis linéaire, ou encore les potentiels  $L_2 - L_0$  qui saturent à partir d'un certain seuil [BZ87]. Il y a alors toujours une équivalence avec un champ de Gibbs et donc une loi de probabilité sur l'objet entier. Bien que très puissants, ces potentiels ont des difficultés de mise en oeuvre [AIG06, GY95, GR92].

**Les modèles à variables cachées.** Ces modèles sont également beaucoup utilisés. Ils consistent à conditionner la probabilité de l'objet  $p(\mathbf{x}|\mathbf{c})$  par d'autres variables  $\mathbf{c}$  qui en fonction de leur valeur modifient la probabilité et donc l'information sur l'objet. Ces modèles sont par exemple utilisés en segmentation [Fér06], où les variables cachées déterminent à quelle classe les pixels appartiennent (nuage, fond, étoile, *etc.*), ou encore en déconvolution impulsionnelle [MIB05, CI93], où les variables cachées déterminent les instants d'apparition d'impulsions.

Les variables cachées ont également été utilisées dans les champs de Markov pour modéliser des variables de lignes qui en fonction de leur valeur viennent changer la structure markovienne par exemple en détectant les ruptures. Cette application est d'ailleurs l'utilisation historique par les frères Geman [GG84] de ces modèles. Black et Rangarajan [BR96] ont d'ailleurs montré qu'il y a une équivalence entre les modèles avec variables de ligne et les modèles de Gibbs-Markov définis par un potentiel particulier, en lien avec la loi de la variable de ligne.

Bien que ces modèles soient décrits pour l'objet *a priori*, il est envisageable de les appliquer à la loi *a priori* pour le bruit. Il serait par exemple possible de modéliser un bruit impulsionnel [FREM04] avec

des potentiels type  $L_1$  ou  $L_0$ . On peut également envisager un bruit corrélé, voir non stationnaire à l'aide de variables cachées.

Nous avons vu dans cette partie que l'approche bayésienne apporte des outils solides et variés pour aborder de nombreux problèmes. La richesse des modèles probabilistes possibles est un grand atout. Nous allons voir dans les deux prochaines parties comment l'approche bayésienne permet également de répondre à deux questions précises abordées dans cette thèse : le réglage du compromis entre les différentes sources d'information et l'estimation de paramètres de la réponse instrumentale lorsque celle-ci est mal connue.

### 1.3.3 Estimation non-supervisée

Nous avons vu partie 1.1.2 que la solution aux problèmes inverses mal-posés reposait toujours d'une manière ou d'une autre sur l'introduction d'information en plus des données. Il se pose ensuite naturellement la question du dosage respectif entre les différentes sources d'information pour mener l'inférence. Historiquement cela était fait manuellement et les méthodes étaient appelées supervisées. Dans le cas des critères composites, par exemple celui vu partie 1.1.2

$$\hat{\phi} \text{ minimise } \|\mathcal{H}(\phi) - \mathbf{y}\|^2 + \alpha \mathcal{F}(\phi),$$

le dosage intervient par l'intermédiaire du paramètre  $\alpha$ . Pour  $\alpha = 0$  l'inférence est effectuée seulement avec le terme d'attache aux données. Avec  $\alpha \rightarrow \infty$  l'inférence est effectuée avec uniquement l'*a priori*.

Dans le cadre bayésien, ce dosage s'effectue au travers des paramètres des lois de probabilité,  $\mathbf{R}_b$  et  $\mathbf{R}_x$  dans les exemples précédents. Pour fixer les choses, illustrons avec le cas gaussien scalaire réel où  $y = x + b$  et les lois de  $x$  et  $b$  sont gaussiennes centrées d'écart-type respectifs  $\sigma_x$  et  $\sigma_b$ , on a alors

$$\begin{aligned} p(x|y, \sigma_x, \sigma_b) &\propto p(y|x, \sigma_x)p(x, \sigma_b) \\ &\propto \frac{1}{\sigma_b \sigma_x} \exp\left[-\frac{1}{2} \frac{(y-x)^2}{\sigma_b^2}\right] \exp\left[-\frac{1}{2} \frac{x^2}{\sigma_x^2}\right]. \end{aligned}$$

Le maximum (égal à la moyenne ici) est obtenu en minimisant l'opposé du logarithme de la loi *a posteriori*

$$J(x) = \frac{(y-x)^2}{\sigma_b^2} + \frac{x^2}{\sigma_x^2}.$$

Cet exemple simple illustre :

- qu'il y a un lien entre les paramètres des lois  $\sigma_b$  et  $\sigma_x$  et le paramètre du critère composite précédent  $\alpha = \sigma_b^2/\sigma_x^2$ . Les deux paramètres de variance pilotent la confiance que l'on accorde à chaque source d'information. Si  $\sigma_* = 0$  une confiance totale est accordée. A l'inverse, si  $\sigma_* \rightarrow \infty$  aucune confiance n'est accordée.
- que les paramètres d'échelle de la loi, qui pilotent le compromis, sont présents dans le facteur de normalisation de la loi (ici  $1/\sigma_b \sigma_x$ ).

Le problème de l'estimation de ces paramètres se pose dans toutes les méthodes faisant appel à différentes sources d'information. De nombreuses approches existent. On peut mentionner la méthode de la courbe en L qui consiste à tracer l'énergie de l'erreur résiduelle par rapport à l'énergie de l'estimée [Han92] en fonction de l'hyper-paramètre. Cette courbe peut avoir l'aspect d'un L et la valeur au coude est choisie. Cette méthode ne fonctionne qu'avec un seul paramètre et la courbe n'a pas nécessairement un coude marqué. De plus le choix du point repose plus sur une interprétation graphique que sur un critère formalisé.

Une autre approche employée est la validation croisée [FD93]. Le principe est de minimiser, par rapport à l'hyper-paramètre, l'erreur de prédiction des données. Tout d'abord on estime l'objet en éliminant une donnée du problème puis celle-ci est prédite à partir de l'objet estimé. Ce processus est itéré sur toutes les données et on cherche la valeur de l'hyper-paramètre qui minimise cette erreur globale. Cette méthode dispose d'équations explicites pour le cas quadratique et s'étend difficilement aux cas plus généraux.

Une méthode très employée est la vraisemblance marginale (ou bayésien empirique) dans un cadre bayésien [MKM99, JBFZ02]. Supposons que nous avons une loi *a posteriori* définie comme

$$\begin{aligned} p(\mathbf{x}|\mathbf{y},\boldsymbol{\theta}) &\propto p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})p(\mathbf{x}|\boldsymbol{\theta}) \\ &\propto p(\mathbf{x},\mathbf{y}|\boldsymbol{\theta}) \end{aligned}$$

où  $\boldsymbol{\theta}$  est l'ensemble des paramètres des lois ( $\sigma_x$  et  $\sigma_b$  dans le cas précédent). Alors en marginalisant par rapport à ce qui est inconnu, c'est-à-dire  $\mathbf{x}$ , on obtient

$$\int p(\mathbf{x},\mathbf{y}|\boldsymbol{\theta}) d\mathbf{x} = p(\mathbf{y}|\boldsymbol{\theta})$$

la vraisemblance (marginale) des paramètres  $\boldsymbol{\theta}$  attachés aux données. On peut alors chercher les paramètres  $\boldsymbol{\theta}$  qui maximisent cette fonction. Elle souffre bien sûr de difficultés. En particulier la vraisemblance n'est pas convexe, y compris dans le cas gaussien sur l'objet, ce qui rend la recherche du maximum parfois difficile. De plus dans les cas non gaussien elle peut ne pas être explicite et il faut faire appel à des outils d'approximation d'intégrale comme l'algorithme EM ou SEM ou encore les MCMC [DMZB99].

Enfin une approche de plus en plus utilisée, et employée dans cette thèse, est l'approche dite entièrement bayésienne. Le principe est simple et cohérent avec l'approche bayésienne : comme les paramètres  $\boldsymbol{\theta}$  sont inconnus alors ils sont placés, au même titre que l'objet, à gauche de la barre de conditionnement dans la loi *a posteriori*, et en appliquant la règle de Bayes on obtient

$$p(\mathbf{x},\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})p(\mathbf{x},\boldsymbol{\theta})}{p(\mathbf{y})}.$$

En séparant les paramètres qui déterminent la vraisemblance  $\boldsymbol{\theta}_b$  et ceux de l'*a priori*  $\boldsymbol{\theta}_x$  et en supposant l'indépendance entre eux alors

$$p(\mathbf{x},\boldsymbol{\theta}_b,\boldsymbol{\theta}_x|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}_b)p(\mathbf{x}|\boldsymbol{\theta}_x)p(\boldsymbol{\theta}_b)p(\boldsymbol{\theta}_x)}{p(\mathbf{y})}.$$

Par rapport à la loi *a posteriori* précédente, sont apparues des lois *a priori* sur ces paramètres. L'intérêt de faire passer ces paramètres du côté gauche du conditionnement, outre la cohérence de l'approche, est qu'il est possible d'utiliser les outils mathématiques des probabilités pour manipuler la loi. Par exemple l'estimation peut être conjointe

$$\left[\hat{\mathbf{x}},\hat{\boldsymbol{\theta}}\right] = \arg \max_{\mathbf{x},\boldsymbol{\theta}} p(\mathbf{x},\boldsymbol{\theta}|\mathbf{y})$$

si l'estimateur du maximum *a posteriori* est choisi. Les paramètres qui ne sont pas des paramètres d'intérêt peuvent aussi être marginalisés, ou intégrés hors du problème,

$$p(\mathbf{x}|\mathbf{y}) = \int p(\mathbf{x},\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta},$$

la loi résultante étant toujours une loi *a posteriori*, qui peut être manipulée comme précédemment.

Le choix des lois *a priori* pour ces nouveaux paramètres doit suivre le même raisonnement que pour la loi sur  $\mathbf{x}$  : il s'agit d'une information à propos du paramètre. Dans l'exemple avec les variances de lois gaussiennes, il est logique d'utiliser des lois *a priori* qui ont un support sur  $[0, +\infty[$ . En pratique, les lois sont choisies pour faciliter les calculs. Dans ces cas il est souvent fait appel aux lois conjuguées [Rob05] : la loi *a posteriori* conditionnelle du paramètre est de la même famille que la loi *a priori*, ces paramètres ayant été mis à jour par la vraisemblance. Ce choix est généralement orienté par des contraintes de calculs. D'autres lois sont les lois dites non-informatives lorsque qu'aucune information n'est considérée disponible [KW96]. Par exemple la loi de Jeffreys, obtenue à l'aide d'outils de mesure d'information [Jay03],

$$p(m, \sigma) = \frac{1}{\sigma}$$

correspond à la loi non-informative pour les paramètres de position  $m$  et d'échelle  $\sigma$  (moyenne et variance pour une loi normale par exemple). C'est une loi uniforme sur tout  $\mathbb{R}$  pour  $m$  et en  $1/\sigma$  pour l'échelle.

Une difficulté concernant cette question est la connaissance du facteur de normalisation. Comme nous l'avons vu dans l'exemple gaussien scalaire, le facteur de normalisation (appelé aussi fonction de partition) dépend de ces paramètres. Pour les estimer il faut connaître la dépendance de la fonction de partition vis-à-vis de ces paramètres. Deux possibilités existent : soit l'expression est explicite, soit il faut calculer une approximation numérique [DMZB99, CGLS08, BMK09]. Les cas où la fonction de partition est explicite pour des lois multi-dimensionnelles sont réduits. Ce sont les cas séparables puisqu'ils se réduisent au produit de lois mono-dimensionnelles et les cas gaussiens corrélés

$$p(\mathbf{x}) = (2\pi)^{-N/2} |\mathbf{R}_{\mathbf{x}}|^{-1/2} \exp \left[ -\frac{1}{2} \mathbf{x}^t \mathbf{R}_{\mathbf{x}} \mathbf{x} \right]$$

si le déterminant de  $\mathbf{R}_{\mathbf{x}}$  [MMK06, FD93] est calculable (on peut également mentionner un cas particulier de champ convexe composite [Gio08]). Dans le cadre de cette thèse nous nous sommes concentrés sur les champs gaussiens corrélés et avons établi une méthode d'estimation conjointe de ces paramètres avec l'objet d'intérêt, décrit dans le chapitre 3.

### 1.3.4 Estimation myope

Une autre question concerne l'estimation de paramètres de la réponse instrumentale. Si le modèle direct est écrit comme

$$\mathbf{y} = \mathcal{H}_{\theta}(\phi) + \mathbf{b}$$

où  $\theta$  est un ensemble de paramètres de la réponse (une largeur de réponse optique, un gain, ...). La valeur de ces paramètres peut parfois être mal connue. Lorsque l'objet observé est également inconnu, il y a un problème d'indétermination puisque l'information présente dans les données peut être expliquée par le modèle d'acquisition ou l'objet observé. La solution repose sur l'ajout d'information dans l'inférence.

Il existe deux façons principales d'introduire de l'information pour lever l'indétermination. Certaines approches cherchent à estimer la réponse  $\mathcal{H}_{\theta}$  entièrement, de façon « non-paramétrique ». Dans le cas de la convolution discrète par exemple, il s'agirait d'estimer tous les coefficients de la réponse impulsionnelle. La solution employée consiste à introduire de l'information par l'intermédiaire de modèles probabilistes [MMK06, BMK09].

À l'inverse, certaines approches exploitent des informations disponibles sur une forme paramétrique de la réponse instrumentale provenant de la physique, de données de calibration, *etc.* La méconnaissance se réduit dans ce cas à quelques paramètres mal connus qui pilotent la forme de la

réponse [Maz05, BMI99, JBF02]. C'est le choix fait dans cette thèse étant donné que le chapitre 2 établit un modèle paramétrique bâti sur la physique. Dans ce cas l'information est introduite par contrôle de dimensions puisque l'espace des solutions est réduit à quelques paramètres.

**Remarque.** *Les modèles physiques, qui font le lien entre quelques paramètres d'intérêt inconnus (une vitesse, une concentration) et une grandeur physique « directement mesurée » comme un spectre, rentrent exactement dans cette modélisation paramétrique  $\mathcal{H}_\theta$ .*

Dans les deux cas (modèle paramétrique ou non), ces dernières années ont vu émerger des méthodes bayésiennes en traitement du signal et de l'image répondant à la question de l'estimation de ces inconnues [BMK09, Gio08, Fér06]. La démarche est la même que pour l'estimation des paramètres des lois. Comme ces grandeurs sont mal connues elles sont placées à gauche de la barre de conditionnement et la règle de Bayes fait apparaître des lois *a priori*

$$p(\phi, \theta | \mathbf{y}) = \frac{p(\mathbf{y} | \phi, \theta)p(\phi, \theta)}{p(\mathbf{y})}.$$

Les lois *a priori* à utiliser dépendent fortement du modèle utilisé. Dans le cas non-paramétrique, le choix est généralement orienté par la facilité de calcul. Dans le cas paramétrique, il y a souvent beaucoup d'information *a priori* sur l'instrument. C'est le cadre des travaux de cette thèse où nous considérons que la valeur des paramètres est dans un intervalle connu (un gain moyen, une largeur de réponse optique typique, *etc.*). Ces informations, entachées d'incertitude, existent. Elles ont notamment été utilisées par exemple pour concevoir l'instrument SPIRE dédié à l'étude d'objets encore jamais observés. Le chapitre 3 présente une méthode pour modéliser ces informations et fournir une estimation. Une difficulté est la dépendance potentiellement non-linéaire des ces paramètres vis-à-vis des données.

### 1.3.5 Conclusion

Nous avons présenté les atouts de l'approche bayésienne, pour l'estimation en général, et pour les problèmes inverses en particulier. Christian Robert avance de nombreux arguments en faveur de l'analyse bayésienne [Rob92], chapitre 10, mais nous ne retiendrons que celui-ci (l'emphase a été ajoutée)

Fondamentalement, la démarche statistique est une *démarche d'inversion*, puisqu'elle vise à remonter des effets aux causes en prenant en compte la nature probabiliste du modèle et l'influence additionnelle de facteurs purement aléatoires (ou inexplicables). Le *Théorème de Bayes*, discret ou continu, formalise cette inversion [...]. De ce point de vue, l'approche bayésienne est donc l'unique système *cohérent* qui respecte ce principe d'inversion.

Pour ma part je préfère parler de nature « incertaine » plutôt que « probabiliste » ou « aléatoire ». Dans la prochaine partie nous présentons les algorithmes classiques pour calculer sur des machines de calculs un estimateur, notamment ceux utilisés dans cette thèse.

## 1.4 Algorithmes de calcul

Comme nous l'avons vu dans la partie précédente, après avoir défini une loi *a posteriori* qui collecte toute l'information disponible et un estimateur ponctuel qui résume cette information, il faut calculer cet estimateur. Nous allons présenter les deux cas les plus généraux : le calcul du maximum *a posteriori* par optimisation et le calcul de la moyenne *a posteriori* par échantillonnage stochastique.

### 1.4.1 Optimisation

Lorsque l'estimateur choisi est le maximum *a posteriori* le problème se ramène à un problème de minimisation d'un critère

$$\begin{aligned}\hat{\phi} &= \arg \max_{\phi} p(\phi|\mathbf{y}) \\ &= \arg \min_{\phi} -f(p(\phi|\mathbf{y}))\end{aligned}$$

pour peu que  $f$  soit une fonction strictement croissante. Par exemple, dans le cas entièrement gaussien l'opposé du logarithme donne

$$\hat{\phi} = \arg \min_{\phi} \|\mathbf{y} - \mathcal{H}(\phi)\|_{\mathbf{R}_b}^2 + \alpha \|\phi\|_{\mathbf{R}_\phi}^2.$$

Pour simplifier le discours nous allons considérer que nous nous ramenons à un problème discret-discret, c'est-à-dire que l'objet, même s'il est défini continûment, s'exprime à l'aide de coefficients de décomposition  $\mathbf{x} \in \mathbb{R}^N$  de telle sorte que le critère à optimiser s'écrit

$$J(\mathbf{x}) = \mathcal{G}(\mathbf{y} - \mathcal{H}'(\mathbf{x})) + \alpha \mathcal{F}(\mathbf{x})$$

avec  $\mathbf{y} \in \mathbb{R}^M$  les données,  $\mathcal{G}$  la mesure de distance aux données, et  $\mathcal{F}$  la mesure de distance à l'*a priori*. Il y a plusieurs cas à envisager.

**Cas linéaire gaussien.** Dans ce cas les distances sont quadratiques et le modèle direct peut se mettre sous la forme d'une matrice  $\mathbf{H}$

$$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_{\mathbf{R}_b}^2 + \frac{\sigma_x^2}{\sigma_b^2} \|\mathbf{x}\|_{\mathbf{R}_x}^2.$$

en ayant factorisé les variances. Le minimiseur de ce critère est explicite [Idi01a], chapitre 4 et s'obtient en annulant la dérivée

$$\hat{\mathbf{x}} = \left( \mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{H} + \frac{\sigma_x^2}{\sigma_b^2} \mathbf{R}_x \right)^{-1} \mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{y}.$$

C'est la solution des moindres carrés régularisés. Cet estimateur est linéaire vis-à-vis des données et nécessite l'inversion d'une matrice de taille  $N$  par  $N$  ce qui est souvent impossible en pratique dans les applications en imagerie. Cependant si les matrices modélisent des opérations invariantes et si les matrices de covariance modélisent des phénomènes stationnaires, alors sous l'hypothèse de circularité les matrices sont diagonalisées dans l'espace de Fourier<sup>2</sup>

$$\hat{\mathbf{x}} = \mathbf{F}^\dagger \left( \mathbf{\Lambda}_H^\dagger \mathbf{\Lambda}_{\mathbf{R}_b}^{-1} \mathbf{\Lambda}_H + \frac{\sigma_x^2}{\sigma_b^2} \mathbf{\Lambda}_{\mathbf{R}_x} \right)^{-1} \mathbf{\Lambda}_H^\dagger \mathbf{\Lambda}_{\mathbf{R}_b}^{-1} \mathbf{F} \mathbf{y}$$

où  $\mathbf{F}$  est la matrice de transformée de Fourier discrète unitaire,  $\dagger$  le symbole du transpose conjugué et  $\mathbf{\Lambda}_H = \mathbf{F} \mathbf{H} \mathbf{F}^\dagger$  une matrice diagonale. La diagonalisation simplifie le problème puisque nous pouvons inverser les matrices directement et mettre en jeu seulement des produits terme à terme. Cet estimateur est appelé filtre de Wiener-Hunt [FHMP88, Hun72, Hun71]. Nous verrons dans le chapitre 3 comment nous avons exploité ces propriétés de diagonalisation.

---

<sup>2</sup>Le cas indépendant  $\mathbf{R}_x = \mathbf{R}_b = I$  est évidemment stationnaire.

**Cas non-invariant et convexe.** Dans le cas où le hessien du critère

$$\mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{H} + \frac{\sigma_x^2}{\sigma_b^2} \mathbf{R}_x$$

ne peut être inversé par des méthodes numériques, ou dans le cas où les distances ne sont pas quadratique (potentiel de Huber ou quadratique tronquée, norme absolue, *etc.*), l'estimateur n'est pas calculable directement ou explicitement. Cependant si les distances sont convexes et que le modèle  $\mathbf{H}$  est linéaire, alors il y a au moins un minimum et tout minimum local est global. Dans ce cas toute méthode itérative de calcul de minimum local converge vers l'estimateur. Une famille très classique correspond aux méthodes de descente.

Le principe est simple, en étant placé quelque part sur un critère convexe et en allant dans la direction opposée du gradient en ce point (et si  $\eta$  est suffisamment petit), le point d'arrivée correspond nécessairement à un critère plus petit. Mis en équation cela donne

$$\mathbf{x}^{(q+1)} = \mathbf{x}^{(q)} - \eta \left. \frac{dJ(\mathbf{x})}{d\mathbf{x}} \right|_{\mathbf{x}^{(q)}}$$

Le paramètre  $\eta > 0$  est le pas et indique la distance à parcourir. Ainsi à convergence, l'itération  $\mathbf{x}^{(k)}$  tend vers le minimiseur. Il existe beaucoup de variantes pour accélérer l'optimisation (gradient conjugué, pré-conditionnement, quasi-newton, pas optimal ou adaptatif, *etc.*). Nous invitons le lecteur à se référer aux livres de références tel que [NW00, Pol71].

**Cas général.** dans le cas général d'un modèle non-linéaire, ou d'un modèle linéaire mais avec des distances non-convexes, le critère peut avoir plusieurs minima. Il est bien sûr toujours possible d'utiliser une méthode de descente mais la convergence vers le minimum global n'est pas assurée.

Pour cela des outils comme le recuit-simulé, inspiré de la physique statistique, ont été introduits en traitement de données [GG84]. L'idée est simple : on fait l'analogie entre des pixels et des particules et on introduit un paramètre de température représentant l'excitation du système. A haute température le système est très excité et passe très facilement d'un état à l'autre. De façon équivalente le critère est très aplati, son exploration est plus facile sans être bloquée dans des minima locaux. Ensuite le système est refroidit et atteint un état stable correspondant à un critère piqué au niveau du minimum global. Sous certaines hypothèses de décroissance de la température on est assuré de converger vers le minimum global.

En pratique la décroissance peut être très lente et donc la mise en œuvre très lourde. C'est un outil qui permet cependant d'employer des modèles très complexes et riches.

Il existe également les méthodes semi-quadratiques [GR92, YBK<sup>+</sup>96, Idi01b]. Elles se placent dans le cadre d'algorithmes itératifs, comme la descente de gradient ou le recuit-simulé, dans les cas de modèles non-quadratiques corrélés. En effet le caractère non-quadratique rend le problème plus complexe, mais la dépendance entre pixels par la corrélation encore plus. Ces algorithmes permettent de transformer certains de ces problèmes en deux sous-problèmes : un quadratique corrélé sur les pixels, et un séparable (donc indépendant) sur des variables auxiliaires en espérant que chacun de ces problèmes soit plus simple.

Nous verrons dans le chapitre 3 que nous avons employé plusieurs de ces approches. En particulier notre modèle direct peut sous certaines hypothèses se séparer en une partie invariante diagonalisable dans Fourier et une autre non-invariante. Nous avons donc employé deux approches : le calcul par filtre de Wiener-Hunt et un algorithme de descente [BCG<sup>+</sup>06] pour le cas général.

### 1.4.2 Intégration

Choisir comme estimateur la moyenne *a posteriori*

$$\hat{\phi} = \int \phi p(\phi|\mathbf{y}) d\phi$$

implique le calcul d'une intégrale. Dans la plupart des cas cette intégrale est multiple et n'est pas explicite. Plusieurs solutions sont possibles pour la calculer.

Tout comme pour les problèmes d'optimisation, dans le cas où le modèle est linéaire et les lois gaussiennes, alors la loi *a posteriori* conditionnelle de l'objet est également gaussienne. Dans ce cas la moyenne est égale au maximum et le calcul se ramène à un problème d'optimisation (voir partie 1.4.1).

Il existe également l'approximation de Laplace qui consiste à approcher la fonction  $f(\phi) = \phi p(\phi|\mathbf{y})$  par une gaussienne centrée sur son maximum (qu'il faut calculer par ailleurs), puis à calculer la surface sous cette gaussienne. Il s'agit en fait de calculer le développement à l'ordre 2 de cette fonction. Dans le cas multidimensionnel cela revient à calculer le hessien autour du maximum de la fonction. Le hessien fournit la matrice de covariance de la gaussienne, dont on calcule le facteur de normalisation pour approcher la moyenne. En pratique cette méthode est peu employée puisqu'elle nécessite déjà le calcul du maximum.

Les approches dites bayésien variationnelles sont une extension de l'approximation de Laplace. Le principe est d'approcher, au sens de la distance de Kullback la loi jointe par une autre loi plus simple [Bea03, MMK06, AMD08]. Ce sont les lois marginales de la loi *a posteriori* qui sont approchées par des lois séparables puis la loi *a posteriori* est approchée par leur produit. Toute la corrélation entre les variables considérées indépendantes est perdue. Cependant cette approximation permet des calculs analytiques possibles. Les marginales peuvent être exprimées avec les paramètres des lois et la minimisation de la distance aboutit à un algorithme déterministe.

Les autres approches consistent en général à calculer par des outils numériques une approximation de l'intégrale. Les outils d'intégration directe sont possibles mais en inférence bayésienne on exploite en général la loi de probabilité en utilisant des générateurs pseudo-aléatoires. Dans ce cas le calcul de l'intégrale est effectué avec une intégration de type Monte-Carlo : si  $\{\mathbf{x}^{(q)}\}$  sont  $Q$  échantillons de la loi cible  $p(\mathbf{x}|\mathbf{y})$ , alors par les théorèmes limites (et sous réserve que la moyenne ait un sens ou existe) la moyenne arithmétique des échantillons tend vers la moyenne de la loi

$$\lim_{Q \rightarrow +\infty} \frac{1}{Q} \sum_{q=0}^{Q-1} \mathbf{x}^{(q)} = \mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}]. \quad (1.2)$$

En toute rigueur d'ailleurs, tous les cumulants de la loi peuvent être calculés et fournissent ainsi une expression de la loi. Donc les outils d'échantillonnage stochastique permettent de calculer une estimation de l'objet (par la moyenne de la loi) mais permettent également d'explorer la loi et d'en extraire l'information qu'elle contient (par exemple l'incertitude sur l'information à l'aide du moment d'ordre deux). Pour une présentation exhaustive de ces outils et de leurs possibilités nous invitons le lecteur à s'orienter vers des ouvrages comme [Rob96, MT93].

Nous voyons donc que si nous sommes capables d'obtenir des échantillons de la loi de probabilité alors nous pouvons calculer la moyenne *a posteriori*. Il existe différents outils pour simuler une loi ou calculer cette moyenne.

**Échantillonnage direct.** Dans certains cas il est possible de simuler directement la loi de probabilité en exploitant certaines propriétés mathématiques. Par exemple si  $x_1$  et  $x_2$  sont indépendamment distribués sous une loi uniforme entre 0 et 1  $\mathcal{U}_{[0,1]}$ , alors les variables,

$y_1 = \sqrt{-2 \ln x_1} \cos 2\pi x_2$  et  $y_2 = \sqrt{-2 \ln x_1} \sin 2\pi x_2$  suivent la loi normale centrée réduite  $\mathcal{N}(0,1)$ . Une méthode plus générale repose sur la connaissance de la fonction de répartition

$$F(x < X) = \int_0^X p(x) dx$$

et plus précisément de son inverse  $F^{-1}$ . En effet si  $u$  est distribuée sous une loi uniforme  $\mathcal{U}_{[0,1]}$ , alors la variable  $x = F^{-1}(u)$  a pour fonction de répartition

$$P(F^{-1}(u) < X) = P(u < F(X)) = F(X).$$

Un exemple est la loi normale à support positif  $\mathcal{N}^+(\mu, \sigma^2)$

$$p(x) = \left( \sqrt{\frac{\pi\sigma^2}{2}} \left[ 1 + \operatorname{erf} \left( \frac{\mu}{\sqrt{2\sigma^2}} \right) \right] \right)^{-1} \exp \left[ -\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2} \right]$$

avec la fonction erreur

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp[-t^2] dt.$$

Cette loi peut être simulée [Dev85] par inversion de fonction de répartition<sup>3</sup> : si  $u \sim \mathcal{U}_{[0,1]}$  alors

$$x = \mu + \sqrt{2\sigma^2} \operatorname{erf}^{-1} \left[ u + \operatorname{erf} \left( \frac{\mu}{\sqrt{2\sigma^2}} \right) \right]$$

est distribuée sous une loi normale tronquée. Ces approches sont les plus efficaces puisque le rendement est de 1. Elles nécessitent cependant la connaissance de l'inverse de la fonction de répartition ce qui n'est souvent pas le cas.

**Échantillonnage d'importance.** Lorsque la simulation d'une loi de probabilité est trop coûteuse pour calculer

$$\int \phi p(\phi) d\phi$$

alors l'expression équivalente

$$\int \phi p(\phi) d\phi = \int \phi \left[ \frac{p(\phi)}{g(\phi)} \right] g(\phi) d\phi$$

permet de simuler des échantillons  $\phi^{(k)} \sim g(\phi)$  et de calculer la moyenne arithmétique

$$\frac{1}{Q} \sum_{q=0}^{Q-1} \phi^{(q)} \frac{p(\phi^{(q)})}{g(\phi^{(q)})}.$$

Le problème de la simulation de  $p(\phi)$  est réduit à la simulation de la loi dite instrumentale  $g(\phi)$ , et de l'évaluation des fonctions  $p$  et  $g$  en  $\phi^{(q)}$ . Il se pose deux difficultés en pratique qui sont le choix de la loi instrumentale  $g$  (qui n'est pas évident pour obtenir un algorithme efficace [RC00]) et la connaissance de la fonction  $p$ , y compris son facteur de normalisation. Or le facteur de normalisation des lois *a posteriori* est souvent inconnu.

**Acceptation-rejet.** S'il est possible d'établir une borne supérieure de la loi cible  $p(\phi) < Mg(\phi)$  à un facteur près  $M$  (donc la loi cible est connue à un facteur près), l'algorithme d'acceptation-rejet consiste en

1. simuler  $\phi \sim g(\phi)$  et  $u \sim \mathcal{U}_{[0,1]}$

<sup>3</sup>Il faut faire attention aux instabilités numériques dû aux approximations des fonctions  $\operatorname{erf}$  et  $\operatorname{erf}^{-1}$  [Maz05].

2. si

$$u < \frac{p(\phi)}{Mg(\phi)} \begin{cases} \phi \text{ est un échantillon de } p \\ \text{sinon recommencer} \end{cases}$$

Tous les échantillons obtenus sont des échantillons de  $p$ . Cependant il peut y avoir un grand nombre de rejets pour obtenir un seul échantillon. L'idée de cet algorithme est simple, elle consiste à déterminer une enveloppe  $g$  de la loi cible  $p$  et d'échantillonner uniformément le couple  $(u, g(\phi))$ , c'est-à-dire la surface sous  $g$ . En ne conservant que les couples  $(u, p(\phi))$  tels que  $u < p(\phi)/Mg(\phi)$  alors c'est la surface sous  $p$  qui est échantillonnée uniformément et la marginalisation par rapport à  $u$  fournit des échantillons de  $p$ . Les difficultés posées ici reposent clairement sur le choix de la loi instrumentale  $g$ . Elle doit approcher au mieux la loi cible pour éviter un trop grand nombre de rejets. Ce peut être le cas par exemple avec des lois cibles très piquées, concentrées autour d'une valeur, et une loi instrumentale très large.

**Monte-Carlo par chaîne de Markov.** Le principe des méthodes MCMC est de construire une chaîne de Markov qui à convergence produit des échantillons distribués sous une loi cible (ici  $p$ ). La simulation d'échantillons de cette chaîne est généralement beaucoup plus facile que la simulation directe de  $p$ . Ces algorithmes peuvent également être plus efficaces pour des problèmes de grande dimension.

La loi d'une chaîne de Markov converge (presque sûrement au sens d'une distance entre mesure de probabilité) vers une loi cible si [Rob96] :

- les échantillons suivants sont sous cette loi cible lorsque la loi cible est atteinte (stationnarité).
- tout l'espace des possibles peut être atteint en un nombre fini d'itérations (irréductibilité).  
En théorie la chaîne n'est pas piégée dans des modes à forte probabilité (ou minima locaux).
- la chaîne n'est pas périodique, ou « déterministe » (apériodicité)
- la chaîne peut passer une infinité de fois aux mêmes endroits (récurrence au sens de Harris, indépendance face à l'initialisation).

et dans ce cas elle est dite ergodique. L'objectif est donc de construire une chaîne de Markov dont la loi converge vers la loi d'intérêt  $p$ . Les algorithmes de Gibbs et Metropolis-Hastings sont des algorithmes génériques, construits à partir de  $p$ , qui assurent l'ergodicité (et donc la convergence vers  $p$ ) de la chaîne.

L'algorithme de Gibbs permet de construire une chaîne de Markov en simulant les lois conditionnelles *a posteriori* plutôt que la loi totale. Par exemple si la loi cible est  $p(x_1, x_2|y)$  alors le schéma d'échantillonnage des lois conditionnelles

1. initialiser  $x_2^{(0)}$
2. simuler  $x_1^{(k+1)} \sim p(x_1|x_2^{(k)}, y)$
3. simuler  $x_2^{(k+1)} \sim p(x_2|x_1^{(k+1)}, y)$
4. retourner en 2.

fournit, à convergence, des échantillons sous la loi  $p(x_1, x_2|y)$ . Ces algorithmes sont très utilisés puisque généralement, par construction de la loi *a posteriori* (entre autre avec des loi *a priori* conjuguées), ces lois conditionnelles sont faciles à simuler. On voit d'ailleurs qu'il n'est pas nécessaire de connaître la fonction de partition de la loi *a posteriori* (mais de connaître celle des lois conditionnelles). Nous verrons dans le chapitre 3 comment nous avons employé cet algorithme pour estimer l'objet d'intérêt ainsi que les paramètres des lois de probabilités.

Si la cible, ou une des lois conditionnelles de l'algorithme de Gibbs, est trop complexe (y compris les lois de scalaires) alors une solution toujours possible est l'algorithme de Metropolis-Hastings. Cette difficulté peut par exemple intervenir si le facteur de normalisation n'est pas connu. Le principe est d'accepter, avec une certaine probabilité, des échantillons générés par

une loi instrumentale  $g$  (ou candidate), ou alors de conserver les anciens (ce qui diffère de l'acceptation-rejet). L'algorithme se décompose comme suit

1. initialiser  $\phi^{(0)}$
2. simuler une proposition  $\phi_p \sim g(\phi)$
3. prendre

$$\phi^{(i+1)} = \begin{cases} \phi_p & \text{avec la probabilité } P \\ \phi^{(i)} & \text{avec la probabilité } (1 - P) \end{cases}$$

où

$$P = \min \left\{ 1, \frac{p(\phi_p)}{g(\phi^{(i)})} \frac{q(\phi^{(i)}|\phi_p)}{g(\phi_p|\phi^{(i)})} \right\}$$

Le grand intérêt de cet algorithme est qu'il est capable de fournir des échantillons de n'importe quelle loi cible  $p$  pour peu que  $p$  puisse être évaluée à *une constante près*, puisque l'algorithme ne fait intervenir que le rapport de la loi. L'inconvénient est que le choix de la loi instrumentale est critique et peut aboutir à un taux d'acceptation très faible. L'exploration de la loi, ou la convergence, est alors très lente. Nous verrons également dans le chapitre 3 comment nous avons utilisé cet algorithme pour estimer les paramètres d'un modèle d'acquisition.

### 1.4.3 Conclusion

Nous avons vu jusqu'ici la problématique posée par l'estimation d'un rayonnement astrophysique à l'aide des données fournies par SPIRE et l'approche bayésienne que nous utiliserons pour résoudre le problème. Nous avons également vu les outils classiques de calcul d'un estimateur. Les deux prochaines parties présentent l'état de l'art dans le contexte de SPIRE, respectivement de la communauté en astrophysique et de la communauté du traitement du signal.

## 1.5 MADmap et SANEPIC

Nous décrivons dans cette partie les méthodes existantes pour l'estimation de carte du ciel à l'aide des données de SPIRE. Nous verrons qu'elles rentrent exactement dans le cadre qui vient d'être présenté.

### 1.5.1 MADmap

MADMAP est le nom donné à la méthode officielle de l'ESA pour l'estimation de carte du ciel [CCB<sup>+</sup>06]. Le modèle d'acquisition considéré pour MADMAP est simple :

- le ciel est monochromatique au sein d'une bande spectrale soit  $\phi(\alpha, \beta, \lambda) \neq 0$  si et seulement si  $\lambda = 250 \mu\text{m}$  par exemple. Ensuite le ciel est constant par morceaux dans l'espace, ou décomposé sur des indicatrices de pixels, soit  $\phi(\alpha, \beta, \lambda_0) = x_{ij}$  si et seulement si  $(\alpha, \beta) \in [\alpha_i, \alpha_{j+1}] \times [\beta_i, \beta_{j+1}]$ . Il s'agit donc d'une image au sens usuel.
- si un bolomètre (un élément sensible des capteurs de SPIRE) pointe à l'instant  $t$  dans un pixel du ciel, la donnée correspondante vaut la valeur du pixel, éventuellement avec une correction de gain considérée connue.
- Nous ne considérerons pas ici, puisqu'elles ne se situent pas sur le même plan, les différentes étapes d'interpolation, d'extrapolation, de lissage, *etc.* qui permettent de rendre plus vraisemblable le modèle ou de modifier les données pour qu'elles correspondent plus aux hypothèses.

En empilant toutes les données de tous les bolomètres à tous les instants, ce modèle d'acquisition s'écrit

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} 0 & \dots & 1 & 0 & \dots \\ 1 & 0 & \dots & & \\ & & \dots & & \\ 0 & \dots & & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

ou en plus condensé

$$\mathbf{y} = \mathbf{H}\mathbf{x}.$$

Donc la matrice  $\mathbf{H}$  est une matrice de pointage qui indique pour chaque donnée  $y_i$  quel pixel était observé. Elle contient donc au plus un seul 1 par ligne puisqu'un bolomètre ne peut voir plusieurs pixels du ciel. Il s'agit donc d'une matrice de troncature. Le modèle ne prend pas en compte la longueur d'onde et les réponses instrumentales (chapitre 2) comme la figure de diffraction du miroir ou la réponse dans le temps du bolomètre (ou encore l'électronique de lecture). De plus le modèle est dégénéré par la présence de colonnes remplies de 0, qui correspondent tout simplement aux pixels non-observés (c'est un modèle à données manquantes). Autrement dit la solution au problème ne sera pas unique : ces pixels pourront prendre n'importe quelle valeur, les données seront toujours parfaitement expliquées.

L'ordre des données dans  $\mathbf{y}$  est important à cause du modèle de bruit considéré. Pour prendre en compte les dérives observées (voir partie 1.1.1), la méthode considère un bruit  $\mathbf{b}$  tel que

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{b}$$

avec comme *a priori* que ce signal  $\mathbf{b}$  a comme loi de probabilité une loi normale centrée corrélée dans le temps avec  $\mathbf{R}_b$  comme matrice de covariance. Autrement dit, le signal produit par chaque bolomètre est dégradé par un bruit stationnaire, indépendant entre bolomètres, mais corrélé dans le temps au sein d'un même bolomètre. Si les données sont ordonnées comme « tous les instants d'un bolomètre, puis tous les instants du suivant », la matrice de covariance est une matrice bloc

$$\mathbf{R}_b = \begin{pmatrix} \mathbf{R}_b^{11} & \mathbf{R}_b^{12} & \dots & \mathbf{R}_b^{1n} \\ \mathbf{R}_b^{12} & \mathbf{R}_b^{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_b^{1n} & \dots & \dots & \mathbf{R}_b^{nn} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_b^{11} & 0 & \dots & 0 \\ 0 & \mathbf{R}_b^{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \mathbf{R}_b^{nn} \end{pmatrix}$$

où  $\mathbf{R}_b^{ij}$  désigne la covariance entre les bolomètre  $i$  et  $j$ . Dans MADMAP, cette matrice  $\mathbf{R}_b$  est considérée connue. Elle est donc estimée par ailleurs avec une méthode qui n'est pas clairement établie.

Avec cette loi de probabilité la vraisemblance est

$$p(\mathbf{y}|\mathbf{x}, \mathbf{R}_b) = (2\pi)^{-N/2} \det[\mathbf{R}_b]^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{H}\mathbf{x})^t \mathbf{R}_b^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \right]$$

et MADMAP recherche le maximum de vraisemblance, c'est-à-dire la maximum de  $p(\mathbf{y}|\mathbf{x}, \mathbf{R}_b)$  lorsque ce sont les données qui sont connues et  $\mathbf{x}$  qui est inconnue. La méthode n'exploite pas d'informations disponibles par ailleurs comme la régularité spatiale. Le maximum de vraisemblance

est la solution des moindres carrés pondérés car

$$\begin{aligned}
\hat{\mathbf{x}} &= \arg \max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}, \mathbf{R}_b) \\
&= \arg \min_{\mathbf{x}} -\log p(\mathbf{y}|\mathbf{x}, \mathbf{R}_b) \\
&= \arg \min_{\mathbf{x}} (\mathbf{y} - \mathbf{H}\mathbf{x})^t \mathbf{R}_b^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \\
&= \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_{\mathbf{R}_b}^2 = \arg \min_{\mathbf{x}} J(\mathbf{x}).
\end{aligned}$$

L'argument du minimum est explicite et s'obtient en annulant le gradient du critère

$$\begin{aligned}
\nabla J(\mathbf{x}) &= \nabla \left( \mathbf{y}^t \mathbf{R}_b^{-1} \mathbf{y} - 2\mathbf{x}^t \mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{y} + \mathbf{x}^t \mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{H} \mathbf{x} \right) \\
&= -2\mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{y} + 2\mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{H} \mathbf{x}
\end{aligned}$$

ce qui aboutit à l'équation normale

$$(\mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{H}) \mathbf{x} = \mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{y}. \quad (1.3)$$

Si la matrice à droite est inversible, ce système linéaire a comme solution

$$\hat{\mathbf{x}} = (\mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{H})^{-1} \mathbf{H}^t \mathbf{R}_b^{-1} \mathbf{y}.$$

Il s'agit de la solution des moindres carrés avec un bruit corrélé (si le bruit est blanc *iid* alors  $\mathbf{R}_b^{-1}$  est l'identité et disparaît des écritures). Le modèle est dégénéré donc il existe une infinité de solutions et parmi cet ensemble est retenue celle à norme minimale, c'est-à-dire que les pixels non-observés sont mis à zéro.

Plusieurs remarques concernant cette méthode peuvent être faites. Tout d'abord elle utilise un modèle simple pour expliquer la formation des données. Elle ignore totalement la formation d'une image dégradée sur la plan focal par la diffraction et la réponse des capteurs. De plus elle considère que le flux lumineux est monochromatique or les réponses de l'instrument dépendent de la longueur d'onde. De plus pour éviter que des pixels ne soient pas observés (et donc avoir des « trous » dans l'image reconstruite) elle est obligée de considérer un pas d'échantillonnage et donc une taille de pixel assez grand. Par ce choix elle limite considérablement la résolution possible de la reconstruction.

D'ailleurs, cette limitation de la résolution de la reconstruction est une forme de régularisation par contrôle de dimension. En effet en limitant la résolution du ciel (ou la dimension par le nombre de pixels), et donc son contenu spectral, elle injecte une information *a priori* sur l'objet. En pratique, le choix de la résolution est autour de 6 secondes d'arc alors que les différents pointages sont espacés de 2 secondes d'arc (au moins dans le protocole *Large map* voir partie 2.2.6).

En conclusion, en se limitant à la seule information provenant des données avec un modèle d'interprétation simple, la méthode MADMAP ne permet pas de restaurer des fréquences qui pourraient l'être. En prenant en compte l'information disponible sur l'instrument et une information disponible sur le ciel (pour limiter l'instabilité) ce serait possible. Pour aller plus loin, cette méthode n'est pas capable de restaurer un ciel qui pourrait être à bande infinie si on injectait comme information *a priori* la présence d'étoiles [GC05].

Enfin le modèle de bruit de cette méthode considère que le signal  $\mathbf{b}$  est indépendant entre chaque bolomètre. L'analyse des données fournie par les bolomètres aveugles (qui ne reçoivent aucun flux directement provenant du ciel), montre une très forte corrélation entre eux. Autrement dit le signal  $\mathbf{b}$  est corrélé dans le temps et l'espace, phénomène mal pris en compte par ce modèle.

### 1.5.2 SANEPIC

La méthode appelée SANEPIC [PAB<sup>+</sup>08] est une extension de MADMAP. Le modèle instrument est le même et l'estimateur également (maximum de vraisemblance). Ce qui change est la loi du bruit qui considère un bruit corrélé dans l'espace également. Plus exactement, la méthode considère que le bruit  $\mathbf{b}_i$  du bolomètre  $i$  est la somme de deux composantes

$$b_i(t) = b^s(t) + b_i^t(t)$$

où  $t$  est le temps. Le signal  $\mathbf{b}^s$  est donc commun à tous les bolomètres et  $\mathbf{b}_i^t$  est différent pour chaque bolomètre. Les deux sont corrélés dans le temps et la corrélation spatiale (de 1) vient du fait qu'il y a un signal commun à tous les bolomètres. Donc finalement chaque signal suit comme loi

$$\mathbf{b}^s \sim \mathcal{N}(0, \mathbf{R}_b^s) \quad (1.4)$$

$$\mathbf{b}_i^t \sim \mathcal{N}(0, \mathbf{R}_b^t) \quad (1.5)$$

et la matrice de covariance de  $\mathbf{b}$  s'écrit

$$\mathbf{R}_b = \begin{pmatrix} \mathbf{R}_b^s + \mathbf{R}_b^t & \mathbf{R}_b^s & \cdots & \mathbf{R}_b^s \\ \mathbf{R}_b^s & \mathbf{R}_b^s + \mathbf{R}_b^t & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_b^s & \cdots & \cdots & \mathbf{R}_b^s + \mathbf{R}_b^t \end{pmatrix}$$

où chaque bloc de la matrice est la covariance entre tous les échantillons de deux bolomètres. Le fait d'utiliser un modèle comme celui-là est une contrainte très forte qui permet de structurer la matrice de covariance et donc de réduire sa complexité. En pratique cette matrice est inconnue et il faut donc utiliser une estimation. Le papier de référence propose une méthode de séparation de sources [PAB<sup>+</sup>08], mais des données de calibration peuvent être employées également. Cette méthode souffre naturellement des mêmes défauts que MADMAP concernant les *a priori*, le modèle de ciel et le modèle instrument. Elle souffre également d'une autre limitation concernant la continuité des données. En effet si certaines données sont considérées manquantes (par un défaut d'acquisition ou parce qu'elles sont trop bruitées) alors la structure de corrélation est brisée. L'approche proposée par SANEPIC repose sur une interpolation (ou remplissage) des données manquantes.

Dans les deux méthodes, les matrices de covariances sont des matrices de corrélation. Donc les signaux sont supposés stationnaires. De plus les deux méthodes supposent les matrices circulantes ce qui permet d'utiliser la transformée de Fourier discrète. Ainsi les estimateurs du maximum de vraisemblance s'écrivent

$$\hat{\mathbf{x}} = (\mathbf{H}^t \mathbf{F}^\dagger \mathbf{\Lambda}_{\mathbf{R}_b}^{-1} \mathbf{F} \mathbf{H})^{-1} \mathbf{H}^t \mathbf{F}^\dagger \mathbf{\Lambda}_{\mathbf{R}_b}^{-1} \mathbf{F} \mathbf{y}.$$

où  $\mathbf{F}$  est la transformée de Fourier discrète unitaire,  $\dagger$  le symbole du transposé conjugué de la matrice et  $\mathbf{\Lambda}$  une matrice diagonale. Ensuite l'estimateur est calculé par un algorithme de descente de gradient (en arrêtant les itérations assez vite semble-t-il [CCB<sup>+</sup>06]).

## 1.6 État de l'art en déconvolution et sur-résolution

Dans une certaine mesure les travaux développés ici sont très proches des thèmes de recherche en déconvolution d'image et sur-résolution. En effet la physique sous-jacente au fonctionnement de l'instrument se modélise avec des convolutions continues décrites dans le chapitre 2. De plus la prise en compte de l'échantillonnage spatial et temporel ainsi que la redondance introduite dans les données par l'intermédiaire du protocole d'observation permet de faire des liens, développés dans la partie 3.3.1, avec les modèles classiquement utilisés en sur-résolution.

### 1.6.1 Déconvolution

La déconvolution est un thème de recherche actif depuis longtemps notamment à travers les travaux fondateurs de Tikhonov [TA76] et de Twomey [Two63]. La problématique rentre généralement dans le cadre présenté plus haut. La solution est souvent définie à travers la minimisation d'un critère sur l'inconnue  $\mathbf{x}$ , le lien avec les données étant effectué au travers d'un modèle convolutif  $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{b}$ .

Ces dernières années ont vu la communauté se concentrer principalement sur les approches bayésiennes pour formaliser l'introduction d'information *a priori*. Ce formalisme permet d'étendre avec une même approche la question de l'estimation d'autres paramètres que seulement l'image d'intérêt comme décrit partie 1.3.3 et 1.3.4. Les questions traitées sont par exemple la déconvolution « aveugle » qui désigne la problématique d'estimation de l'image conjointement à la réponse impulsionnelle. Une autre classe de problème concerne l'estimation des hyper-paramètres ou des paramètres des lois dans le formalisme bayésien.

Dans ce cadre on peut mentionner les travaux de Molina *et al.* [MMK06] qui propose une méthode bayésienne de déconvolution « aveugle » et non-supervisée. Le modèle  $\mathbf{H}$  est linéaire et convolutif. Le modèle probabiliste pour l'image est quadratique corrélé sur l'image ainsi que la réponse impulsionnelle dont tous les pixels, au sein d'un support potentiellement restreint, sont estimés. Les paramètres de variance de la loi de l'image, la loi de la réponse et la loi du bruit blanc gaussien sont estimés. Les lois *a priori* de ces paramètres sont des lois conjuguées gamma. Les auteurs proposent ensuite un algorithme bayésien variationnel pour explorer la loi (voir partie 1.4.2). Les auteurs se placent dans un cas où peu d'information est disponible sur la réponse à part une régularité spatiale ce qui n'est pas le cas ici où un modèle instrument paramétré est disponible.

Plus récemment, Babacan, Katsaggelos et Molina [BMK09] ont fait évoluer leurs approches pour exploiter un *a priori variation total*  $\sqrt{|\mathbf{D}\mathbf{x}|^2}$  où  $\mathbf{D}$  est un opérateur de différence. Pour mener les calculs plusieurs approximations sont effectuées. Tout d'abord la fonction de partition, inconnue dorénavant, est approchée par  $\gamma_{\mathbf{x}}^{N/2}$  où  $\gamma_{\mathbf{x}}$  est la précision d'une gaussienne. Cette approximation permet d'exploiter un *a priori* gamma sur l'hyper-paramètre. Enfin le potentiel *variation total*  $\sqrt{|u|^2}$  est approché par l'inégalité  $\sqrt{u} \leq (u+z)/(2\sqrt{z})$ . Cela permet de faire apparaître un problème avec une loi quadratique sur les pixels  $u$  et « augmenté » de variables auxiliaires  $z$  séparables. Ces variables permettent d'éloigner le modèle d'une pénalisation quadratique. Malheureusement aucun lieu n'est fait avec les approches semi-quadratiques.

### 1.6.2 Sur-résolution

La sur-résolution (ou super-résolution) consiste à reconstruire à partir de plusieurs images basse-résolution d'une même scène une image haute résolution [PPK03]. Le modèle d'acquisition peut être plus ou moins complexe et dépend fortement du contexte. Le plus souvent il consiste à modéliser l'acquisition dans le temps d'une scène fixe avec des translations inférieures à la taille du capteur. Dans ce cas il se modélise comme  $\mathbf{y} = \mathbf{P}\mathbf{H}\mathbf{x} + \mathbf{b}$  où  $\mathbf{H}$  est une matrice de convolution et  $\mathbf{P}$  une matrice de décalage et décimation. Une difficulté apparaît puisque dans ce cas le modèle n'est plus globalement invariant.

La sur-résolution pose des problèmes variés et complexes comme le choix des modèles *a priori*, l'estimation des décalages qui peuvent être inconnus, le choix d'un modèle d'acquisition permettant de prendre en compte des phénomènes plus complexes comme la rotation [RCLBG06] ou encore le déplacement d'objets dans la scène [SS96]. Les travaux présentés dans cette thèse peuvent être rattachés à la sur-résolution. Dans notre cas la scène n'évolue pas (ou à des échelles de temps trop importantes) et les décalages entre les images sont connus.

Concernant les méthodes proposées dans la littérature on peut citer les travaux de Farsiu *et al.* [FREM04] qui proposent une minimisation d'un critère robuste ( $L_1$ ) sur l'attache aux données et robuste sur le gradient de l'image pour la régularisation. Le calcul de l'estimateur est fait par un algorithme de descente de gradient. Chantas *et al.* [CGW07] propose de mener l'inférence avec une approche bayésienne. Dans ces travaux le modèle *a priori* pour l'image est une gaussienne mono-dimensionnelle, avec un paramètre de variance inconnu, pour chaque pixel du gradient de l'image. Les mêmes auteurs proposeront [CGLS08] une interprétation de leur modèle *a priori* sous forme de lois de *t*-Student exprimées comme mélange par échelle de gaussiennes (*Scale Mixture of Gaussian*). De nouveau aucun lien n'est fait avec les approches semi-quadratiques dans le format Geman et Reynolds, lien établi notamment dans [CI04].

## 1.7 Conclusion

Nous avons montré quels sont les problèmes posés par l'estimation du ciel à partir des données de l'instrument SPIRE. Nous savons déjà que l'information présente dans l'objet recherché a été dégradée de façon telle que l'inversion posera des difficultés. En pratique l'inversion du modèle d'acquisition sera mal conditionnée et une approche naïve donnera des solutions instables.

Nous avons exposé notre choix de mener l'inversion dans un cadre bayésien et avons argumenté dans ce sens, en exposant notamment les apports de cette approche ainsi que les difficultés impliquées par chaque choix. Enfin, nous avons présenté les différents algorithmes de calcul si le maximum (optimisation) ou la moyenne (intégration) *a posteriori* sont choisis comme estimateur.

Enfin les méthodes existantes MADMAP et SANEPIC, ont été décrites. Elles sont suffisamment formalisées pour permettre une analyse des hypothèses sous-jacentes et donc de leurs potentiels et de leurs limites. Des travaux en déconvolution et sur-résolution présentés permettent de situer les travaux de cette thèse.

Le prochain chapitre est une étude de l'instrument et construit un modèle d'acquisition. Nous avons vu dans les parties sur l'optimisation 1.4.1 et l'échantillonnage 1.4.2 que beaucoup de ces algorithmes ont un caractère itératif. Ils exploitent en effet une comparaison entre des données reproduites par le modèle et les données acquises (une mesure de vraisemblance). Par conséquent les temps de calculs ne doivent pas être excessifs. Le modèle est en partie bâti dans ce sens tout en décrivant au plus près les propriétés de l'instrument.

Le chapitre 3 est consacré à la méthode d'inversion proprement dite. Elle se place dans un cadre bayésien et exploite le modèle direct linéaire non-invariant, résultat du chapitre 2. Le résultat est un algorithme MCMC qui permet l'estimation conjointe du ciel d'intérêt, de la dérive thermique, des hyper-paramètres et des paramètres instrumentaux.

Enfin le chapitre 4 présente une étude et une validation expérimentale des résultats du chapitre 3. L'étude met en avant les possibilités de restauration de composantes spatiales hautes-fréquences du ciel. Une analyse plus préliminaire porte sur l'estimation des hyper-paramètres.

## Chapitre 2

# Modèle d'acquisition pour SPIRE

Ce chapitre décrit le modèle direct, c'est-à-dire le modèle du processus d'acquisition en vue de reproduire les données observées. La part la plus importante concerne la modélisation de SPIRE mais une partie plus générale sur le miroir et les cornets concerne plus globalement le télescope Herschel.

L'objectif du modèle est de décrire le fonctionnement de l'instrument. Plus la description sera fine et pertinente, mieux l'information présente dans les données sera traitée. Par conséquent, la méthode de traitement sera plus à même de faire la distinction entre l'information provenant de l'objet d'intérêt et la dégradation introduite par l'instrument.

Face à ce principe de fidélité à la réalité que doit suivre le modèle, il convient d'opposer une contrainte pratique. En effet, le modèle est destiné à être inclus dans une méthode de traitement. Or ces méthodes, comme celle décrite dans le chapitre suivant ou celles décrites en introduction, font de nombreux appels au modèle pour reproduire des données et calculer des termes d'attache aux données. C'est une des distinctions avec un simulateur [SCG09, SG06] destiné à être utilisé une seule fois. Il faut donc faire des choix ou approximations permettant des temps de calculs raisonnables.

### 2.1 Introduction

SPIRE est un instrument complexe qui fait partie de l'observatoire Herschel ayant lui-même ses particularités. Globalement le modèle peut se mettre sous la forme d'un schéma bloc illustré figure 2.1. Chaque bloc sera décrit ultérieurement.

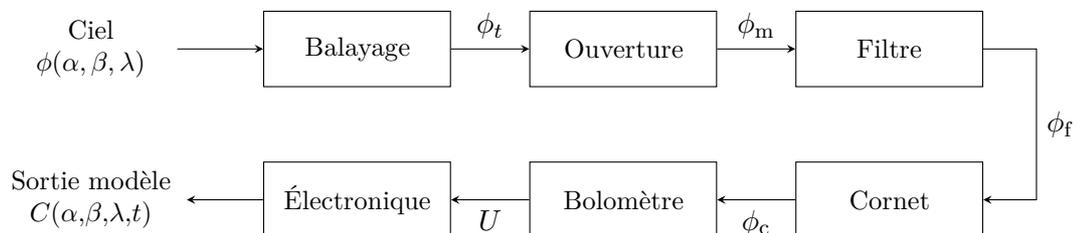


Figure 2.1 – Schéma bloc du modèle.

L'entrée  $\phi$ , l'objet d'intérêt, est un rayonnement. Il est caractérisé par deux dimensions spatiales  $(\alpha, \beta)$  et une spectrale  $\lambda$ . Dans la description du modèle direct il est considéré fixé et l'on cherche à reproduire les données connaissant cette entrée. Pour le moment on peut se limiter à une définition

fonctionnelle comme

$$\begin{aligned} \phi \in L_2 : \mathbb{R}^2 \times \mathbb{R}_+^* &\rightarrow \mathbb{R}_+ \\ (\alpha, \beta, \lambda) &\rightarrow \phi(\alpha, \beta, \lambda). \end{aligned} \tag{2.1}$$

Après l'entrée se trouve une succession de systèmes qui modélisent différents étages de l'instrument modifiant le signal [Gri06]. Chaque système est caractérisé par une relation entrée–sortie et se succède jusqu'à la reproduction des données appelées sortie modèle.

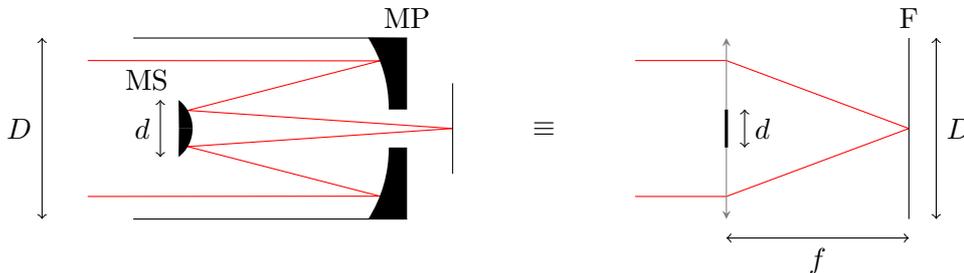
Le premier système correspond au balayage. Ce n'est pas un système physique proprement dit mais il modifie le signal d'entrée en ajoutant une dépendance temporelle. Ensuite se trouve l'optique, constituée principalement des miroirs. L'optique collecte le flux lumineux et focalise les rayons sur le plan focal. On trouve ensuite des filtres en longueurs d'ondes qui sélectionnent les bandes spectrales d'intérêt. L'étage suivant est constitué de cornets qui sont des antennes jouant le rôle d'intégrateurs spatiaux et qui concentrent le rayonnement sur les bolomètres. Les bolomètres sont les détecteurs sensibles à la chaleur, donc au flux incident, et génèrent un signal électrique. Le dernier étage, l'électronique de lecture, produit les données numériques transmises au sol (par modulation, filtrage, échantillonnage, ...).

## 2.2 Modèles physiques

Cette partie présente les modèles et équations de la physique utilisés pour déterminer les réponses des systèmes précédents. Les parties traitant de l'optique étant généralement plus connues sont décrites moins en détail que la physique du bolomètre dont la physique, décrite dans l'annexe A est moins répandue.

### 2.2.1 Ouverture

L'ouverture correspond principalement au miroir du télescope. Cette optique focalise le flux lumineux  $\phi(\alpha, \beta, \lambda)$  sur le plan détecteur et forme une image du ciel. Le miroir de type Ritchey-Chrétien, illustré sur la figure 2.2, est une succession de deux miroirs : le primaire et le secondaire. Par rapport à une lentille de longueur focale équivalente cette conception a l'avantage d'être techniquement réalisable et réduit l'encombrement. Cependant la présence du miroir secondaire diminue la quantité de flux observé et modifie la réponse du système.



**Figure 2.2** – l'ouverture du télescope : Les rayons sont réfléchis par le miroir primaire (MP) puis le miroir secondaire (MS), respectivement de diamètre  $D$  et  $d$ . Les parties grisées de la lentille laissent passer les rayons tandis que la partie noire est opaque au flux. La lentille a une longueur focale équivalente  $f$  et concentre les rayons sur le plan focal F.

En considérant le modèle de la lentille pour le télescope, les grandeurs caractéristiques sont le diamètre de la lentille  $D$ , le diamètre du trou central  $d$ , la longueur focale équivalente  $f$  et

la transmittance  $t(\alpha, \beta)$  de la lentille. Les valeurs de ces paramètres dans le cas d'Herschel sont reportées annexe D page 167. En supposant les dimensions du système grandes devant les longueurs d'onde  $\lambda$  typiques de l'infrarouge ( $D \gg \lambda$ ,  $d \gg \lambda$  et  $f \gg \lambda$ ) et sous l'hypothèse des angles faibles  $\sin \alpha \approx \alpha$ , la sortie  $\phi_m$  du système s'écrit comme la convolution à deux dimensions [Goo72] suivant  $(\alpha, \beta)$  de l'entrée avec la réponse  $h_m$  du système

$$\begin{aligned}\phi_m(\alpha', \beta', \lambda) &= \int_{\alpha} \int_{\beta} \phi(\alpha, \beta, \lambda) h_m(\alpha' - \alpha, \beta' - \beta, \lambda) d\alpha d\beta \\ &= \int_{\alpha} \int_{\beta} \phi(\alpha' - \alpha, \beta' - \beta, \lambda) h_m(\alpha, \beta, \lambda) d\alpha d\beta.\end{aligned}$$

La réponse  $h_m$  de cette optique est décrite par la théorie de la diffraction de Fresnel et s'obtient avec la transformée de Fourier de la transmittance de la lentille [Goo72]. Si on néglige les distorsions, ou d'autres effets optiques provenant des miroirs, la transmittance s'écrit comme une indicatrice à symétrie circulaire

$$t_c(r) = \text{rect}\left(\frac{r}{D/2}\right) - \text{rect}\left(\frac{r}{d/2}\right)$$

avec  $r$  le rayon et  $\text{rect}(r)$  l'indicatrice sur  $[0, 1]$ . La réponse impulsionnelle spatiale du miroir est donc l'image de la transformée de Fourier d'un disque ou d'un anneau. La transformée de Fourier pour des fonctions à symétrie circulaire est la transformée dite de Hankel. Avec la fréquence spatiale  $\rho$  et le nombre d'onde  $k = 2\pi/\lambda$ ,

$$\begin{aligned}h(\rho) &= \frac{k}{f} \int_0^{+\infty} t_c(r') J_0(2\pi\rho r') r' dr' \\ &= \frac{k}{f} \int_0^{D/2} J_0(2\pi\rho r') r' dr' - \frac{k}{f} \int_0^{d/2} J_0(2\pi\rho r') r' dr'.\end{aligned}\quad (2.2)$$

Les deux intégrales sont identiques, seules les bornes changent. L'expression de l'intégrale est connue et s'écrit

$$\frac{k}{f} \int_0^{D/2} J_0(2\pi\rho r) r dr = \frac{k}{f} \left(\frac{D}{2}\right)^2 \frac{J_1(\pi D\rho)}{\pi D\rho}\quad (2.3)$$

où  $J_1$  est la fonction de Bessel de premier ordre  $n = 1$  et première espèce  $J$ . L'expression précédente (2.3) correspond à la réponse du miroir primaire sans trou central. En soustrayant la deuxième intégrale de l'équation (2.2) à (2.3) on obtient

$$h(\rho) = \left(\frac{kD^2}{8f}\right) \left[2\frac{J_1(\pi D\rho)}{\pi D\rho}\right] - \left(\frac{kd^2}{8f}\right) \left[2\frac{J_1(\pi d\rho)}{\pi d\rho}\right].\quad (2.4)$$

L'abscisse du premier zéro se situe en

$$\Delta\theta = \Delta r/f \approx 1,22 \frac{\lambda}{D}.\quad (2.5)$$

L'équation (2.4) correspond à la réponse du miroir. Cependant il s'agit de la réponse à une onde et non simplement à son intensité. Il faudrait dans ce cas utiliser l'expression exacte de l'onde à la sortie du miroir ce qui ferait intervenir des termes de phases [Goo72]. Cependant, comme nous le verrons plus tard (dans la partie 2.2.4 sur les bolomètres), les éléments sensibles du capteur sont sensibles à l'énergie du rayonnement. Généralement dans ce cas on ne considère que la réponse en intensité, c'est-à-dire la réponse en énergie obtenue à l'aide du produit scalaire

$$h_m(\rho) = \langle h, h^* \rangle = \left( \left( \frac{kD^2}{8f} \right) \left[ 2 \frac{J_1(\pi D\rho)}{\pi D\rho} \right] - \left( \frac{kd^2}{8f} \right) \left[ 2 \frac{J_1(\pi d\rho)}{\pi d\rho} \right] \right)^2.\quad (2.6)$$

L'expression en deux dimensions s'obtient en posant la fréquence spatiale  $\rho$  [Goo72]

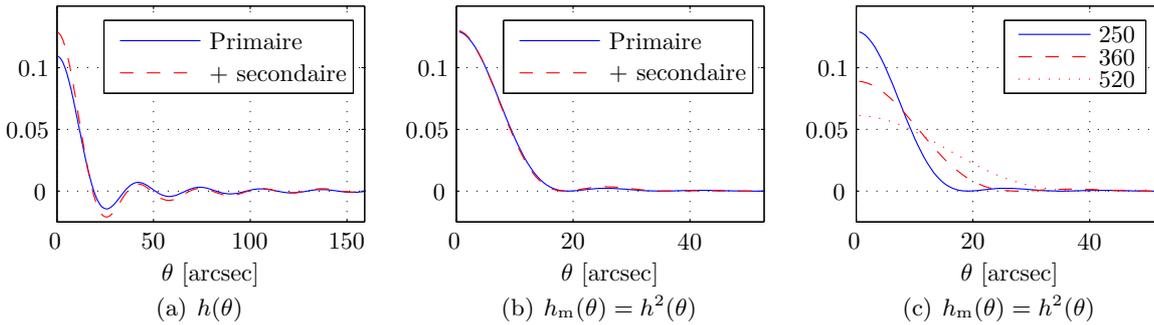
$$\begin{aligned}\rho &= r/\lambda f \\ &= f \sin(\theta)/\lambda f \approx \theta/\lambda \\ &\approx \sqrt{\alpha^2 + \beta^2}/\lambda\end{aligned}$$

avec la formule des lentilles et en supposant que l'angle d'incidence  $\theta = \sqrt{\alpha^2 + \beta^2}$  est petit (le champ de vue est petit). Dans ce cas la réponse spatiale s'écrit

$$h_m(\alpha, \beta, \lambda) = \left( \left( \frac{kD^2}{8f} \right) \frac{J_1 \left( kD\sqrt{\alpha^2 + \beta^2}/2 \right)}{kD\sqrt{\alpha^2 + \beta^2}/2} - \left( \frac{kd^2}{8f} \right) \frac{J_1 \left( kd\sqrt{\alpha^2 + \beta^2}/2 \right)}{kd\sqrt{\alpha^2 + \beta^2}/2} \right)^2. \quad (2.7)$$

Il s'agit de l'expression de la figure d'*Airy*. Le profil de cette réponse, illustrée sur la figure 2.3(c), ressemble beaucoup à une gaussienne mais avec des lobes secondaires.

Ainsi le miroir est modélisé comme un système linéaire invariant dont la réponse impulsionnelle dépend de la longueur d'onde par  $k = 2\pi/\lambda$ . La réponse devient plus ou moins large comme illustré sur la figure 2.3(c) et l'équation (2.5). La réponse peut également se modéliser sans le trou central. Dans ce cas le diamètre de ce dernier est  $d = 0$  et la différence est illustrée sur la figure 2.3(b).



**Figure 2.3** – Illustration de la réponse optique. La figure 2.3(a) représente la réponse de l'optique face à une onde lorsque l'on prend en compte la présence du miroir secondaire ou non. La figure d'*Airy* est représentée figure 2.3(b). La dépendance de la réponse en fonction de la longueur d'onde est illustrée figure 2.3(c) avec les trois longueurs d'onde centrales de SPIRE.

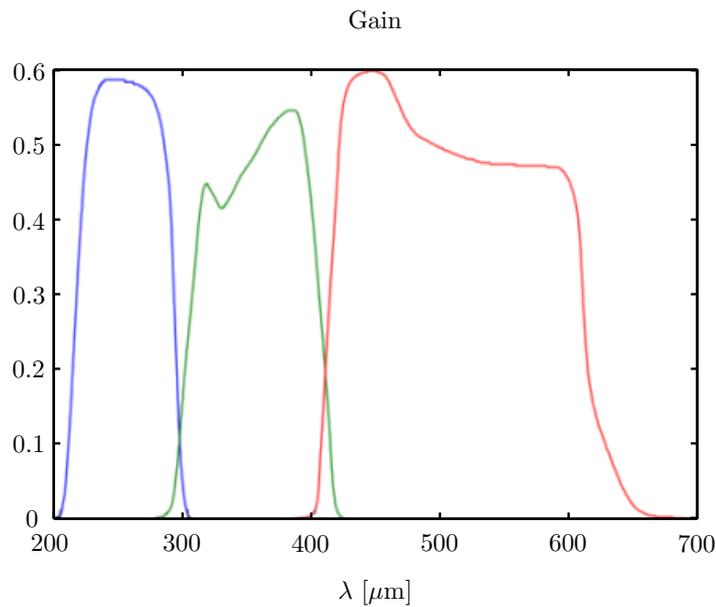
## 2.2.2 Filtres

Des filtres sont placés sur le trajet optique pour sélectionner les longueurs d'ondes [ESA07]. Il y en a trois, un pour chaque matrice de bolomètres et ils ont tous les trois un profil passe-bande centré sur les longueurs d'ondes 250, 360 et 520  $\mu\text{m}$ . Actuellement il n'y a pas d'expression analytique pour modéliser le profil des filtres. Cependant des mesures, illustrées par la figure 2.4, sont disponibles.

Leur action est modélisée par une fonction  $h_f(\lambda)$  où  $f = 1, 2$  ou  $3$  est l'indice de la bande. C'est une simple fonction qui décrit la transmission des filtres en fonction de la longueur d'onde. La sortie  $\phi_f$  du système

$$\phi_f(\alpha, \beta, \lambda) = h_f(\lambda)\phi_m(\alpha, \beta, \lambda) \quad (2.8)$$

est obtenue par pondération de l'objet d'entrée.



**Figure 2.4** – Gain des trois filtres en longueurs d’onde obtenu par mesure en laboratoire.

### 2.2.3 Cornets

Les cornets sont des antennes en forme de cône à ouverture circulaire placées sur le plan focal, leur rôle étant de concentrer le flux incident. Derrière chaque cornet est placé un bolomètre, les éléments sensibles de SPIRE. Ces bolomètres absorbent la très grande majorité du flux provenant des cornets. Les cornets ont donc le rôle d’intégrateurs spatiaux, comme la surface d’éléments CCD. La sortie n’est cependant pas un signal électrique mais reste un flux lumineux.

Utiliser des cornets à la place des bolomètres, beaucoup plus petits, placés directement sur le plan focal, augmente la sensibilité globale de l’instrument puisque la surface de collecte du flux est plus grande [GBG02]. En contre partie, puisque la surface de collecte est plus importante, on perd en résolution spatiale.



**Figure 2.5** – Une des matrices de cornets à gauche et la représentation schématique d’une coupe transversale. La courbe en cloche rouge représente la réponse impulsionnelle des cornets  $h_c$ .

Les cornets ont une réponse qui n’est pas uniforme sur leur surface mais gaussienne. De plus, tout comme le miroir, cette réponse obéit aux lois de la diffraction. La largeur de la gaussienne

varie donc linéairement avec la longueur d'onde et le diamètre de l'ouverture [Fer07], au moins au premier ordre et dans une bande spectrale donnée. Plus précisément, la largeur augmente avec la longueur d'onde mais diminue avec le diamètre. Donc une augmentation du diamètre diminue la fenêtre d'intégration. Cependant l'augmentation du diamètre augmente la surface de collecte. Il y a donc un compromis sur le diamètre à trouver pour optimiser la quantité de flux collecté et donc la sensibilité de l'instrument [GBG02]. Ce compromis dépend de la longueur d'onde puisque la largeur de la réponse augmente avec  $\lambda$ . Donc les cornets ont des diamètres différents en fonction de la longueur d'onde observée (suivant  $f = 1, 2$  ou  $3$ ). Les travaux de [GBG02] établissent que la plus grande sensibilité est atteinte lorsque le diamètre du cornet vaut  $2\lambda/D$  où  $D$  est le diamètre du miroir primaire. Autrement dit, un cornet est deux fois plus large que la largeur à mi-hauteur d'une figure d'*Airy*.

Concrètement, en se basant sur la théorie de la diffraction [Goo72, Fer07], le système est considéré comme linéaire invariant. La sortie s'obtient donc par convolution entre l'entrée  $\phi_f$  et la réponse impulsionnelle

$$\phi_c(\alpha', \beta', \lambda) = \int_{\alpha} \int_{\beta} \phi_f(\alpha, \beta, \lambda) h_c(\alpha - \alpha', \beta - \beta', \lambda) d\alpha d\beta$$

où la réponse  $h_c$  d'un cornet dans une bande spectrale  $f$  s'écrit

$$h_c(\alpha, \beta, \lambda) = \frac{1}{2\pi\sigma_f^2} \exp\left[-\frac{1}{2} \frac{\alpha^2 + \beta^2}{\sigma_c^2}\right]. \quad (2.9)$$

La dépendance en longueur d'onde s'exprime à travers l'écart-type  $\sigma_c(\lambda)$  dont les valeurs typiques pour SPIRE sont présentées annexe D.

Placer un ensemble de cornets côte à côte au niveau du plan focal permet de faire une matrice [CGB<sup>+</sup>03] illustrée sur la photo figure 2.5. On observe ainsi toute une étendue spatiale correspondant à une région du ciel. Sur la photo on peut voir que, étant donné la forme circulaire des cornets, pour minimiser la surface non observée ceux-ci sont disposés de manière hexagonale.

Une matrice différente est placée à la suite des trois filtres en longueur d'onde. Ces trois matrices observent exactement la même portion de ciel de 8 minutes d'arc sur 4 (le même flux est en fait divisé en trois au moyen d'éléments optiques). Une représentation schématique de ces trois matrices placées côte à côte est montrée sur la figure 2.6.

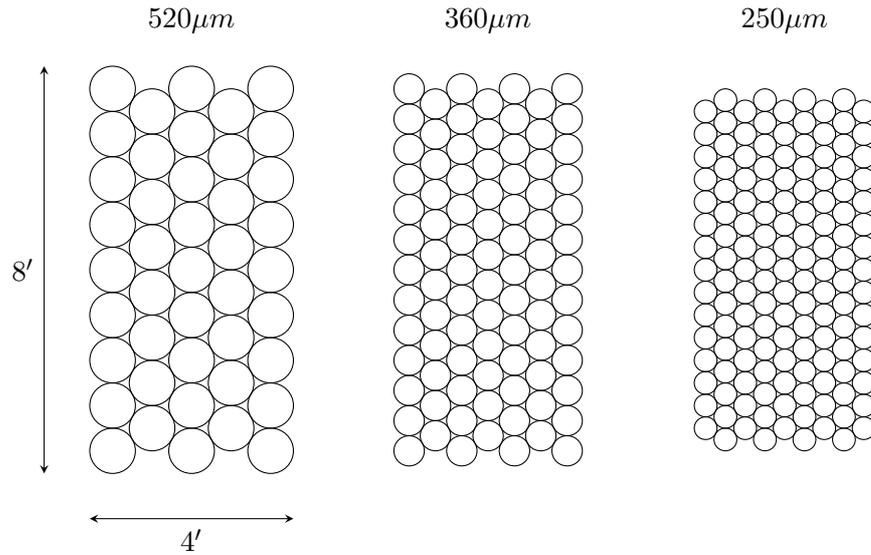
## 2.2.4 Bolomètres

Les bolomètres sont des détecteurs thermiques qui convertissent un rayonnement incident en chaleur puis en tension électrique. Certaines caractéristiques du bolomètre changent avec la température : on accède ainsi à une mesure du flux incident. La figure 2.8, prise au microscope, présente un bolomètre de SPIRE.

Pour établir le modèle des bolomètres nous utilisons le modèle thermique Sudiwala *et al.* [SGW02]. Ce modèle est lui-même une amélioration des modèles thermiques classiquement utilisés pour modéliser les bolomètres [Ban05, Cam02, Tor92]. Les détails de calcul sont dans l'annexe A. Nous présentons ici uniquement les principes généraux et le modèle obtenu.

### Principes et conception

Les bolomètres, dont un schéma est présenté sur la figure 2.7, sont constitués de deux parties aux rôles distincts.



**Figure 2.6** – Position spatiale des trois matrices de cornets. Comme la même portion de ciel est observée, les trois matrices peuvent être vues comme superposées.

1. Une première partie a pour rôle d'absorber une puissance incidente  $P$ . La puissance absorbée modifie sa température  $T$ . L'évolution de la température du matériau est caractérisée par une capacité calorifique  $C$  en  $[\text{JK}^{-1}]$ .

Une fois cette puissance absorbée, un équilibre thermique s'installe par conduction à travers le lien thermique dans l'environnement. Ainsi pour permettre au bolomètre d'être sensible à de très faibles énergies, celui-ci est placé dans un environnement maintenu à très basse température de l'ordre de  $T_0 = 0,3 \text{ K}$  par un bain thermique.

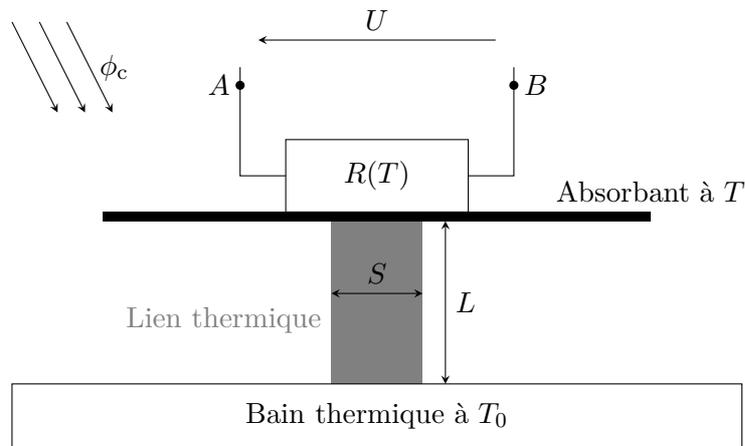
2. La deuxième partie consiste à lire la température du bolomètre à l'aide d'une thermistance. Il s'agit d'un composant dont la résistivité  $R(T)$  dépend de la température. C'est ainsi que l'on a accès à  $T$ .

Pour mesurer la valeur de la résistance on utilise une polarisation. Elle s'effectue avec un courant  $I_p$  qui génère une différence de tension  $U$  aux bornes. C'est précisément cette tension qui est mesurée à l'aide d'un étage électronique en aval.

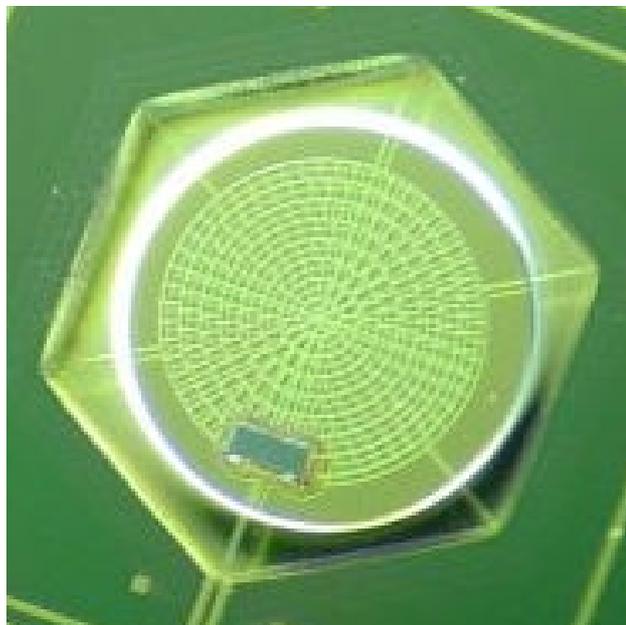
Sur la figure 2.8 on peut observer quelques détails. Tout d'abord l'absorbant est en forme de « toile d'araignée ». Il y a plusieurs intérêts technologiques à cela :

- la capacité thermique  $C$  du bolomètre est diminuée et par conséquent la vitesse de réaction du bolomètre dans le temps augmente ;
- la sensibilité du bolomètre aux particules cosmiques est réduite puisqu'il y a moins de surface efficace. Les particules passent à travers ;
- Le rayonnement aux longueurs d'onde inférieures à une longueur caractéristique du bolomètre et de cette structure, qui proviennent de l'environnement et non du flux incident, n'est pas absorbé. Le bolomètre est ainsi moins sensible aux signaux parasites.

On peut également voir sur cette photo la résistance, petit rectangle gris en bas du disque, avec les deux fils de lecture correspondant aux points  $A$  et  $B$  de la figure 2.7.



**Figure 2.7** – Schéma d'un bolomètre. Le rayonnement incident  $\phi_c$  est convertit en chaleur par l'absorbant (en noir) dont la température  $T$  se modifie. La chaleur est évacuée à travers le lien thermique (en gris) dans le bain thermique. La lecture de la température se fait avec la résistance dont la valeur change en fonction de  $T$ .



**Figure 2.8** – Bolomètre araignée de SPIRE. Le flux provenant du cornet est concentré sur ce disque. On peut voir le rectangle gris dans le bas du disque correspondant à la thermistance ainsi que les deux fils de lecture qui partent vers le bas.

## Réponse du bolomètre

Les bolomètres de SPIRE ont une structure en toile d'araignée. Seules les longueurs d'ondes plus grandes qu'une longueur caractéristique sont absorbées. Cependant, les bolomètres ont été dimensionnés pour absorber tout le flux passé à travers les filtres en longueurs d'onde des étages précédents sans variation de réponse. De plus le bolomètre a une certaine extension spatiale, cependant il n'existe pas actuellement de modèle pour décrire le fonctionnement spatial du bolomètre. On considère donc que la réponse spatiale du bolomètre est homogène, c'est-à-dire qu'elle ne change pas en fonction de la position. De ces deux remarques on déduit que le bolomètre intègre la puissance lumineuse incidente sur toutes les longueurs d'ondes et toute la surface  $\mathcal{S}$  soit

$$P(t) = \int_{\lambda} \int_{\mathcal{S}} \phi_c(\alpha, \beta, \lambda, t) d\alpha d\beta d\lambda. \quad (2.10)$$

On remarque que les dimensions spatiales et spectrales ont disparu. Il reste ensuite à établir la réponse du bolomètre dans le temps, établie dans l'annexe A. La réponse  $T(t)$  à l'entrée  $P(t)$  du bolomètre s'obtient avec un bilan de puissance qui fournit l'équation différentielle

$$C \frac{dT}{dt} - \frac{R(T)V_p^2}{R_c^2} + \frac{G_0}{T_0^\beta(\beta+1)} (T^{\beta+1} - T_0^{\beta+1}) = P(t). \quad (2.11)$$

La littérature ne contient pas pour le moment de résolution explicite de cette équation. La démarche adoptée ici, qui est celle de la littérature [SGW02, Tor92, Cam02], consiste généralement à linéariser cette équation à l'aide d'un développement au premier ordre

$$C \frac{dT}{dt} + \phi(\bar{T}) + \left. \frac{d\phi}{dT} \right|_{\bar{T}} \tilde{T} = \bar{P} + \tilde{P} \quad (2.12)$$

en considérant de petites variations  $\tilde{T}$  autour d'un point de fonctionnement constant  $\bar{T}$  et où

$$\phi(T) = \frac{G_0}{T_0^\beta(\beta+1)} (T^{\beta+1} - T_0^{\beta+1}) - I_p^2 R_g \exp \left[ \left( \frac{T_g}{T} \right)^n \right].$$

Le lecteur qui souhaiterait les détails des calculs les trouvera dans l'annexe A. L'important est que ce développement au premier ordre modélise le bolomètre comme un système linéaire invariant avec comme réponse temporelle un filtre linéaire passe-bas du premier ordre dont la réponse impulsionnelle est

$$h_b(t) = S' \exp \left[ -\frac{t}{\tau_e} \right] \quad (2.13)$$

où le gain et la constante de temps dépendent des paramètres physiques

$$\left\{ \begin{array}{l} S' = \nu \bar{R} I_p / C \\ \tau_e = C / (G_d - \nu \bar{R} I_p^2) \\ \bar{R} = R(\bar{T}) \\ \nu = n T_g^n / \bar{T}^{n+1} \\ \bar{T} = \phi^{-1}(\bar{P}) \\ G_d = G_0 \bar{T}^\beta T_0^\beta \\ \bar{P} = \int \bar{\phi}_c d\alpha d\beta d\lambda. \end{array} \right. \quad (2.14)$$

En pratique seulement le gain  $S'$  et la constante de temps  $\tau_e$  sont utilisés. Finalement la sortie en tension en fonction du flux provenant des cornets s'écrit

$$\begin{aligned} U(t) &= \bar{R}I_p + \int_t \tilde{P}(t)h_b(t' - t)dt \\ &= \bar{R}I_p + \int_t \int_{(\alpha,\beta) \in \mathcal{S}} \int_{\lambda \in \mathbb{R}_+} \tilde{\phi}_c(\alpha,\beta,\lambda,t)h_b(t' - t)dt d\alpha d\beta d\lambda \end{aligned} \quad (2.15)$$

Toutes les constantes sont définies à partir du point de fonctionnement déterminé par  $\bar{P}$ . Le domaine de validité de ce modèle dépend de  $T$ . Physiquement il paraît raisonnable de supposer que de petites variations de  $T$  soient dues à de petites variations de signal incident  $P$ . Si les variations sont importantes, alors l'approximation d'un comportement modélisé au premier ordre est trop sévère. Une solution pourrait être de modifier la position du point de fonctionnement au cours du temps (et le considérer localement constant) ou encore de considérer une réponse qui ne soit plus invariante. Il est également possible d'envisager une méthode de résolution numérique du bilan de puissance.

Dans le cas de SPIRE il est admis que la majorité des observations seront effectuées dans un régime linéaire [Gri06]. On considère donc que le développement au premier ordre suffit pour modéliser correctement le comportement du bolomètre. De plus les paramètres du modèle sont modifiables et peuvent être adaptés (voire estimés) en fonction de ce qui est observé.

### 2.2.5 Électronique de lecture

L'électronique de lecture [Gri07] permet de mesurer et de numériser la valeur de la tension aux bornes de la thermistance. Elle est décomposée en plusieurs étages.

**Polarisation.** Le générateur de tension produit une excitation sinusoïdale. Cette modulation est présente à cause de l'étage d'amplification. Cette amplification est faite à l'aide de transistors et introduit un signal basse-fréquence qui vient dégrader le signal produit par les bolomètres [Gri06]. La modulation d'amplitude permet de décaler le spectre utile dans une bande de fréquence plus élevée que la bande occupée par ce signal parasite.

**Amplification.** Le signal  $U(t)$  aux bornes de la résistance est amplifié par un étage à transistors. L'objectif est d'avoir la puissance nécessaire à la circulation du courant sur la longueur de câble qui suit l'adaptation.

**Filtre passe-bande.** A la suite de l'amplification un filtre passe-bande sélectionne le signal utile, c'est-à-dire le spectre modulé avec le résidu filtré du signal basse-fréquence.

**Démodulateur.** Pour démoduler, le signal est multiplié par un signal carré en phase avec la porteuse du signal à l'entrée. Il s'agit d'une démodulation classique où le spectre du signal utile est à nouveau reproduit sur des bandes différentes. Une des reproductions, qui est celle qui nous intéresse, se trouve centrée sur la fréquence nulle.

**Filtre passe-bas.** Le filtre passe-bas qui suit permet de récupérer le signal utile. Il est mis en œuvre à l'aide d'un filtre de Bessel d'ordre cinq. La fréquence de coupure du filtre est de l'ordre de 5 Hz [Gri07]. On obtient ainsi un signal correspondant au signal utile avant la modulation provoquée par la polarisation. L'ensemble de ces étapes permet de minimiser la corruption du signal utile par le signal basse-fréquence introduit par l'étage d'adaptation d'impédance.

**Numérisation.** À la suite de cette chaîne de traitement, présente pour chaque bolomètre, on trouve un multiplexeur unique. Ensuite un convertisseur analogique numérique (CAN) fournit un signal échantillonné et quantifié qui peut être ainsi stocké, traité en vol puis transmis au sol. Selon le simulateur officiel de Bruce Sibthorpe (aucune autre donnée fiable n'a pu être

malheureusement trouvée pour le moment), la fréquence d'échantillonnage est de l'ordre de 15 Hz.

De tous ces étages ceux qui ont le plus d'influence sur le signal utile sont l'étage d'amplification qui introduit un signal parasite et le filtre passe-bas à 5 Hz dont la fréquence de coupure est sensiblement plus importante que la fréquence de coupure du bolomètre. Leur effet semble cependant négligeable pour le moment par rapport aux autres éléments comme l'optique. L'augmentation de la complexité du modèle n'encourage pas à prendre en compte ces phénomènes dans le modèle. Cependant les équations sont disponibles et il s'agit de systèmes linéaires invariants donc leur intégration future est possible.

Finalement, il ne reste que l'échantillonnage régulier dans le temps qui est nécessaire. Il peut être modélisé par la multiplication du signal utile avec un peigne de Dirac soit

$$C(nT_e) = \int_t U(t)\delta(t - nT_e) dt \quad (2.16)$$

où  $T_e$  est la période d'échantillonnage. On obtient avec cette modélisation, une reproduction des données  $C(i,j,k)$  en fonction du signal présent à l'entrée  $\phi$  et d'un modèle physique de l'instrument. Ce modèle prend en compte, la réponse de l'optique, l'intégration spatiale et la réponse dans le temps des bolomètres.

### 2.2.6 Mode d'observation

La position spatiale pointée par le télescope n'est pas fixe dans le temps. Son pointage évolue suivant plusieurs protocoles possibles [ESA07], appelés modes d'observation. Dans tous les cas, les seuls déplacements possibles du télescope, au cours d'une séquence d'observation, sont des translations dans le repère du ciel. En effet, à la fois pour se protéger du rayonnement solaire, mais également à cause du coût important en énergie pour faire tourner le télescope, les rotations ne sont pas permises. En conclusion, la position pointée par le télescope s'écrit sous la forme

$$\begin{array}{ll} p_\alpha : \mathbb{R} \rightarrow \mathbb{R} & p_\beta : \mathbb{R} \rightarrow \mathbb{R} \\ t \mapsto p_\alpha(t) & t \mapsto p_\beta(t) \end{array}$$

où  $t$  est le temps en seconde et  $p_{\alpha,\beta}$  est l'angle pointé dans la dimension  $\alpha$  ou  $\beta$  à l'instant  $t$ , en radians.

Pour modéliser les translations du télescope, on se place dans le repère défini par l'instrument et on considère que le signal présent à l'entrée dépend du temps. Ce signal dépend du ciel et du pointage et s'écrit

$$\phi_t(\alpha,\beta,\lambda,t) = \phi(\alpha - p_\alpha(t),\beta - p_\beta(t),\lambda). \quad (2.17)$$

Vu du télescope donc, le signal observé s'obtient par un décalage par rapport à l'origine correspondant à la position courante.

Jusqu'à maintenant nous n'avons pas précisé la structure des fonctions  $p_\alpha$  et  $p_\beta$ . Or le télescope suit des trajectoires déterminées [ESA07]. Pour SPIRE il existe trois modes d'observation.

**Point source et Small map.** Le mode *Point source* est conçu pour observer des sources ponctuelles comme des étoiles. Le télescope pointe sept fois autour de la source avec des décalages inférieurs à la distance entre deux centres de cornets (décalage sub-pixelique).

Le mode *Small map* permet de faire des cartes de  $4 \times 4$  minutes d'arc (la moitié de la taille d'une matrice de bolomètre). En moyenne le télescope est centré sur un point. Le pointage se

décale 64 fois. Les deux modes précédents se modélisent comme

$$p_\alpha(t) = c_\alpha^1 \mathbb{1}_{[t_1 \ t_2[}(t) + c_\alpha^2 \mathbb{1}_{[t_2 \ t_3[}(t) + \dots = \sum_{q=1}^Q c_\alpha^q \mathbb{1}_{[t_q \ t_{q+1[}(t) \quad (2.18)$$

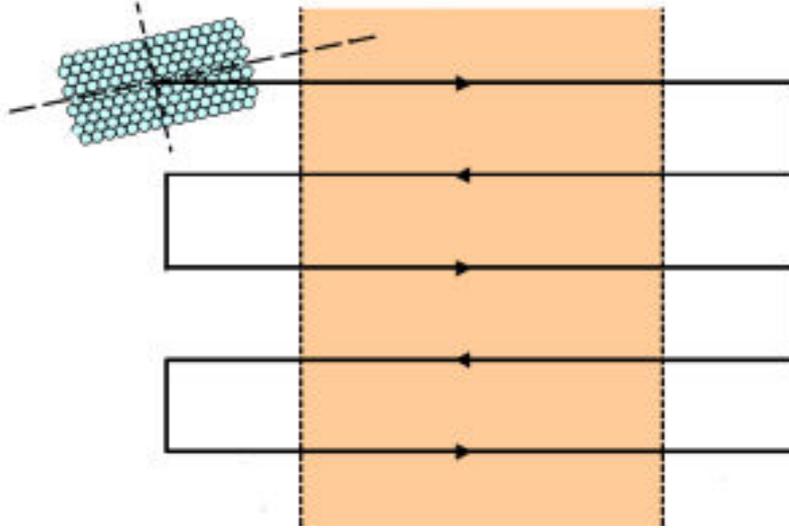
et

$$p_\beta(t) = \sum_{q=1}^Q c_\beta^q \mathbb{1}_{[t_q \ t_{q+1[}(t) \quad (2.19)$$

où les  $(c_\alpha^q, c_\beta^q)$  sont les positions pointées successivement au cours du temps, indexées par  $q$ , pendant les périodes  $\mathbb{1}_{[t_q \ t_{q+1[}(t)$ .

Ces deux modes permettent d'introduire de la redondance spatiale dans les données. Les décalages sub-pixeliques permettent de lever par un traitement adapté la résolution limitée du capteur.

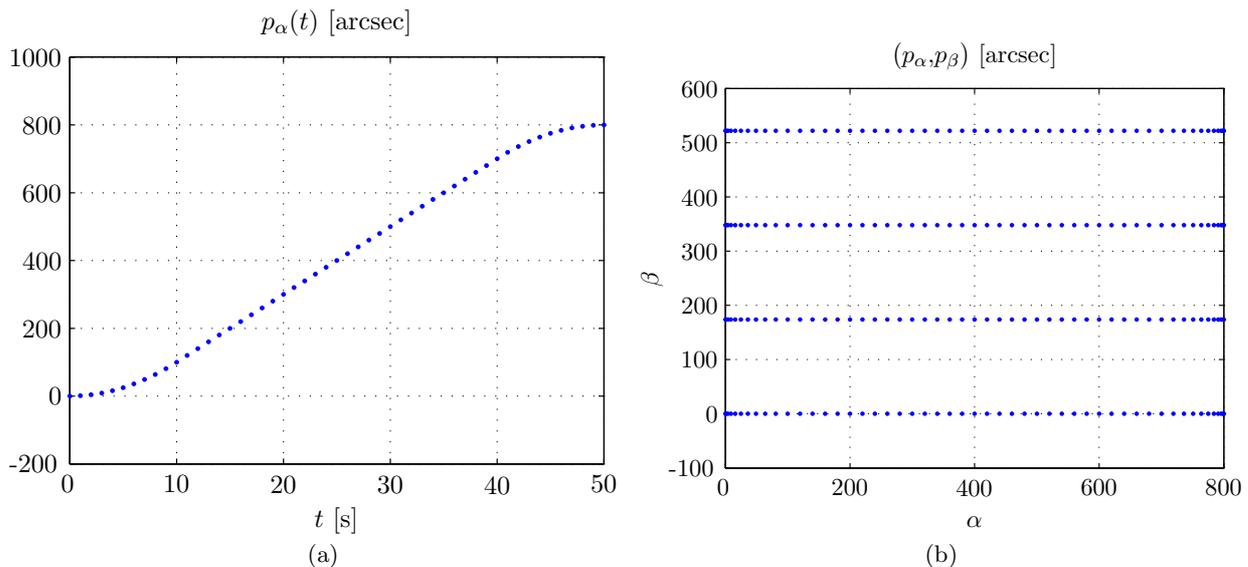
**Large map.** Ce mode permet un pointage qui balaye le ciel sur des distances définies par l'utilisateur. Il est prévu pour permettre la construction de cartes à grand champ du ciel. Pour cela le télescope balaye le ciel successivement dans certaines directions comme illustré sur la figure 2.9. Chaque balayage se fait dans une direction fixe. L'écart entre deux balayages est fixé par le protocole et est de 348 secondes d'arc. Les signaux à la sortie des bolomètres sont mesurés tout au long de la séquence.



**Figure 2.9** – Illustration du balayage *Large Map*. L'angle entre la matrice et la direction du balayage est fixe ainsi que l'écart entre deux balayages. La zone d'intérêt est la zone centrale colorée.

Pour que deux bolomètres ne pointent pas au même endroit, la direction de balayage fait un angle précis de 42,4 degrés avec le capteur. Pendant chaque balayage, des phases d'accélération, de balayage à vitesse constante et de décélération se succèdent comme illustrées sur la figure 2.10. La phase à vitesse constante est la principale et en pratique l'observateur a le choix entre deux vitesses : 30" ou 60"/s. Les phases d'accélération et de décélération sont placées en dehors de la zone d'intérêt, le contrôle du télescope étant jugé peu sûr pendant ces périodes.

Une séquence complète d'observation est constituée de balayage dans un sens et dans l'autre. Puis le télescope peut balayer dans une direction quasi perpendiculaire de 88 degrés. Pendant



**Figure 2.10** – Évolution du pointage dans le balayage *Large Map* selon les équations (2.21). La figure 2.10(a) illustre la phase d’accélération et de décélération en  $\alpha$ . La figure 2.10(b) montre la même chose en  $(\alpha, \beta)$  avec un décalage dans la direction perpendiculaire entre chaque balayage.

toute une séquence d’observation il n’y a pas de rotation. Les fonctions de pointage pour un balayage  $q$  s’écrivent

$$\begin{aligned} p_\alpha(t) &= a_\alpha^q t^2 + v_\alpha^q t + c_\alpha^q \\ p_\beta(t) &= a_\beta^q t^2 + v_\beta^q t + c_\beta^q \end{aligned} \quad (2.20)$$

durant les phases d’accélération ou de décélération où  $t \in [t_1, t_2]$ , et

$$\begin{aligned} p_\alpha(t) &= v_\alpha^q t + c_\alpha^q \\ p_\beta(t) &= v_\beta^q t + c_\beta^q \end{aligned} \quad (2.21)$$

pendant les phases de balayage à vitesse constante où  $t \in [t_2, t_3]$ . La précision du pointage, en relatif et absolue, est de l’ordre de quelques secondes d’arc.

### 2.2.7 Équation complète du modèle

Dans cette partie nous mettons bout à bout tous les éléments décrits précédemment pour fournir l’équation qui modélise la chaîne d’acquisition. Nous ne considérons ici que la partie variable du flux. La partie constante qui définit le point de fonctionnement, c’est-à-dire un scalaire, ne rentre pas dans les écritures. La description de cette partie définit en fait un jeu d’équations parallèles décrites dans l’annexe A. Par la suite, la valeur de ce point de fonctionnement est ignorée sauf à travers les paramètres de la réponse des bolomètres.

**Le pointage.** Le pointage fournit un signal observé

$$\phi_t(\alpha, \beta, \lambda, t) = \phi(\alpha - p_\alpha(t), \beta - p_\beta(t), \lambda).$$

**L’optique.** Le miroir convole le signal observé

$$\phi_m(\alpha - p_\alpha(t), \beta - p_\beta(t), \lambda) = \iint_{\mathbb{R}^2} \phi(\alpha_t - p_\alpha(t), \beta_t - p_\beta(t), \lambda) h_m(\alpha_t - \alpha, \beta_t - \beta, \lambda) d\alpha_t d\beta_t$$

**Les filtres.** Les filtres en longueur d'onde pondèrent le flux provenant du miroir

$$\phi_f(\alpha - p_\alpha(t), \beta - p_\beta(t), \lambda) = h_f(\lambda)\phi_m(\alpha - p_\alpha(t), \beta - p_\beta(t), \lambda).$$

**Les cornets.** La taille relative des bolomètres par rapport à celles des cornets est bien plus petite. On considère donc que l'intégration spatiale s'effectue au niveau des cornets est que les bolomètres n'ont par conséquent pas d'étendue spatiale, soit

$$\phi_c^{lm}(t, \lambda) = \iint_{\mathbb{R}^2} \phi_f(\alpha - p_\alpha(t), \beta - p_\beta(t), \lambda) h_c(\alpha - \alpha_{lm}, \beta - \beta_{lm}, \lambda) d\alpha d\beta$$

où  $(\alpha_{lm}, \beta_{lm})$  représente la position spatiale du centre du cornet identifié par les indices  $l$  et  $m$ .

**Les bolomètres.** Enfin les bolomètres intègrent le flux sur les longueurs d'onde et réagissent dans le temps en convoluant le signal d'entrée avec une réponse impulsionnelle

$$y^{lm}(nT_e) = \int_t \int_\lambda \phi_c^{lm}(t, \lambda) h_b(t - nT_e) d\lambda dt$$

où  $nT_e$  représente l'instant d'échantillonnage par l'électronique.

Donc finalement pour chacun des trois capteurs à 250, 360 ou 520 $\mu\text{m}$ , le signal temporel échantillonné du bolomètre  $(l, m)$  s'écrit

$$y^{lm}(nT_e) = \int_t \int_\lambda h_f(\lambda) \iint_{\mathbb{R}^2} \iint_{\mathbb{R}^2} \phi(\alpha_t - p_\alpha(t), \beta_t - p_\beta(t), \lambda) h_m(\alpha_t - \alpha, \beta_t - \beta, \lambda) d\alpha_t d\beta_t h_c(\alpha - \alpha_{lm}, \beta - \beta_{lm}, \lambda) d\alpha d\beta h_b(t - nT_e) d\lambda dt. \quad (2.22)$$

Cette équation fait intervenir six intégrales : quatre provenant d'une convolution et d'une intégration spatiales, une correspondant à l'intégration spectrale et une convolution temporelle. Il s'agit de l'équation fondamentale du modèle direct puisqu'elle modélise l'acquisition de tous les bolomètres des trois capteurs au cours du temps. Nous verrons dans la partie 2.4 comment faire l'économie de calculs numériques lourds avec des calculs analytiques explicites.

On peut remarquer tout d'abord que toutes les opérations sont linéaires. Donc le modèle direct est linéaire malgré sa complexité. Ensuite certains éléments sont invariants, comme les réponses spatiales  $h_m$  et  $h_c$  et temporelle  $h_b$ . Par contre le comportement de l'instrument en longueur d'onde n'est plus invariant puisque les réponses dépendent de  $\lambda$ .

La partie suivante est dédiée à la description du modèle direct dans l'espace de Fourier. Cela permettra de faire ressortir certaines propriétés intéressantes du modèle. On montrera également pourquoi les données acquises par SPIRE permettent de faire une estimation du ciel à une résolution plus importante que la résolution naturelle du capteur (indépendamment des informations disponibles par ailleurs).

## 2.3 Analyse fréquentielle et possibilités de sur-résolution

Cette partie étudie le modèle d'acquisition équation (2.22) dans l'espace de Fourier suivant les variables spatiales et temporelle. On se place pour ce travail à longueur d'onde fixée, elle ne sera donc pas notée dans la suite de cette partie.

Le modèle fait intervenir des convolutions et des intégrations, spatiales et temporelles, deux opérations qui peuvent s'écrire dans l'espace de Fourier. La particularité de ce modèle se situe

au niveau du temps, introduit par les modes d'observation mais qui ne définit pas le ciel. Le fait d'introduire une dimension supplémentaire permettant la redondance, fait apparaître des caractéristiques intéressantes dans les fréquences spatiales. Une interprétation de ces caractéristiques pour le cas *Large map* montre que les données sont moins repliées grâce à la redondance ce qui constitue un point de vue nouveau sur la sur-résolution [ORG09].

### 2.3.1 Ciel et pointage

Le ciel est défini comme une fonction  $\phi \in L_2$

$$\begin{aligned} \phi : \mathbb{R}^2 &\rightarrow \mathbb{R}^+ \\ (\alpha, \beta) &\mapsto \phi(\alpha, \beta) \end{aligned}$$

dont la transformée de Fourier spatiale par rapport aux variables  $(\alpha, \beta)$  est

$$\mathring{\phi}(f_\alpha, f_\beta) = \iint_{\mathbb{R}^2} \phi(\alpha, \beta) e^{-2j\pi(\alpha f_\alpha + \beta f_\beta)} d\alpha d\beta.$$

La transformée de Fourier par rapport aux variables d'espaces  $(\alpha, \beta)$  du signal observé par le télescope (2.17), est

$$\begin{aligned} \mathring{\phi}_t(f_\alpha, f_\beta, t) &= \iint_{\mathbb{R}^2} \phi(\alpha - p_\alpha(t), \beta - p_\beta(t)) e^{-2j\pi(\alpha f_\alpha + \beta f_\beta)} d\alpha d\beta \\ &= \mathring{\phi}(f_\alpha, f_\beta) e^{-2j\pi(f_\alpha p_\alpha(t) + f_\beta p_\beta(t))}, \end{aligned} \quad (2.23)$$

c'est-à-dire la transformée de Fourier spatiale du ciel déphasée, avec un déphasage qui dépend du temps. Il est intéressant de remarquer que quel que soit le pointage, l'information temporelle ne se situe que dans le déphasage. C'est cette information qui va structurer le contenu temporel des données, le ciel étant défini uniquement par des variables spatiales.

### 2.3.2 Convolution et échantillonnage spatial

On sait que le premier système est l'optique formée par le miroir et les cornets. Ces deux systèmes se modélisent par une convolution pour le miroir et une intégration spatiale par les cornets. Cela correspond à une convolution totale et un échantillonnage. Dans ce cas, comme le nombre d'échantillons est limité, on introduira une fenêtre d'observation.

Le signal à la sortie du télescope  $\phi_m$  s'obtient à l'aide d'une convolution spatiale entre  $\phi_t$  et la réponse  $h_m$

$$\phi_m(\alpha, \beta, t) = \iint_{\mathbb{R}^2} \phi(\alpha - p_\alpha(t), \beta - p_\beta(t)) h_m(\alpha_t - \alpha, \beta_t - \beta) d\alpha_t d\beta_t.$$

La transformée de Fourier du signal à la sortie du télescope s'écrit

$$\begin{aligned} \mathring{\phi}_m(f_\alpha, f_\beta, t) &= \iint_{\mathbb{R}^2} \iint_{\mathbb{R}^2} \phi_t(\alpha_t, \beta_t, t) h_m(\alpha_t - \alpha, \beta_t - \beta') d\alpha_t d\beta_t e^{-2j\pi(f_\alpha \alpha + f_\beta \beta)} d\alpha d\beta \\ &= \mathring{\phi}(f_\alpha, f_\beta) \mathring{H}_m(f_\alpha, f_\beta) e^{-2j\pi(f_\alpha p_\alpha(t) + f_\beta p_\beta(t))} \end{aligned}$$

qui correspond au spectre du ciel, filtré dans l'espace par la fonction de transfert  $\mathring{H}_m$  du miroir et déphasé par le pointage. En appliquant exactement le même raisonnement à la réponse du cornet on obtient

$$\mathring{\phi}_c(f_\alpha, f_\beta, t) = \mathring{\phi}(f_\alpha, f_\beta) \mathring{H}_m(f_\alpha, f_\beta) \mathring{H}_c(f_\alpha, f_\beta) e^{-2j\pi(f_\alpha p_\alpha(t) + f_\beta p_\beta(t))}. \quad (2.24)$$

L'expression est sensiblement identique sauf que la fonction de transfert totale correspond au produit de celle du miroir  $\mathring{H}_m$  et du cornet  $\mathring{H}_c$ .

Les cornets sont des intégrateurs spatiaux. Ils se modélisent par une convolution suivie d'un échantillonnage aux positions  $(\alpha_{lm}, \beta_{lm})$ . Cet échantillonnage s'écrit comme la multiplication de la sortie de la convolution avec un peigne de Dirac.

La grille d'échantillonnage est une grille hexagonale, décrite dans l'annexe B.1 page 151, et s'écrit formellement comme une somme de Dirac

$$\Delta_{T_{\alpha\beta}}(\alpha, \beta) = \sum_{l,m} \delta(\alpha - \alpha_{lm}, \beta - \beta_{lm}).$$

Les points d'échantillonnage  $(\alpha_{lm}, \beta_{lm})$  placés sur une grille hexagonale peuvent également se voir comme des points de deux grilles rectangulaires en quinconce (annexe B.1). Il se trouve donc que la transformée de Fourier d'un peigne de Dirac hexagonal est également un peigne de Dirac hexagonal de période inverse et avec une rotation de 90 degrés. Donc l'expression du signal échantillonné dans l'espace de Fourier est

$$\sum_{l,m} \mathring{\phi}_c(f_\alpha - f_\alpha^{lm}, f_\beta - f_\beta^{lm}, t)$$

et correspond tout simplement à la duplication du spectre obtenu à la sortie du cornet. Chaque duplication du spectre  $\mathring{\phi}_c$  est centrée en  $(f_\alpha^{lm}, f_\beta^{lm})$ , correspondant à des points du plan fréquence placés de façon hexagonale (voir annexe B.1).

Le capteur possède un nombre limité de bolomètres. Il faut donc prendre en compte également ce fait avec une multiplication par une fonction porte

$$\mathbb{1}_E(\alpha, \beta) = \mathbb{1}_{L_\alpha}(\alpha) \mathbb{1}_{L_\beta}(\beta) = 1 \quad \text{si} \quad (\alpha, \beta) \in [-L_\alpha/2, L_\alpha/2] \times [-L_\beta/2, L_\beta/2]$$

avec  $L_\alpha$  et  $L_\beta$  la taille du capteur dans les dimensions  $\alpha$  et  $\beta$  respectivement. Dans l'espace de Fourier, la multiplication de l'espace direct s'écrit comme une convolution

$$\mathring{\mathbb{1}}_E(f_\alpha, f_\beta) = L_\alpha L_\beta \text{sinc}_\pi(L_\alpha f_\alpha) \text{sinc}_\pi(L_\beta f_\beta)$$

soit

$$\mathring{\phi}_c^e(f_\alpha, f_\beta, t) = \sum_{l,m} \iint \mathring{\phi}_c(f'_\alpha - f_\alpha^{lm}, f'_\beta - f_\beta^{lm}, t) \mathring{\mathbb{1}}_E(f'_\alpha - f_\alpha, f'_\beta - f_\beta) df'_\alpha df'_\beta \quad (2.25)$$

qui est la convolution entre le spectre dupliqué, filtré et déphasé, par un sinus cardinal. Nous avons donc l'expression du contenu fréquentiel du signal à la sortie des cornets. Fondamentalement, le spectre du ciel est filtré par la fonction de transfert de l'optique et le résultat est rendu périodique par l'échantillonnage. Ensuite il y a une convolution avec des sinus-cardinaux pour prendre en compte le fait que le champ de vue est limité.

À la suite de ce système se trouvent les bolomètres. En quelque sorte nous allons effectuer les mêmes opérations : les bolomètres convoluent le signal d'entrée, l'échantillonnent et comme le nombre d'échantillons temporels est limité nous allons multiplier par une fenêtre. Une différence cependant est que la dimension  $t$  ne se trouve que dans le déphasage (2.23). Le résultat des transformées de Fourier sera donc un peu particulier.

### 2.3.3 Convolution et échantillonnage temporel

Tout comme pour les cornets, les bolomètres introduisent une intégration capteur avec la réponse  $h_b$ . Les opérations se font dans la dimension temporelle cette fois. De nouveau on peut écrire

l'intégration capteur comme une convolution avec  $h_b$ , suivie d'une multiplication avec un peigne de Dirac  $\Delta_{T_e}$  de période  $T_e$  et d'une multiplication par une fenêtre temporelle  $\mathbf{1}_{\mathcal{T}}$  de temps  $\mathcal{T}$

$$y(\alpha, \beta, t) = \mathbf{1}_{\mathcal{E}}(\alpha, \beta) \mathbf{1}_{\mathcal{T}}(t) \Delta_{T_{\alpha\beta}}(\alpha, \beta) \Delta_{T_e}(t) \int_{t'} \phi_c(\alpha, \beta, t') h_b(t' - t) dt'$$

Cette fenêtre d'observation  $\mathcal{T}$  correspond par exemple à une phase du protocole *Large map* ou un pointage précis dans les protocoles *Small map* ou *Point source*.

Comme précédemment, la transformée de Fourier de cette expression fera apparaître une convolution avec un sinus cardinal, ainsi qu'une répllication provoquée par l'échantillonnage. La transformée de Fourier des données, par rapport aux variable  $(\alpha, \beta)$  et  $t$ , s'écrit donc

$$\begin{aligned} \mathring{Y}(f_\alpha, f_\beta, f) &= \iiint y(\alpha, \beta, t) e^{-2j\pi(\alpha f_\alpha + f_\alpha \beta + ft)} d\alpha d\beta dt \\ &= \iint_{t, t'} \iint_{\alpha, \beta} \mathbf{1}_{\mathcal{E}}(\alpha, \beta) \Delta_{T_{\alpha\beta}}(\alpha, \beta) \phi_c(\alpha, \beta, t') e^{-2j\pi(\alpha f_\alpha + f_\alpha \beta)} d\alpha d\beta \\ &\quad h_b(t' - t) dt' \mathbf{1}_{\mathcal{T}}(t) \Delta_{T_e}(t) e^{-2j\pi ft} dt \end{aligned} \quad (2.26)$$

On voit apparaître dans la première partie (2.26) la transformée de Fourier  $\mathring{\phi}_c$  du flux échantillonné à la sortie des cornets, obtenue précédemment équation (2.25). En remplaçant l'expression on obtient

$$\begin{aligned} \mathring{Y}(f_\alpha, f_\beta, f) &= \sum_{l, m} \iint_{f'_\alpha, f'_\beta} \mathring{\mathbf{1}}_{\mathcal{E}}(f'_\alpha - f_\alpha, f'_\beta - f_\beta) \iint_{t, t'} \mathring{\phi}_c(f'_\alpha - f_\alpha^{lm}, f'_\beta - f_\beta^{lm}, t) \\ &\quad h_b(t' - t) dt' \mathbf{1}_{\mathcal{T}}(t) \Delta_{T_e}(t) e^{-2j\pi ft} dt df'_\alpha df'_\beta \\ &= \sum_{l, m} \iint_{f'_\alpha, f'_\beta} \mathring{\mathbf{1}}_{\mathcal{E}}(f'_\alpha - f_\alpha, f'_\beta - f_\beta) \mathring{\phi}(f'_\alpha - f_\alpha^{lm}, f'_\beta - f_\beta^{lm}) \mathring{H}_{mc}(f'_\alpha - f_\alpha^{lm}, f'_\beta - f_\beta^{lm}) \\ &\quad \left[ \iint_{t, t'} h_b(t' - t) \mathbf{1}_{\mathcal{T}}(t) \Delta_{T_e}(t) e^{-2j\pi((f'_\alpha - f_\alpha^{lm})p_\alpha(t') + (f'_\beta - f_\beta^{lm})p_\beta(t') + ft)} dt dt' \right] df'_\alpha df'_\beta \end{aligned} \quad (2.27)$$

$$\left[ \iint_{t, t'} h_b(t' - t) \mathbf{1}_{\mathcal{T}}(t) \Delta_{T_e}(t) e^{-2j\pi((f'_\alpha - f_\alpha^{lm})p_\alpha(t') + (f'_\beta - f_\beta^{lm})p_\beta(t') + ft)} dt dt' \right] df'_\alpha df'_\beta \quad (2.28)$$

avec  $\mathring{H}_{mc}(f_\alpha, f_\beta) = \mathring{H}_m(f_\alpha, f_\beta) \mathring{H}_c(f_\alpha, f_\beta)$ , l'expression de  $\mathring{\phi}_c$  de l'équation (2.24) et en réarrangeant les intégrations. La première partie (2.27) correspond au spectre du ciel filtré et convolué avec le sinus cardinal. La deuxième partie (2.28) correspond aux opérations temporelles que nous allons détailler.

À l'origine le balayage était lié au signal d'entrée  $\phi$ . On peut voir maintenant que le déphasage provenant de l'observation n'est présent que dans la deuxième partie concernant le temps. L'information spatiale introduite à l'aide du balayage dans l'espace est, dans les données, portée par le temps. Il y a donc un lien entre les dimensions temporelle et spatiales qui dépend des fonctions  $p_\alpha$  et  $p_\beta$ . Pour continuer la description il faut calculer (2.28)

$$\begin{aligned} I &= \iint_{t, t'} h_b(t' - t) \mathbf{1}_{\mathcal{T}}(t) \Delta_{T_e}(t) e^{-2j\pi((f'_\alpha - f_\alpha^{lm})p_\alpha(t') + (f'_\beta - f_\beta^{lm})p_\beta(t') + ft)} dt dt' \\ I &= \iint_{t, t'} h_b(t' - t) \mathbf{1}_{\mathcal{T}}(t) \Delta_{T_e}(t) e^{-2j\pi ft} dt e^{-2j\pi((f'_\alpha - f_\alpha^{lm})p_\alpha(t') + (f'_\beta - f_\beta^{lm})p_\beta(t'))} dt' \end{aligned}$$

Si on extrait seulement l'intégrale suivant  $t$  nous avons

$$\begin{aligned} \int_t h_b(t' - t) \mathbb{1}_{\mathcal{T}}(t) \Delta_{T_e}(t) e^{-2j\pi ft} dt &= \mathring{H}_b(f) e^{-2i\pi ft'} \star \int_t \mathbb{1}_{\mathcal{T}}(t) \Delta_{T_e}(t) dt \\ &= \mathring{H}_b(f) e^{-2i\pi ft'} \star \sum_k \mathring{\mathbb{1}}_{\mathcal{T}}(f - f^k) \\ &= \sum_k \int \mathring{H}_b(f' - f^k) e^{-2i\pi(f' - f^k)t'} \mathring{\mathbb{1}}_{\mathcal{T}}(f' - f) df' \end{aligned}$$

avec  $f^k = k/T_e$  et  $\star$  l'opérateur de convolution. En injectant cette expression dans (2.28), le spectre des données s'écrit

$$\begin{aligned} \mathring{Y}(f_\alpha, f_\beta, f) &= \sum_{l,m,k} \iiint \mathring{\mathbb{1}}_{\mathcal{T}}(f' - f) \mathring{\mathbb{1}}_{\mathcal{E}}(f'_\alpha - f_\alpha, f'_\beta - f_\beta) \left[ \mathring{\phi}(f'_\alpha - f_\alpha^{lm}, f'_\beta - f_\beta^{lm}) \mathring{H}_b(f' - f^k) \right. \\ &\quad \left. \mathring{H}_{mc}(f'_\alpha - f_\alpha^{lm}, f'_\beta - f_\beta^{lm}) \int_{t'} e^{-2j\pi((f'_\alpha - f_\alpha^{lm})p_\alpha(t') + (f'_\beta - f_\beta^{lm})p_\beta(t') + (f' - f^k)t')} dt' \right] df'_\alpha df'_\beta df'. \quad (2.29) \end{aligned}$$

On remarque que l'ensemble est cohérent : tout d'abord nous avons trois convolutions suivant  $f'_\alpha, f'_\beta$  et  $f'$  par des sinus cardinaux dues à la limitation du champ de vue. Ensuite on trouve ce qui forme le contenu fréquentiel des données formées par le spectre du ciel et les trois fonctions de transfert du miroir, des cornets et des bolomètres. Il y a cependant une transformation du spectre formé par le protocole d'observation  $(p_\alpha, p_\beta)$ .

Pour obtenir le spectre des données il faut calculer l'intégrale par rapport à  $t'$  qui dépend du protocole d'observation. Avant de faire les calculs et pour alléger les écritures nous allons considérer que le champ de vue du capteur est infini et qu'il observe pendant un temps infini. Dans ce cas les sinus cardinaux  $\mathring{\mathbb{1}}_{\mathcal{E}}$  et  $\mathring{\mathbb{1}}_{\mathcal{T}}$  sont remplacés par des Dirac et on obtient

$$\begin{aligned} \mathring{Y}(f_\alpha, f_\beta, f) &= \sum_{l,m,k} \mathring{\phi}(f_\alpha - f_\alpha^{lm}, f_\beta - f_\beta^{lm}) \mathring{H}_{mc}(f_\alpha - f_\alpha^{lm}, f_\beta - f_\beta^{lm}) \mathring{H}_b(f - f^k) \\ &\quad \int_{t'} e^{-2j\pi((f_\alpha - f_\alpha^{lm})p_\alpha(t') + (f_\beta - f_\beta^{lm})p_\beta(t') + (f - f^k)t')} dt' \quad (2.30) \end{aligned}$$

qui fait apparaître le spectre du ciel filtré et répliqué, le filtre des bolomètres et l'échantillonnage temporel ainsi qu'une intégrale qui dépendant du protocole d'observation. Nous avons donc deux cas à envisager en fonction du protocole d'observation.

### 2.3.4 Cas des protocoles *Small map* et *Point source*

Ces protocoles consistent en différentes positions pointées indexées par  $q \in \{1, 2, \dots, Q\}$ . D'après (2.18) et (2.19) page 48 l'expression de l'intégrale (2.30) est

$$\begin{aligned} \int_{t'} e^{-2j\pi((f'_\alpha - f_\alpha^{lm}) \sum c_\alpha^q \mathbb{1}_{[t_q, t_{q+1}[}(t') + (f'_\beta - f_\beta^{lm}) \sum c_\beta^q \mathbb{1}_{[t_q, t_{q+1}[}(t') + (f - f^k)t')} dt' = \\ \int_{t'} e^{-2j\pi(f - f^k)t'} \prod_q e^{-2j\pi(f'_\alpha - f_\alpha^{lm})c_\alpha^q \mathbb{1}_{[t_q, t_{q+1}[}(t')} e^{-2j\pi(f'_\beta - f_\beta^{lm})c_\beta^q \mathbb{1}_{[t_q, t_{q+1}[}(t')} dt'. \quad (2.31) \end{aligned}$$

En remarquant que l'exponentiation d'une fonction constante par morceaux (le pointage) est également une fonction constante par morceaux

$$e^{-2j\pi c_\alpha^q \mathbb{1}_{[t_q, t_{q+1}[}(t)} = \begin{cases} e^{-2j\pi c_\alpha^q} & \text{si } t \in [t_q, t_{q+1}[ \\ 1 & \text{sinon} \end{cases}$$

et comme les intervalles  $[t_q \ t_{q+1}[$  sont deux à deux disjoints, on peut réécrire le produit en une somme

$$\begin{aligned}
 \prod_q e^{-2j\pi c_\alpha^q} \mathbb{1}_{[t_q \ t_{q+1}[}(t) &= \begin{cases} e^{-2j\pi c_\alpha^q} & \text{si } t \in [t_q \ t_{q+1}[ \\ 1 & \text{si } t \notin [t_1 \ t_Q] \end{cases} \\
 &= \sum_q \mathbb{1}_{[t_q \ t_{q+1}[}(t) e^{-2j\pi c_\alpha^q} + \mathbb{1}_{[-\infty \ t_1[}(t) + \mathbb{1}_{[t_Q \ +\infty[}(t). \\
 &= 1 + \sum_q \mathbb{1}_{[t_q \ t_{q+1}[}(t) \left( e^{-2j\pi c_\alpha^q} - 1 \right) \\
 &= 1 - \mathbb{1}_{\mathcal{T}}(t - t_m) + \sum_q \mathbb{1}_{[t_q \ t_{q+1}[}(t) e^{-2j\pi c_\alpha^q} \quad (2.32)
 \end{aligned}$$

où  $\mathcal{T} = t_Q - t_1$  est le temps total d'observation de l'acquisition et  $t_m = (t_1 + t_Q)/2$ . En exploitant (2.32) l'intégrale (2.31) se réécrit

$$\begin{aligned}
 I &= \int_{t'} \prod_q e^{-2j\pi \left( (f'_\alpha - f_\alpha^{lm})c_\alpha^q + (f'_\beta - f_\beta^{lm})c_\beta^q \right)} \mathbb{1}_{[t_q \ t_{q+1}[}(t') + (f - f^k)t' \ dt' \\
 &= \int_{t'} e^{-2j\pi (f - f^k)t'} \left( 1 - \mathbb{1}_{\mathcal{T}}(t - t_m) + \sum_q \mathbb{1}_{[t_q \ t_{q+1}[}(t') e^{-2j\pi \left( (f'_\alpha - f_\alpha^{lm})c_\alpha^q + (f'_\beta - f_\beta^{lm})c_\beta^q \right)} \right) dt' \\
 &= \int_{t'} e^{-2j\pi (f - f^k)t'} dt' - \int_{t'} \mathbb{1}_{\mathcal{T}}(t - t_m) e^{-2j\pi (f - f^k)t'} dt' + \\
 &\quad \sum_q \int_{t'} \mathbb{1}_{[t_q \ t_{q+1}[}(t') e^{-2j\pi \left( (f'_\alpha - f_\alpha^{lm})c_\alpha^q + (f'_\beta - f_\beta^{lm})c_\beta^q + (f - f^k)t' \right)} dt'.
 \end{aligned}$$

Les trois intégrales sont explicites soit

$$\begin{aligned}
 \int_{t'} e^{-2j\pi (f - f^k)t'} dt' &= \delta(f - f^k) \\
 \int_{t'} \mathbb{1}_{\mathcal{T}}(t - t_m) e^{-2j\pi (f - f^k)t'} dt' &= \mathcal{T} \operatorname{sinc} \left( \mathcal{T}(f - f^k) \right) e^{-2j\pi t_m}
 \end{aligned}$$

et

$$\begin{aligned}
 \sum_q e^{-2j\pi \left( (f'_\alpha - f_\alpha^{lm})c_\alpha^q + (f'_\beta - f_\beta^{lm})c_\beta^q \right)} \int_{t'} \mathbb{1}_{[t_q \ t_{q+1}[}(t') e^{-2j\pi (f - f^k)t'} dt' &= \\
 \sum_q \Delta T_q \operatorname{sinc} \left( \Delta T_q (f - f^k) \right) e^{-2j\pi \left( (f'_\alpha - f_\alpha^{lm})c_\alpha^q + (f'_\beta - f_\beta^{lm})c_\beta^q + T_q \right)} &
 \end{aligned}$$

où  $T_q = (t_q + t_{q+1})/2$  et  $\Delta T_q = t_{q+1} - t_q$ . En reportant ce résultat dans l'équation de la transformée de Fourier des données (2.30) le spectre des données est explicite soit

$$\begin{aligned}
 \mathring{Y}(f_\alpha, f_\beta, f) &= \sum_{l,m,k} \mathring{\phi}(f_\alpha - f_\alpha^{lm}, f_\beta - f_\beta^{lm}) \mathring{H}_{\text{mc}}(f_\alpha - f_\alpha^{lm}, f_\beta - f_\beta^{lm}) \mathring{H}_{\text{b}}(f - f^k) \\
 &\quad \left( \delta(f - f^k) - \mathcal{T} \operatorname{sinc} \left( \mathcal{T}(f - f^k) \right) e^{-2j\pi t_m} + \right. \\
 &\quad \left. \sum_q \Delta T_q \operatorname{sinc} \left( \Delta T_q (f - f^k) \right) e^{-2j\pi T_q} e^{-2j\pi \left( (f_\alpha - f_\alpha^{lm})c_\alpha^q + (f_\beta - f_\beta^{lm})c_\beta^q \right)} \right) \quad (2.33)
 \end{aligned}$$

où  $(f_\alpha^{lm}, f_\beta^{lm})$  correspondent aux positions des spectres dupliqués sur les axes  $f_\alpha$  et  $f_\beta$  dues à l'échantillonnage spatial,  $f^k = k/T_e$  correspondent aux positions des spectres dupliqués sur l'axe

$f$  dues à l'échantillonnage temporel,  $\mathcal{T}$  la durée totale d'observation et  $\Delta T_q$  la durée d'observation sur chaque pointage aux coordonnées  $(c_\alpha^q, c_\beta^q)$ . On peut remarquer que chaque réplication centrée en  $(f_\alpha^{lm}, f_\beta^{lm})$  du spectre  $\mathring{\phi}$  filtré possède une phase différente venant de

$$\exp -2j\pi((f_\alpha - f_\alpha^{lm})c_\alpha^q + (f_\beta - f_\beta^{lm})c_\beta^q).$$

Autrement dit les fréquences repliées ne sont plus parfaitement identiques entre les répliques, et la dégénérescence est levée, au moins en partie. On peut en exploitant correctement ces données espérer lever le repliement de spectre. Cette possibilité est plus illustrée et démontrée dans le cas *Large map*.

### 2.3.5 Cas du protocole *Large map*

Dans ce cas pour une phase particulière indexée par  $q$ , nous avons

$$p_\alpha(t) = a_\alpha^q t^2 + v_\alpha^q t + c_\alpha^q \quad \text{et} \quad p_\beta(t) = a_\beta^q t^2 + v_\beta^q t + c_\beta^q$$

pour les phases d'accélération, et pendant les phases à vitesse constante

$$p_\alpha(t) = v_\alpha^q t + c_\alpha^q \quad \text{et} \quad p_\beta(t) = v_\beta^q t + c_\beta^q.$$

Dans le cas de la phase d'accélération l'intégrale de (2.30) n'est pas explicite. En revanche dans la phase à vitesse constante nous avons

$$\int_{t'} e^{-2j\pi((f_\alpha - f_\alpha^{lm})(v_\alpha^q t' + c_\alpha^q) + (f_\beta - f_\beta^{lm})(v_\beta^q t' + c_\beta^q) + (f - f^k)t')} dt' = \delta\left((f_\alpha - f_\alpha^{lm})v_\alpha^q + (f_\beta - f_\beta^{lm})v_\beta^q + f - f^k\right) e^{-2j\pi((f_\alpha - f_\alpha^{lm})c_\alpha^q + (f_\beta - f_\beta^{lm})c_\beta^q)}.$$

Donc le spectre des données dans le cas *Large map* pendant la phase à vitesse constante est

$$\boxed{\mathring{Y}(f_\alpha, f_\beta, f) = \sum_{l, m, k} \mathring{\phi}(f_\alpha - f_\alpha^{lm}, f_\beta - f_\beta^{lm}) \mathring{H}_{\text{mc}}(f_\alpha - f_\alpha^{lm}, f_\beta - f_\beta^{lm}) \mathring{H}_{\text{b}}(f - f^k) \delta\left((f_\alpha - f_\alpha^{lm})v_\alpha^q + (f_\beta - f_\beta^{lm})v_\beta^q + f - f^k\right) e^{-2j\pi((f_\alpha - f_\alpha^{lm})c_\alpha^q + (f_\beta - f_\beta^{lm})c_\beta^q)}} \quad (2.34)$$

où  $(f_\alpha^{lm}, f_\beta^{lm})$  sont les positions des spectres dupliqués sur les axes  $f_\alpha$  et  $f_\beta$  dues à l'échantillonnage spatial et  $f^k = k/T_e$  sont les positions des spectres dupliqués sur l'axe  $f$  des périodes dues à l'échantillonnage temporel.

### 2.3.6 Interprétation des spectres dans le cas *Large map*

Nous proposons de mettre en évidence dans le cas *Large map* les possibilités de sur-résolution, c'est-à-dire la possibilité de restaurer des fréquences spatiales au-delà de la limite d'échantillonnage spatiale du détecteur. Pour faciliter l'analyse nous allons considérer un signal mono-dimensionnel, la généralisation au cas à deux dimensions étant immédiate. De plus nous allons ignorer pour le moment l'échantillonnage temporel en ne considérant qu'une réplication des motifs sur l'axe  $f$ . Dans ce cas le spectre des données est plus simple et s'écrit

$$\mathring{Y}(f_\alpha, f) = \sum_l \mathring{\phi}(f_\alpha - f_\alpha^l) \mathring{H}_{\text{mc}}(f_\alpha - f_\alpha^l) \mathring{H}_{\text{b}}(f) \delta\left((f_\alpha - f_\alpha^l)v_\alpha^q + f\right) e^{-2j\pi(f_\alpha - f_\alpha^l)c_\alpha^q}. \quad (2.35)$$

On peut tout d'abord noter que :

- le spectre du ciel  $\overset{\circ}{\phi}$  est filtré par la combinaison du miroir et des cornets  $\overset{\circ}{H}_{\text{mc}}$ . De plus le spectre résultant est dupliqué en  $f_{\alpha}^l$ ,
- la dimension duale du temps  $f$  n'apparaît que dans  $\overset{\circ}{H}_{\text{b}}$  et le Dirac,
- la position d'origine du balayage se retrouve dans le déphasage.

Finalement le plus intéressant est la présence du Dirac qui opère un lien entre les dimensions temporelle et spatiale. On peut en effet se poser la question de l'influence du bolomètre, qui ne réagit que dans le temps, vis-à-vis du signal d'intérêt, le ciel, qui lui n'est défini que dans l'espace. Le lien entre le temps et l'espace dans les données s'exprime à travers le Dirac en notant que celui-ci est non nul si et seulement si

$$f = - (f_{\alpha} - f_{\alpha}^l) v_{\alpha}^q.$$

Il faut donc prendre en compte les deux axes fréquentiels conjointement pour que le spectre des données soit non nul. Cela permet d'ailleurs d'écrire le spectre des données suivant l'axe des fréquences spatiales  $f_{\alpha}$

$$\overset{\circ}{Y}_s(f_{\alpha}) = \sum_l \overset{\circ}{\phi}(f_{\alpha} - f_{\alpha}^l) \overset{\circ}{H}_{\text{mc}}(f_{\alpha} - f_{\alpha}^l) \overset{\circ}{H}_{\text{b}}(- (f_{\alpha} - f_{\alpha}^l) v_{\alpha}^q) e^{-2j\pi((f_{\alpha} - f_{\alpha}^l) c_{\alpha}^q)} \quad (2.36)$$

ou suivant l'axe des fréquences temporelles  $f$

$$\overset{\circ}{Y}_t(f) = \sum_l \overset{\circ}{\phi}\left(-\frac{f}{v_{\alpha}^q}\right) \overset{\circ}{H}_{\text{mc}}\left(-\frac{f}{v_{\alpha}^q}\right) \overset{\circ}{H}_{\text{b}}(f) e^{-2j\pi(-f c_{\alpha}^q / v_{\alpha}^q)}. \quad (2.37)$$

On peut également remarquer que

$$\frac{f}{v_{\alpha}^q} \quad \text{en} \quad \left[ \frac{1 \text{ s}}{\text{s m}} \right] = [\text{m}^{-1}]$$

est bien homogène à une fréquence spatiale et que

$$(f_{\alpha} - f_{\alpha}^l) v_{\alpha}^q \quad \text{en} \quad \left[ \frac{1 \text{ m}}{\text{m s}} \right] = [\text{s}^{-1}]$$

est bien homogène à une fréquence duale du temps. Les équations (2.36) et (2.37) illustrent comment les fonctions de transfert spatiales affectent le contenu spectral des données d'un point de vue temporel et inversement. L'équation (2.37) modélise l'influence de la fonction de transfert du bolomètre (définie dans le temps) sur les fréquences spatiales.

### Mise en regard des fonctions de transfert

Étant donné que l'influence des différentes fonctions de transfert peuvent être comparées il est possible de les mettre en regard les unes avec les autres. Supposons que nous puissions définir pour chaque fonction de transfert une fréquence de coupure spatiale caractéristique  $f_{\alpha}^s$  pour la fonction de transfert  $\overset{\circ}{H}_{\text{mc}}$  et une fréquence de coupure temporelle  $f^t$  pour la fonction de transfert temporelle  $\overset{\circ}{H}_{\text{b}}$ . Dans ce cas d'après (2.36) la fréquence de coupure de  $\overset{\circ}{H}_{\text{mc}}(f_{\alpha})$  est  $f_{\alpha}^s$  et celle de  $\overset{\circ}{H}_{\text{b}}(-f_{\alpha} v_{\alpha}^q)$  est  $f^t / v_{\alpha}^q$ . Les relations suivantes peuvent être déduites :

- si la vitesse tend vers l'infini ( $v_{\alpha}^q \rightarrow +\infty$ ), alors la fréquence de coupure de  $\overset{\circ}{H}_{\text{b}}(-f_{\alpha} v_{\alpha}^q)$  tend vers zéro et tout est filtré. Il s'agit du cas où l'instrument se déplace trop rapidement pour que le capteur ait le temps de réagir.
- si la vitesse tend vers zéro ( $v_{\alpha}^q \rightarrow 0$ ) alors la fréquence de coupure tend vers l'infini. La réponse temporelle du capteur n'a plus aucune influence. Cependant l'instrument ne bouge plus il n'y a donc effectivement plus d'information temporelle.

- si  $f^t/v_\alpha^q < f_\alpha^s$  alors la réponse temporelle filtre plus que la réponse spatiale. On se déplace trop rapidement. Et inversement si  $f^t/v_\alpha^q > f_\alpha^s$  alors la réponse spatiale filtre plus que la réponse temporelle et on se déplace trop lentement.
- si  $f^t/v_\alpha^q = f_\alpha^s$  alors la vitesse est telle qu'aucune des deux réponses ne filtre plus que l'autre.

Ces remarques sont importantes dans le sens où, en astronomie, beaucoup d'efforts sont mis dans la construction de miroirs plus grands possédant ainsi une fréquence de coupure plus grande. Si le capteur et la vitesse sont tels que la réponse dans le temps filtre plus que la réponse de l'optique, un miroir de plus petite taille n'introduirait pas de perte supplémentaire. En quelque sorte il n'était pas nécessaire de construire un miroir de cette taille.

Dans le cas de SPIRE la fréquence de coupure du télescope (de la figure d'*Airy*) est le diamètre sur la longueur d'onde  $D/\lambda$ . À la longueur d'onde la plus courte (considéré à  $213 \mu\text{m}$ ) la fréquence spatiale de coupure du miroir est

$$f_\alpha = \frac{D}{\lambda} \approx 0,075 \text{ arcsec}^{-1}.$$

Pour la longueur d'onde la plus grande ( $611 \mu\text{m}$ ), la fréquence de coupure est

$$f_\alpha = \frac{D}{\lambda} \approx 0,026 \text{ arcsec}^{-1}.$$

La fréquence de coupure du bolomètre à 3dB est, voir annexe D,

$$f^t \approx 51\text{Hz}$$

donc nous avons

$$\frac{f^t}{v_\alpha^q} = \frac{51}{30} \approx 1,7 \text{ arcsec}^{-1}$$

ou

$$\frac{f^t}{v_\alpha^q} = \frac{51}{60} \approx 0,85 \text{ arcsec}^{-1}.$$

Autrement dit quelque soit la vitesse de balayage, le miroir filtre plus que le bolomètre. Avec une vitesse de  $60 \text{ arcsec.s}^{-1}$  les fréquences de coupure sont presque égales. En quelque sorte, le déplacement pourrait être plus rapide sans plus de perte de fréquence que celles induites par le miroir.

Pour une description plus fine il faut cependant prendre en compte le filtre de l'électronique de lecture qui a une fréquence de coupure à 3dB à 5 Hz (pour être exact, il s'agit d'un filtre d'ordre cinq, la comparaison est donc limitée). Dans ce cas nous avons

$$\frac{f^t}{v_\alpha^q} = \frac{5}{30} \approx 0,16 \text{ arcsec}^{-1}$$

ou

$$\frac{f^t}{v_\alpha^q} = \frac{5}{60} \approx 0,083 \text{ arcsec}^{-1}.$$

Donc avec un balayage à vitesse de  $60 \text{ arcsec.s}^{-1}$ , les fréquences de coupure du filtre et du miroir sont équivalentes à  $213 \mu\text{m}$  et le miroir filtre plus au-delà de  $213 \mu\text{m}$ . Avec une vitesse de  $30 \text{ arcsec.s}^{-1}$  la fréquence de coupure du bolomètre est largement au delà de la fréquence de coupure du miroir. On peut en conclure qu'il ne sert à rien d'effectuer un balayage à une vitesse inférieure à  $60 \text{ arcsec.s}^{-1}$  puisque le miroir a déjà supprimé toute l'information que l'on cherche à préserver en balayant à cette vitesse. Les données ne porteront pas plus d'information. Il est possible par conséquent de gagner en temps d'observation ou en surface parcourue, ainsi qu'en volume de données à traiter.

Deux arguments tempèrent fortement cette conclusion. Tout d'abord, à surface équivalente, à vitesse réduite plus de données sont collectées permettant une meilleure estimation de l'objet (le rapport signal sur bruit est plus favorable). De plus l'étude précédente néglige l'effet de l'échantillonnage et une vitesse plus réduite rend possible plus facilement la restauration des hautes fréquences, comme présenté dans la partie suivante, puisque la distance entre deux instants d'échantillonnage est plus petite.

### Possibilité de sur-résolution

Nous allons décrire comment un balayage à vitesse constante permet de lever la dégénérescence introduite par un repliement spectral. Nous allons également retrouver certaines propriétés sur la vitesse de balayage décrites précédemment. Un aspect important dans l'expression de la transformée de Fourier des données

$$\mathring{Y}(f_\alpha, f) = \sum_l \mathring{\phi}(f_\alpha - f_\alpha^l) \mathring{H}_{\text{mc}}(f_\alpha - f_\alpha^l) \mathring{H}_b(f) \delta((f_\alpha - f_\alpha^l)v_\alpha^q + f) e^{-2j\pi(f_\alpha - f_\alpha^l)c_\alpha^q} \quad (2.38)$$

est la transformation effectuée sur un signal défini dans l'espace  $\mathring{\phi}(f_\alpha)$  pour obtenir un signal défini dans l'espace et le temps  $\mathring{Y}(f_\alpha, f)$ . Le lien entre les deux dimensions est fait par le Dirac  $\delta$ .

Si le spectre spatial de l'image échantillonnée

$$\sum_l \mathring{\phi}(f_\alpha - f_\alpha^l) \mathring{H}_{\text{mc}}(f_\alpha - f_\alpha^l)$$

présente un repliement spectral on obtient le schéma de la figure 2.11(a) qui représente le support de chaque réplique. De même en représentant le spectre contenant seulement des Diracs

$$\sum_l \delta((f_\alpha - f_\alpha^l)v_\alpha^q + f)$$

on obtient la figure 2.11(b). Chaque Dirac définit une droite, dans le plan, d'équation

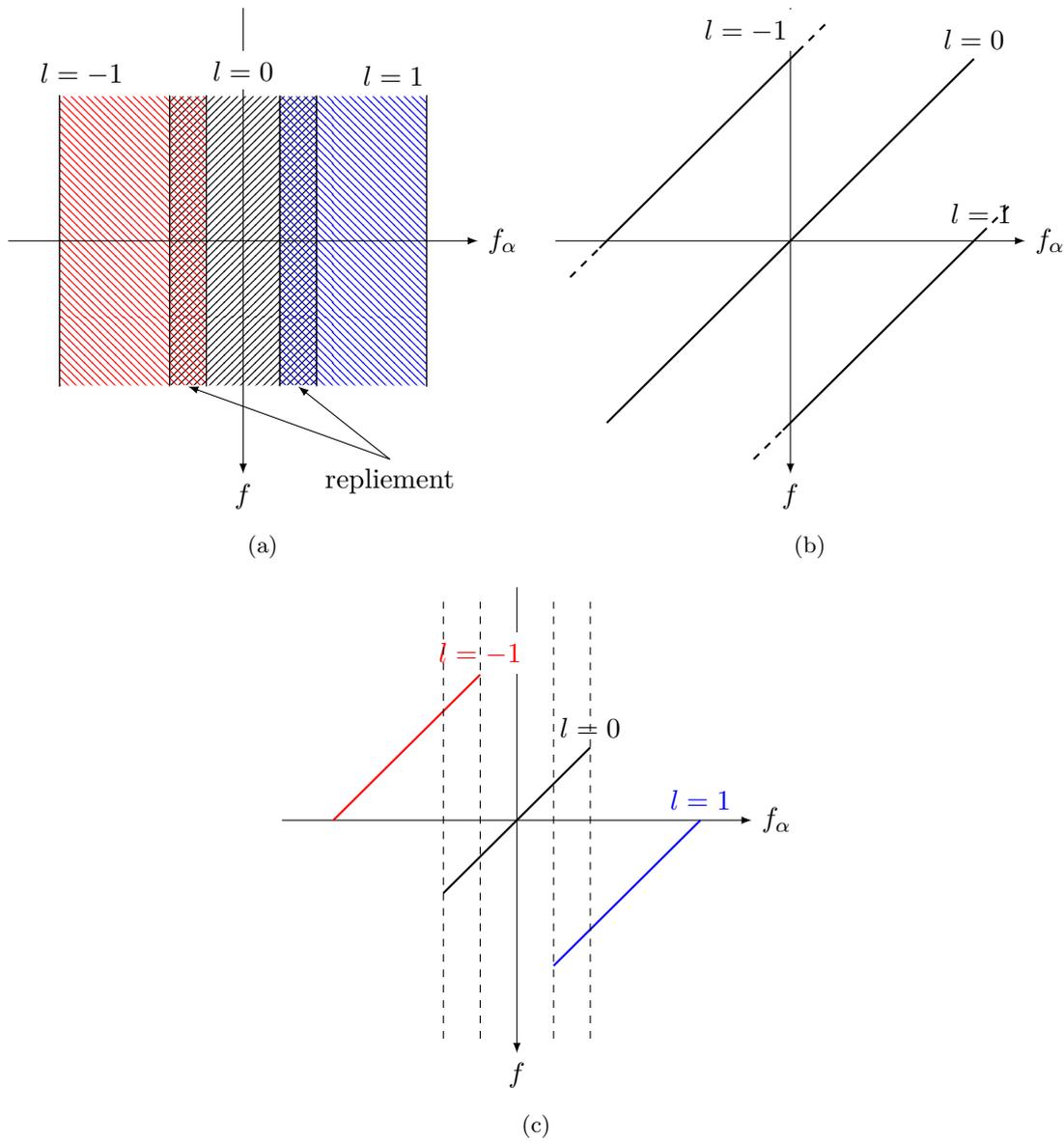
$$f = -\frac{f_\alpha}{v_\alpha^q} + \frac{f_\alpha^l}{v_\alpha^q}$$

en dehors de laquelle le spectre des données est nul. Toutes ces droites sont parfaitement parallèles, leur pente étant donnée par la vitesse. Lorsque l'on procède à la multiplication des duplications deux à deux

$$\mathring{Y}(f_\alpha, f) = \sum_l \mathring{\phi}(f_\alpha - f_\alpha^l) \mathring{H}_{\text{mc}}(f_\alpha - f_\alpha^l) \delta((f_\alpha - f_\alpha^l)v_\alpha^q + f)$$

on obtient un spectre dont le support est illustré figure 2.11(c) et qui correspond exactement au spectre des données (2.38) sans le filtre  $\mathring{H}_b$  et le déphasage. Autrement dit, les périodes provenant du spectre du ciel sont parfaitement séparées, il n'y a pas de repliement. Cela veut dire que le spectre des données acquises pendant un balayage est moins replié que le spectre de données qui auraient été acquises en observant directement le signal, sans mouvement. On peut de plus faire quelques remarques :

- si la vitesse tend vers l'infini ( $v_\alpha^q \rightarrow +\infty$ ), le coefficient des droites  $1/v_\alpha^q$  tend vers 0 ainsi que l'origine  $f_\alpha^l/v_\alpha^q$ . Les Diracs sont confondus avec l'axe  $f_\alpha$  : il n'y a pas de séparation entre les périodes.
- si la vitesse tend vers zéro ( $v_\alpha^q \rightarrow 0$ ) alors les droites ont pour expression  $f = 0$ . Il n'y a plus de sur-résolution possible, les répliques sont de nouveau repliées.



**Figure 2.11** – Illustration du support des spectres. La figure 2.11(a) représente le support de trois duplications du spectre du ciel filtré et échantillonné. Ce spectre est défini seulement sur  $f_\alpha$  et est étendu le long de l'axe  $f$ . La figure 2.11(b) représente le support des Diracs. Le support du spectre des données est illustré sur le figure 2.11(c). C'est la somme des produits deux à deux entre chaque période  $l$  présente dans le signal et les Diracs décrit par (2.38).

Cette présentation est une démonstration des possibilités de sur-résolution en exploitant correctement les propriétés des données et la manière dont elles ont été acquises. Par ailleurs nous avons fait plusieurs hypothèses pour cet exposé. Premièrement la duplication provoquée par l'échantillonnage temporel n'a pas été incluse. Ensuite les champs de vue spatiaux et temporels ne sont pas infinis ce qui introduit des sinus cardinaux à la place de Dirac (il n'y a donc plus de séparation parfaite en les périodes  $l$ ). Certains développements dans ce sens ont fait l'objet d'une publication [ORG09]. Enfin l'étude n'a été menée que dans le cas du protocole *Large map*. Il serait intéressant d'étudier les autres cas.

La partie suivante est dédiée aux calculs explicites d'un modèle d'acquisition. L'expression du ciel est explicitée ainsi que l'expression des réponses. Le tout est combiné pour obtenir l'expression d'une fonction la plus explicite possible pour alléger les calculs numériques.

## 2.4 Calcul explicite du modèle d'acquisition

Les équations qui relient l'inconnue  $\phi$  aux données permettent d'exploiter l'information présente dans celles-ci au travers d'un terme de vraisemblance. Étant donné que nous utiliserons un ensemble fini de donnée l'image de l'estimateur sera également de dimension fini et la fonction  $\phi$  sera exprimée à l'aide de coefficients. Dans ce cas les algorithmes d'optimisation et d'échantillonnage stochastique en dimensions finies sont utilisables. Cette partie présente les calculs qui peuvent être menés pour obtenir un modèle d'acquisition explicite permettant de réduire les calculs numériques.

### 2.4.1 Modèle de ciel

Le ciel est défini comme une fonction continue sur trois variables  $\alpha, \beta$  et  $\lambda$ . La question est de déterminer un modèle de ciel capable de représenter l'information à restaurer. Nous pouvons classer l'information suivant deux types : l'information spatiale définie suivant  $(\alpha, \beta)$  et l'information spectrale suivant  $\lambda$ . Étant donnée l'analyse précédente de l'instrument partie 2.2, il apparaît que ce dernier ne contient pas dans les données la même quantité d'information dans un cas et dans l'autre.

- Dans le cas spectral seuls trois points d'échantillonnage sont disponibles, un par bande d'observation (250, 360 et 520  $\mu\text{m}$ ). Il y a donc très peu d'informations dans les données. Il serait pourtant dommage de négliger les longueurs d'onde car les astronomes s'avèrent très intéressés par l'information présente dans cette dimension au travers, entre autres, de la loi physique des corps noirs qui prédit le profil spectral d'objets à une température donnée. A contrario alors que nous n'avons que trois points d'échantillonnages il serait mal adapté d'utiliser un modèle en longueur d'onde capable de représenter une information riche suivant cette dimension. A cela se rajoute une difficulté concernant la réponse instrument qui n'est pas invariante en longueur d'onde : la largeur de la figure d'*Airy* change par exemple.
- Dans le cas spatial de nombreux points d'échantillonnage sont disponibles et dans les trois protocoles d'observation. L'écart de pointage entre les différentes acquisitions est même très inférieur à la taille d'un capteur. Donc à l'inverse de la dimension spectrale  $\lambda$ , beaucoup d'informations sur le comportement spatial du ciel sont disponibles dans les données et un modèle adapté à sa représentation est pertinent.

Il existe de nombreux moyens de représenter une fonction à l'aide de coefficients  $\mathbf{x} \in \mathbb{R}^N$ . L'idée est généralement de choisir une représentation adaptée à l'information traitée. On peut mentionner

les développements limités

$$f(\alpha) = \sum_{k=0}^K x_k (\alpha - \alpha_0)^k + o((\alpha - \alpha_0)^K),$$

les décompositions par combinaisons linéaires

$$f(\alpha) = \sum_{i \in I} x_i \psi^i(\alpha)$$

où les fonctions  $\psi^i$  peuvent être placées régulièrement ou non, séparables ou non dans le cas multidimensionnel, *etc.* On peut également citer les ondelettes, les splines ou encore les coefficients de Fourier pour les fonctions périodiques de période  $T$

$$f(\alpha) = \sum_{n=-\infty}^{\infty} x_n e^{2i\pi \frac{n}{T} \alpha}.$$

Pour exprimer la fonction  $\phi$  nous avons choisi d'exploiter deux représentations différentes. Premièrement nous considérons une approximation du ciel en longueur d'onde par un développement à l'ordre  $K$ , pour chaque bande spectrale  $[\lambda_f^m \lambda_f^M]$  de longueur d'onde centrale  $\lambda_0$ ,

$$\phi(\alpha, \beta, \lambda) = \sum_{k=0}^K \phi^k(\alpha, \beta) (\lambda - \lambda_0)^k \quad (2.39)$$

où  $K$  est l'ordre du développement,  $\phi^k$  les coefficients à l'ordre  $k$  et qui dépendent des variables d'espace. Bien qu'il soit possible de développer la fonction jusqu'à un ordre élevé, l'idée est de diminuer la complexité du modèle en se limitant à un ordre petit tout en ayant une description assez juste. Au regard de ce que la physique prédit comme profil spectral au travers de la loi des corps noirs (linéaire sur une certaine plage en log-log), un développement à l'ordre 1 semble adapté si la bande spectrale *est assez étroite*, tout en restant très simple. Autrement dit, le profil spectral du ciel est considéré linéaire par morceaux sur une bande étroite correspondant au bande de SPIRE

$$\phi(\alpha, \beta, \lambda) = \phi^0(\alpha, \beta) + \phi^1(\alpha, \beta)(\lambda - \lambda_0), \forall \lambda \in [\lambda_f^m \lambda_f^M] \quad (2.40)$$

avec une origine ou un niveau moyen  $\phi^0$  sur toute la bande, et une pente  $\phi^1$ . Ces deux fonctions représentent l'information spatiale présente dans le ciel. Elles sont donc naturellement considérées comme appartenant aux fonctions à énergie finie  $L_2$  et à bande limitée.

Deuxièmement la représentation spatiale ( $\phi^0$  et  $\phi^1$ ) est effectuée avec une combinaison linéaire, ou une décomposition sur une famille de fonctions. L'idée est que l'information présente dans les données (et celle disponible par ailleurs) permettront d'obtenir une grande résolution spatiale du ciel (ou du moins supérieure à la résolution du capteur). La décomposition sur une famille de fonctions

$$\phi^k(\alpha, \beta) = \sum_{ij} x^k(i, j) \psi^{ij}(\alpha, \beta) \quad (2.41)$$

est adaptée pour cette représentation avec un choix judicieux sur le nombre de coefficients et des positions spatiales des fonctions  $\psi^{ij}$ . Il faut ensuite faire des choix sur les fonctions  $\psi^{ij}$  : placées régulièrement, appartenant à une famille unique, formant une base de  $L_2$ , séparables, ... ? Fondamentalement ces choix ne sont pas importants puisque chaque décomposition permet une représentation adaptée du ciel. Le choix est surtout d'ordre pratique puisqu'avec un choix judicieux, les calculs peuvent être plus simples et plus rapides.

Nous avons choisi dans la suite des fonctions  $\psi^{ij}$  séparables, placées régulièrement et appartenant à une seule famille : des gaussiennes pour des raisons pratiques de calculs explicites et des sinus cardinaux car ils forment une base de  $L_2$  des fonctions à bande limitée. La prochaine partie présente les calculs explicites qui peuvent être menés avec cette expression pour le ciel.

### 2.4.2 Calculs explicites

Pour calculer une sortie modèle nous avons vu qu'en exploitant l'expression du modèle (2.22) page 50 et du ciel (2.39)–(2.41) il faut calculer pour chaque donnée

$$y^{lm}(nT_e) = \int_t \int_\lambda h_f(\lambda) \iint_{\mathbb{R}^2} \iint_{\mathbb{R}^2} \sum_{k=0}^K \sum_{ij} x^k(i,j) \psi^{ij}(\alpha_t - p_\alpha(t), \beta_t - p_\beta(t)) (\lambda - \lambda_0)^k h_m(\alpha_t - \alpha, \beta_t - \beta, \lambda) d\alpha_t d\beta_t h_c(\alpha - \alpha_{lm}, \beta - \beta_{lm}, \lambda) d\alpha d\beta h_b(t - nT_e) d\lambda dt$$

qui peut se réécrire comme

$$y^{lm}(nT_e) = \sum_{k=0}^K \sum_{ij} x^k(i,j) \int_\lambda h_f(\lambda) (\lambda - \lambda_0)^k \int_t \iiint \psi^{ij}(\alpha_t - p_\alpha(t), \beta_t - p_\beta(t)) h_m(\alpha_t - \alpha, \beta_t - \beta, \lambda) d\alpha_t d\beta_t h_c(\alpha - \alpha_{lm}, \beta - \beta_{lm}, \lambda) d\alpha d\beta h_b(t - nT_e) dt d\lambda. \quad (2.42)$$

Donc pour obtenir la contribution d'un coefficient sur une donnée il faut calculer six intégrales. Soit une expression explicite est disponible soit l'intégrale doit être discrétisée (ce qui sur dix points pour chaque intégrale revient à faire un million d'évaluations de la fonction à intégrer). Étant donné les choix effectués sur les modèles physiques et le ciel, les deux cas sont présents. La suite décrit comment expliciter au maximum ces intégrales pour éviter la discrétisation, lourde en calculs numériques.

#### Approximation de la figure d'*Airy*

Pour expliciter les intégrales, une difficulté se situe au niveau de la figure d'*Airy* (2.7). Cette réponse fait intervenir une fonction de Bessel de première espèce du premier ordre qui s'avère difficile à manipuler en général. Il n'est pas possible ici d'expliciter une intégrale faisant intervenir cette fonction de Bessel. Pour lever cette difficulté la figure d'*Airy* est approchée par une gaussienne à deux dimensions

$$h_m(\alpha, \beta, \lambda) \approx \frac{1}{2\pi\sigma_m^2} \exp -\frac{1}{2} \frac{\alpha^2 + \beta^2}{\sigma_m^2} = \bar{h}(\alpha, \beta, \lambda). \quad (2.43)$$

Cette approximation est illustrée figure 2.12. La largeur de la figure d'*Airy* varie linéairement avec la longueur d'onde, voir équation (2.4). Pour obtenir une expression de la variance en fonction de la longueur d'onde, la largeur peut être estimée avec une minimisation de

$$\hat{\sigma}_m = \arg \min_{\sigma_m} \|h_m - \bar{h}(\sigma_m)\|^2$$

par simple recherche exhaustive sous contrainte que les deux fonctions  $h_m$  et  $\bar{h}$  aient le même maximum. Comme la largeur de la figure d'*Airy* varie linéairement avec la longueur d'onde, on déduit que dans le cas d'Herschel

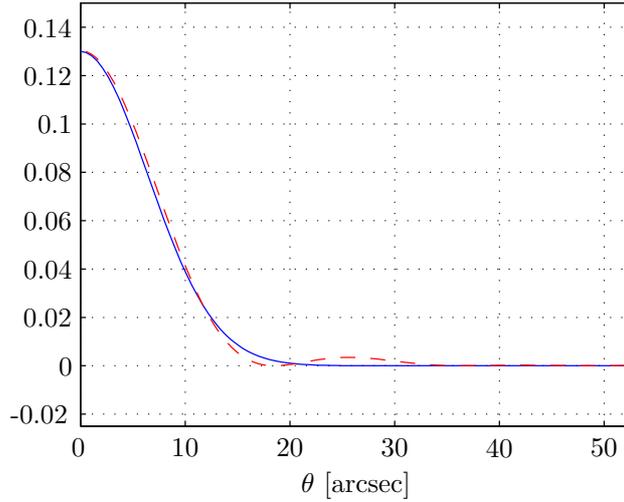
$$\hat{\sigma}_m \approx \frac{\lambda}{8} \text{ radians}$$

fournit une bonne approximation (avec  $\lambda$  en mètre). L'erreur relative introduite est de

$$\int (h_m - \bar{h}(\hat{\sigma}_m))^2 d\alpha d\beta / \int h_m^2 d\alpha d\beta = 0,25\%, \text{ et}$$

$$\int |h_m - \bar{h}(\hat{\sigma}_m)| d\alpha d\beta / \int |h_m| d\alpha d\beta = 8,19\%.$$

Nous avons ainsi une approximation gaussienne de la réponse du miroir avec une expression de l'écart-type fonction linéaire de la longueur d'onde.



**Figure 2.12** – Approximation de la figure d’Airy par une gaussienne. La courbe en rouge pointillé correspond à la figure d’Airy pour une longueur d’onde  $\lambda = 250\mu\text{m}$  avec la présence d’un miroir secondaire. L’écart type de la gaussienne en bleu vaut  $\sigma_m \approx \lambda/8$ .

### Réponse complète de l’optique

Avec cette approximation il apparait une convolution spatiale à deux dimensions de deux gaussiennes, celle du miroir et celle du cornet

$$h_{\text{mc}}(\alpha_t - \alpha_{lm}, \beta_t - \beta_{lm}, \lambda) = \iint h_{\text{m}}(\alpha_t - \alpha, \beta_t - \beta, \lambda) h_{\text{c}}(\alpha - \alpha_{lm}, \beta - \beta_{lm}, \lambda) d\alpha d\beta.$$

Le résultat, décrit dans l’annexe C.1 équation (C.4) vaut donc

$$h_{\text{mc}}(\alpha_t - \alpha_{lm}, \beta_t - \beta_{lm}, \lambda) = \frac{1}{2\pi\sigma_{\text{mc}}^2} \exp\left[-\frac{1}{2} \frac{(\alpha_t - \alpha_{lm})^2 + (\beta_t - \beta_{lm})^2}{\sigma_{\text{mc}}^2}\right]$$

avec  $\sigma_{\text{mc}}^2 = \sigma_m^2 + \sigma_c^2$ , et dans ce cas le signal à l’entrée des filtres fait intervenir l’expression

$$\phi_c(\alpha_{lm}, \beta_{lm}, \lambda, t) = \iint \psi^{ij}(\alpha_t - p_\alpha(t), \beta_t - p_\beta(t)) h_{\text{mc}}(\alpha_t - \alpha_{lm}, \beta_t - \beta_{lm}, \lambda) d\alpha_t d\beta_t \quad (2.44)$$

d’une convolution de la fonction de décomposition spatiale et de la réponse globale de l’optique (miroir et cornet). Comme mentionné précédemment les fonctions  $\psi$  sont des gaussiennes ou des sinus cardinaux. Dans les deux cas les deux intégrales précédentes sont explicites. Cependant le cas avec des sinus cardinaux sera plus complexe et ne permettra pas d’expliciter les intégrales suivantes dans tous les cas.

En ce qui concerne les autres éléments de la chaîne d’acquisition, il n’y a pas de problème particulier. Nous verrons d’ailleurs dans certains cas que nous pourrons combiner les réponses optiques et la réponse du bolomètre en une réponse globale explicite. Deux calculs sont présentés avec une décomposition sur des gaussiennes et une décomposition sur des sinus cardinaux.

### Décomposition sur des gaussiennes

Dans ce cas le ciel est décomposé dans l’espace sur des gaussiennes séparables

$$\psi^{ij}(\alpha, \beta) = A \exp\left[-\frac{1}{2} \frac{(\alpha - \alpha^{ij})^2}{\sigma_\alpha^2}\right] \exp\left[-\frac{1}{2} \frac{(\beta - \beta^{ij})^2}{\sigma_\beta^2}\right]$$

où  $(\alpha^{ij}, \beta^{ij})$  sont les positions des gaussiennes et  $A = 1$  ou  $1/2\pi\sigma_\beta\sigma_\alpha$  si l'on choisit des gaussiennes normalisées ou non. Dans la suite les calculs sont faits avec des gaussiennes normalisées mais on peut souhaiter utiliser des gaussiennes non normalisées pour approcher au mieux un sinus cardinal. Dans ce cas, la convolution entre  $\psi$  et la réponse de l'optique (2.44) est

$$\begin{aligned}\phi_c^{ij}(\alpha_{lm}, \beta_{lm}, \lambda, t) &= \iint \psi^{ij}(\alpha_t - p_\alpha(t), \beta_t - p_\beta(t)) h_{mc}(\alpha_t - \alpha_{lm}, \beta_t - \beta_{lm}, \lambda) d\alpha_t d\beta_t \\ &= \frac{1}{2\pi\Sigma_\alpha\Sigma_\beta} \exp\left[-\frac{1}{2}\frac{(p_\alpha(t) + \alpha^{ij} - \alpha_{lm})^2}{\Sigma_\alpha^2}\right] \exp\left[-\frac{1}{2}\frac{(p_\beta(t) + \beta^{ij} - \beta_{lm})^2}{\Sigma_\beta^2}\right]\end{aligned}$$

avec

$$\Sigma_\alpha^2 = \sigma_\alpha^2 + \sigma_m^2 + \sigma_c^2 \quad \text{et} \quad \Sigma_\beta^2 = \sigma_\beta^2 + \sigma_m^2 + \sigma_c^2.$$

Finalement toutes les convolutions spatiales sont explicites. La sortie modèle dans le cas où tout est gaussien s'écrit

$$\begin{aligned}y^{lm}(nT_e) &= \frac{1}{2\pi} \sum_{k=0}^K \sum_{ij} x^k(i, j) \int_\lambda h_f(\lambda) \frac{1}{\Sigma_\alpha\Sigma_\beta} (\lambda - \lambda_0)^k \\ &\quad \int_t \exp\left[-\frac{1}{2}\frac{(p_\alpha(t) + \alpha^{ij} - \alpha_{lm})^2}{\Sigma_\alpha^2}\right] \exp\left[-\frac{1}{2}\frac{(p_\beta(t) + \beta^{ij} - \beta_{lm})^2}{\Sigma_\beta^2}\right] h_b(t - nT_e) dt d\lambda. \quad (2.45)\end{aligned}$$

qui fait intervenir la convolution temporelle d'une gaussienne à deux dimensions, qui dépend du temps au travers du pointage, et de la réponse des bolomètres.

### Décomposition sur des sinus cardinaux

Dans cette partie le ciel est décomposé sur des sinus cardinaux séparables

$$\psi^{ij}(\alpha, \beta) = T_\alpha^{ij} T_\beta^{ij} \text{sinc}_\pi(T_\alpha^{ij}(\alpha - \alpha^{ij})) \text{sinc}_\pi(T_\beta^{ij}(\beta - \beta^{ij}))$$

où  $(\alpha^{ij}, \beta^{ij})$  sont les positions des sinus cardinaux et  $\text{sinc}_\pi(x) = \sin(\pi x)/\pi x$ . Le choix des sinus cardinaux est motivé par le fait qu'ils forment une base pour les signaux de  $L_2$  à bande limitée avec tous les avantages que cela apporte pour l'exploration de cet espace. Étant donné que nous sommes dans une problématique de sur-résolution, dans le sens où nous nous posons la question de la restauration d'une image à une fréquence spatiale plus élevée que la résolution naturelle du capteur, on sera naturellement amené à considérer une fréquence limite pour le ciel  $\phi$ . Le choix de cette fréquence limite posera naturellement la base correspondante avec l'espacement et la largeur des sinus cardinaux.

Avec ce choix il faut calculer l'intégrale

$$\begin{aligned}\phi_f^{ij}(\alpha_{lm}, \beta_{lm}, \lambda) &= \iint T_\alpha^{ij} T_\beta^{ij} \text{sinc}_\pi(T_\alpha^{ij}(\alpha_t - p_\alpha(t) - \alpha^{ij})) \\ &\quad \text{sinc}_\pi(T_\beta^{ij}(\beta_t - p_\beta(t) - \beta^{ij})) h_{mc}(\alpha_t - \alpha_{lm}, \beta_t - \beta_{lm}, \lambda) d\alpha_t d\beta_t\end{aligned}$$

qui fait intervenir deux convolutions d'un sinus cardinal et d'une gaussienne mono-variée (puisque les fonctions de décompositions gaussiennes sont séparables). Le résultat de cette intégrale est explicite et est décrit dans l'annexe C.2. La solution fait intervenir une fonction spéciale  $\Phi$  qui est la fonction hypergéométrique dégénérée voir [GR80] page 1058 et s'écrit, avec  $\Im$  la partie imaginaire,

$$\begin{aligned}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma_{mc}} \exp\left[-\frac{1}{2}\frac{(\alpha_t - \alpha_{lm})^2}{\sigma_{mc}^2}\right] \text{sinc}(\pi T_\alpha(\alpha_t - c)) d\alpha_t = \\ \frac{2}{\sigma_{mc}} \exp\left[-\frac{1}{2}\frac{(c - m)^2}{\sigma_{mc}^2}\right] \Im\left[\left(\frac{c - \alpha_{lm}}{\sigma_{mc}} + i\pi T_\alpha \sigma_{mc}\right) \Phi\left(\frac{1}{2}, \frac{3}{2}, \frac{1}{2}\left(\frac{c - \alpha_{lm}}{\sigma_{mc}} + i\pi T_\alpha \sigma_{mc}\right)^2\right)\right]\end{aligned}$$

noté  $\text{esinc}_{T_\alpha, \sigma_{mc}}(c - \alpha_{lm})$  pour simplifier. En reportant cette expression on obtient

$$\phi_c^{ij}(\alpha_{lm}, \beta_{lm}, \lambda, t) = T_\alpha^{ij} T_\beta^{ij} \text{esinc}_{T_\alpha^{ij}, \sigma_{mc}}(p_\alpha(t) - \alpha^{ij} - \alpha_{lm}) \text{esinc}_{T_\beta^{ij}, \sigma_{mc}}(p_\beta(t) - \beta^{ij} - \beta_{lm})$$

qui est l'expression explicite du signal à la sortie des cornets (donc avec la décomposition du ciel, la réponse du miroir et du cornet) et la sortie du modèle instrument s'écrit

$$y^{lm}(nT_e) = \sum_{k=0}^K \sum_{ij} x^k(i, j) T_\alpha^{ij} T_\beta^{ij} \int_{\lambda} h_f(\lambda) (\lambda - \lambda_0)^k \int_t \text{esinc}_{T_\alpha^{ij}, \sigma_{mc}}(p_\alpha(t) - \alpha^{ij} - \alpha_{lm}) \text{esinc}_{T_\beta^{ij}, \sigma_{mc}}(p_\beta(t) - \beta^{ij} - \beta_{lm}) h_b(t - nT_e) dt d\lambda. \quad (2.46)$$

qui fait intervenir la convolution de l'expression obtenue précédemment et de la réponse du bolomètre.

Des calculs explicites supplémentaires feront intervenir le temps  $t$  et donc le protocole d'observation. Dans ce cas certains calculs seront possibles et d'autres non. Plus précisément c'est le cas *Large map* avec des sinus cardinaux qui posera problème.

### 2.4.3 Cas du protocole *Large map*

Deux cas sont à envisager : celui où le ciel est décomposé sur des gaussiennes et celui où il est décomposé sur des sinus cardinaux. Le cas avec des sinus cardinaux est impossible à calculer puisque la fonction  $\text{esinc}$  est trop complexe pour être manipulée. Dans le cas du protocole *Large map* il y a des phases d'accélération et des phases à vitesse constante. Les phases à vitesse constante sont les phases d'intérêt pour les astronomes. De plus, les équations des phases d'accélérations sont trop complexes pour être calculées analytiquement. Donc seulement les phases à vitesses constantes seront analysées.

Pour calculer l'intégrale suivant le temps  $t$ , noté  $f$ , de (2.45) il faut remplacer les expressions  $p_\alpha(t)$  et  $p_\beta(t)$  voir (2.21) page 49, pour un balayage à vitesse constante soit

$$f(nT_e) = \frac{1}{2\pi} \frac{1}{\Sigma_\alpha \Sigma_\beta} \int_t \exp \left[ -\frac{1}{2} \frac{(v_\alpha t + c_\alpha + \alpha^{ij} - \alpha_{lm})^2}{\Sigma_\alpha^2} \right] \exp \left[ -\frac{1}{2} \frac{(v_\beta t + c_\beta + \beta^{ij} - \beta_{lm})^2}{\Sigma_\beta^2} \right] h_b(t - nT_e) dt$$

avec

$$h_b(t - nT_e) = \mathbf{1}_{[0, +\infty[}(t - nT_e) S' \exp \left[ -\frac{t - nT_e}{\tau_e} \right].$$

Dans la phase d'accélération le calcul est plus complexe voir impossible puisqu'il fait intervenir un facteur  $t$  à la puissance 4. Ensuite dans le calcul avec des sinus cardinaux il faudrait remplacer les exponentielles par des fonctions  $\text{esinc}$ , ce qui rend le calcul impossible. Dans le cas ici de gaussiennes avec un balayage à vitesse constante le calcul est

$$f(nT_e) = \frac{1}{2\pi} \frac{S'}{\Sigma_\alpha \Sigma_\beta} \exp \left[ \frac{nT_e}{\tau_e} \right] \int_{nT_e}^{+\infty} \exp \left[ -\frac{1}{2} \frac{(v_\alpha t + o_\alpha)^2}{\Sigma_\alpha^2} \right] \exp \left[ -\frac{1}{2} \frac{(v_\beta t + o_\beta)^2}{\Sigma_\beta^2} \right] \exp \left[ \frac{-t}{\tau_e} \right] dt \quad (2.47)$$

avec

$$o_\alpha = c_\alpha + \alpha^{ij} - \alpha_{lm} \quad \text{et} \quad o_\beta = c_\beta + \beta^{ij} - \beta_{lm}.$$

Il s'agit de l'intégration partielle d'une gaussienne puisqu'il s'agit d'une exponentielle dont l'argument a une forme quadratique en  $t$  tronqué par une indicatrice. Ce calcul peut être mené avec la fonction erreur erf. Pour cela il faut faire apparaître la forme quadratique en  $t$  dans l'exponentielle pour l'intégration. Les calculs sont présentés dans l'annexe B.2. L'expression obtenue est

$$\begin{aligned}
 f(nT_e) &= \frac{1}{2\pi} S' \frac{\sqrt{\pi\tau_e\Sigma}}{\sqrt{2}} \exp \left[ \frac{nT_e}{\tau_e} - \frac{1}{2} \frac{b-a^2}{\Sigma^2\Sigma_\alpha^2\Sigma_\beta^2\tau_e} \right] \left( 1 - \operatorname{erf} \left( \frac{nT_e + a}{\sqrt{2\tau_e\Sigma\Sigma_\alpha\Sigma_\beta}} \right) \right) \\
 &= \frac{S'}{2\sqrt{2\pi}\Sigma_v} \exp \left[ -\frac{(o_\alpha + nT_e v_\alpha)^2}{2\Sigma_\alpha^2} - \frac{(o_\beta + nT_e v_\beta)^2}{2\Sigma_\beta^2} \right] \\
 &\quad \operatorname{erfcx} \left( \frac{\Sigma_\beta v_\alpha (o_\alpha + nT_e v_\alpha)}{\sqrt{2}\Sigma_\alpha\Sigma_v} + \frac{\Sigma_\alpha v_\beta (o_\beta + nT_e v_\beta)}{\sqrt{2}\Sigma_\beta\Sigma_v} + \frac{\Sigma_\alpha\Sigma_\beta}{\sqrt{2\tau_e\Sigma_v}} \right) \quad (2.48)
 \end{aligned}$$

avec

$$\begin{aligned}
 \Sigma_{\alpha,\beta}^2 &= \sigma_{\alpha,\beta}^2 + \sigma_m^2 + \sigma_c^2 & \text{et} & \quad \Sigma_v = \sqrt{\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2}, \\
 o_\alpha &= c_\alpha + \alpha^{ij} - \alpha_{lm} & \text{et} & \quad o_\beta = c_\beta + \beta^{ij} - \beta_{lm}.
 \end{aligned}$$

Cette équation (2.48) permet de connaître la contribution d'un coefficient  $(i,j)$  du ciel, à une longueur d'onde fixée, sur le bolomètre  $(l,m)$  à l'instant  $n$ . Pour connaître toutes les contributions sur toutes les données (à une longueur d'onde fixée), il suffit d'évaluer cette fonction aux bonnes valeurs pour tous les  $(i,j,l,m,n)$ .

Cette fonction est indépendante du temps puisqu'il est intégré, sauf au travers de l'indice  $nT_e$ . Cependant cet indice est multiplié par une vitesse pour donner une distance angulaire. Donc cette fonction s'exprime uniquement en fonction de la position spatiale. De plus la partie gaussienne correspond à une gaussienne centrée sur une position spatiale combinant tous les facteurs de décalage, y compris le décalage régulier dans le temps

$$o_\alpha + nT_e v_\alpha = c_\alpha + \alpha^{ij} - \alpha_{lm} + nT_e v_\alpha$$

où  $c_\alpha$  est la position au départ du balayage,  $\alpha^{ij}$  est la position de la gaussienne de décomposition,  $\alpha_{lm}$  la position du cornet  $lm$  et  $nT_e v_\alpha$  la position pointée à l'instant  $n$ , respectivement pour  $\beta$ . Nous aurions obtenu la même chose si la réponse du bolomètre avait été modélisée par un Dirac en  $nT_e$ , puisque dans ce cas nous aurions procédé à un simple décalage et le terme  $\operatorname{erfcx}$  ne serait pas apparu.

Le seul facteur où la constante de temps du bolomètre  $\tau_e$  intervient est dans le terme

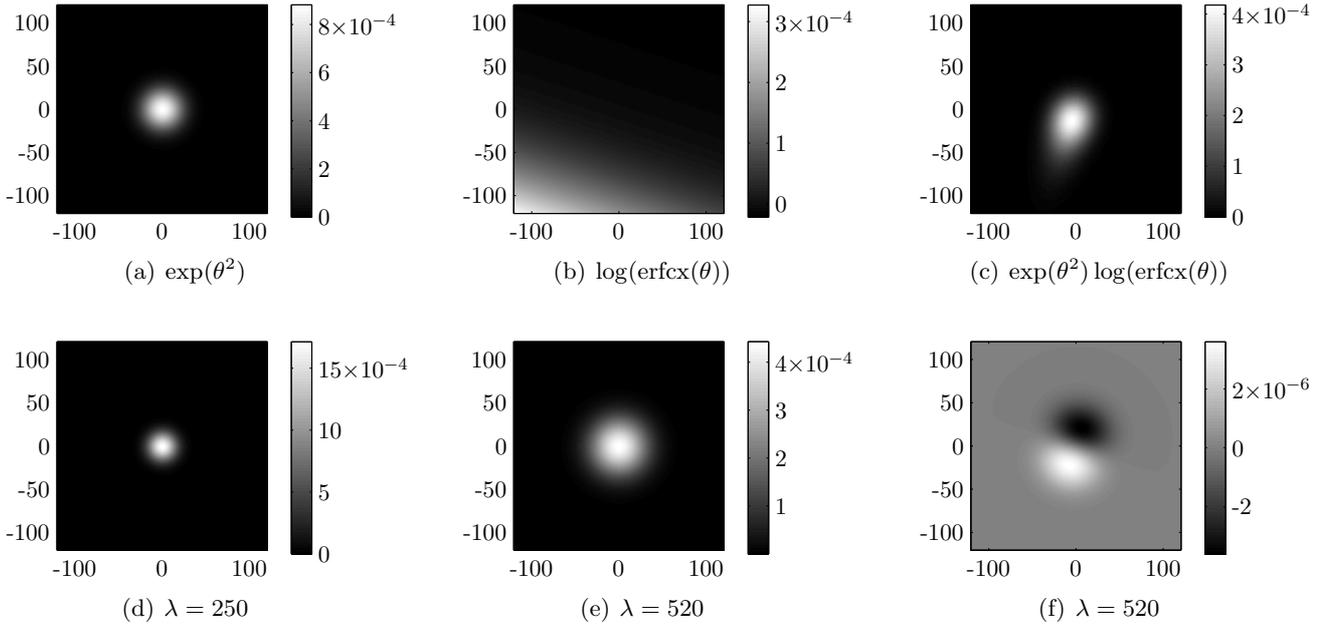
$$\frac{\Sigma_\alpha\Sigma_\beta}{\sqrt{2\tau_e\Sigma_v}}$$

de la fonction  $\operatorname{erfcx}$ . C'est donc la fonction  $\operatorname{erfcx}$  qui permet de prendre en compte l'influence du bolomètre sur le contenu *spatial* du signal. Pour illustrer cela, la figure 2.13 représente la fonction

$$f(\alpha,\beta) = \frac{S'}{2\sqrt{2\pi}\Sigma_v} \exp \left[ -\frac{\alpha^2}{2\Sigma_\alpha^2} - \frac{\beta^2}{2\Sigma_\beta^2} \right] \operatorname{erfcx} \left( \frac{\Sigma_\beta v_\alpha \alpha}{\sqrt{2}\Sigma_\alpha\Sigma_v} + \frac{\Sigma_\alpha v_\beta \beta}{\sqrt{2}\Sigma_\beta\Sigma_v} + \frac{\Sigma_\alpha\Sigma_\beta}{\sqrt{2\tau_e\Sigma_v}} \right) \quad (2.49)$$

avec comme paramètre  $v_\alpha = 15 \text{ arcsec.s}^{-1}$ ,  $v_\beta = 5 \text{ arcsec.s}^{-1}$ . Les écart-types  $\Sigma_\alpha = \Sigma_\beta$  correspondent à la longueur d'onde  $\lambda = 250 \text{ }\mu\text{m}$  avec des gaussiennes de décomposition de largeur  $\sigma = 1 \text{ arcsec}$ . Pour bien illustrer l'influence du bolomètre la constante de temps a été choisie arbitrairement pour faire apparaître la traînée. La figure 2.13(a) représente la partie exponentielle de la fonction, celle indépendante du bolomètre qui vient uniquement de l'optique et du ciel.

La figure 2.13(b) représente seulement le logarithme de la fonction  $\operatorname{erfcx}$  pour faire apparaître sa structure. Malgré son expression à deux dimensions, elle ne varie que dans la direction indiquée par la vitesse  $(v_\alpha, v_\beta)$ . Elle croît de façon monotone très rapidement à l'opposé du sens indiqué par la vitesse (c'est la causalité du filtre passe-bas). C'est cette croissance très rapide qui compense la décroissance de la gaussienne et fait apparaître une traînée. Cette traînée est illustrée sur la figure 2.13(c) qui correspond à la fonction complète qui prend la forme d'une goutte.



**Figure 2.13** – Combinaison entre la réponse de l'optique et des bolomètres dans le cas d'une décomposition gaussienne pour le ciel et du protocole *Large map* à vitesse constante. La figure 2.13(a) correspond à la partie exponentielle de la fonction (2.49). La figure 2.13(b) correspond à la partie  $\operatorname{erfcx}$  qui provoque la traînée. La figure 2.13(c) correspond à la fonction complète. Les figures 2.13(d) et 2.13(e) représentent la fonction pour les deux longueurs d'onde 250 et 520  $\mu\text{m}$  avec les paramètres instruments réels. La figure 2.13(f) représente la différence entre la réponse à 520  $\mu\text{m}$  où le bolomètre est négligé et la réponse avec la prise en compte du bolomètre.

Pour voir la traînée nous avons considéré une constante de temps arbitrairement grande. Dans le cas de SPIRE celle-ci est bien plus faible (voir annexe D) et dans ce cas on observe les figures 2.13(d) et 2.13(e). La traînée est moins visible. La figure 2.13(f) est la différence entre deux réponses : la réponse avec le bolomètre et la réponse de l'optique uniquement (comme si le bolomètre avait été négligé). La différence est de l'ordre de deux ordres de grandeurs.

Il serait tentant de négliger la réponse temporelle des bolomètres. C'est une option possible qu'il faut mettre en regard de la souplesse plus grande d'un modèle qui prend en compte cette réponse. En effet, la constante de temps peut s'avérer être plus importante que prévu. De plus, certaines particules à haute énergie présentes dans l'espace sont absorbées par les bolomètres sans être influencées par les autres éléments de l'instrument comme l'optique. Ces particules ne sont donc vues *que* par les bolomètres et laissent une traînée bien visible dans les données si leur énergie est importante.

#### 2.4.4 Cas des protocoles *Point source* et *Small map*

Tout comme pour le cas *Large map* deux cas sont à envisager : celui où le ciel est décomposé sur des gaussiennes et celui où il est décomposé sur des sinus cardinaux. Dans les deux cas le calcul est possible car la variable  $t$  n'intervient que dans l'expression de la réponse du bolomètre. En effet nous avons

$$f(nT_e) = \int_t g(p_\alpha(t) + \alpha^{ij} - \alpha_{lm}) g(p_\beta(t) + \beta^{ij} - \beta_{lm}) h_b(t - nT_e) dt$$

où  $g$  est soit une fonction gaussienne soit une fonction esinc. Dans le cas des protocoles *Point source* et *Small map* les pointages permettent des calculs plus simples mais la causalité du filtre bolomètre implique des conditions particulières. Les calculs sont présentés dans l'annexe B.3. Si le ciel est décomposé sur des gaussiennes alors nous avons

$$f(nT_e) = \frac{\tau_e S'}{2\pi \Sigma_\alpha \Sigma_\beta} \sum_q \exp \left[ -\frac{1}{2} \frac{\alpha_q + \alpha^{ij} - \alpha_{lm}}{\Sigma_\alpha^2} \right] \exp \left[ -\frac{1}{2} \frac{\beta_q + \beta^{ij} - \beta_{lm}}{\Sigma_\beta^2} \right] \left( \exp \left[ \frac{nT_e - \max(t_q, nT_e)}{\tau_e} \right] - \exp \left[ \frac{nT_e - t_{q+1}}{\tau_e} \right] \right) \quad (2.50)$$

et dans le cas des sinus cardinaux

$$f(nT_e) = \tau_e \sum_q \text{esinc}_{T_\alpha^{ij}, \sigma_{mc}}(\alpha_q - \alpha^{ij} - \alpha_{lm}) \text{esinc}_{T_\beta^{ij}, \sigma_{mc}}(\beta_q - \beta^{ij} - \beta_{lm}) \left( \exp \left[ \frac{nT_e - \max(t_q, nT_e)}{\tau_e} \right] - \exp \left[ \frac{nT_e - t_{q+1}}{\tau_e} \right] \right). \quad (2.51)$$

Finalement, dans le cas des protocoles *Small map* et *Point source* l'expression de la convolution temporelle de la réponse du bolomètre, avec des gaussiennes et des sinus cardinaux est explicite.

#### 2.4.5 Expression complète des modèles

Nous récapitulons ici les différentes équations permettant de calculer les sorties modèle en fonction du protocole d'acquisition pour une bande spectrale donnée. Dans le cas *Large map*, pour un balayage à vitesse constante et avec une décomposition sur des gaussiennes le modèle s'écrit

$$y^{lm}(nT_e) = \frac{S'}{2\sqrt{2\pi}\Sigma_v} \sum_{k=0}^K \sum_{ij} x^k(i,j) \int_\lambda h_f(\lambda) (\lambda - \lambda_0)^k \exp \left[ -\frac{(o_\alpha + nT_e v_\alpha)^2}{2\Sigma_\alpha^2} - \frac{(o_\beta + nT_e v_\beta)^2}{2\Sigma_\beta^2} \right] \text{erfcx} \left( \frac{\Sigma_\beta v_\alpha (o_\alpha + nT_e v_\alpha)}{\sqrt{2}\Sigma_\alpha \Sigma_v} + \frac{\Sigma_\alpha v_\beta (o_\beta + nT_e v_\beta)}{\sqrt{2}\Sigma_\beta \Sigma_v} + \frac{\Sigma_\alpha \Sigma_\beta}{\sqrt{2}\tau_e \Sigma_v} \right) d\lambda \quad (2.52)$$

avec

$$\Sigma_{\alpha,\beta}^2 = \sigma_{\alpha,\beta}^2 + \sigma_m^2(\lambda) + \sigma_c^2(\lambda), \quad \Sigma_v = \sqrt{\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2}$$

et

$$o_\alpha = c_\alpha + \alpha^{ij} - \alpha_{lm}, \quad o_\beta = c_\beta + \beta^{ij} - \beta_{lm}.$$

Rien ne peut être extrait de l'intégrale à cause de la dépendance en longueurs d'onde des réponses.

Dans les cas des protocoles *Small map* et *Point source* le modèle s'écrit

$$y^{lm}(nT_e) = \frac{\tau_e S'}{2\pi \Sigma_\alpha \Sigma_\beta} \sum_{k=0}^K \sum_{ij} \sum_q x^k(i,j) \left( \exp \left[ \frac{nT_e - \max(t_q, nT_e)}{\tau_e} \right] - \exp \left[ \frac{nT_e - t_{q+1}}{\tau_e} \right] \right) \int_{\lambda} h_f(\lambda) (\lambda - \lambda_0)^k \exp \left[ -\frac{1}{2} \frac{(\alpha_q + \alpha^{ij} - \alpha_{lm})^2}{\Sigma_\alpha(\lambda)^2} \right] \exp \left[ -\frac{1}{2} \frac{(\beta_q + \beta^{ij} - \beta_{lm})^2}{\Sigma_\beta(\lambda)^2} \right] d\lambda \quad (2.53)$$

si le ciel est décomposé sur des gaussiennes et

$$y^{lm}(nT_e) = \tau_e K' T_\alpha^{ij} T_\beta^{ij} \sum_{k=0}^K \sum_{ij} \sum_q c^k(i,j) \left( \exp \left[ \frac{nT_e - \max(t_q, nT_e)}{\tau_e} \right] - \exp \left[ \frac{nT_e - t_{q+1}}{\tau_e} \right] \right) \int_{\lambda} h_f(\lambda) (\lambda - \lambda_0)^k \text{esinc}_{T_\alpha^{ij}, \sigma_{mc}}(\alpha_q - \alpha^{ij} - \alpha_{lm}) \text{esinc}_{T_\beta^{ij}, \sigma_{mc}}(\beta_q - \beta^{ij} - \beta_{lm}) d\lambda \quad (2.54)$$

si le ciel est décomposé sur des sinus cardinaux.

Dans les trois cas l'intégration explicite suivant  $\lambda$  n'est pas possible. Cependant si l'on se limite à l'ordre 1 pour le développement limité en  $\lambda$ , alors nous n'avons que deux intégrales à discrétiser, alors que nous en avons douze précédemment : six pour l'ordre zéro et six pour l'ordre un. L'annexe B.4 rassemble d'autres équations pour d'autres modèles.

## 2.5 Conclusion

Les trois équations (2.52), (2.53) et (2.54) modélisent l'acquisition d'une donnée à l'instant  $n$  d'un bolomètre  $(l,m)$  obtenue à partir de tous les coefficients  $x^k(i,j)$  définissant le ciel. Dans tous les cas on peut écrire ces équations comme

$$y^{lm}(nT_e) = \sum_{k=0}^{K-1} \sum_{ij} x^k(i,j) \mathcal{H}(i,j,k,l,m,n,q). \quad (2.55)$$

Les équations (2.52), (2.53) et (2.54) permettent de calculer  $\mathcal{H}$  pour tous les  $i,j,k,l,m$  et  $n$  possibles en fonction des hypothèses choisies qui sont les protocoles ou les fonctions de décomposition par exemple.

Comme nous avons des modèles linéaires, l'équation (2.55) ne fait intervenir que des sommes et des produits. Cette somme de produits terme à terme se met sous la forme d'un produit entre deux vecteurs

$$y^{lm}(nT_e) = \mathbf{h}^t \mathbf{x} \quad (2.56)$$

où  $\mathbf{h}$  est un vecteur colonne pour les indices courant  $l,m,n$  de la donnée et qui contient les valeurs de  $\mathcal{H}$  en faisant varier  $i,j,k$ ,  $\mathbf{h}^t$  son transposé et  $\mathbf{x}$  le vecteur colonne des coefficients. L'ensemble des données s'écrit comme

$$\mathbf{y} = \mathbf{H} \mathbf{x} \quad (2.57)$$

où  $\mathbf{H}$  est une matrice de  $LMN$  lignes et  $IJK$  colonnes et dont chaque ligne est obtenue à partir des vecteur  $\mathbf{h}$  en faisant varier les indices  $l,m$  et  $n$ . On obtient donc le très classique modèle linéaire. Ce formalisme permet de se dégager des contraintes spécifiques au modèle ou des contraintes de mise en œuvre sur machines de calculs pour se concentrer sur le formalisme d'inversion présenté dans le prochain chapitre.

En conclusion les travaux menés ont permis de bâtir un modèle instrument construit sur la physique. Les réponses du miroir, des cornets et des bolomètres ont été prises en compte. De

plus le modèle repose sur une description continue du ciel dans les trois dimensions, y compris en longueur d'onde. Par rapport aux modèles présents dans SANEPIC ou MADMAP, le modèle présenté ici est plus fidèle à la réalité. Généralement les méthodes précédentes ne prennent en compte que l'information de pointage, le ciel étant souvent considéré implicitement monochromatique.

Par rapport au simulateur de Bruce Sibthorpe [SCG09], il existe plusieurs points de divergence concernant les modèles physiques. Tout d'abord le modèle du simulateur ne prend pas en compte la longueur d'onde pour la modélisation de la réponse du miroir. Le ciel est considéré monochromatique. De plus le modèle pour la physique des bolomètres est légèrement différent dans le simulateur puisqu'il utilise un modèle non-linéaire et qu'il néglige la réponse temporelle des bolomètres. Ensuite les objectifs des deux modèles sont très différents. Le simulateur est destiné à reproduire le plus fidèlement possible tous les comportements de l'instrument, y compris les caractéristiques attendues du bruit, de l'électronique ou de l'évolution de la température du cryostat. Il n'est utilisé qu'une seule fois et non pas dans une perspective d'inversion à base de méthodes itératives. Nous verrons dans le chapitre suivant que le modèle développé dans ce chapitre peut être calculé en quelques secondes sous certaines hypothèses.

Nous avons également étudié en détail, à l'aide de ce modèle, les caractéristiques de l'instrument et de ses capacités en terme de résolution spatiale. Cela a permis d'illustrer les capacités de sur (ou super!)-résolution mais également d'obtenir des critères quantitatifs d'analyse. De plus, malgré la complexité de la physique nous avons obtenu des expressions explicites des modèles prenant en compte les réponses ainsi que les protocoles d'observation. C'est une caractéristique importante pour éviter la discrétisation des intégrales et obtenir des calculs aussi rapides que possible.

Le prochain chapitre décrit en détail la méthodologie employée pour estimer un ciel inconnu à partir de données observées. Il décrit en détail comment le modèle établi ici est utilisé dans une vraisemblance.



# Chapitre 3

## Inversion

Ce chapitre présente l’approche développée pour estimer le ciel observé par SPIRE. L’inférence est menée dans un cadre bayésien décrit en introduction. Une part de ce travail, particulièrement la partie 3.3, se consacre essentiellement au protocole *Large map*. Le travail est également concentré sur une décomposition spatiale du ciel sur des gaussiennes. Trois points principaux sont abordés dans ce chapitre.

Tout d’abord on établit les modèles probabilistes utilisés pour construire la loi *a posteriori*. On présente ainsi le modèle pour le ciel permettant l’inversion d’un problème bien-posé malgré le mauvais conditionnement du modèle d’acquisition. On présente également le modèle pour la composante basse fréquence, provenant de la dérive thermique, voir partie 1.1.1 et 2.2.5. Le modèle proposé permet d’exploiter directement les données des bolomètres aveugles ainsi que les informations statistiques *a priori* disponibles. Enfin les lois *a priori* pour les paramètres supplémentaires inconnus, comme les puissances de bruit, sont posées, permettant une estimation dans un cadre entièrement bayésien et surtout un réglage automatique du compromis entre les différentes sources d’information.

Ensuite on présente l’algorithme permettant de calculer l’estimateur de la moyenne *a posteriori* pour toutes les inconnues conjointement. On présente la mise en œuvre d’un algorithme MCMC et on étudie en particulier l’échantillonnage d’une loi gaussienne corrélée possédant une matrice de covariance de grande taille à la structure complexe.

Enfin, les algorithmes font appel à un calcul d’une sortie modèle et du transposé dont le temps processeur représente la grande majorité du temps total. On présente comment avec une certaine approximation le modèle peut être écrit en une partie invariante et une partie non invariante, modèle classique en sur-résolution, permettant une mise en œuvre partielle par transformée de Fourier. Enfin on décrira une deuxième approximation permettant d’avoir un modèle entièrement convolutif.

### 3.1 Modèles probabilistes

Cette partie décrit la loi *a posteriori* utilisée pour mener l’inférence. La vraisemblance est définie avec le modèle linéaire  $\mathbf{H}$  décrit dans le chapitre précédent, la composante basse fréquence notée  $\mathbf{c}$  ainsi que le modèle pour le bruit. Les *a priori* seront définis en fonction des informations disponibles ou voulues ainsi que des contraintes de calcul. L’objectif est donc de construire une loi

$$p(\mathbf{x}, \mathbf{c}, \boldsymbol{\theta} | \mathbf{y}) = \frac{p(\mathbf{y} | \mathbf{x}, \mathbf{c}, \boldsymbol{\theta}) p(\mathbf{x}, \mathbf{c}, \boldsymbol{\theta})}{p(\mathbf{y})} \quad (3.1)$$

où  $\mathbf{x}$  est l'objet d'intérêt,  $\mathbf{c}$  la composante basse fréquence,  $\boldsymbol{\theta}$  les paramètres des lois et du modèle instrument, et  $\mathbf{y}$  les données.

Dans la suite du chapitre on note les coefficients de décomposition  $\mathbf{x} \in \mathbb{R}^{I \times J \times K}$  où  $K$  est l'ordre du développement de Taylor (voir 2.4.1) dans la dimension spectrale et  $I, J$  le nombre de coefficients de décomposition. Les coefficients de l'ordre  $k$  sont notés  $\mathbf{x}_k \in \mathbb{R}^{I \times J}$ . Les données sont notées  $\mathbf{y} \in \mathbb{R}^P$  où  $P = (B + 2)N$  est le nombre total de données obtenues à partir de  $B$  bolomètres et des deux bolomètres aveugles à  $N$  instants temporels. Les données produites par un bolomètre en particulier sont notées  $\mathbf{y}_b$  et le modèle d'acquisition correspondant est noté  $\mathbf{H}_b$ .

### 3.1.1 Modélisation de la composante basse fréquence

Une composante basse fréquence  $\mathbf{c}$ , provenant de la dérive thermique et présentée partie 1.1.1 et 2.2.5, s'ajoute aux données des bolomètres. Cette partie présente une modélisation de cette composante. Le chapitre précédent établit le modèle d'acquisition des données comme un modèle linéaire

$$\mathbf{y}_b = \mathbf{H}_b \mathbf{x} + \mathbf{n}_b$$

où  $\mathbf{y}_b$  correspond aux données du bolomètre 'b',  $\mathbf{H}_b$  à la partie du modèle correspondante et  $\mathbf{n}_b$  au bruit affectant ce bolomètre. D'après les études menées [CCB<sup>+</sup>06, PAB<sup>+</sup>08, SBL<sup>+</sup>08] sur cette composante  $\mathbf{n}_b$  il apparaît que ce signal peut être vu comme la somme de deux parties. Une partie corrélée dans le temps commune à tous les bolomètres (ce que les astronomes appellent le bruit en  $1/f$  corrélé). Une autre, également corrélée dans le temps, est indépendante entre bolomètre (appelé bruit en  $1/f$  non-corrélé). L'interprétation physique la plus admise est que la partie commune, dominante, provient des fluctuations en température de l'ensemble de l'instrument (par une instabilité du cryostat par exemple, voir partie 2.2.4) et que la partie indépendante provient majoritairement de l'électronique de lecture [SCG09, SG06], voir partie 2.2.5.

Ces considérations invitent à modéliser ce signal perturbant comme la somme de deux composantes soit

$$\mathbf{n}_b = \mathbf{c} + \mathbf{b}_b$$

où  $\mathbf{c}$  est la composante corrélée dans le temps commune à tous les bolomètres alors que  $\mathbf{b}_b$  dépend du bolomètre. Dans ce cas la formation des données s'écrit

$$\mathbf{y}_b = \mathbf{H}_b \mathbf{x} + \mathbf{c} + \mathbf{b}_b.$$

Les données fournies par les deux bolomètres aveugles  $\mathbf{y}_B$  et  $\mathbf{y}_{B+1}$  s'écrivent quant à elles

$$\begin{aligned} \mathbf{y}_B &= \mathbf{c} + \mathbf{b}_B \\ \mathbf{y}_{B+1} &= \mathbf{c} + \mathbf{b}_{B+1}. \end{aligned}$$

L'ensemble des données s'écrit sous la forme

$$\begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_B \\ \mathbf{y}_{B+1} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_0 & \mathbf{I} \\ \mathbf{H}_1 & \mathbf{I} \\ \vdots & \vdots \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix} + \begin{pmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_B \\ \mathbf{b}_{B+1} \end{pmatrix} \quad (3.2)$$

où  $\mathbf{I}$  est la matrice identité de taille  $N$  par  $N$  et  $\mathbf{0}$  la matrice nulle de taille  $N$  par  $IJK$ . Pour modéliser les données aberrantes observées aux extrémités des balayages, une solution est de considérer un modèle à données manquantes. Le signal  $\mathbf{c}$  n'est pas observé en continu sur les données des

bolomètres non aveugles mais seulement une sous-partie. Cette troncature des données se modélise comme

$$\mathbf{y} = \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix} + \mathbf{b}. \quad (3.3)$$

où  $\mathbf{T}$  est une matrice de 1 ou de 0 indiquant si un échantillon de  $\mathbf{c}$  est observé (conjointement à la sortie  $\mathbf{H}\mathbf{x}$ ). La matrice  $\mathbf{T}^t$  est une matrice de bourrage de zéros. Pour les données des bolomètres non aveugles il n'y a pas de raison *a priori* de négliger certaines données (sauf cas particulier comme les événements impulsionnels tels que les *glitches*), on conserve donc l'identité pour des soucis de simplification de la présentation.

Ce modèle « augmenté » permet de prendre en compte la composante basse fréquence dans la formation des données. Pour établir la vraisemblance il faut en plus établir les informations disponibles sur ces signaux. Tout d'abord le signal  $\mathbf{c}$  ainsi que les signaux  $\mathbf{b}_b$  sont majoritairement basse fréquence et *a priori* indépendants entre eux. On suppose également qu'ils sont de moyenne nulle. Enfin on suppose que tous les signaux  $\mathbf{b}_b$  suivent la même loi à un facteur de dispersion  $\gamma_{b_b}$  près. Sans introduire d'information supplémentaire, alors la loi la moins informative est la loi normale. On suppose donc que les signaux  $\mathbf{c}$  et  $\mathbf{b}_b$  suivent

$$\mathbf{c} \sim \mathcal{N}(0, \Sigma_c) \quad (3.4)$$

$$\mathbf{b}_b \sim \mathcal{N}(0, \gamma_{b_b}^{-1} \Sigma_b). \quad (3.5)$$

Les matrices de covariances modélisent la connaissance que l'on a sur le comportement temporel de ces signaux. S'ils sont supposés *a priori* stationnaires, alors les matrices de covariances sont des matrices de corrélations et en supposant la circularité, elles se diagonalisent dans l'espace de Fourier soit

$$\begin{aligned} \Lambda_c &= \mathbf{F}^\dagger \Sigma_c \mathbf{F} \\ \gamma_{b_b}^{-1} \Lambda_b &= \gamma_{b_b}^{-1} \mathbf{F}^\dagger \Sigma_b \mathbf{F} \end{aligned}$$

où  $\mathbf{F}$  est la matrice de transformée de Fourier unitaire,  $\dagger$  le symbole pour la transposée conjuguée et  $\Lambda$  une matrice diagonale. Les hypothèses de circularité des matrices sont entre autres utilisées dans les méthodes MADMAP [CCB<sup>+</sup>06] et SANEPIC [PAB<sup>+</sup>08] (partie 1.5).

La loi des données  $\mathbf{y}$  connaissant  $\mathbf{x}$  et  $\mathbf{c}$  permet d'établir la vraisemblance. Comme toutes les lois sont gaussiennes et que le modèle est linéaire, alors la loi des données est une loi normale

$$\mathbf{y} | \mathbf{x}, \mathbf{c}, \theta \sim \mathcal{N}(\mathbf{m}_y, \Sigma_y)$$

de moyenne

$$\mathbf{m}_y = \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix}$$

et de matrice de covariance

$$\Sigma_y = \begin{pmatrix} \gamma_{b_0}^{-1} \Sigma_b & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \gamma_{b_1}^{-1} \Sigma_b & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \dots & \dots & \gamma_{b_{B+1}}^{-1} \Sigma_b \end{pmatrix}$$

puisque les signaux  $\mathbf{b}_b$  sont indépendants entre eux. Ce modèle est très proche de celui de MADMAP à ceci près que les bolomètres aveugles et la composante  $\mathbf{c}$  ont été inclus dans le modèle.

Il reste à définir plus précisément la structure des matrices de covariances  $\Sigma_{\mathbf{c}}$  et  $\Sigma_{\mathbf{b}}$ . En ce qui concerne MADMAP, la méthode n'exploite que  $\Sigma_{\mathbf{b}}$  et considère que ce sont des matrices de corrélations circulantes connues. Dans le cas de SANEPIC, les deux matrices sont circulantes et sont estimées [PSCD03] à partir de données de calibration, ou à partir de données acquises en observant des régions du ciel possédant un faible flux incident. L'approche proposée ici repose sur l'exploitation d'une connaissance *a priori* sur les densités spectrales de puissance (DSP) de ces composantes. En effet les recherches [PAB<sup>+</sup>08, SBL<sup>+</sup>08] semblent indiquer que la forme des DSP  $\mathring{\Psi}$  s'exprime en trois parties :

- une moyenne généralement considérée inconnue et qui définit un plateau jusqu'à une certaine fréquence  $f_c$

$$\mathring{\Psi}(f) = \gamma_{\mathbf{c}}, \forall f \in [0, f_c]$$

- une descente linéaire en échelle log–log avec un certain coefficient négatif  $d$  jusqu'à une certaine fréquence  $f'_c$

$$\log(\mathring{\Psi}(f)) = \gamma_{\mathbf{c}} + d \log(f), \forall f \in [f_c, f'_c]$$

- un plateau considéré comme une composante blanche

$$\mathring{\Psi}(f) = \gamma_{\mathbf{b}_b}, \forall f \in [f'_c, +\infty[.$$

Ces informations peuvent être modélisées avec une loi de puissance très employée dans la littérature [CE07, CLBK09]

$$\mathring{\Psi}(f) = \frac{\gamma_{\mathbf{c}} - \gamma_{\mathbf{b}_b}}{1 + (f/f_c)^d} + \gamma_{\mathbf{b}_b}$$

qui se comporte comme  $\gamma_{\mathbf{c}}$  vers 0 et comme  $\gamma_{\mathbf{b}_b}$  vers  $+\infty$  avec une décroissance en puissance  $d$ . La première fréquence de coupure est déterminée par  $f_c$ , la seconde par  $f_c, d$  et  $\gamma_{\mathbf{b}_b}$ . Cette loi de puissance est représentée sur la figure 3.1. Sans composante blanche elle s'écrit

$$\mathring{\Psi}(f) = \frac{\gamma_{\mathbf{c}}}{1 + (f/f_c)^d}. \quad (3.6)$$

En pratique, il s'avère que la composante  $\mathbf{b}_b$  possède une composante basse fréquence très faible par rapport à la composante commune  $\mathbf{c}$  [SBL<sup>+</sup>08]. Étant donnée l'incertitude autour de cette modélisation, et pour faciliter les calculs, il semble raisonnable en première approximation de considérer que la composante indépendante sur chaque bolomètre est blanche et stationnaire soit

$$\Sigma_{\mathbf{b}} = \mathbf{I}.$$

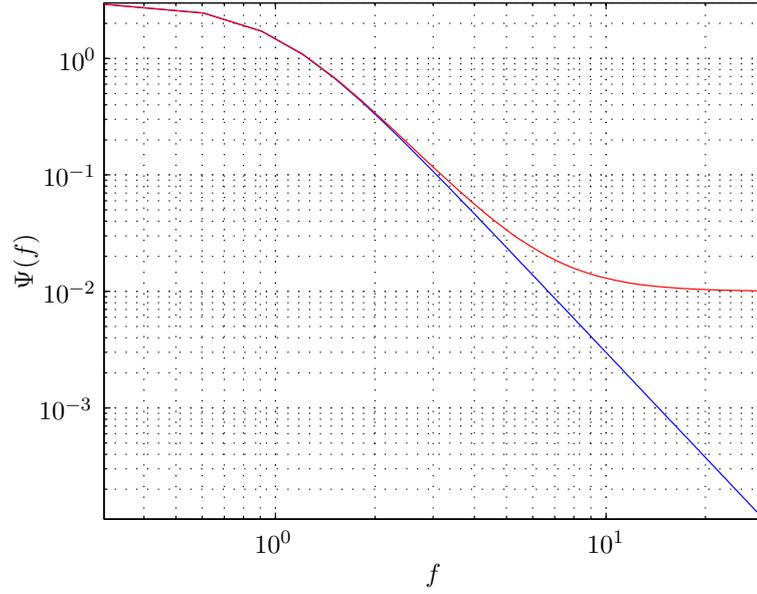
En ce qui concerne la composante  $\mathbf{c}$  nous utiliserons un modèle corrélé stationnaire *a priori* avec pour densité spectrale de puissance celle décrite équation (3.6) sans composante blanche déjà modélisée au travers de  $\mathbf{b}_b$ . La loi *a priori* pour cette composante est

$$p(\mathbf{c} | \gamma_{\mathbf{c}}, f_c, d) = (2\pi)^{-N/2} \gamma_{\mathbf{c}}^{N/2} \det[\Sigma_{\mathbf{c}}]^{-1/2} \exp \left[ -\frac{1}{2} \gamma_{\mathbf{c}} \mathbf{c}^t \Sigma_{\mathbf{c}}^{-1} \mathbf{c} \right] \quad (3.7)$$

et la corrélation  $\Sigma_{\mathbf{c}}$  est caractérisée par un paramètre d'échelle  $\gamma_{\mathbf{c}}$  et deux paramètres de forme  $f_c$  et  $d$ .

### 3.1.2 Modélisation de la régularité spatiale du ciel

Nous considérons ici que l'inconnue correspond à « l'émission étendue » en astronomie, objet principalement basse fréquence dans les dimensions spatiales. Le ciel est modélisé comme une



**Figure 3.1** – Densité spectrale de puissance *a priori* pour le bruit. La courbe rouge correspond à la DSP *a priori*. La courbe bleue correspond à la DSP *a priori* sans composante blanche. Les paramètres choisis sont  $\gamma_c = 3$ ,  $f_c = 1$ ,  $d = 3$ ,  $\gamma_{b_b} = 0,01$ .

fonction continue, voir équation (2.39) et (2.41) page 62,

$$\begin{aligned}\phi(\alpha, \beta, \lambda) &= \sum_{k=0}^K \phi^k(\alpha, \beta) (\lambda - \lambda_0)^k \\ \phi^k(\alpha, \beta) &= \sum_{ij} x^k(i, j) \psi^{ij}(\alpha, \beta).\end{aligned}$$

Une solution naturelle pour mesurer la régularité spatiale est l'énergie de la dérivée première des fonctions  $\phi^k$ , dans les deux directions, soit dans la direction  $\alpha$

$$\begin{aligned}\left\| \frac{\partial \phi^k(\alpha, \beta)}{\partial \alpha} \right\|^2 &= \iint_{\mathbb{R}^2} \left( \sum_{i'j'} x^k(i', j') \frac{\partial}{\partial \alpha} \psi^{i'j'}(\alpha, \beta) \right) \left( \sum_{ij} x^k(i, j) \frac{\partial}{\partial \alpha} \psi^{ij}(\alpha, \beta) \right) d\alpha d\beta \\ &= \sum_{i'j'} x^k(i', j') \sum_{ij} x^k(i, j) \iint_{\mathbb{R}^2} \left( \frac{\partial}{\partial \alpha} \psi^{i'j'}(\alpha, \beta) \right) \left( \frac{\partial}{\partial \alpha} \psi^{ij}(\alpha, \beta) \right) d\alpha d\beta\end{aligned}$$

l'expression dans la direction  $\beta$  étant similaire. Les fonctions  $\psi^{ij}$  sont décalées dans l'espace soit

$$\psi^{ij}(\alpha, \beta) = \psi(\alpha - T_\alpha^i, \beta - T_\beta^j).$$

En notant la dérivée

$$\frac{\partial}{\partial \alpha} \psi(\alpha - T_\alpha^i, \beta - T_\beta^j) = \psi'_\alpha(\alpha - T_\alpha^i, \beta - T_\beta^j)$$

l'énergie de la dérivée première s'écrit

$$\begin{aligned}\left\| \frac{\partial \phi^k(\alpha, \beta)}{\partial \alpha} \right\|^2 &= \sum_{i'j'} x^k(i', j') \sum_{ij} x^k(i, j) \iint_{\mathbb{R}^2} \psi'_\alpha(\alpha - T_\alpha^{i'}, \beta - T_\beta^{j'}) \psi'_\alpha(\alpha - T_\alpha^i, \beta - T_\beta^j) d\alpha d\beta \\ &= \sum_{i'j'} x^k(i', j') \sum_{ij} x^k(i, j) [\psi'_\alpha \star \psi'_\alpha](T_\alpha^{i'} - T_\alpha^i, T_\beta^{j'} - T_\beta^j)\end{aligned}$$

et fait intervenir l'auto-corrélation  $\Psi_\alpha = [\psi'_\alpha \star \psi'_\alpha]$  de la dérivée première de la fonction de décomposition. Comme il y a un nombre fini de coefficients  $(i, j)$ , la mesure se met sous la forme d'une distance quadratique

$$\left\| \frac{\partial \phi^k(\alpha, \beta)}{\partial \alpha} \right\|^2 = \mathbf{x}_k^t \mathbf{D}_\alpha \mathbf{x}_k$$

où la matrice  $\mathbf{D}_\alpha$  est obtenue à partir de l'auto-corrélation de  $\psi'_\alpha$  et le vecteur  $\mathbf{x}_k$  contient les coefficients  $x^k$  de l'ordre  $k$ . La mesure de régularité quadratique sur la fonction continue s'exprime donc à travers une mesure de régularité quadratique sur les coefficients.

La matrice  $\mathbf{D}_\alpha$  est indépendante de l'ordre de développement de Taylor en longueur d'onde (si  $\psi$  l'est) bien qu'il soit possible d'adapter la mesure à l'ordre (par exemple en utilisant l'énergie de la fonction et non de sa dérivée). De plus, bien que les éléments de  $\mathbf{D}_\alpha$  soient obtenus à partir de l'auto-corrélation, la matrice n'a pas nécessairement une structure stationnaire puisque les coefficients peuvent être disposés irrégulièrement dans l'espace.

### Densité spectrale d'énergie

La transformée de Fourier de l'auto-corrélation est la densité spectrale d'énergie et correspond au module au carré de la transformée de Fourier de  $\psi'_\alpha$  soit

$$\begin{aligned} \hat{\Psi}_\alpha(f_\alpha, f_\beta) &= \iint_{\mathbb{R}^2} \Psi_\alpha(\alpha, \beta) e^{-2j\pi(\alpha f_\alpha + \beta f_\beta)} d\alpha d\beta \\ &= \left| \iint_{\mathbb{R}^2} \psi'_\alpha(\alpha, \beta) e^{-2j\pi(\alpha f_\alpha + \beta f_\beta)} d\alpha d\beta \right|^2 \\ &= \left| 2j\pi f_\alpha \hat{\psi}(f_\alpha, f_\beta) \right|^2 \\ &= 4\pi^2 f_\alpha^2 \left| \hat{\psi}(f_\alpha, f_\beta) \right|^2 \end{aligned}$$

où  $\hat{\psi}(f_\alpha, f_\beta)$  est la transformée de Fourier de la fonction de décomposition  $\psi$ . Ce calcul est intéressant puisqu'il fait apparaître la structure fréquentielle introduite *a priori* sur le ciel en fonction des fonctions de décomposition choisies. De plus avec l'introduction d'une mesure sur la deuxième dimension  $\beta$  il apparaît tout de suite que

$$\left\| \frac{\partial \phi^k(\alpha, \beta)}{\partial \alpha} \right\|^2 + \left\| \frac{\partial \phi^k(\alpha, \beta)}{\partial \beta} \right\|^2 = \mathbf{x}_k^t \mathbf{D}_\alpha \mathbf{x}_k + \mathbf{x}_k^t \mathbf{D}_\beta \mathbf{x}_k$$

et donc la densité spectrale d'énergie *a priori* pour le ciel est naturellement à symétrie circulaire

$$\begin{aligned} \hat{\Psi}(f_\alpha, f_\beta) &= \hat{\Psi}_\alpha(f_\alpha, f_\beta) + \hat{\Psi}_\beta(f_\alpha, f_\beta) \\ &= 4\pi^2 (f_\alpha^2 + f_\beta^2) \left| \hat{\psi}(f_\alpha, f_\beta) \right|^2. \end{aligned} \tag{3.8}$$

La structure de cette équation fait apparaître déjà que les basses fréquences ne seront pas affectées par cette régularisation puisque la forme quadratique  $f_\alpha^2 + f_\beta^2$  aura tendance à annuler  $\hat{\Psi}$  autour de zéro. C'est une structure passe-haut, cohérente avec une mesure de régularité par la dérivée première de la fonction. De plus le comportement autour de zéro est linéaire avec la fréquence (la densité  $\hat{\Psi}$  étant le module au carré), cohérent également avec le fait qu'il s'agit de la dérivée première.

La stationnarité de  $\mathbf{D}_\alpha$  est obtenue si le ciel est décomposé sur des fonctions  $\psi$  décalées régulièrement, ce qui est le choix effectué en pratique ici. Dans ce cas la matrice  $\mathbf{D}_\alpha$  est de Toeplitz et nous avons

$$\left\| \frac{\partial \phi^k(\alpha, \beta)}{\partial \alpha} \right\|^2 = \sum_{i'j'} x^k(i', j') \sum_{ij} x^k(i, j) \Psi_\alpha((i' - i)T_\alpha, (j' - j)T_\beta). \quad (3.9)$$

Ainsi le calcul

$$\mathbf{D}_\alpha \mathbf{x} = \sum_{ij} x^k(i, j) \Psi_\alpha((i' - i)T_\alpha, (j' - j)T_\beta)$$

est une convolution discrète, la deuxième somme sur les indices  $i'$  et  $j'$  de l'équation (3.9) étant une somme de produits terme à terme. Si on suppose la circularité de la matrice  $\mathbf{D}_\alpha$  alors

$$\mathbf{x}_k^t \mathbf{D}_\alpha \mathbf{x}_k = \mathbf{x}_k^t \mathbf{F}^\dagger \mathbf{\Lambda}_{\mathbf{D}_\alpha} \mathbf{F} \mathbf{x}_k = \mathring{\mathbf{x}}_k^\dagger \mathbf{\Lambda}_{\mathbf{D}_\alpha} \mathring{\mathbf{x}}_k$$

avec  $\mathbf{\Lambda}_{\mathbf{D}_\alpha}$  une matrice diagonale. Le calcul de la pénalisation peut donc se faire dans l'espace de Fourier. Dans ce cas la matrice  $\mathbf{\Lambda}_{\mathbf{D}_\alpha}$  est obtenue à partir de la fonction  $\mathring{\Psi}_\alpha$ .

### Décomposition sur des gaussiennes

Dans ce travail, la fonction  $\phi$  est décomposée dans les dimensions  $(\alpha, \beta)$  soit sur des gaussiennes soit sur des sinus cardinaux décalés régulièrement, voir 2.4.1. Les sinus cardinaux sont entièrement déterminés par les pas d'échantillonnages  $T_\alpha$  et  $T_\beta$ . Dans le cas des gaussiennes il faut en plus déterminer la largeur. Pour cela elle a été choisie en minimisant, au sens des moindres carrés, la largeur à mi-hauteur du sinus cardinal de même pas, soit

$$\sigma_{\alpha, \beta} \approx 0.6T_{\alpha, \beta}. \quad (3.10)$$

Dans le cas des gaussiennes la densité spectrale d'énergie s'écrit

$$\mathring{\Psi}(f_\alpha, f_\beta) = 4\pi^2 (f_\alpha^2 + f_\beta^2) \exp \left[ 2\pi \left( -\sigma_\alpha^2 f_\alpha^2 - \sigma_\beta^2 f_\beta^2 \right) \right] \quad (3.11)$$

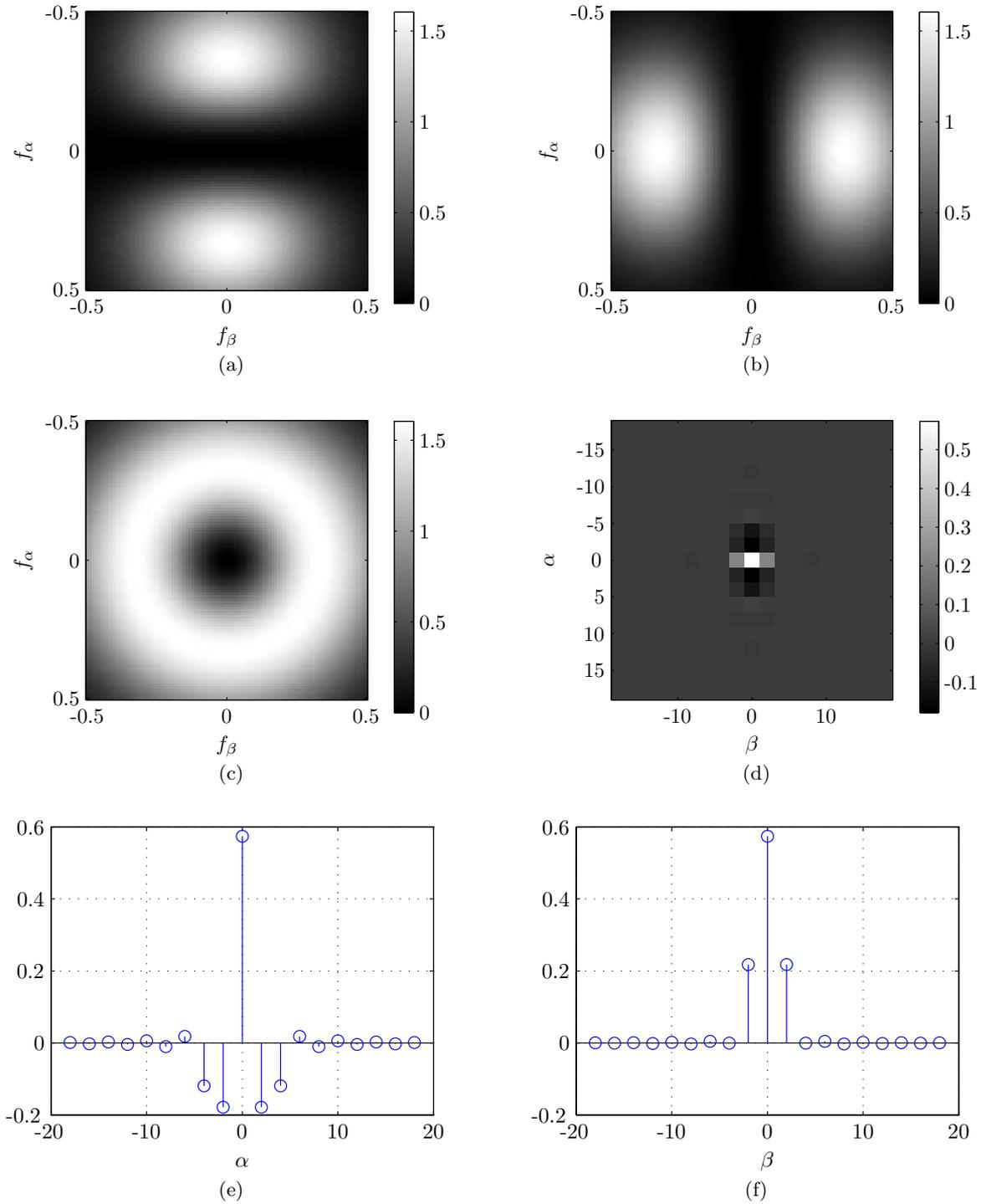
et l'auto-corrélation [Rod] pour la direction  $\alpha$  et  $\beta$  respectivement, s'écrit

$$\Psi_\alpha(\alpha, \beta) = \frac{1}{8\pi\sigma_\beta^5\sigma_\alpha} \left( \sigma_\beta^2 - \frac{T_\beta}{2} (j - j')^2 \right) \exp \left[ -\frac{1}{2} \frac{(i - i')^2 T_\alpha^2}{2\sigma_\alpha^2} - \frac{1}{2} \frac{(j - j')^2 T_\beta^2}{2\sigma_\beta^2} \right], \quad (3.12)$$

$$\Psi_\beta(\alpha, \beta) = \frac{1}{8\pi\sigma_\alpha^5\sigma_\beta} \left( \sigma_\alpha^2 - \frac{T_\alpha}{2} (i - i')^2 \right) \exp \left[ -\frac{1}{2} \frac{(i - i')^2 T_\alpha^2}{2\sigma_\alpha^2} - \frac{1}{2} \frac{(j - j')^2 T_\beta^2}{2\sigma_\beta^2} \right]. \quad (3.13)$$

La figure 3.3 est une illustration de ces fonctions avec les paramètres décrits précédemment (soit  $\sigma_{\alpha, \beta} \approx 0.6T_{\alpha, \beta}$ ). La figure la plus intéressante pour l'étude est la figure 3.3(c) correspondant à la somme des deux densités spectrales.

Tout d'abord elle est à symétrie circulaire comme prédit. Mais surtout elle a une structure passe-bande. La mesure de régularité ne pourra pas s'effectuer sur les très hautes fréquences des coefficients de décompositions. Il s'agit d'une illustration du fait que les gaussiennes ne sont pas capables, en fonction des paramètres donnés, de représenter toutes les fréquences possibles au sein de la bande de fréquences déterminée par  $T_\alpha$ . En conséquence, la mesure de régularité ne se fait pas sur les très hautes fréquences. Par conséquent lors de l'inversion, le système étant mal conditionné dans les très hautes fréquences malgré la régularisation, la solution pourra contenir des fréquences provenant d'une amplification du bruit. Cependant la solution recherchée est le ciel  $\phi$  et non les coefficients de sa décomposition.



**Figure 3.2** – Corrélation et densité spectrale d'énergie pour un ciel décomposé sur des gaussiennes. La figure 3.2(a) correspond à la densité spectrale d'énergie pour une dérivation dans la dimension  $\alpha$ , respectivement figure 3.2(b) pour  $\beta$ . La figure 3.2(c) correspond à la somme des deux. La figure 3.2(d) correspond au noyau de convolution discret à appliquer sur les coefficients pour une dérivation dans la direction  $\alpha$ . Les figures 3.2(e) et 3.2(f) correspondent aux coupes passant par l'origine. Les paramètres choisis sont  $\sigma_\alpha = \sigma_\beta = 0.6T_\alpha$  avec  $T_\alpha = 2$  arcsecs.

Des choix de paramètres peuvent être faits pour que la décroissance de la gaussienne dans l'espace de Fourier ne soit importante qu'après la fréquence limite déterminée par  $(T_\alpha, T_\beta)$ . Dans le cas choisi ici, utilisé en pratique pour les résultats d'inversion, la décroissance est proche de 0,5 en fréquence réduite.

### Décomposition sur des sinus cardinaux

Dans le cas des sinus cardinaux normalisés (qui n'ont pas été employés en pratique), la densité spectrale d'énergie s'écrit

$$\mathring{\Psi}(f_\alpha, f_\beta) = 4\pi^2 T_\alpha^2 T_\beta^2 (f_\alpha^2 + f_\beta^2) \text{rect}(f_\alpha T_\alpha) \text{rect}(f_\beta T_\beta) \quad (3.14)$$

où  $\text{rect}(f) = 1$  si et seulement si  $|f| < 1/2$ , 0 sinon. La figure 3.3 illustre ces fonctions.

Tout d'abord la présence de fonction porte illustre que les fonctions estimées sont à bande limitée. De plus le filtre équivalent est bien linéaire en fonction de la fréquence (puisque la densité est le module au carré) et ce régime n'est pas modifié contrairement au cas avec des gaussiennes. La densité spectrale d'énergie est bien à symétrie circulaire.

Le plus intéressant se situe sur les figures 3.3(e) et 3.3(f) qui représentent le noyau de convolution pour la dérivée dans la direction  $\alpha$ . Bien que les fonctions de base soient des sinus cardinaux, le filtre ressemble à une différence finie  $\mathbf{D}^t \mathbf{D} = [-1, 2, -1]$  sur les coefficients. Par contre, contrairement au cas gaussien où la dérivée dans une direction influence l'énergie dans l'autre, voir figure 3.2(f), ici il y a une séparation parfaite : le noyau de convolution est mono-dimensionnel.

### Loi *a priori*

En conclusion, la mesure de régularité de la fonction s'effectue avec l'auto-corrélation de la fonction de décomposition. Le nombre de coefficients de décomposition étant fini, la pénalisation s'exprime comme

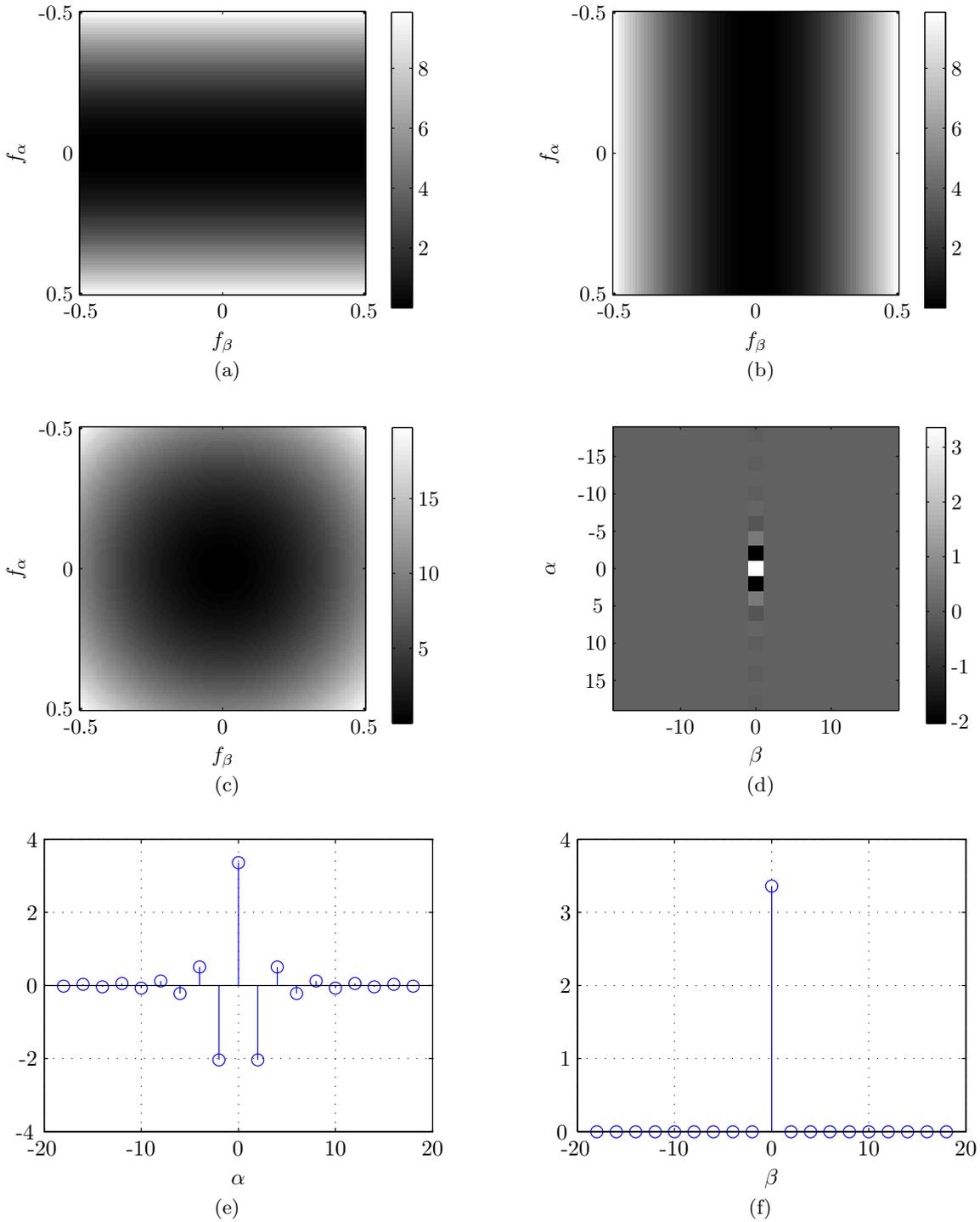
$$\begin{aligned} \left\| \frac{\partial \phi^k(\alpha, \beta)}{\partial \alpha} \right\|^2 + \left\| \frac{\partial \phi^k(\alpha, \beta)}{\partial \beta} \right\|^2 &= \mathbf{x}_k^t (\mathbf{D}_\alpha + \mathbf{D}_\beta) \mathbf{x}_k \\ &= \mathring{\mathbf{x}}_k^\dagger (\mathbf{\Lambda}_{\mathbf{D}_\alpha} + \mathbf{\Lambda}_{\mathbf{D}_\beta}) \mathring{\mathbf{x}}_k = \mathring{\mathbf{x}}_k^\dagger \mathbf{\Lambda}_{\mathbf{D}} \mathring{\mathbf{x}}_k \end{aligned} \quad (3.15)$$

où  $\mathbf{\Lambda}_{\mathbf{D}}$  est une matrice diagonale obtenue à partir de la densité spectrale équation (3.11) dans le cas gaussien, étudié ici.

On sait que le ciel est plutôt basse fréquence. Pour permettre la restauration de fréquences dégradées par la réponse instrumentale sans restaurer des fréquences provenant du bruit, nous allons utiliser cette mesure de régularité pour introduire comme information que les grandes variations sont peu probables. Par conséquent la loi *a priori* pour les coefficients d'un ordre  $k$  de développement Taylor en longueur d'onde est une loi normale classique qui s'exprime

$$p(\mathbf{x}_k | \gamma_{\mathbf{x}_k}) = (2\pi)^{-(IJ-1)/2} \gamma_{\mathbf{x}_k}^{(IJ-1)/2} \det[\mathbf{D}]^{1/2} \exp \left[ -\frac{1}{2} \gamma_{\mathbf{x}_k} \mathbf{x}_k^t \mathbf{D} \mathbf{x}_k \right] \quad (3.16)$$

où  $\gamma_{\mathbf{x}_k}$  est un facteur de puissance qui pilote le degré de régularité. C'est une loi potentiellement dégénérée (nulle partout) puisque l'opérateur  $\mathbf{D}$  est un opérateur de différence et ne parle pas de la fréquence nulle (la première valeur propre de  $\mathbf{\Lambda}_{\mathbf{D}}$  est nulle). On peut cependant employer cet *a priori* corrigé de la fréquence nulle pour obtenir une loi *a posteriori* propre, puis faire tendre l'*a priori* sur la moyenne vers zéro [Gio08, MKB92], pour peu que la vraisemblance se prononce sur la



**Figure 3.3** – Corrélation et densité spectrale d'énergie pour un ciel décomposé sur des sinus cardinaux. La figure 3.3(a) correspond à la densité spectrale d'énergie pour une dérivation dans la dimension  $\alpha$ , respectivement figure 3.3(b) pour  $\beta$ . La figure 3.3(c) correspond à la somme des deux. La figure 3.3(d) correspond au noyau de convolution discret, obtenue par transformée de Fourier discrète inverse. Les figures 3.3(e) et 3.3(f) correspondent aux coupes passant par l'origine.

moyenne. L'article de l'annexe F étudie cela en détail. Enfin, en pratique l'opérateur de régularité de l'équation (3.12) est calculé sur un support limité alors qu'il est défini sur un support infini. Une conséquence est que la transformée de Fourier n'est pas nulle à la fréquence nulle. Donc en pratique la loi *a priori* est propre. Nous considérerons donc dans la suite que cet *a priori* est directement utilisé sans se soucier de la dégénérescence, la loi *a posteriori* étant propre.

Un seul paramètre  $\gamma_{\mathbf{x}_k}$  est utilisé pour piloter les différences en  $\alpha$  et en  $\beta$ . En effet les deux grandeurs  $\mathbf{D}_\alpha \mathbf{x}$  et  $\mathbf{D}_\beta \mathbf{x}$  ne sont pas indépendantes. On ne peut donc pas considérer une loi *a priori*

$$p(\mathbf{x}_k) = p(\mathbf{D}_\alpha \mathbf{x}_k)p(\mathbf{D}_\beta \mathbf{x}_k)$$

permettant de définir deux hyper-paramètres sans ajouter une approximation supplémentaire d'indépendance (entre les différences en ligne et les différences en colonnes en quelque sorte). Dans un cadre classique de régularisation énergétique à hyper-paramètres fixés, ce n'est pas un problème. Dans une approche bayésienne à hyper-paramètres fixés non plus. Cependant dans une interprétation bayésienne pour estimer  $\gamma_{\mathbf{x}_k}$ , la connaissance de la fonction de partition est nécessaire. Dans ce cas, la fonction de partition est connue et c'est l'inverse de la racine carrée du déterminant de la matrice de covariance

$$\det [\mathbf{D}_\alpha + \mathbf{D}_\beta]^{1/2}.$$

Cependant même si la fonction de partition est connue, la loi *a posteriori* pour les hyper-paramètres n'en est pas moins complexe même avec une diagonalisation dans l'espace de Fourier. En effet, en factorisant un facteur de puissance que l'on souhaiterait estimer, alors le déterminant est

$$\begin{aligned} \det [\gamma_\alpha \mathbf{D}_\alpha + \gamma_\beta \mathbf{D}_\beta]^{1/2} &= \det [\gamma_\alpha \boldsymbol{\Lambda}_{\mathbf{D}_\alpha} + \gamma_\beta \boldsymbol{\Lambda}_{\mathbf{D}_\beta}]^{1/2} \\ &= \prod_i \left( \gamma_\alpha \lambda_{\mathbf{D}_\alpha}^i + \gamma_\beta \lambda_{\mathbf{D}_\beta}^i \right)^{1/2} \end{aligned}$$

et les facteurs de puissances ne peuvent pas, dans le cas général, être séparés. Dans ce cas toute loi pour  $\gamma_\alpha$  ou  $\gamma_\beta$  obtenue à partir de la loi jointe est complexe et ne s'identifie pas à une loi classique. Dans une perspective d'échantillonnage stochastique cela alourdira les calculs numériques. On peut mentionner un cas particulier. Si les matrices  $\lambda_{\mathbf{D}_\alpha}$  et  $\lambda_{\mathbf{D}_\beta}$  sont des pavés disjoints du plan fréquence alors le déterminant devient séparable puisque les matrices ne parlent pas des même fréquences : les lois sont *indépendantes*. Dans ce cas le déterminant s'écrit

$$\begin{aligned} \det [\gamma_\alpha \mathbf{D}_\alpha + \gamma_\beta \mathbf{D}_\beta]^{1/2} &= \prod_i \left( \gamma_\alpha \lambda_{\mathbf{D}_\alpha}^i + \gamma_\beta \lambda_{\mathbf{D}_\beta}^i \right)^{1/2} \\ &= \prod_{i \in \mathcal{I}} \left( \gamma_\alpha \lambda_{\mathbf{D}_\alpha}^i \right)^{1/2} \prod_{j \in \mathcal{J}} \left( \gamma_\beta \lambda_{\mathbf{D}_\beta}^j \right)^{1/2} \end{aligned}$$

avec  $\mathcal{I}$  et  $\mathcal{J}$  deux ensembles disjoints d'indices des fréquences. Ce qui correspond au déterminant obtenu par le produit des deux lois *a priori* indépendantes.

Dans notre cas les *a priori* ne sont pas indépendants comme le montrent les figures 3.2 et 3.3. Pour éviter la difficulté de l'estimation de deux paramètres de puissances différents, nous avons considéré un unique paramètre qui pilote toutes les fréquences. De plus cela modélise également un caractère *a priori* isotrope de l'inconnue.

Pour finir on suppose l'indépendance entre les ordres de développement de Taylor. Ainsi, avec la vraisemblance définie dans la partie précédente et la loi pour le ciel, une première loi *a posteriori* peut être bâtie

$$p(\mathbf{x}, \mathbf{c} | \mathbf{y}, \boldsymbol{\theta}) = \frac{p(\mathbf{y} | \mathbf{x}, \mathbf{c}, \boldsymbol{\theta}) p(\mathbf{x} | \boldsymbol{\theta}) p(\mathbf{c} | \boldsymbol{\theta})}{p(\mathbf{y}, \boldsymbol{\theta})}$$

où  $\boldsymbol{\theta}$  collecte tous les paramètres nécessaires à la définition de la loi c'est-à-dire  $\gamma_{\mathbf{x}_k}, \boldsymbol{\Sigma}_{\mathbf{c}} \dots$

La prochaine partie présente les deux stratégies possibles pour inférer avec la composante de nuisance  $\mathbf{c}$ .

### 3.1.3 Estimation jointe et marginalisation de dérive thermique

Pour simplifier le discours dans la suite toutes les données sont utilisées pour l'estimation. Autrement dit il n'y a pas de données manquantes ou aberrantes qui seraient négligées. La méthodologie reste identique et la généralisation à la présence de données manquantes ou aberrantes est immédiate. Cette hypothèse est faite pour ne pas casser la structure de corrélation temporelle des signaux, en particulier  $\mathbf{c}$  mais également  $\mathbf{y}$ .

La loi *a posteriori* jointe pour l'objet d'intérêt  $\mathbf{x}$  et la composante basse fréquence  $\mathbf{c}$  s'écrit

$$p(\mathbf{x}, \mathbf{c} | \mathbf{y}, \boldsymbol{\theta}) = \frac{p(\mathbf{y} | \mathbf{x}, \mathbf{c}, \boldsymbol{\theta}) p(\mathbf{x} | \boldsymbol{\theta}) p(\mathbf{c} | \boldsymbol{\theta})}{p(\mathbf{y}, \boldsymbol{\theta})}$$

et deux stratégies sont possibles. Soit  $\mathbf{x}$  et  $\mathbf{c}$  sont estimés conjointement, ce qui correspond à l'estimation d'un signal  $\mathbf{c}$ , soit  $\mathbf{x}$  est estimé à partir de la loi *a posteriori* marginale

$$\begin{aligned} p(\mathbf{x} | \mathbf{y}, \boldsymbol{\theta}) &= \int \frac{p(\mathbf{y} | \mathbf{x}, \mathbf{c}, \boldsymbol{\theta}) p(\mathbf{x} | \boldsymbol{\theta}) p(\mathbf{c} | \boldsymbol{\theta})}{p(\mathbf{y}, \boldsymbol{\theta})} d\mathbf{c} \\ &= \frac{1}{p(\mathbf{y}, \boldsymbol{\theta})} p(\mathbf{x} | \boldsymbol{\theta}) \int p(\mathbf{y}, \mathbf{c} | \mathbf{x}, \boldsymbol{\theta}) d\mathbf{c} \end{aligned}$$

ce qui revient à considérer une autre vraisemblance

$$p(\mathbf{y} | \mathbf{x}, \boldsymbol{\theta}) = \int p(\mathbf{y}, \mathbf{c} | \mathbf{x}, \boldsymbol{\theta}) d\mathbf{c}$$

qui inclut la statistique *a priori* de  $\mathbf{c}$  sans l'estimer.

En vérité, comme le modèle  $\mathbf{H}$  est linéaire et que toutes les lois sont gaussiennes, alors la loi *a posteriori* jointe est gaussienne

$$p(\mathbf{x}, \mathbf{c} | \mathbf{y}, \boldsymbol{\theta}) \sim \mathcal{N} \left( \begin{pmatrix} \mathbf{m}_{\mathbf{x}} \\ \mathbf{m}_{\mathbf{c}} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{x}} & \boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{c}} \\ \boldsymbol{\Sigma}_{\mathbf{c}, \mathbf{x}} & \boldsymbol{\Sigma}_{\mathbf{c}, \mathbf{c}} \end{pmatrix} \right)$$

et les estimateurs *pour l'objet d'intérêt*  $\mathbf{x}$  du maximum et de la moyenne *a posteriori*, dans les deux stratégies, sont rigoureusement identiques. En effet dans la stratégie jointe le maximum *a posteriori*

$$\begin{aligned} [\hat{\mathbf{x}}_{\text{MAP}}, \hat{\mathbf{c}}_{\text{MAP}}] &= \arg \max_{\mathbf{x}, \mathbf{c}} p(\mathbf{x}, \mathbf{c} | \mathbf{y}, \boldsymbol{\theta}) \\ &= [\mathbf{m}_{\mathbf{x}}, \mathbf{m}_{\mathbf{c}}] \\ &= [\hat{\mathbf{x}}_{\text{EAP}}, \hat{\mathbf{c}}_{\text{EAP}}] \end{aligned}$$

est la moyenne et dans la stratégie conjointe, comme la loi *a posteriori* marginale s'écrit

$$p(\mathbf{x} | \mathbf{y}, \boldsymbol{\theta}) \sim \mathcal{N}(\mathbf{m}_{\mathbf{x}}, \boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{x}}),$$

alors de nouveau

$$[\hat{\mathbf{x}}_{\text{MAP}}] = [\hat{\mathbf{x}}_{\text{EAP}}] = \mathbf{m}_{\mathbf{x}}.$$

Le choix de la stratégie repose donc sur un intérêt pratique de calcul de l'estimateur. La différence entre les stratégies repose sur la loi  $p(\mathbf{y}, \mathbf{c} | \mathbf{x}) = p(\mathbf{y} | \mathbf{x}, \mathbf{c}) p(\mathbf{c})$  ou sa marginale par rapport à  $\mathbf{c}$ .

**Estimation conjointe.** Toutes les lois sont gaussiennes alors la loi des données  $\mathbf{y}$  et de  $\mathbf{c}$  sont également gaussiennes. Elle est définie à partir de ses deux premiers moments. La moyenne de la loi  $p(\mathbf{y}, \mathbf{c} | \mathbf{x})$  est

$$\mathbf{m}_{\mathbf{y}, \mathbf{c}} = \mathbb{E} \left[ \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_B \\ \mathbf{y}_{B+1} \\ \mathbf{c} \end{pmatrix} \right] = \mathbb{E} \left[ \begin{pmatrix} \mathbf{H}_0 \mathbf{x} + \mathbf{c} + \mathbf{b}_0 \\ \mathbf{H}_1 \mathbf{x} + \mathbf{c} + \mathbf{b}_1 \\ \vdots \\ \mathbf{c} + \mathbf{b}_B \\ \mathbf{c} + \mathbf{b}_{B+1} \\ \mathbf{c} \end{pmatrix} \right] = \begin{pmatrix} \mathbf{H}_0 \mathbf{x} \\ \mathbf{H}_1 \mathbf{x} \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \quad (3.17)$$

puisque les lois *a priori* de  $\mathbf{b}_b$  et de  $\mathbf{c}$  sont à moyenne nulle. La matrice de covariance de la loi  $p(\mathbf{y}, \mathbf{c} | \mathbf{x})$  est

$$\begin{aligned} \Sigma_{\mathbf{y}, \mathbf{c}} &= \mathbb{E} \left[ \left( \begin{pmatrix} \mathbf{y} \\ \mathbf{c} \end{pmatrix} - \mathbf{m}_{\mathbf{y}, \mathbf{c}} \right) \left( \begin{pmatrix} \mathbf{y} \\ \mathbf{c} \end{pmatrix} - \mathbf{m}_{\mathbf{y}, \mathbf{c}} \right)^t \right] = \mathbb{E} \left[ \begin{pmatrix} \mathbf{c} + \mathbf{b}_0 \\ \mathbf{c} + \mathbf{b}_1 \\ \vdots \\ \mathbf{c} + \mathbf{b}_B \\ \mathbf{c} + \mathbf{b}_{B+1} \\ \mathbf{c} \end{pmatrix} \begin{pmatrix} \mathbf{c} + \mathbf{b}_0 \\ \mathbf{c} + \mathbf{b}_1 \\ \vdots \\ \mathbf{c} + \mathbf{b}_B \\ \mathbf{c} + \mathbf{b}_{B+1} \\ \mathbf{c} \end{pmatrix}^t \right] \\ &= \begin{pmatrix} \gamma_{\mathbf{b}_0}^{-1} \Sigma_{\mathbf{b}} + \Sigma_{\mathbf{c}} & \Sigma_{\mathbf{c}} & \dots & \Sigma_{\mathbf{c}} & \Sigma_{\mathbf{c}} \\ \Sigma_{\mathbf{c}} & \gamma_{\mathbf{b}_1}^{-1} \Sigma_{\mathbf{b}} + \Sigma_{\mathbf{c}} & & \vdots & \Sigma_{\mathbf{c}} \\ \vdots & & \ddots & \vdots & \vdots \\ \Sigma_{\mathbf{c}} & \dots & \dots & \gamma_{\mathbf{b}_{B+1}}^{-1} \Sigma_{\mathbf{b}} + \Sigma_{\mathbf{c}} & \vdots \\ \Sigma_{\mathbf{c}} & \Sigma_{\mathbf{c}} & \dots & \dots & \Sigma_{\mathbf{c}} \end{pmatrix} \end{aligned}$$

puisque les signaux  $\mathbf{b}_b$  et  $\mathbf{c}$  sont mutuellement indépendants. On voit ici l'importance de l'hypothèse présentée en début de partie qui établissait que toutes les données étaient utilisées. Les blocs auraient été  $\mathbf{T} \Sigma_{\mathbf{c}}^{-1} \mathbf{T}^t$  au lieu de simplement  $\Sigma_{\mathbf{c}}^{-1}$ .

La structure de  $\Sigma_{\mathbf{y}, \mathbf{c}}$  indique que la corrélation entre deux bolomètres est identique quels que soient les bolomètres et correspond à la corrélation de  $\mathbf{c}$ . L'auto-corrélation d'un même bolomètre fait également intervenir une corrélation supplémentaire  $\Sigma_{\mathbf{b}}$ .

Cette loi est intéressante mais n'est pas nécessaire pour connaître la solution des moindres carrés régularisés. En effet en posant la loi *a priori* jointe

$$p(\mathbf{x}, \mathbf{c}) \sim \mathcal{N} \left( \mathbf{0}, \begin{pmatrix} \Sigma_{\mathbf{x}} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\mathbf{c}} \end{pmatrix} \right)$$

alors la solution des moindres carrés régularisés conjointe est

$$\begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{c}} \end{pmatrix} = \left( \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}^t \Sigma_{\mathbf{y} | \mathbf{x}, \mathbf{c}}^{-1} \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} + \begin{pmatrix} \Sigma_{\mathbf{x}} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\mathbf{c}} \end{pmatrix}^{-1} \right)^{-1} \begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \\ \mathbf{T}^t & \mathbf{I} & \mathbf{I} \end{pmatrix} \Sigma_{\mathbf{y} | \mathbf{x}, \mathbf{c}}^{-1} \begin{pmatrix} \mathbf{y} \\ \mathbf{y}_B \\ \mathbf{y}_{B+1} \end{pmatrix}$$

où  $\mathbf{y}$  correspond aux données de tous les bolomètres non aveugles et

$$\Sigma_{\mathbf{y} | \mathbf{x}, \mathbf{c}} = \begin{pmatrix} \gamma_{\mathbf{b}_0}^{-1} \Sigma_{\mathbf{b}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \gamma_{\mathbf{b}_1}^{-1} \Sigma_{\mathbf{b}} & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \dots & \dots & \gamma_{\mathbf{b}_{B+1}}^{-1} \Sigma_{\mathbf{b}} \end{pmatrix}.$$

Nous savons déjà que la matrice  $\mathbf{H}$  sera complexe et de grande taille, la solution ne pourra donc pas être calculée directement.

Bien que l'expression pour le calcul de  $\hat{\mathbf{x}}$  fasse intervenir

$$\begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{y}_B \\ \mathbf{y}_{B+1} \end{pmatrix}$$

l'estimation dépend bien des données des bolomètres aveugles par la présence de l'inverse du hessien. Pour comprendre supposons que la matrice  $\Sigma_{\mathbf{y}|\mathbf{x},c}^{-1}$  est l'identité. Alors la solution des moindres carrés s'écrit

$$\begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{c}} \end{pmatrix} = \begin{pmatrix} \mathbf{H}^t \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} & \mathbf{H}^t \mathbf{T} \\ \mathbf{T}^t \mathbf{H} & \mathbf{T}^t \mathbf{T} + 2\mathbf{I} + \Sigma_{\mathbf{c}}^{-1} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{H}^t \mathbf{y} \\ \mathbf{T}^t \mathbf{y} + \mathbf{y}_B + \mathbf{y}_{B+1} \end{pmatrix}.$$

On pourrait obtenir une expression explicite de l'inverse du hessien à l'aide du lemme d'inversion de matrices partitionnées mais ce ne serait pas très instructif ici. L'important est de remarquer que la partie en haut à droite du hessien composé de  $\mathbf{H}^t \mathbf{T}$  n'est pas nulle et donc que l'estimation  $\hat{\mathbf{x}}$  dépend des données des bolomètres aveugles. Cette caractéristique réapparaîtra dans l'algorithme d'optimisation.

**Estimation par marginalisation.** La loi marginale se déduit de la loi jointe. La moyenne et la matrice de covariance de la loi sont

$$\mathbf{m}_{\mathbf{y}|\mathbf{x}} = \begin{pmatrix} \mathbf{H}_0 \mathbf{x} \\ \mathbf{H}_1 \mathbf{x} \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad (3.18)$$

et

$$\Sigma_{\mathbf{y}|\mathbf{x}} = \begin{pmatrix} \gamma_{b_0}^{-1} \Sigma_{\mathbf{b}} + \Sigma_{\mathbf{c}} & \Sigma_{\mathbf{c}} & \dots & \Sigma_{\mathbf{c}} \\ \Sigma_{\mathbf{c}} & \gamma_{b_1}^{-1} \Sigma_{\mathbf{b}} + \Sigma_{\mathbf{c}} & & \vdots \\ \vdots & & \ddots & \vdots \\ \Sigma_{\mathbf{c}} & \dots & \dots & \gamma_{b_{B+1}}^{-1} \Sigma_{\mathbf{b}} + \Sigma_{\mathbf{c}} \end{pmatrix}. \quad (3.19)$$

qui correspond au modèle de bruit de SANEPIC. Dans ce cas la solution des moindres carrés est

$$\hat{\mathbf{x}} = \left( \begin{pmatrix} \mathbf{H} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}^t \Sigma_{\mathbf{y}|\mathbf{x}}^{-1} \begin{pmatrix} \mathbf{H} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} + \Sigma_{\mathbf{x}}^{-1} \right)^{-1} \begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \end{pmatrix} \Sigma_{\mathbf{y}|\mathbf{x}}^{-1} \begin{pmatrix} \mathbf{y} \\ \mathbf{y}_{a1} \\ \mathbf{y}_{a2} \end{pmatrix}$$

La présence de zéros pourrait de nouveau laisser penser que dans ce cas l'estimation  $\hat{\mathbf{x}}$  ne dépend plus des données des bolomètres aveugles mais uniquement de la covariance  $\Sigma_{\mathbf{c}}$ . En fait c'est toujours le cas et de nouveau à cause de l'inversion du hessien qui mélange les données entre elles puisque la matrice n'est pas bloc diagonale.

Pour bien comprendre étudions un cas dégénéré scalaire qui nous permettra de mener les calculs. Considérons un bolomètre qui fournit une donnée scalaire  $y$  à propos d'un scalaire  $x$ . Un autre bolomètre, aveugle, fournit une donnée  $y_a$ . Le bruit est modélisé par un bruit blanc de variance  $\sigma_b^2$ . Une composante  $c$  est observée par tous les bolomètres et suit une loi normale  $\mathcal{N}(0, \sigma_c^2)$ . Pour simplifier nous considérons qu'il n'y a pas d'*a priori* sur l'objet et choisissons

l'estimateur du maximum de vraisemblance. Comme nous considérons la loi marginale alors le modèle est

$$\begin{aligned} y &= hx + b, \\ y_a &= b_a, \end{aligned}$$

avec une loi du bruit

$$p(b, b_a) \sim \mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_b^2 + \sigma_c^2 & \sigma_c^2 \\ \sigma_c^2 & \sigma_b^2 + \sigma_c^2 \end{pmatrix} \right).$$

Cette matrice de covariance peut être inversée. Pour cela on pose

$$\Sigma = \begin{pmatrix} \sigma_b^2 + \sigma_c^2 & \sigma_c^2 \\ \sigma_c^2 & \sigma_b^2 + \sigma_c^2 \end{pmatrix} = \bar{r} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

avec

$$\bar{r} = \sigma_b^2 + \sigma_c^2 \quad \text{et} \quad \rho = \sigma_c^2 / \bar{r}$$

alors

$$\Sigma^{-1} = \begin{pmatrix} \sigma_b^2 + \sigma_c^2 & \sigma_c^2 \\ \sigma_c^2 & \sigma_b^2 + \sigma_c^2 \end{pmatrix} = \frac{1}{\bar{r}(1 - \rho^2)} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}.$$

La solution des moindres carrés est donc

$$\begin{aligned} \hat{x} &= \left( \begin{pmatrix} h & 0 \end{pmatrix} \Sigma^{-1} \begin{pmatrix} h \\ 0 \end{pmatrix} \right)^{-1} \begin{pmatrix} h & 0 \end{pmatrix} \Sigma^{-1} \begin{pmatrix} y \\ y_a \end{pmatrix} \\ &= \left( \begin{pmatrix} h & 0 \end{pmatrix} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} h \\ 0 \end{pmatrix} \right)^{-1} \begin{pmatrix} h & 0 \end{pmatrix} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} y \\ y_a \end{pmatrix} \\ &= h^{-1}(y - \rho y_a). \end{aligned}$$

La solution consiste à retrancher la donnée observée par le bolomètre aveugle, *corrigée* par un certain facteur  $\rho$ , puis d'effectuer la division par le gain  $h$ .

Dans un cas plus général nous avons bien sûr les contributions des autres bolomètres, aveugles ou non, tous corrigés des facteurs adaptés dont la connaissance demande le calcul direct de la solution des moindres carrés.

Il est intéressant de remarquer que les bolomètres aveugles sont très importants puisqu'ils apportent énormément d'information sur la composante  $c$ . Cependant la soustraction directe des données aveugles aux données non aveugles n'est pas une stratégie efficace puisqu'il apparaît, avec ce modèle, qu'il faut les corriger d'un facteur

$$\rho = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_b^2}.$$

L'interprétation de ce facteur  $\rho$  est simple. Tout d'abord  $\rho = 0$  si la puissance de la composante corrélée  $\sigma_c^2$  est nulle ou si la puissance du bruit blanc  $\sigma_b^2$  est infinie, et dans ces deux cas les données aveugles ne servent évidemment à rien. Ensuite  $\rho = 1$  si la puissance du bruit blanc  $\sigma_b^2$  est nulle. Dans ce cas la solution est triviale puisque nous sommes dans un système à deux équations deux inconnues : la même réalisation du bruit<sup>1</sup>, réduit à  $c$ , est observée. La solution est donc directement  $x = (y - y_a)/h$ .

<sup>1</sup>Le sens usuel de bruit n'est plus très adapté dans ce cas.

En conclusion, les deux stratégies sont équivalentes et gagnent toutes les deux à exploiter les données des bolomètres aveugles (ce qui n'est pas le cas de SANEPIC actuellement). Bien que toutes les hypothèses ne soient pas avancées :

- les méthodes proposées par Kevin Xu [SBL<sup>+</sup>08] qui modélisent et estiment explicitement une composante  $\mathbf{c}$ , puis effectuent des opérations linéaires avec les données,
  - et les méthodes qui proposent une modélisation avec une structure de covariance contrainte comme SANEPIC et qui effectuent également des opérations linéaires avec les données,
- proposent en quelque sorte le même estimateur  $\mathbf{m}_{\mathbf{x}|\mathbf{y}}$ . Les différences se situent au niveau des modèles utilisés (matrice de corrélation, splines, *etc.*) et des paramètres. La différence se situe également au niveau des algorithmes correspondants pour calculer l'estimateur. Il y a deux désavantages à considérer une stratégie par marginalisation comme SANEPIC.

- Tout d'abord il faut procéder à l'inversion de la matrice de covariance  $\Sigma_{\mathbf{y}|\mathbf{x}}$  équation (3.19) ce qui, même en considérant un caractère circulant pour  $\Sigma_{\mathbf{b}}$  et  $\Sigma_{\mathbf{c}}$ , n'est pas direct, sauf à considérer que la matrice  $\Sigma_{\mathbf{y}|\mathbf{x}}$  est circulante globalement.
- Le fait d'avoir comme modèle un bruit corrélé pose problème dès qu'il y a des données manquantes ou aberrantes qui cassent la structure de stationnarité. Les outils comme SANEPIC, pour gérer ces données manquantes, doivent utiliser des méthodes et approximations supplémentaires pour « boucher » les trous.

L'estimation conjointe permet de faire apparaître la composante  $\mathbf{c}$  comme un *signal* observé à travers un modèle quelconque qui permettra de modéliser ces données manquantes directement dans le problème. Il est également possible de faire un lien entre le signal  $\mathbf{c}$  et le comportement physique de l'instrument comme la dérive de la température du cryostat dont on sait qu'elle est l'origine principale de cette composante (voir partie 2.2.4).

Dans ce travail nous avons fait le choix d'utiliser une stratégie conjointe pour la prise en compte de la composante corrélée  $\mathbf{c}$ . Toute la corrélation du bruit ayant été modélisée dans la composante  $\mathbf{c}$ , la loi du bruit  $\mathbf{b}$  a été choisie gaussienne blanche *iid* soit

$$p(\mathbf{y}|\mathbf{x}, \mathbf{c}, \gamma_{\mathbf{b}}) \sim \mathcal{N}(\mathbf{m}_{\mathbf{y}}, \Sigma_{\mathbf{y}}) \quad (3.20)$$

avec

$$\mathbf{m}_{\mathbf{y}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \mathbf{x} + \begin{pmatrix} \mathbf{T}_0 \\ \mathbf{T}_1 \\ \vdots \\ \mathbf{I} \\ \mathbf{I} \end{pmatrix} \mathbf{c} \quad (3.21)$$

où  $\mathbf{H}_b$  et  $\mathbf{T}_b$  correspond au modèle d'acquisition des données du bolomètre  $b$  à tous les instants d'échantillonnage temporel considéré valide. Les deux modèles « vide »  $\mathbf{0}$  correspondent aux deux bolomètres aveugles. La matrice de covariance est

$$\Sigma_{\mathbf{y}} = \begin{pmatrix} \gamma_{b_0}^{-1} \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \gamma_{b_1}^{-1} \mathbf{I} & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \gamma_{b_{B+1}}^{-1} \mathbf{I} \end{pmatrix} \quad (3.22)$$

où  $\mathbf{I}$  correspond à la matrice identité de taille  $N$  par  $N$ . En conclusion la vraisemblance du problème s'écrit

$$p(\mathbf{y}|\mathbf{x}, \mathbf{c}, \gamma_{\mathbf{b}}) = (2\pi)^{(B+2)N/2} \prod_{b=0}^{B+1} \gamma_{b_b}^{N/2} \exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{m}_{\mathbf{y}})^t \Sigma_{\mathbf{y}}^{-1} (\mathbf{y} - \mathbf{m}_{\mathbf{y}}) \right]. \quad (3.23)$$

### 3.1.4 Loi *a posteriori* étendue

Les modèles probabilistes pour le ciel  $\mathbf{x}$  et la composante basse fréquence  $\mathbf{c}$  dépendent de paramètres. De plus, la vraisemblance dépend également de paramètres instruments  $\boldsymbol{\theta}_i$ . A part pour les paramètres instruments qui peuvent être considérés connus, ce n'est jamais le cas des paramètres des lois. La stratégie utilisée pour les estimer, présentée en introduction partie 1.3.3, repose sur une loi *a posteriori* étendue

$$p(\mathbf{x}, \mathbf{c}, \boldsymbol{\theta} | \mathbf{y}) = \frac{p(\mathbf{y} | \mathbf{x}, \mathbf{c}, \boldsymbol{\theta}) p(\mathbf{x}, \mathbf{c}, \boldsymbol{\theta})}{p(\mathbf{y})}$$

en considérant la structure hiérarchique pour l'*a priori*

$$p(\mathbf{x}, \mathbf{c}, \boldsymbol{\theta}) = p(\mathbf{x} | \boldsymbol{\theta}) p(\mathbf{c} | \boldsymbol{\theta}) p(\boldsymbol{\theta}).$$

Le ciel  $\mathbf{x}$  et la composante basse fréquence  $\mathbf{c}$  sont indépendants *a priori* (la connaissance de l'un n'apporte rien sur l'autre *a priori*) et dépendent des paramètres  $\boldsymbol{\theta}$ . Le choix des paramètres supplémentaires estimés est libre. Dans ce travail les paramètres potentiellement estimés sont :

- les puissances de la corrélation *a priori* sur l'objet  $\gamma_{\mathbf{x}}$ ,
- les puissances de la composante blanche  $\gamma_{\mathbf{b}}$ , potentiellement différentes entre les bandes voire entre bolomètres,
- la corrélation *a priori* sur la composante continue  $\boldsymbol{\Sigma}_{\mathbf{c}}$ . D'après la partie 3.1.1, la connaissance des paramètres  $f_{\mathbf{c}}$ ,  $\gamma_{\mathbf{c}}$  et  $d$  suffit pour déterminer entièrement cette corrélation.
- des paramètres de la réponse instrumentale  $\boldsymbol{\theta}_i$ . On peut envisager la largeur de la réponse optique gaussienne  $\sigma_m$ , la constante de temps du bolomètre  $\tau_e$ , ou encore les gains des bolomètres  $S'$ .

Ces paramètres peuvent être classés en fonction de l'existence ou non de lois conjuguées. Tout d'abord les paramètres des puissances des différentes lois gaussiennes possèdent des lois conjuguées. Les paramètres  $\gamma_{\mathbf{x}}$  pour l'objet,  $\gamma_{\mathbf{b}}$  pour la composante blanche et  $\gamma_{\mathbf{c}}$  pour la composante corrélée sont des paramètres de précision des gaussiennes. La loi conjuguée, utilisée avec succès dans la littérature [Fér06, BMK09, Rob05] en problèmes inverses et en statistique bayésienne, est la loi gamma

$$p(\gamma_{\mathbf{x}_k}) \sim \mathcal{G}(\alpha_{\mathbf{x}}, \beta_{\mathbf{x}}) = \frac{1}{\beta_{\mathbf{x}}^{\alpha_{\mathbf{x}}} \Gamma(\alpha_{\mathbf{x}})} \gamma_{\mathbf{x}_k}^{\alpha_{\mathbf{x}}-1} \exp \left[ -\frac{\gamma_{\mathbf{x}_k}}{\beta_{\mathbf{x}}} \right] \quad (3.24)$$

$$p(\gamma_{\mathbf{b}_b}) \sim \mathcal{G}(\alpha_{\mathbf{b}}, \beta_{\mathbf{b}}) = \frac{1}{\beta_{\mathbf{b}}^{\alpha_{\mathbf{b}}} \Gamma(\alpha_{\mathbf{b}})} \gamma_{\mathbf{b}_b}^{\alpha_{\mathbf{b}}-1} \exp \left[ -\frac{\gamma_{\mathbf{b}_b}}{\beta_{\mathbf{b}}} \right], \quad (3.25)$$

et

$$p(\gamma_{\mathbf{c}}) \sim \mathcal{G}(\alpha_{\mathbf{c}}, \beta_{\mathbf{c}}) = \frac{1}{\beta_{\mathbf{c}}^{\alpha_{\mathbf{c}}} \Gamma(\alpha_{\mathbf{c}})} \gamma_{\mathbf{c}}^{\alpha_{\mathbf{c}}-1} \exp \left[ -\frac{\gamma_{\mathbf{c}}}{\beta_{\mathbf{c}}} \right]. \quad (3.26)$$

où  $(\alpha_i, \beta_i)$  sont des paramètres de la loi. La loi gamma présente deux avantages.

Tout d'abord il s'agit d'une loi conjuguée pour ces paramètres donc la loi conditionnelle *a posteriori* est également une loi gamma scalaire. Deuxièmement, en fonction des paramètres de la loi *a priori* on peut obtenir deux lois dites non-informatives : la loi uniforme  $p(\gamma) = \mathcal{U}_{[0, +\infty]}(\gamma)$  avec  $(\alpha_i = 0, \beta_i = +\infty)$  et la loi de Jeffreys  $p(\gamma) = 1/\gamma$  avec  $(1, +\infty)$ . Ce sont deux lois impropres (non intégrables). Cependant la loi de Jeffreys est considérée comme une loi non-informative [KW96] et elle est invariante par transformation de la forme  $\eta = \gamma^n$ , autrement dit si la variance suit une loi de Jeffreys, c'est également le cas de la précision, de l'écart-type, etc. L'information injectée ne dépend pas de la paramétrisation du problème.

L'impropreté de la loi *a priori* est un problème puisque nous ne sommes plus en mesure d'assurer la propreté de la loi *a posteriori*. En pratique, le nombre d'inconnues pilotées par les paramètres de

puissance  $\gamma$  est très important (un seul paramètre  $\gamma_{\mathbf{x}_k}$  influence tous les coefficients d'un ordre de Taylor). La vraisemblance apporte donc énormément d'information sur ces paramètres ce qui rassure quant à la propriété de l'*a posteriori*. De plus, il est toujours possible de choisir des paramètres  $(\alpha, \beta)$  aussi proches que l'on souhaite de la loi de Jeffreys tout en n'étant pas à la divergence.

En ce qui concerne les autres paramètres, il n'existe pas de loi conjuguée, ni pour les paramètres de forme des densités spectrales *a priori*  $f_c$  et  $d$ , ni pour les paramètres de forme de la réponse instrumentale  $\theta_i$ . Le choix est donc uniquement orienté par des informations d'experts. Si nous avons à disposition une information sur le fait que les paramètres évoluent dans un intervalle, la loi uniforme est adaptée. On peut également disposer d'une valeur moyenne et d'une incertitude qui conduisent à une loi normale. En pratique nous avons choisi des gaussiennes pour chacun des paramètres. La moyenne et la variance ont été fournies à chaque fois par les experts de l'instrument ou alors fixées de manière *ad-hoc* pour les besoins de l'étude. De plus nous avons considéré que tous les paramètres sont *a priori* indépendants entre eux.

Toutes les lois nécessaires pour construire la loi jointe sont définies et donc la loi *a posteriori* peut être établie. Aucune information *a priori* n'a été introduite sur un lien entre les trois bandes spectrales de SPIRE, que ce soit par l'intermédiaire d'un modèle pour le ciel ou d'un modèle probabiliste. De plus les données d'une bande spectrale ne contiennent que très peu d'information sur les bandes adjacentes, voir partie 2.2.2. L'inférence s'effectue donc sur un problème où les trois bandes spectrales pour le ciel sont indépendantes.

Le protocole considéré est le protocole *Large map*. Toutes les données acquises considérées valides sont utilisées pour résoudre le problème. La solution employée repose sur un modèle d'acquisition à données manquantes c'est-à-dire que le modèle  $\mathbf{H}$  ne reproduit pas ces données.

La loi jointe pour une bande spectrale s'écrit

$$p(\mathbf{x}, \mathbf{c}, \gamma_{\mathbf{x}}, \gamma_{\mathbf{b}}, \gamma_{\mathbf{c}}, f_c, d, \theta_i, \mathbf{y}) = p(\mathbf{y} | \mathbf{x}, \mathbf{c}, \gamma_{\mathbf{b}}, \theta_i) p(\mathbf{x} | \gamma_{\mathbf{x}}) p(\mathbf{c} | \gamma_{\mathbf{c}}, f_c, d) p(\gamma_{\mathbf{x}}) p(\gamma_{\mathbf{b}}) p(\gamma_{\mathbf{c}}) p(f_c) p(d) p(\theta_i) \quad (3.27)$$

et la loi *a posteriori*

$$p(\mathbf{x}, \mathbf{c}, \gamma_{\mathbf{x}}, \gamma_{\mathbf{b}}, \gamma_{\mathbf{c}}, f_c, d, \theta_i | \mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{c}, \gamma_{\mathbf{x}}, \gamma_{\mathbf{b}}, \gamma_{\mathbf{c}}, f_c, d, \theta_i, \mathbf{y})}{p(\mathbf{y})}. \quad (3.28)$$

où

- $\mathbf{x} \in \mathbb{R}^{I \times J \times K}$  sont les coefficients de décomposition du ciel avec  $K$  l'ordre de développement de Taylor en longueur d'onde,  $I$  et  $J$  le nombre de coefficients de décomposition dans les dimensions spatiales. La loi *a priori* pour un développement de Taylor sur l'objet est déjà connue (3.16) page 81.
- $\mathbf{c} \in \mathbb{R}^N$  est la composante basse-fréquence et  $N$  le nombre d'échantillons temporels. La loi de la composante  $\mathbf{c}$  a été établie partie 3.1.1 et est une loi normale centrée de matrice de covariance  $\gamma_{\mathbf{c}}^{-1} \Sigma_{\mathbf{c}}$  voir équation (3.7) page 76.
- $\gamma_{\mathbf{x}} \in \mathbb{R}^K$  contient les  $K$  paramètres de puissances de la corrélation spatiale *a priori* pour l'objet d'intérêt. La loi *a priori* est une loi gamma, tout comme les paramètres de puissance suivants.
- $\gamma_{\mathbf{b}} \in \mathbb{R}^{B+2}$  correspond aux  $B+2$  paramètres de puissance du bruit blanc des  $B+2$  bolomètres d'un photomètre.
- $\gamma_{\mathbf{c}} \in \mathbb{R}$  est la puissance de la corrélation temporelle *a priori* de la composante basse-fréquence. La fréquence de coupure est  $f_c \in \mathbb{R}$  et la décroissante  $d \in \mathbb{R}$ .
- $\theta_i$  regroupe les paramètres de la réponse instrumentale que l'on souhaite estimer. Les lois *a priori* de ces paramètres sont des lois gaussiennes mono-variées.

– et  $\mathbf{y} \in \mathbb{R}^{(B+2) \times N}$  rassemble toutes les données produites par les  $B + 2$  bolomètres à  $N$  instants temporels. La loi des données sachant le reste est une loi gaussienne dont la matrice de covariance est bloc stationnaire, voir équation (3.23) page 88

Finalement, à partir des équations (3.7), (3.16), (3.23), et (3.24) à (3.26), et sans les termes constants de normalisation, la loi jointe s'écrit

$$\begin{aligned}
 p(\mathbf{x}, \mathbf{c}, \gamma_{\mathbf{x}}, \gamma_{\mathbf{b}}, \gamma_{\mathbf{c}}, f_c, d, \boldsymbol{\theta}_i, \mathbf{y}) &\propto \gamma_{\mathbf{c}}^{\alpha_{\mathbf{c}} + N/2 - 1} \det[\boldsymbol{\Sigma}_{\mathbf{c}}]^{-1/2} \prod_{b=0}^{B+1} \gamma_{\mathbf{b}_b}^{\alpha_{\mathbf{b}} + N/2 - 1} \prod_{k=0}^{K-1} \gamma_{\mathbf{x}_k}^{\alpha_{\mathbf{x}} + (IJ-1)/2 - 1} \\
 &\exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{m}_{\mathbf{y}})^t \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} (\mathbf{y} - \mathbf{m}_{\mathbf{y}}) \right] \exp \left[ -\frac{1}{2} \sum_{k=0}^K \gamma_{\mathbf{x}_k} \mathbf{x}_k^t \mathbf{D}_k \mathbf{x}_k \right] \exp \left[ -\frac{1}{2} \gamma_{\mathbf{c}} \mathbf{c}^t \boldsymbol{\Sigma}_{\mathbf{c}}^{-1} \mathbf{c} \right] \\
 &\exp \left[ -\sum_{k=0}^{K-1} \frac{\gamma_{\mathbf{x}_k}}{\beta_{\mathbf{x}}} - \sum_{b=0}^{B+1} \frac{\gamma_{\mathbf{b}_b}}{\beta_{\mathbf{b}}} - \frac{\gamma_{\mathbf{c}}}{\beta_{\mathbf{c}}} \right] \exp \left[ -\frac{(f_c - m_{f_c})^2}{2\sigma_{f_c}^2} - \frac{(d - m_d)^2}{2\sigma_d^2} \right] \prod_i \exp \left[ -\frac{(\theta_i - m_{\theta_i})^2}{2\sigma_{\theta_i}^2} \right].
 \end{aligned} \tag{3.29}$$

Cette loi jointe, dont on déduit la loi *a posteriori* en fixant  $\mathbf{y}$  aux données réellement collectées, rassemble toute l'information pour estimer les paramètres inconnus. Les algorithmes et les estimateurs présentés par la suite reposent entièrement sur cette loi. Malgré la lourdeur de l'expression elle est relativement simple. On trouve tout d'abord les termes de normalisation qui dépendent des paramètres. Ensuite les lois gaussiennes des données, de l'objet et de  $\mathbf{c}$  puis les lois des paramètres  $\boldsymbol{\theta}$ .

Cette loi est propre par construction, entre autres concernant l'objet. En effet comme décrit partie 3.1.2, les termes  $\mathbf{D}_k$  sont obtenus numériquement par transformée de Fourier discrète sur des noyaux à support limité. Étant de support infini par définition, le repliement fait que la fréquence nulle est pénalisée très faiblement, relativement aux autres fréquences. Toutes les lois composant la loi jointe sont donc intégrables. Il en est donc de même de la loi jointe et de la loi *a posteriori*.

Dans la prochaine partie nous présentons principalement deux algorithmes d'estimation de l'objet d'intérêt. Le premier correspond à l'estimation, au sens de la moyenne *a posteriori*, de tous les paramètres inconnus. Le second, cas particulier du premier, correspond aux moindres carrés régularisés pour l'objet  $\mathbf{x}$  et la composante  $\mathbf{c}$  lorsque les autres paramètres sont considérés connus.

## 3.2 Estimateur et algorithme de calcul

La loi jointe équation (3.29) est trop complexe pour être analysée dans son ensemble. Tout d'abord elle est définie sur des inconnues de dimension importante (le ciel est de taille  $IJK$ ). De plus elle ne peut pas être identifiée à une forme connue puisqu'elle est définie également sur les paramètres  $\boldsymbol{\gamma}$  et  $\boldsymbol{\theta}_i$ .

Le principe consiste à extraire l'information essentielle au travers d'un estimateur ponctuel. Le choix s'est porté sur la moyenne *a posteriori* pour des raisons algorithmiques. En effet les lois *a priori* définissent une structure de loi *a posteriori* qui peut être simulée facilement à l'aide d'un échantillonneur de Gibbs (voir partie 1.4.2).

Tout d'abord à partir de la loi jointe nous présenterons l'algorithme de calcul de la moyenne *a posteriori* pour tous les paramètres. En deuxième partie nous décrirons plus en détails la solution des moindres carrés régularisés pour  $\mathbf{x}$  et  $\mathbf{c}$  lorsque les autres paramètres sont fixés. Cette solution correspond en effet à l'étape d'échantillonnage de  $\mathbf{x}$  et  $\mathbf{c}$  dans l'échantillonneur de Gibbs.

### 3.2.1 Algorithme MCMC proposé

L'échantillonneur de Gibbs consiste à simuler des échantillons des lois *a posteriori* conditionnelles connaissant les échantillons précédents des autres paramètres. Il faut choisir un découpage de l'ensemble des inconnues pour déterminer les lois conditionnelles. En pratique le choix est très simple et consiste à simuler la conditionnelle de chaque paquet de paramètres.

**Échantillonnage de  $\mathbf{x}$  et  $\mathbf{c}$ .** Nous avons décidé de procéder à l'échantillonnage des deux conjointement pour des raisons algorithmiques qui seront présentées dans la prochaine partie.

La loi de ces paramètres conditionnellement aux autres à l'itération  $q$  est

$$p(\mathbf{x}, \mathbf{c} | \mathbf{y}, \gamma_{\mathbf{x}}^{(q)}, \gamma_{\mathbf{b}}^{(q)}, \gamma_{\mathbf{c}}^{(q)}, f_{\mathbf{c}}^{(q)}, \boldsymbol{\theta}_i^{(q)}) \propto \exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{m}_{\mathbf{y}})^t \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} (\mathbf{y} - \mathbf{m}_{\mathbf{y}}) \right] \exp \left[ -\frac{1}{2} \sum_{k=0}^K \gamma_{\mathbf{x}_k} \mathbf{x}_k^t \mathbf{D}_k \mathbf{x}_k \right] \exp \left[ -\frac{1}{2} \gamma_{\mathbf{c}} \mathbf{c}^t \boldsymbol{\Sigma}_{\mathbf{c}}^{-1} \mathbf{c} \right] \quad (3.30)$$

avec

$$\mathbf{m}_{\mathbf{y}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \mathbf{x} + \begin{pmatrix} \mathbf{T}_0 \\ \mathbf{T}_1 \\ \vdots \\ \mathbf{I} \\ \mathbf{I} \end{pmatrix} \mathbf{c} = \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix} \quad (3.31)$$

où  $\mathbf{H}$  dépend de  $\boldsymbol{\theta}_i^{(q)}$  et

$$\boldsymbol{\Sigma}_{\mathbf{y}}^{-1} = \begin{pmatrix} \gamma_{\mathbf{b}_0}^{(q)} \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \gamma_{\mathbf{b}_1}^{(q)} \mathbf{I} & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \gamma_{\mathbf{b}_{B+1}}^{(q)} \mathbf{I} \end{pmatrix}. \quad (3.32)$$

C'est une loi gaussienne puisque la vraisemblance est gaussienne ainsi que les deux lois *a priori*. On connaît donc sa moyenne

$$\mathbf{m}_{\mathbf{x}, \mathbf{c}} = \left( \begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \\ \mathbf{T}^t & \mathbf{I} & \mathbf{I} \end{pmatrix} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{x}}^{(q)} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{c}}^{(q)} \end{pmatrix}^{-1} \right)^{-1} \begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \\ \mathbf{T}^t & \mathbf{I} & \mathbf{I} \end{pmatrix} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{y}$$

ainsi que sa matrice de covariance

$$\boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{c}} = \left( \begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \\ \mathbf{T}^t & \mathbf{I} & \mathbf{I} \end{pmatrix} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{x}}^{(q)} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{c}}^{(q)} \end{pmatrix}^{-1} \right)^{-1}.$$

Même si les écritures sont un peu chargées, la principale difficulté concernant l'échantillonnage de cette loi se situe au niveau du modèle instrument  $\mathbf{H}$  qui n'est pas nécessairement invariant. Dans le cas contraire nous aurions pu diagonaliser l'opérateur  $\mathbf{H}$  dans l'espace de Fourier et toutes les composantes de  $\mathbf{F}\mathbf{x}$  auraient été indépendantes.

Une solution existant pour l'échantillonnage de loi normale multivariée repose sur la décomposition de Cholesky pour des matrices symétriques définies positives. Ce n'est pas une solution ici de par la grande dimension de  $\boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{c}}$  sans parler du calcul de la moyenne. Une autre solution

classique repose sur une échantillonneur de Gibbs composante par composante qui s'avererait bien trop long.

La solution retenue pour obtenir un échantillon repose sur un calcul d'optimisation pour des valeurs de paramètres de loi perturbés. Cette étape est décrite dans la prochaine partie 3.2.2 et repose sur l'optimisation d'un système perturbé.

**Échantillonnage des paramètres de puissances  $\gamma_{\mathbf{x}}$ ,  $\gamma_{\mathbf{b}}$  et  $\gamma_{\mathbf{c}}$ .** La loi conditionnelle des paramètres  $\gamma_{\mathbf{x}}$  s'écrit

$$\begin{aligned} p(\gamma_{\mathbf{x}_k} | \mathbf{y}, \mathbf{x}, \mathbf{c}, \mathbf{g}, \gamma_{\mathbf{c}}, f_{\mathbf{c}}, d, \boldsymbol{\theta}_i) &= p(\gamma_{\mathbf{x}_k} | \mathbf{y}, \mathbf{x}^{(q+1)}) \\ &\propto \gamma_{\mathbf{x}_k}^{\alpha_{\mathbf{x}} + (IJ-1)/2-1} \exp\left[-\frac{1}{2}\gamma_{\mathbf{x}_k} \mathbf{x}_k^t \mathbf{D}_k \mathbf{x}_k\right] \exp\left[-\frac{\gamma_{\mathbf{x}_k}}{\beta_{\mathbf{x}}}\right] \\ &\propto \gamma_{\mathbf{x}_k}^{\alpha_{\mathbf{x}} + (IJ-1)/2-1} \exp\left[-\gamma_{\mathbf{x}_k} \left(\frac{1}{2} \mathbf{x}_k^t \mathbf{D}_k \mathbf{x}_k + \frac{1}{\beta_{\mathbf{x}}}\right)\right] \end{aligned} \quad (3.33)$$

qui est comme attendu une loi gamma

$$p(\gamma_{\mathbf{x}_k}) \sim \mathcal{G}\left(\alpha_{\mathbf{x}} + \frac{IJ-1}{2}, \left(\frac{1}{2} \mathbf{x}_k^t \mathbf{D}_k \mathbf{x}_k + \frac{1}{\beta_{\mathbf{x}}}\right)^{-1}\right). \quad (3.34)$$

Comme nous avons fait le choix de décomposer le ciel sur des gaussiennes espacées régulièrement, le calcul de  $\mathbf{x}_k^t \mathbf{D}_k \mathbf{x}_k$  peut se faire soit par convolution soit par transformée de Fourier discrète. Le même raisonnement s'applique aux autres paramètres  $\gamma_{\mathbf{b}}$  et  $\gamma_{\mathbf{c}}$  soit

$$p(\gamma_{\mathbf{b}_b}) \sim \mathcal{G}\left(\alpha_{\mathbf{b}} + \frac{N}{2}, \left(\frac{1}{2} \|\mathbf{y}_{\mathbf{b}_b} - \mathbf{m}_{\mathbf{y}_{\mathbf{b}_b}}\|^2 + \frac{1}{\beta_{\mathbf{b}}}\right)^{-1}\right) \quad (3.35)$$

où  $\mathbf{y}_{\mathbf{b}}$  correspond aux données d'un bolomètre  $b$  particulier et

$$\mathbf{m}_{\mathbf{y}_{\mathbf{b}}} = \begin{cases} \mathbf{H}_{\mathbf{b}} \mathbf{x}^{(q+1)} + \mathbf{T}_{\mathbf{b}} \mathbf{c}^{(q+1)} & \text{si c'est un bolomètre standard} \\ \mathbf{c}^{(q+1)} & \text{si c'est un bolomètre aveugle} \end{cases}$$

et

$$p(\gamma_{\mathbf{c}}) \sim \mathcal{G}\left(\alpha_{\mathbf{c}} + \frac{N}{2}, \left(\frac{1}{2} \gamma_{\mathbf{c}} \mathbf{c}^t \boldsymbol{\Sigma}_{\mathbf{c}}^{-1} \mathbf{c} + \frac{1}{\beta_{\mathbf{c}}}\right)^{-1}\right). \quad (3.36)$$

Si les lois *a priori* sont des lois de Jeffreys alors

$$\begin{aligned} p(\gamma_{\mathbf{x}_k}) &\sim \mathcal{G}\left(\frac{IJ-1}{2}, \frac{2}{\|\mathbf{D}_k^{1/2} \mathbf{x}_k\|^2}\right) \\ p(\gamma_{\mathbf{b}_b}) &\sim \mathcal{G}\left(\frac{N}{2}, \frac{2}{\|\mathbf{y}_{\mathbf{b}_b} - \mathbf{m}_{\mathbf{y}_{\mathbf{b}_b}}\|^2}\right) \\ p(\gamma_{\mathbf{c}}) &\sim \mathcal{G}\left(\frac{N}{2}, \frac{2}{\|\boldsymbol{\Sigma}_{\mathbf{c}}^{-1/2} \mathbf{c}\|^2}\right). \end{aligned}$$

**Échantillonnage des paramètres aux lois complexes ( $f_{\mathbf{c}}$ ,  $d$  et  $\boldsymbol{\theta}_i$ ).** La loi conditionnelle de la fréquence de coupure  $f_{\mathbf{c}}$  est

$$\begin{aligned} p(f_{\mathbf{c}} | \mathbf{y}, \mathbf{x}, \mathbf{c}, \gamma_{\mathbf{x}}, \gamma_{\mathbf{b}}, \gamma_{\mathbf{c}}, d, \boldsymbol{\theta}_i) &\propto p(f_{\mathbf{c}} | \mathbf{c}^{(q+1)}, \gamma_{\mathbf{c}}^{(q+1)}) \\ &\propto \det[\boldsymbol{\Sigma}_{\mathbf{c}}^{-1/2}] \exp\left[-\frac{1}{2} \gamma_{\mathbf{c}} \mathbf{c}^t \boldsymbol{\Sigma}_{\mathbf{c}}^{-1} \mathbf{c}\right] \exp\left[-\frac{1}{2} \frac{(f_{\mathbf{c}} - m_{f_{\mathbf{c}}})^2}{\sigma_{f_{\mathbf{c}}}^2}\right] \end{aligned} \quad (3.37)$$

et n'est pas une loi usuelle puisque le paramètre  $f_c$ , ainsi que  $d$ , apparaît de façon complexe dans la matrice  $\Sigma_c$ . De même pour les paramètres  $\theta_i$  dont la loi conditionnelle

$$\begin{aligned} p(\theta_i | \mathbf{y}, \mathbf{x}, \mathbf{c}, \gamma_x, \gamma_b, \gamma_c, f_c) &\propto p(\theta_i | \mathbf{y}, \mathbf{x}^{(q+1)}, \mathbf{c}^{(q+1)}, \gamma_b^{(q+1)}) \\ &\propto \exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{m}_y)^t \Sigma_y^{-1} (\mathbf{y} - \mathbf{m}_y) \right] \exp \left[ -\frac{1}{2} \frac{(\theta_i - m_{\theta_i})^2}{\sigma_{\theta_i}^2} \right] \end{aligned} \quad (3.38)$$

n'est pas non plus une loi usuelle. Le paramètre  $\theta_i$  est en effet dissimulé au sein de la matrice  $\mathbf{H}$  dans la moyenne

$$\mathbf{m}_y = \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix}.$$

L'algorithme proposé est un algorithme de Metropolis-Hastings. Différentes formes sont possibles. La forme utilisée ici est la forme indépendante avec pour loi de proposition (ou la loi instrumentale) la loi *a priori*. L'algorithme proposé pour le paramètre  $f_c$  se décompose en plusieurs étapes.

1. Proposer un échantillon  $f_c^p$  sous la loi *a priori*  $f_c^p \sim \mathcal{N}(m_{f_c}, \sigma_{f_c}^2)$  et en déduire  $\Sigma_c^p$ .
2. Calculer le rapport de vraisemblance (la loi instrumentale étant la loi *a priori*)

$$P = \frac{\det [\Sigma_c^p]^{-1/2} \exp \left[ -\frac{1}{2} \gamma_c^{(q+1)} \mathbf{c}^{(q+1)t} \Sigma_c^{p-1} \mathbf{c}^{(q+1)} \right]}{\det [\Sigma_c^{(q)}]^{-1/2} \exp \left[ -\frac{1}{2} \gamma_c^{(q+1)} \mathbf{c}^{(q+1)t} \Sigma_c^{(q)-1} \mathbf{c}^{(q+1)} \right]} \quad (3.39)$$

3. Simuler  $\eta \sim \mathcal{U}_{[0, 1]}(\eta)$  et prendre

$$f_c^{(q+1)} = \begin{cases} f_c^p & \text{si } \eta < \min\{1, P\} \\ f_c^{(q)} & \text{sinon.} \end{cases}$$

Si  $\Sigma_c$  est circulante alors

$$\det [\Sigma_c^{-1/2}] = \det[\mathbf{F}^\dagger] \det [\Lambda_c^{1/2}] \det[\mathbf{F}] = [\Lambda_c^{1/2}] = \prod_{i=0}^{N-1} \lambda_{c_i}^{1/2}$$

le déterminant est obtenu à l'aide du produit des valeurs propres. De plus l'expression des coefficients de la matrice  $\Lambda_c$  est explicite (voir partie 3.1.1), quoique sans duplication dû à l'échantillonnage. En ce qui concerne l'échantillonnage de  $d$ , l'algorithme est en tout point similaire excepté la loi instrumentale qui est la loi *a priori*. Par contre le calcul du rapport de vraisemblance  $P$  est identique.

La solution pour obtenir des échantillons de  $\theta_i$  repose également sur un algorithme de Metropolis-Hastings. L'algorithme est de nouveau similaire au cas de l'échantillonnage de  $f_c$ , le rapport étant cette fois

$$P = \frac{\exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{m}_y^p)^t \Sigma_y^{-1} (\mathbf{y} - \mathbf{m}_y^p) \right]}{\exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{m}_y^{(q)})^t \Sigma_y^{-1} (\mathbf{y} - \mathbf{m}_y^{(q)}) \right]}.$$

Le calcul de ce rapport peut s'avérer coûteux puisqu'il faut non seulement calculer l'opérateur instrument  $\mathbf{H}^p$  mais également calculer une sortie modèle correspondante (bien que dans certains cas les deux soient effectués en même temps). La partie 3.3 présente sous certaines approximations des moyens pour réduire le coût en calculs en étudiant finement l'invariance du modèle. Le gain est cependant relatif puisqu'il peut prendre plusieurs secondes pour chaque étape sur un problème de taille assez réduite comme nous verrons dans le chapitre 4.

**Calcul de la moyenne *a posteriori*.** A l'aide des échantillons fournis la moyenne *a posteriori* peut être approchée par

$$\hat{\mathbf{x}}_{\text{EAP}} \approx \frac{1}{Q} \sum_{q=0}^{Q-1} \mathbf{x}^{(q)}.$$

Les autres cumulants peuvent être également approchés. Le plus souvent on s'intéresse à la diagonale de la matrice de covariance *a posteriori* marginale

$$\mathbf{V}[\mathbf{x}]_{\text{EAP}} = \mathbb{E} \left[ (\mathbf{x} - \mathbb{E}[\mathbf{x}]) (\mathbf{x} - \mathbb{E}[\mathbf{x}])^t \right]$$

pour fournir une incertitude sur l'estimation, une « carte d'erreur ». Cette diagonale peut être calculée par

$$\hat{\mathbf{V}}[\mathbf{x}]_{\text{EAP}} \approx \frac{1}{Q} \sum_{q=0}^{Q-1} \left( \mathbf{x}^{(q)} \odot \mathbf{x}^{(q)} \right) - \hat{\mathbf{x}}_{\text{EAP}}^2$$

où  $\odot$  est le produit terme à terme de chaque composante du vecteur. Nous avons vu que l'étape d'échantillonnage de la loi conditionnelle pour  $\mathbf{x}$  et  $\mathbf{c}$  est complexe à cause de la dimension du problème et de la non invariance du modèle instrument. La prochaine partie présente une solution basée sur un problème d'optimisation.

L'étape la plus complexe ici est clairement l'échantillonnage de l'objet sous une loi gaussienne multivariée dont la corrélation fait intervenir le modèle instrument et l'*a priori*. Les deux prochaines parties décrivent en détail l'approche utilisée pour résoudre cette difficulté.

### 3.2.2 Échantillonnage de l'objet par optimisation

L'essentiel de cette partie est tiré de la thèse d'Olivier Féron [Fér06], Annexe C. L'échantillonnage de la loi conditionnelle

$$p(\mathbf{x}, \mathbf{c} | \mathbf{y}, \gamma_{\mathbf{x}}^{(q)}, \gamma_{\mathbf{b}_0}^{(q)}, \gamma_{\mathbf{c}}^{(q)}, f_{\mathbf{c}}^{(q)}, \theta_1^{(q)}) \sim \mathcal{N}(\mathbf{m}_{\mathbf{x}, \mathbf{c}}, \boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{c}})$$

avec

$$\mathbf{m}_{\mathbf{x}, \mathbf{c}} = \boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{c}} \begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \\ \mathbf{T}^t & \mathbf{I} & \mathbf{I} \end{pmatrix} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{y}$$

et

$$\boldsymbol{\Sigma}_{\mathbf{x}, \mathbf{c}} = \left( \begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \\ \mathbf{T}^t & \mathbf{I} & \mathbf{I} \end{pmatrix} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{x}}^{(q)} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{c}}^{(q)} \end{pmatrix}^{-1} \right)^{-1}$$

est une gaussienne multivariée avec une structure de covariance complexe. En considérant une vraisemblance générale de la forme  $p(\mathbf{y} | \mathbf{x}) \sim \mathcal{N}(\mathbf{m}_{\mathbf{y}}, \boldsymbol{\Sigma}_{\mathbf{y}})$  où  $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{b}$  et une loi *a priori*  $p(\mathbf{x}) \sim \mathcal{N}(\mathbf{m}_{\mathbf{x}}, \boldsymbol{\Sigma}_{\mathbf{x}})$ , alors on sait que la loi *a posteriori* a pour covariance

$$\boldsymbol{\Sigma}_{\mathbf{x} | \mathbf{y}} = \left( \mathbf{H}^t \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{H} + \boldsymbol{\Sigma}_{\mathbf{x}}^{-1} \right)^{-1}$$

et moyenne

$$\hat{\mathbf{x}} = \left( \mathbf{H}^t \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{H} + \boldsymbol{\Sigma}_{\mathbf{x}}^{-1} \right)^{-1} \left( \mathbf{H}^t \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{y} + \boldsymbol{\Sigma}_{\mathbf{x}}^{-1} \mathbf{m}_{\mathbf{x}} \right).$$

Dans notre cas  $\boldsymbol{\Sigma}_{\mathbf{y}}$  est proche de l'identité et  $\mathbf{m}_{\mathbf{x}}$  est nulle. La moyenne peut être calculée par un algorithme de minimisation comme une descente de gradient (voir partie 1.4.1). L'idée est de faire en sorte que le calcul fournisse un échantillon de la loi *a posteriori* cible et non la moyenne  $\hat{\mathbf{x}}$ .

Pour cela il suffit de « perturber le système » en introduisant des fluctuations dans  $\mathbf{y}$  et  $\mathbf{m}_x$ . Considérons les nouvelles variables « perturbées »

$$\begin{aligned}\tilde{\mathbf{y}} &= \mathbf{y} + \delta_{\mathbf{y}} \\ \widetilde{\mathbf{m}}_x &= \mathbf{m}_x + \delta_{\mathbf{m}_x}\end{aligned}$$

où  $\delta_{\mathbf{y}} \sim \mathcal{N}(0, \Sigma_{\mathbf{y}})$  et  $\delta_{\mathbf{m}_x} \sim \mathcal{N}(0, \Sigma_{\mathbf{x}})$ . Alors la variable

$$\tilde{\mathbf{x}} = \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} \right)^{-1} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \tilde{\mathbf{y}} + \Sigma_{\mathbf{x}}^{-1} \widetilde{\mathbf{m}}_x \right), \quad (3.40)$$

qui peut être calculée par optimisation du critère

$$J(\mathbf{x}) = (\tilde{\mathbf{y}} - \mathbf{H}\mathbf{x})^t \Sigma_{\mathbf{y}}^{-1} (\tilde{\mathbf{y}} - \mathbf{H}\mathbf{x}) + (\mathbf{x} - \widetilde{\mathbf{m}}_x)^t \Sigma_{\mathbf{x}}^{-1} (\mathbf{x} - \widetilde{\mathbf{m}}_x),$$

suit une loi gaussienne (puisque tout est gaussien), de moyenne

$$\begin{aligned}\mathbb{E}[\tilde{\mathbf{x}}] &= \mathbb{E} \left[ \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} \right)^{-1} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \tilde{\mathbf{y}} + \Sigma_{\mathbf{x}}^{-1} \widetilde{\mathbf{m}}_x \right) \right] \\ &= \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} \right)^{-1} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbb{E}[\tilde{\mathbf{y}}] + \Sigma_{\mathbf{x}}^{-1} \mathbb{E}[\widetilde{\mathbf{m}}_x] \right) \\ &= \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} \right)^{-1} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{y} + \Sigma_{\mathbf{x}}^{-1} \mathbf{m}_x \right) \\ &= \Sigma_{\mathbf{x}|\mathbf{y}} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{y} + \Sigma_{\mathbf{x}}^{-1} \mathbf{m}_x \right)\end{aligned}$$

qui est bien la moyenne de la loi cible. De même le deuxième moment s'obtient

$$\begin{aligned}\mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^t] &= \mathbb{E} \left[ \Sigma_{\mathbf{x}|\mathbf{y}} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \tilde{\mathbf{y}} + \Sigma_{\mathbf{x}}^{-1} \widetilde{\mathbf{m}}_x \right) \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \tilde{\mathbf{y}} + \Sigma_{\mathbf{x}}^{-1} \widetilde{\mathbf{m}}_x \right)^t \Sigma_{\mathbf{x}|\mathbf{y}} \right] \\ &= \Sigma_{\mathbf{x}|\mathbf{y}} \mathbb{E} \left[ \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \tilde{\mathbf{y}} \tilde{\mathbf{y}}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \tilde{\mathbf{y}} \widetilde{\mathbf{m}}_x^t \Sigma_{\mathbf{x}}^{-1} + \right. \\ &\quad \left. \Sigma_{\mathbf{x}}^{-1} \widetilde{\mathbf{m}}_x \tilde{\mathbf{y}}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} \widetilde{\mathbf{m}}_x \widetilde{\mathbf{m}}_x^t \Sigma_{\mathbf{x}}^{-1} \right] \Sigma_{\mathbf{x}|\mathbf{y}} \\ &= \Sigma_{\mathbf{x}|\mathbf{y}} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbb{E}[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^t] \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} \mathbb{E}[\widetilde{\mathbf{m}}_x \widetilde{\mathbf{m}}_x^t] \Sigma_{\mathbf{x}}^{-1} \right) \Sigma_{\mathbf{x}|\mathbf{y}}\end{aligned}$$

puisque  $\tilde{\mathbf{y}}$  et  $\widetilde{\mathbf{m}}_x$  sont indépendants. Or

$$\begin{aligned}\mathbb{E}[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^t] &= \mathbb{E} \left[ (\mathbf{y} + \delta_{\mathbf{y}}) (\mathbf{y} + \delta_{\mathbf{y}})^t \right] \\ &= \mathbf{y}\mathbf{y}^t + \mathbb{E}[\delta_{\mathbf{y}}\delta_{\mathbf{y}}^t] \\ &= \mathbf{y}\mathbf{y}^t + \Sigma_{\mathbf{y}}\end{aligned}$$

et de même

$$\mathbb{E}[\widetilde{\mathbf{m}}_x \widetilde{\mathbf{m}}_x^t] = \mathbf{m}_x \mathbf{m}_x^t + \Sigma_{\mathbf{x}}.$$

Donc le deuxième moment s'écrit

$$\begin{aligned}\mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^t] &= \Sigma_{\mathbf{x}|\mathbf{y}} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{y}\mathbf{y}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} \mathbf{m}_x \mathbf{m}_x^t \Sigma_{\mathbf{x}}^{-1} \right) \Sigma_{\mathbf{x}|\mathbf{y}} + \Sigma_{\mathbf{x}|\mathbf{y}} \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} \Sigma_{\mathbf{x}|\mathbf{y}} + \Sigma_{\mathbf{x}|\mathbf{y}} \Sigma_{\mathbf{x}}^{-1} \Sigma_{\mathbf{x}|\mathbf{y}} \\ &= \mathbb{E}[\tilde{\mathbf{x}}] \mathbb{E}[\tilde{\mathbf{x}}]^t + \Sigma_{\mathbf{x}|\mathbf{y}} \left( \mathbf{H}^t \Sigma_{\mathbf{y}}^{-1} \mathbf{H} + \Sigma_{\mathbf{x}}^{-1} \right) \Sigma_{\mathbf{x}|\mathbf{y}} \\ &= \mathbb{E}[\tilde{\mathbf{x}}] \mathbb{E}[\tilde{\mathbf{x}}]^t + \Sigma_{\mathbf{x}|\mathbf{y}} \Sigma_{\mathbf{x}|\mathbf{y}} \Sigma_{\mathbf{x}|\mathbf{y}}\end{aligned}$$

et la matrice de covariance

$$\mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^t] - \mathbb{E}[\tilde{\mathbf{x}}] \mathbb{E}[\tilde{\mathbf{x}}]^t = \Sigma_{\mathbf{x}|\mathbf{y}}$$

correspond bien à la matrice de covariance de la loi cible. Donc le calcul du minimum de

$$J(\mathbf{x}) = (\tilde{\mathbf{y}} - \mathbf{H}\mathbf{x})^t \Sigma_{\mathbf{y}}^{-1} (\tilde{\mathbf{y}} - \mathbf{H}\mathbf{x}) + (\mathbf{x} - \widetilde{\mathbf{m}}_x)^t \Sigma_{\mathbf{x}}^{-1} (\mathbf{x} - \widetilde{\mathbf{m}}_x), \quad (3.41)$$

par optimisation fournit bien un échantillon de la loi *a posteriori*. Il s'agit donc de calculer la solution des moindres carrés régularisés avec des paramètres perturbés. Si le système n'est pas perturbé, la solution calculée correspond au moindres carrés régularisés à hyper-paramètres fixés.

### 3.2.3 Critère et gradients pour l'optimisation

Pour appliquer cet algorithme d'échantillonnage il faut considérer les variables perturbées

$$\begin{aligned}\tilde{\mathbf{y}} &= \mathbf{y} + \delta_{\mathbf{y}} \\ \widetilde{\mathbf{m}}_{\mathbf{x}} &= \delta_{\mathbf{x}},\end{aligned}$$

et

$$\widetilde{\mathbf{m}}_{\mathbf{c}} = \delta_{\mathbf{c}}$$

où  $\delta_{\mathbf{x}} \sim \mathcal{N}(0, \Sigma_{\mathbf{x}})$  et  $\delta_{\mathbf{c}} \sim \mathcal{N}(0, \gamma_{\mathbf{c}}^{-1} \Sigma_{\mathbf{c}})$ . La perturbation consiste donc à échantillonner la loi *a priori*. Il faut ensuite calculer l'optimum du critère (3.41)

$$\begin{aligned}J(\mathbf{x}, \mathbf{c}) &= \left( \tilde{\mathbf{y}} - \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix} \right)^t \Sigma_{\mathbf{y}}^{-1} \left( \tilde{\mathbf{y}} - \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix} \right) + \\ &\quad (\mathbf{x} - \widetilde{\mathbf{m}}_{\mathbf{x}})^t \Sigma_{\mathbf{x}}^{-1} (\mathbf{x} - \widetilde{\mathbf{m}}_{\mathbf{x}}) + \gamma_{\mathbf{c}} (\mathbf{c} - \widetilde{\mathbf{m}}_{\mathbf{c}})^t \Sigma_{\mathbf{c}}^{-1} (\mathbf{c} - \widetilde{\mathbf{m}}_{\mathbf{c}}).\end{aligned}\quad (3.42)$$

On peut réécrire ce critère en prenant compte du fait que  $\Sigma_{\mathbf{y}}$  est composée de blocs proportionnels à l'identité soit

$$\begin{aligned}J(\mathbf{x}, \mathbf{c}) &= \sum_{b=0}^{B-1} \gamma_{b_b} \|\tilde{\mathbf{y}}_{b_b} - \mathbf{H}_{b_b} \mathbf{x} - \mathbf{T}_{b_b} \mathbf{c}\|^2 + \sum_{b=B}^{B+1} \gamma_{b_b} \|\tilde{\mathbf{y}}_{b_b} - \mathbf{c}\|^2 + \\ &\quad \sum_{k=0}^K \gamma_{x_k} (\mathbf{x}_k - \widetilde{\mathbf{m}}_{x_k})^t \mathbf{D}_k (\mathbf{x}_k - \widetilde{\mathbf{m}}_{x_k}) + \gamma_{\mathbf{c}} (\mathbf{c} - \widetilde{\mathbf{m}}_{\mathbf{c}})^t \Sigma_{\mathbf{c}}^{-1} (\mathbf{c} - \widetilde{\mathbf{m}}_{\mathbf{c}})\end{aligned}\quad (3.43)$$

où  $\tilde{\mathbf{y}}_{b_b}$  correspond aux données perturbées du bolomètre 'b',  $\mathbf{H}_{b_b}$  au modèle d'acquisition correspondant,  $\mathbf{x}$  à tous les coefficients du ciel avec  $K$  ordre  $\mathbf{x}_k$  de corrélation respective  $\mathbf{D}_k$  et  $\mathbf{c}$  la composante basse fréquence de corrélation  $\Sigma_{\mathbf{c}}$ . Les termes de régularisation peuvent également se calculer dans l'espace de Fourier. La solution explicite correspondant à (3.40) n'est pas calculable directement. Pour cela nous allons utiliser un algorithme de descente de gradient pour calculer l'argument du minimum.

L'algorithme du gradient, décrit partie 1.4.1, a besoin du gradient  $\mathbf{g}$  de ce critère. Pour cela on pose

$$\tilde{\mathbf{H}} = \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \quad \mathbf{x}_e = \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix}, \quad \mathbf{m}_{\mathbf{x}_e} = \begin{pmatrix} \widetilde{\mathbf{m}}_{\mathbf{x}} \\ \widetilde{\mathbf{m}}_{\mathbf{c}} \end{pmatrix}, \quad \Sigma_{\mathbf{x}_e}^{-1} = \begin{pmatrix} \Sigma_{\mathbf{x}}^{-1} & \mathbf{0} \\ \mathbf{0} & \gamma_{\mathbf{c}} \Sigma_{\mathbf{c}}^{-1} \end{pmatrix}$$

et le critère à dériver devient

$$J(\mathbf{x}, \mathbf{c}) = (\tilde{\mathbf{y}} - \tilde{\mathbf{H}} \mathbf{x}_e)^t \Sigma_{\mathbf{y}}^{-1} (\tilde{\mathbf{y}} - \tilde{\mathbf{H}} \mathbf{x}_e) + (\mathbf{x}_e - \mathbf{m}_{\mathbf{x}_e})^t \Sigma_{\mathbf{x}_e}^{-1} (\mathbf{x}_e - \mathbf{m}_{\mathbf{x}_e}).$$

La dérivée par rapport  $\mathbf{x}_e$  est connue

$$\frac{\partial}{\partial \mathbf{x}_e} J(\mathbf{x}, \mathbf{c}) = 2 \tilde{\mathbf{H}}^t \Sigma_{\mathbf{y}}^{-1} (\tilde{\mathbf{H}} \mathbf{x}_e - \tilde{\mathbf{y}}) + 2 \Sigma_{\mathbf{x}_e}^{-1} (\mathbf{x}_e - \mathbf{m}_{\mathbf{x}_e}) = \mathbf{g}.$$

On remplace alors par les définitions respectives et on obtient le gradient du critère (3.43)

$$\begin{pmatrix} \mathbf{g}_x \\ \mathbf{g}_c \end{pmatrix} = 2 \begin{pmatrix} \mathbf{H}^t & \mathbf{0} & \mathbf{0} \\ \mathbf{T}^t & \mathbf{I} & \mathbf{I} \end{pmatrix} \Sigma_{\mathbf{y}}^{-1} \left( \begin{pmatrix} \mathbf{H} & \mathbf{T} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{c} \end{pmatrix} - \tilde{\mathbf{y}} \right) + 2 \begin{pmatrix} \Sigma_{\mathbf{x}}^{-1} & \mathbf{0} \\ \mathbf{0} & \gamma_{\mathbf{c}} \Sigma_{\mathbf{c}}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{x} - \widetilde{\mathbf{m}}_{\mathbf{x}} \\ \mathbf{c} - \widetilde{\mathbf{m}}_{\mathbf{c}} \end{pmatrix}.$$

La matrice  $\Sigma_{\mathbf{y}}$  est bloc diagonale et chaque bloc est proportionnel à l'identité, les calculs du gradient pour l'objet  $\mathbf{g}_{\mathbf{x}}$  et pour la composante  $\mathbf{g}_{\mathbf{c}}$  peuvent être aisément séparés. Pour cela en détaillant on obtient

$$\begin{pmatrix} \mathbf{g}_{\mathbf{x}} \\ \mathbf{g}_{\mathbf{c}} \end{pmatrix} = 2 \begin{pmatrix} \mathbf{H}_0^t & \mathbf{H}_1^t & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_0^t & \mathbf{T}_1^t & \dots & \mathbf{I} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \gamma_{b_0}(\mathbf{H}_0 \mathbf{x} + \mathbf{T}_0 \mathbf{c} - \tilde{\mathbf{y}}_0) \\ \gamma_{b_1}(\mathbf{H}_1 \mathbf{x} + \mathbf{T}_1 \mathbf{c} - \tilde{\mathbf{y}}_1) \\ \vdots \\ \gamma_{b_B}(\mathbf{c} - \tilde{\mathbf{y}}_B) \\ \gamma_{b_{B+1}}(\mathbf{c} - \tilde{\mathbf{y}}_{B+1}) \end{pmatrix} + 2 \begin{pmatrix} \Sigma_{\mathbf{x}}^{-1} & \mathbf{0} \\ \mathbf{0} & \gamma_{\mathbf{c}} \Sigma_{\mathbf{c}}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{x} - \tilde{\mathbf{m}}_{\mathbf{x}} \\ \mathbf{c} - \tilde{\mathbf{m}}_{\mathbf{c}} \end{pmatrix}.$$

En conclusion, pour obtenir un échantillon de la loi conditionnelle *a posteriori* pour l'objet et la composante basse fréquence nous disposons de l'algorithme suivant.

1. Simuler pour chaque ordre de développement de Taylor

$$\tilde{\mathbf{m}}_{\mathbf{x}_k} = \delta_{\mathbf{x}_k} = \gamma_{\mathbf{x}_k}^{-1/2} \mathbf{D}_k^{-1/2} \mathbf{n}$$

avec  $\delta_{\mathbf{x}_k} \sim \mathcal{N}(0, \gamma_{\mathbf{x}_k}^{-1} \mathbf{D}_k^{-1})$  et  $\mathbf{n} \sim \mathcal{N}(0, \mathbf{I})$ . Ce terme nécessite le calcul de  $\mathbf{D}_k^{-1/2}$ . Si la matrice est circulante cette inversion et le calcul de  $\delta_{\mathbf{x}_k}$  peut se faire aisément dans l'espace de Fourier.

2. Simuler

$$\tilde{\mathbf{m}}_{\mathbf{c}} = \delta_{\mathbf{c}} = \gamma_{\mathbf{c}}^{-1/2} \Sigma_{\mathbf{c}}^{1/2} \mathbf{n}$$

avec  $\delta_{\mathbf{c}} \sim \mathcal{N}(0, \gamma_{\mathbf{c}}^{-1} \Sigma_{\mathbf{c}})$  et  $\mathbf{n} \sim \mathcal{N}(0, \mathbf{I})$ . De nouveau cette étape est facile si  $\Sigma_{\mathbf{c}}$  a une structure Toeplitz ou circulante.

3. Simuler

$$\tilde{\mathbf{y}} = \mathbf{y} + \delta_{\mathbf{y}} = \mathbf{y} + \Sigma_{\mathbf{y}}^{1/2} \mathbf{b}$$

qui est très facile puisque nous avons une hypothèses de bruit blanc *iid* pour chaque bolomètre.

4. Initialiser  $\mathbf{x}^{(0)}$  et  $\mathbf{c}^{(0)}$ .
5. Itérer jusqu'à convergence

$$\begin{aligned} \mathbf{x}^{(q+1)} &= \mathbf{x}^{(q)} - \eta^{(q)} \mathbf{g}_{\mathbf{x}}^{(q)} \\ \mathbf{c}^{(q+1)} &= \mathbf{c}^{(q)} - \eta^{(q)} \mathbf{g}_{\mathbf{c}}^{(q)} \end{aligned}$$

où

$$\mathbf{g}_{\mathbf{x}}^{(q)} = 2 \sum_{b=0}^{B-1} \gamma_{b_b} \mathbf{H}_b^t \left( \mathbf{H}_b \mathbf{x}^{(q)} + \mathbf{T}_b \mathbf{c}^{(q)} - \tilde{\mathbf{y}}_b \right) + 2 \begin{pmatrix} \gamma_{\mathbf{x}_0} \mathbf{D}(\mathbf{x}_0 - \tilde{\mathbf{m}}_{\mathbf{x}_0}) \\ \vdots \\ \gamma_{\mathbf{x}_K} \mathbf{D}(\mathbf{x}_K - \tilde{\mathbf{m}}_{\mathbf{x}_K}) \end{pmatrix} \quad (3.44)$$

et

$$\mathbf{g}_{\mathbf{c}}^{(q)} = 2 \sum_{b=0}^{B-1} \gamma_{b_b} \mathbf{T}_b^t \left( \mathbf{H}_b \mathbf{x}^{(q)} + \mathbf{T}_b \mathbf{c}^{(q)} - \tilde{\mathbf{y}}_b \right) + 2 \sum_{b=B}^{B+1} \gamma_{b_b} (\mathbf{c}^{(q)} - \tilde{\mathbf{y}}_b) + 2 \gamma_{\mathbf{c}} \Sigma_{\mathbf{c}}^{-1} (\mathbf{c}^{(q)} - \tilde{\mathbf{m}}_{\mathbf{c}}) \quad (3.45)$$

où le pas  $\eta^{(q)} > 0$ . En pratique ce n'est pas l'algorithme du gradient qui est utilisé mais une stratégie plus complexe de gradient conjugué avec un pas adaptatif que nous ne décrirons pas dans ce manuscrit mais qui nécessite de toute façon les équations (3.44), (3.45) et (3.43) mis en œuvre dans le logiciel GPAC [BCG<sup>+</sup>06].

L'algorithme d'optimisation converge vers l'argument du minimum du critère puisque tout est gaussien et convexe. A convergence, le calcul fournit un échantillon de la loi cible (3.30) page 92. Pour  $\tilde{\mathbf{y}} = \mathbf{y}$ ,  $\tilde{\mathbf{m}}_{\mathbf{x}} = 0$  et  $\mathbf{m}_c = 0$ , alors l'algorithme calcule la moyenne de la loi, c'est à dire la solution des moindres carrés régularisés aux hyper-paramètres fixés.

Dans cette partie nous avons décrit l'algorithme utilisé pour estimer les paramètres inconnus. Il repose sur un échantillonneur de Gibbs qui fournit des échantillons de la loi *a posteriori*. Pour l'ensemble des paramètres nous disposons soit de lois conditionnelles usuelles comme les lois gamma pour les puissances, soit de forme inconnues simulées par un algorithme de type Metropolis-Hastings.

Le cas de la simulation de l'objet est particulier puisqu'il s'agit d'une loi gaussienne dont l'échantillonnage est rendu très difficile pour trois raisons :

- la grande dimension de l'inconnue.
- la corrélation *a posteriori* introduite par le modèle instrument et la loi *a priori*.
- la non invariance du modèle instrument qui ne permet pas la simulation en passant dans l'espace de Fourier qui aurait permis l'indépendance.

La solution proposée repose sur l'optimisation d'un système perturbé.

La résolution de ce problème nécessite le calcul d'un modèle direct  $\mathbf{H}\mathbf{x}$  et de son transposé  $\mathbf{H}^t\mathbf{e}$ . C'est clairement l'échantillonnage de cette loi pour  $\mathbf{x}$  qui prend le plus de temps. La prochaine partie présente les travaux effectués sur l'analyse de la structure de ce modèle. Du modèle avec une matrice pleine présentée à la fin du chapitre 2 sont présentés des modèles approchés mais qui s'exécutent bien plus rapidement.

### 3.3 Calculs efficaces d'une sortie modèle et du transposé

Cette partie est dédiée à l'analyse de la structure de  $\mathbf{H}$  dans l'objectif d'accélérer l'algorithme d'optimisation. Nous partons tout d'abord de la structure « pleine » c'est à dire que les éléments de  $\mathbf{H}$  sont calculés à partir de l'équation (2.52) page 69

$$\mathcal{H}(i,j,k,l,m,n,q) = \frac{S'}{2\sqrt{2\pi}\Sigma_v} \int_{\lambda} h_f(\lambda) (\lambda - \lambda_0)^k \exp \left[ -\frac{(o_{\alpha} + nT_e v_{\alpha})^2}{2\Sigma_{\alpha}^2} - \frac{(o_{\beta} + nT_e v_{\beta})^2}{2\Sigma_{\beta}^2} \right] \operatorname{erfcx} \left( \frac{\Sigma_{\beta} v_{\alpha} (o_{\alpha} + nT_e v_{\alpha})}{\sqrt{2}\Sigma_{\alpha}\Sigma_v} + \frac{\Sigma_{\alpha} v_{\beta} (o_{\beta} + nT_e v_{\beta})}{\sqrt{2}\Sigma_{\beta}\Sigma_v} + \frac{\Sigma_{\alpha}\Sigma_{\beta}}{\sqrt{2}\tau_e\Sigma_v} \right) d\lambda$$

où  $(i,j)$  sont les indices spatiaux des coefficients de décomposition du ciel,  $k$  l'ordre de développement de Taylor,  $(l,m)$  les indices de l'échantillonnage spatial des bolomètres,  $n$  l'indice de l'échantillonnage temporel. Les lignes de  $\mathbf{H}$  sont déterminées par les indices  $l,m,n$  et les colonnes par les indices  $i,j,k$ . Comme toutes les fonctions utilisées pour le modèle ont une étendue spatiale infinie, cette matrice est pleine, il n'y a pas de zéros.

Le travail présenté ici montre, à l'aide d'approximations, que la matrice pleine  $\mathbf{H}$  peut tout d'abord avoir une approximation creuse. Puis on montre que le modèle peut se séparer en une partie invariante et une partie non invariante mais très peu coûteuse en calculs pour ressembler à un modèle classique de sur-résolution. Enfin une troisième approximation consiste à rendre le modèle entièrement invariant.

#### 3.3.1 Économies de mémoire et de calculs

En analysant la structure du modèle et sans faire d'approximation supplémentaire, il est possible de faire ressortir des propriétés intéressantes. Il y a deux aspects. Tout d'abord un aspect creux de

la matrice  $\mathbf{H}$  vient du fait que l'étendue observée est limitée. Deuxièmement un aspect invariant discret, reliquat de l'invariance continue venant des convolutions.

### Aspects creux

Les réponses et les gaussiennes de décomposition spatiale ont un support infini donc la matrice  $\mathbf{H}$  est pleine. Cependant, pour effectuer une mise en œuvre sur une machine de calculs il faut considérer une décomposition sur un nombre fini d'échantillons  $\mathbf{x}$ . On néglige donc la contribution d'une portion de ciel trop lointaine en considérant que l'énergie de la réponse est négligeable au delà d'une certaine étendue spatiale.

Si cette hypothèse est employée pour le bord du ciel, lorsque l'on fait le calcul pour les extrémités, il n'y a aucune raison à ne pas considérer la même hypothèse pour les contributions du centre du ciel ce qui revient à considérer que la réponse est négligeable au delà d'une certaine étendue.

De même lorsque l'on reproduit les données acquises lorsque le centre du ciel est observé, il n'y a aucune raison à ne pas considérer également ici que la réponse est négligeable au delà d'une certaine étendue. On néglige les contributions du ciel aux extrémités.

En conclusion, une donnée produite par un bolomètre dans une région n'est quasiment pas influencée par une région du ciel relativement éloignée dans le temps et l'espace ce qui revient à tronquer l'étendue spatiale et temporelle de la réponse.

Dans la partie 2.4.3 nous avons vu que la dimension temporelle se mêlait aux dimensions spatiales. La réponse était définie seulement par les dimensions spatiales. La réponse temporelle introduisait une traînée dans la réponse spatiale mais très peu visible. Pour effectuer une troncature nous allons donc considérer uniquement la taille du lobe principal calculée à partir de la largeur de la gaussienne de décomposition, du miroir et des cornets, avant convolution par la réponse du bolomètre. Cette réponse centrée a pour expression

$$\frac{1}{2\pi(\Sigma_\alpha + \Sigma_\beta)} \exp \left[ -\frac{1}{2} \left( \frac{\alpha^2}{\Sigma_\alpha^2} + \frac{\beta^2}{\Sigma_\beta^2} \right) \right]$$

et sa largeur est caractérisée par deux écart-types qui dépendent de la longueur d'onde. Dans ce cas une mise à zéro pour des valeurs supérieures ou inférieures à  $\pm 3(\Sigma_\alpha + \Sigma_\beta)/2$  conserverait environ 99% de l'énergie.

Ce principe peut être appliqué lors du calcul de la matrice  $\mathbf{H}$ . Pour quantifier le gain nous allons considérer un problème de taille similaire à ce qui peut être observé par SPIRE. La carte recherchée possède une étendue de 60 arcminutes de côté. Si on souhaite un échantillonnage de 2 secondes d'arc, correspondant à la distance parcourue par le télescope entre deux échantillons temporels, il faut alors environ 1800 coefficients de décomposition dans l'espace dans une dimension. Dans ce cas la matrice  $\mathbf{H}$  possède 3 240 000 colonnes par ordre de développement de Taylor. De plus pour un balayage du photomètre possédant le plus de bolomètres cela correspond également environ à 1 800 échantillons temporels pour 139 bolomètres. Donc la matrice  $\mathbf{H}$  possède environ 250 000 lignes. Autrement dit la matrice en double précision fait environ  $250\,000 \times 1\,800^2 \times 8/10^9 = 6\,480$  giga bits (GB) pour un ordre de développement de Taylor et un balayage. C'est une taille importante même sur un support de stockage de type disque dur magnétique.

Si on considère le caractère creux, c'est à dire que l'on tronque la réponse à  $\pm 3\sigma \approx 200$  secondes d'arc, voir annexe D, alors il n'y a plus que 100 coefficients significatifs environs par ligne de  $\mathbf{H}$ . Nous avons donc une réduction de plus de 300 ce qui fait une matrice de 20 GB. Cette matrice peut être mise dans la mémoire d'un serveur de calcul mais il s'agit de la matrice pour un ordre de développement de Taylor et un balayage. Les dimensions même réduites sont toujours considérables. Une étude plus fine du modèle et notamment son caractère invariant permettra de réduire les calculs.

### Aspects invariants

Le modèle fait initialement intervenir des convolutions spatiales continues et donc invariantes. On peut se poser la question de l'exploitation de ces propriétés d'invariance continue après échantillonnage. La question est de savoir comment se transmet la structure d'invariance continue dans le formalisme discret.

Le modèle de l'équation (2.52) page 69 peut s'écrire

$$\mathcal{H}(i,j,k,l,m,n,q) = \mathcal{H}(o_\alpha + nT_e v_\alpha, o_\beta + nT_e v_\beta)$$

où

$$o_\alpha = c_\alpha + \alpha^{ij} - \alpha_{lm}, \quad \text{et} \quad o_\beta = c_\beta + \beta^{ij} - \beta_{lm}.$$

L'invariance continue apparaît dans le fait que  $\mathcal{H}$  dépend de *la somme* des différentes positions spatiales des gaussiennes, réponses instrument, *etc.* La réponse ne dépend que du décalage entre ces positions. Comme le problème est formulé en discret  $(i,j)$  il faudrait faire apparaître une invariance par décalage *quantifié* dans les expressions

$$c_\alpha + \alpha^{ij} - \alpha_{lm} + nT_e v_\alpha \quad \text{et} \quad c_\beta + \beta^{ij} - \beta_{lm} + nT_e v_\beta.$$

Les variables  $c_\alpha$  et  $c_\beta$  sont négligées dans la suite puisqu'elles sont fixées pour un balayage. Comme nous avons choisi de décomposer le ciel sur une famille de gaussiennes régulièrement espacées alors

$$iT_\alpha - \alpha_{lm} + nT_e v_\alpha \quad \text{et} \quad jT_\beta - \beta_{lm} + nT_e v_\beta.$$

La position des bolomètres  $(\alpha_{lm}, \beta_{lm})$  peut s'exprimer à l'aide d'une grille hexagonale et d'une rotation (voir annexe B.1). Cependant ce n'est pas très intéressant ici et nous allons voir pourquoi. Le seul degré de liberté est sur les paramètres  $T_\alpha$  et  $T_\beta$  le reste étant fixé par l'acquisition. On peut donc choisir

$$T_\alpha = T_e v_\alpha \quad \text{et} \quad T_\beta = T_e v_\beta$$

c'est à dire que le pas d'échantillonnage des gaussiennes correspond au déplacement effectué par l'instrument entre deux échantillonnages temporels (ou un multiple de ce pas). Ainsi

$$\begin{aligned} (i+n)T_\alpha - \alpha_{lm} \\ (j+n)T_\beta - \beta_{lm} \end{aligned}$$

et nous obtenons une invariance quantifiée. Donc une partie des calculs peut se faire par convolution discrète. Il y a donc un gain en occupation mémoire et en quantité de calculs. Le choix aurait pu se porter sur les positions des bolomètres plutôt que les décalages dans un balayage. Cependant nous cherchons à gagner en temps de calculs. Or nous avons au maximum 139 bolomètres alors que nous pouvons avoir plusieurs centaines d'échantillons temporels pour un balayage. Le gain est donc bien plus important.

En pratique nous avons considéré précédemment que nous avons 1 800 échantillons temporels par bolomètre pour un balayage, chacun correspondant à une ligne de la matrice  $\mathbf{H}$ . Comme nous avons une invariance discrète, les coefficients de cette ligne sont les mêmes décalés et le bloc est de Toeplitz. Nous n'avons donc plus qu'une matrice de taille 20 GB/1 800  $\approx$  12 MB pour calculer la sortie modèle pour un ordre de développement de Taylor en  $\lambda$  et un balayage. La taille est cette fois bien plus raisonnable.

Le gain est double. Tout d'abord les matrices sont de tailles réduites, calculables et peuvent être stockées. Ensuite le calcul de tous les échantillons temporels pour un bolomètre et un balayage peut se faire par convolution.

En conclusion nous avons montré deux pistes pour obtenir un calcul efficace sans faire d'approximation supplémentaire. En effet le caractère invariant est présent dans le problème et le caractère creux est inhérent aux calculs avec un nombre fini de coefficients de décomposition.

Nous verrons dans la prochaine partie qu'il est possible d'établir l'invariance discrète sur tous les paramètres, au prix d'une approximation supplémentaire.

### 3.3.2 Structure invariante et modèle de la littérature en sur-résolution

La position des bolomètres s'exprime à l'aide d'une grille hexagonale. Cette grille peut se décomposer en deux sous-grilles rectangulaires. Pour clarifier l'exposé on suppose que le capteur n'a pas de rotation par rapport au ciel. On considère pour cette description uniquement une des deux sous-grilles de pas  $P_\alpha$  et  $P_\beta$ , l'extension au cas hexagonal étant immédiate.

La position des bolomètres en fonction de leurs indices devient

$$\begin{aligned} o_\alpha + nT_e v_\alpha &= c_\alpha + iT_\alpha - lP_\alpha - mP_\beta + nT_e v_\alpha \\ o_\beta + nT_e v_\beta &= c_\beta + jT_\beta - lP_\alpha - mP_\beta + nT_e v_\beta. \end{aligned}$$

Si tous les facteurs sont multiples d'un autre facteur  $T$  alors on obtient

$$\begin{aligned} c_\alpha &= n_1 T & P_\beta &= n_4 T \\ c_\beta &= n_2 T & T_e v_\alpha &= n_5 T \\ P_\alpha &= n_3 T & T_e v_\beta &= n_6 T \\ T_\alpha &= n_7 T & T_\beta &= n_8 T. \end{aligned}$$

En pratique il semble judicieux de choisir  $T = T_\alpha = T_\beta$  alors le décalage continu s'écrit

$$\begin{aligned} o_\alpha + nT_e v_\alpha &= (n_1 + in_7 - ln_3 - mn_4 + nn_5)T \\ o_\beta + nT_e v_\beta &= (n_2 + jn_8 - ln_3 - mn_4 + nn_6)T. \end{aligned}$$

Ceci revient à tronquer les décalages. La littérature en super-résolution [RCLBG06, FREM04, EF97] considère souvent que les décalages sont des multiples de la taille d'un pixel de l'image haute résolution que l'on cherche à reconstruire. Il s'agit ici d'une approximation similaire à un multiple près. La position réelle des bolomètres a été décalée, ou les données ont été interpolées au plus proche voisin.

Dans ce cas la sortie modèle s'écrit

$$y^{lm}(nT_e) = \sum_{k=0}^K \sum_{ij} x^k(i,j) \mathcal{H}\left((n_1 + in_7 - ln_3 - mn_4 + nn_5)T, (n_2 + jn_8 - ln_3 - mn_4 + nn_6)T, k\right)$$

Autrement dit, en calculant la convolution discrète

$$y(i',j') = \sum_{k=0}^K \sum_{ij} x^k(i,j) \mathcal{H}\left((i-i')T, (j-j')T, k\right) \quad (3.46)$$

alors

$$y^{lm}(nT_e) = y(i',j')$$

si et seulement si

$$i - i' = n_7 i + ln_3 + mn_4 - nn_5 - n_1 \quad (3.47)$$

$$j - j' = n_8 j + ln_3 + mn_4 - nn_6 - n_2. \quad (3.48)$$

Toutes les données  $\mathbf{y}$ , pour une direction de balayage, peuvent être calculées par convolution discrète (3.46) suivie d'un prélèvement de certaines valeurs définies par (3.47) et (3.48).

Il y a plusieurs remarques à faire. Tout d'abord il suffit de stocker la réponse impulsionnelle, c'est à dire uniquement la première ligne de la matrice et non toute la matrice et ce pour chaque direction de balayage.

De plus l'occupation mémoire et les pré-calculs sont considérablement réduits (l'information de pointage est de la taille des données et quatre réponses impulsionnelles sont nécessaires pour le cas *Large map*). Enfin l'aspect invariant permet également de passer dans l'espace de Fourier et permet un calcul de la convolution par transformée de Fourier discrète. Enfin le prélèvement des bonnes valeurs est une opération très simple et rapide.

Il y a deux inconvénients à cette approximation. Premièrement des calculs sont faits en trop puisque la convolution est sur tout  $\mathbf{x}$  et que seulement une partie du résultat est prélevée. Le ratio dépend du choix de la finesse de la grille pour le ciel relativement à l'échantillonnage des bolomètres et de leurs décalages dans le temps. En pratique, l'excès de calcul est réduit car nous choisissons un pas d'échantillonnage proche du décalage entre deux instants temporels. Presque tous les résultats de la convolution sont observés.

Le deuxième inconvénient est qu'une approximation sur la position des bolomètres à tous les instants est nécessaire. Or ces décalages sont importants puisque ce sont eux qui contiennent l'information nécessaire pour la super-résolution. Cependant nous avons choisi un pas  $T$  proche du décalage minimum des bolomètres ce qui correspond au plus à  $P_\alpha/15$  pour le pas d'échantillonnage du capteur le plus petit soit environ  $T \approx 2$  arc-secondes. L'erreur introduite est donc réduite étant donné les niveaux de bruits attendus sur les données. De plus, les erreurs de pointage du télescope sont du même ordre de grandeur soit 2 arc-secondes.

Finalement le modèle très général utilisé pour décrire l'algorithme dans la partie précédente peut se mettre sous la forme

$$\mathbf{y} = \mathbf{H}\mathbf{x} = \mathbf{P}\mathbf{H}_c\mathbf{x}$$

où  $\mathbf{H}_c$  est une matrice de convolution et  $\mathbf{P}$  une matrice de troncature qui prélève les valeurs observées après la convolution (c'est la matrice de pointage des méthodes SANEPIC et MADMAP, voir partie 1.5). C'est une matrice avec un seul 1 par ligne puisque chaque donnée ne peut provenir que d'une position. On en déduit déjà que la matrice  $\mathbf{P}$  est une matrice creuse. De plus il suffit pour connaître  $\mathbf{H}_c$  de connaître les réponses impulsionnelles.

Il est possible enfin de réduire encore un peu l'occupation mémoire. Tout d'abord les réponses impulsionnelles sont obtenues avec la formule équation (2.55) page 70

$$\mathcal{H}(i,j,k,v_\alpha,v_\beta) = \frac{S'}{2\sqrt{2\pi}\Sigma_v} \int_{\lambda} h_f(\lambda) (\lambda - \lambda_0)^k \exp \left[ -\frac{i^2 T^2}{2\Sigma_\alpha^2} - \frac{j^2 T^2}{2\Sigma_\beta^2} \right] \operatorname{erfcx} \left( \frac{iT\Sigma_\beta v_\alpha}{\sqrt{2}\Sigma_\alpha \Sigma_v} + \frac{jT\Sigma_\alpha v_\beta}{\sqrt{2}\Sigma_\beta \Sigma_v} + \frac{\Sigma_\alpha \Sigma_\beta}{\sqrt{2}\tau_e \Sigma_v} \right) d\lambda.$$

Il y a, par ordre  $k$  de développement de Taylor, quatre réponses impulsionnelles, une par vitesse  $v$ . Cependant on peut remarquer qu'un balayage sur deux correspond à une vitesse dans l'autre (voir partie 2.2.6) sens soit

$$\mathcal{H}(i,j,k,v_\alpha,v_\beta) = \mathcal{H}(-i,-j,k,-v_\alpha,-v_\beta).$$

Donc nous n'avons besoin que de deux réponses impulsionnelles par ordre de Taylor puisque la matrice  $\mathbf{H}_c$  correspondant au balayage dans la direction  $(v_\alpha,v_\beta)$ , est la transposée de la matrice correspondant à la direction  $(-v_\alpha,-v_\beta)$  (les réponses impulsionnelles sont retournées).



- faciles à pré-calculer pour le cas de  $\mathbf{H}_c$  et de  $\mathbf{P}$ , ou à remplacer par un code de calcul trivial pour le cas de  $\mathbf{D}$  et  $\Sigma$
- très creuses, en particulier pour  $\mathbf{P}$ . De plus il suffit de ne stocker seulement la première ligne de  $\mathbf{H}_c$ .

**Remarque.** Les matrices  $\mathbf{P}_b^s$  sont assez polyvalentes pour modéliser des données non observées que l'on ne souhaite pas utiliser pour l'inférence. En effet les bolomètres acquièrent des données en continu, y compris aux extrémités des balayages où le télescope ne balaye pas à vitesse constante, change de direction, etc. pour démarrer un nouveau balayage (voir partie 2.2.6). Pendant ces périodes il n'y a pas assez de confiance dans les données, on considère que l'instrument est dans un état trop incertain. Dans ce cas pour ne pas prendre en compte ces données il suffit d'enlever les lignes correspondantes dans  $\mathbf{P}$  et dans les données. L'avantage se situe au niveau de l'estimation de  $\mathbf{c}$  puisque l'on peut au contraire exploiter les données des bolomètres aveugles pendant ces périodes.

On peut déduire enfin la structure de la transposée du modèle qui s'écrit simplement

$$\mathbf{H}^t = \mathbf{D}^t \mathbf{H}_c^t \Sigma^t \mathbf{P}^t \quad (3.49)$$

comme une matrice de bourrage de zéros  $\mathbf{P}^t$  (addition des données qui possèdent le même pointage et mise à zéro des autres coefficients), d'une matrice de duplication des ordres de développement de Taylor  $\Sigma^t$ , d'une convolution par réponses impulsionnelles retournées puis d'une sommation de chaque balayage.

En conclusion nous avons, à l'aide d'une approximation sur la position des bolomètres, séparé le modèle  $\mathbf{H}$  avec une structure très non invariante en deux sous-modèles  $\mathbf{H} = \mathbf{P}\mathbf{H}_c$  où  $\mathbf{H}_c$  est invariant et  $\mathbf{P}$  contient la structure invariante. L'intérêt est bien sûr que le coût en calcul de  $\mathbf{P}$  ou de  $\mathbf{P}\mathbf{x}$  est minime. De nouveau il s'agit d'un modèle d'acquisition classique utilisé en super-résolution [EF97, FREM04]. Une différence cependant se situe sur la connaissance des décalages et le fait qu'il y ait plusieurs convolutions différentes.

On peut remarquer d'ailleurs que la matrice  $\mathbf{P}$  correspond à la matrice de pointage des méthodes MADMAP et SANEPIC. C'est en effet cette matrice qui contient toute l'information de pointage. Nous avons finalement ajouté tout un modèle  $\mathbf{H}_c$  qui contient la structure en longueur d'onde du ciel ainsi que la réponse instrumentale.

Dans la prochaine partie nous allons encore plus loin dans l'approximation du pointage pour obtenir un modèle entièrement invariant. Cela permettra une mise en œuvre rapide par filtrage de Wiener–Hunt ainsi qu'une approximation du hessien.

### 3.3.3 Approximation du hessien

Nous nous intéressons à l'estimation de  $\mathbf{x}$  à hyper-paramètres fixés et  $\mathbf{c}$  fixé (que nous ignorerons dans les écritures). Dans ce cas on sait qu'il s'agit de la solution des moindres carrés régularisés décrite partie 3.2.1. Nous ferons quelques simplifications dans les notations par souci de clarté et pour se concentrer sur l'essentiel. Dans ce cas la solution explicite est connue et s'écrit

$$\hat{\mathbf{x}} = \left( \mathbf{H}^t \Sigma_y^{-1} \mathbf{H} + \Sigma_x^{-1} \right)^{-1} \mathbf{H}^t \Sigma_y^{-1} \mathbf{y}.$$

Ici nous considérons que le bruit est blanc *iid* avec un facteur  $\gamma_{b_b}$  différent pour chaque bolomètre. En exploitant le modèle de la partie précédente, la solution s'exprime

$$\hat{\mathbf{x}} = \left( \mathbf{D}^t \mathbf{H}_c^t \Sigma^t \mathbf{P}^t \mathbf{P} \Sigma \mathbf{H}_c \mathbf{D} + \Sigma_x^{-1} \right)^{-1} \mathbf{D}^t \mathbf{H}_c^t \Sigma^t \mathbf{P}^t \mathbf{y}.$$

Le calcul de l'inverse du hessien n'est pas possible directement. Cependant on sait que les matrices  $\mathbf{D}$  et  $\mathbf{\Sigma}$  sont composées de matrices identités donc diagonales. De plus  $\mathbf{H}_c$  est composée de matrices de convolution donc diagonalisables dans l'espace de Fourier. Par contre la matrice  $\mathbf{P}$  est plus complexe. La question est donc : peut-on trouver une approximation telle que l'ensemble soit composé de matrices diagonalisables dans l'espace de Fourier, permettant d'inverser le hessien efficacement ?

En fait, comme  $\mathbf{P}$  est une matrice de troncature, la matrice  $\mathbf{P}^t\mathbf{P}$  est une matrice diagonale. Plus précisément, les éléments diagonaux correspondent au nombre de fois qu'un coefficient (à la sortie de la convolution) est observé. La diagonale est donc composée de 0, de 1, *etc.* la trace étant égale au nombre de données. L'idée est d'approcher cette matrice par l'identité et ainsi d'obtenir l'expression

$$\hat{\mathbf{x}} = \left( \mathbf{D}^t \mathbf{H}_c^t \mathbf{\Sigma}^t \mathbf{\Sigma} \mathbf{H}_c \mathbf{D} + \mathbf{\Sigma}_x^{-1} \right)^{-1} \mathbf{D}^t \mathbf{H}_c^t \mathbf{\Sigma}^t \mathbf{P}^t \mathbf{y}.$$

qui pourra certainement avoir une expression diagonale dans l'espace de Fourier.

Il faut donc déterminer sous quelles approximations la matrice  $\mathbf{P}^t\mathbf{P}$  est l'identité. On peut essayer d'établir la meilleure matrice au sens d'un critère tel que  $\mathbf{P}^t\mathbf{P} \propto \mathbf{I}$ . Donc pour que  $\mathbf{P}$  soit l'identité, il suffit que :

- là où il y avait des zéros, qui correspondent à des pixels non observés, il y ait maintenant un 1. Cela sous-entend soit l'existence de données non acquises en réalité, soit de considérer qu'une donnée est obtenue par combinaison linéaire de pixels.
- là où les coefficients sont plus grand que 1, on les positionne à 1, ce qui sous-entend que ces pixels ont été vus une seule fois. Donc soit certaines données sont ignorées ou alors on peut considérer une moyenne de celles-ci.

Une autre possibilité pour s'affranchir du problème des zéros consisterait à diminuer la résolution du ciel, c'est à dire les pas  $T_\alpha$  et  $T_\beta$ . C'est la solution envisagée naturellement par les solutions qui n'exploitent pas d'information *a priori* tel que MADMAP ou SANEPIC. On perd potentiellement en résolution spatiale.

Ces considérations ne sont pas fondamentales. En effet l'intérêt repose plus sur la possibilité d'avoir une approximation relativement correcte pour appliquer un filtre de Wiener–Hunt. Si nous sommes capables d'avoir un filtre de Wiener–Hunt mis en œuvre par transformée de Fourier nous pouvons ainsi fournir une première estimation du ciel grossière, une initialisation pour l'algorithme d'optimisation voir une estimation des hyper-paramètres avec l'algorithme présenté partie 3.2 mais beaucoup plus rapide puisque l'échantillonnage de  $\mathbf{x}$  se ferait dans l'espace de Fourier.

Une autre possibilité, employée dans certaines approches, considère que  $\mathbf{P}^t\mathbf{P}$  est proportionnelle à l'identité et fabrique des données intermédiaires avec une rétro-projection et une moyenne dans l'espace ciel

$$\tilde{\mathbf{y}} = \mathbf{P}^t (\mathbf{P}\mathbf{P}^t)^{-1} \mathbf{y}$$

Il s'agit d'ailleurs de la méthode appelée *Shift and add* dans la littérature en astronomie ou en super-résolution lorsqu'une étape de déconvolution est incluse [PPK03].

L'approche proposée consiste à calculer l'optimum du critère

$$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{P}\mathbf{x}\|^2 + \lambda \|\mathbf{D}\mathbf{x}\|^2$$

ce qui revient à estimer au préalable un ciel doux par pénalisation de hautes fréquences  $\mathbf{D}\mathbf{x}$  avec un modèle d'acquisition correspondant à  $\mathbf{P}$ . Autrement dit il n'y a pas de déconvolution puisque  $\mathbf{P}$  est une matrice de prélèvement et  $\mathbf{P}^t$  une matrice de bourrage de zéros. En fonction de la résolution choisie il peut y avoir des pixels non observés. L'*a priori* de douceur  $\|\mathbf{D}\mathbf{x}\|^2$ , où en pratique  $\mathbf{D}$  est le gradient discret, comble alors naturellement les sites où les données sont manquantes. L'optimum

est calculé par une algorithmme de descente de gradient relativement rapide puisque les calculs  $\mathbf{P}\mathbf{x}$  et  $\mathbf{D}\mathbf{x}$  sont très simples. C'est une approche notamment proposée par Farsiu *et al.* [FREM04].

La question maintenant est de savoir comment la solution s'exprime et se diagonalise dans l'espace de Fourier avec l'approximation  $\mathbf{P}^t\mathbf{P} = \mathbf{I}$ . La solution des moindres carrés régularisés s'écrit maintenant

$$\hat{\mathbf{x}} = \left( \mathbf{D}^t \mathbf{H}_c^t \boldsymbol{\Sigma}^t \boldsymbol{\Sigma} \mathbf{H}_c \mathbf{D} + \boldsymbol{\Sigma}_x^{-1} \right)^{-1} \mathbf{D}^t \mathbf{H}_c^t \boldsymbol{\Sigma}^t \tilde{\mathbf{y}}. \quad (3.50)$$

Les calculs sont dans l'annexe C.4. En particulier si nous avons un ordre unique, c'est à dire l'ordre zéro, alors la solution est simplement

$$\mathbf{x}_0 = \mathbf{F}^\dagger \left( \sum_v |\Lambda_0^v|^2 + |\Lambda_{\mathbf{D}0}|^2 \right)^{-1} \left( \sum_v \Lambda_0^{1\dagger \frac{\circ}{v}} \tilde{\mathbf{y}} \right). \quad (3.51)$$

Nous avons obtenu à l'aide d'une approximation sur la redondance de l'observation, une solution des moindres carrés régularisés par filtre de Wiener–Hunt, c'est à dire diagonalisée dans l'espace de Fourier grâce à l'inversion explicite du hessien. Cette inversion peut également servir dans l'algorithme d'optimisation, sans que l'approximation affecte la solution. On obtient alors un algorithme avec pré-conditionneur où l'itération suivante est calculée par

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \eta \mathbf{H}_e^{-1} \mathbf{g}^{(k)}.$$

Si l'inverse du hessien  $\mathbf{H}_e^{-1}$  est disponible, alors avec un pas  $\eta = 1$  et une initialisation  $\mathbf{x}^{(0)} = \mathbf{0}$  la solution des moindres carrés régularisés est directement obtenue. L'idée ici n'est pas d'utiliser l'inverse exacte puisqu'elle n'est pas calculable en pratique, mais l'approximation obtenue. L'objectif est d'accélérer la descente en corrigeant les directions du gradient  $\mathbf{g}$ .

### 3.4 Conclusion

Ce chapitre présente la démarche proposée pour l'estimation d'un signal astrophysique à l'aide des données fournies par l'instrument SPIRE. L'approche proposée est nouvelle dans le sens où il s'agit de la seule, pour les données de SPIRE, incluant explicitement le processus d'acquisition complet dans le problème d'inférence, le tout avec une approche globale. Cela permet, conjointement à l'estimation du signal de nuisance  $\mathbf{c}$ , d'estimer un ciel sur-résolu. Elle propose également de déterminer explicitement, en reposant sur un critère formalisé, le compromis entre les différentes sources d'informations et les paramètres instruments.

Tout d'abord nous avons défini la loi *a posteriori* collectant toute l'information injectée dans le problème. Elle est construite en partie avec une vraisemblance qui repose sur le modèle linéaire non-invariant défini dans le chapitre 2 et une loi gaussienne indépendante stationnaire pour le bruit affectant les bolomètres. Elle est construite également pour une autre part sur les informations *a priori* sur l'objet, considéré régulier dans les dimensions spatiales, sur la composante basse fréquence régulière dans le temps, ainsi que sur des lois *a priori* sur les paramètres de ces lois.

L'estimateur proposé est la moyenne *a posteriori* calculée à l'aide d'un échantillonneur de Gibbs. Cet algorithme fournit une estimation de l'objet ainsi que des paramètres des lois permettant ainsi un réglage automatique du compromis entre les différentes sources d'information. L'approche permet également d'estimer des paramètres instruments et peut potentiellement prendre en compte une dérive thermique.

Le point difficile concerne l'échantillonnage de la loi conditionnelle *a posteriori* de l'objet qui est une loi normale multivariée. La structure non-invariante du modèle rend en effet impossible l'échantillonnage dans l'espace de Fourier ce qui aurait permis l'échantillonnage des coefficients

de Fourier indépendants entre eux. Malgré la dépendance complexe entre pixels, le calcul d'un échantillon peut s'effectuer en optimisant un système perturbé.

Enfin, plusieurs moyens d'accélérer les calculs en analysant la structure du modèle sont présentés. Tout d'abord nous exposons comment une partie de l'invariance continue est préservée dans l'espace discret, ce qui permet de réduire considérablement la taille dans la mise en œuvre sur machines de calcul. Nous montrons ensuite comment avec une approximation, avec un domaine de validité important, nous pouvons séparer le problème en une partie invariante et une partie non invariante. Enfin nous poussons l'approximation plus loin, permettant d'avoir un calcul faisable de l'inverse du hessien. Cela permet une mise en œuvre rapide par transformée de Fourier. Cette approximation permet également une correction des directions de descente, avec un préconditionneur, dans l'algorithme d'optimisation.

Le prochain chapitre présente une étude expérimentale sur le potentiel de la méthode concernant la restauration de fréquences spatiales et les possibilités de sur-résolution. L'étude porte également sur les aspects non-supervisés et des résultats sur l'estimation de paramètres instruments sont présentés dans l'article de l'annexe F. La prise en compte de la composante basse fréquence n'est pas étudiée.

## Chapitre 4

# Étude expérimentale

Le chapitre précédent présente les résultats méthodologiques pour l'inversion de données SPIRE avec une approche globale. L'étude expérimentale menée dans ce chapitre se concentre principalement sur les capacités de restauration de fréquences spatiales ou de structures fines du ciel au moyen de la méthode proposée. La prise en compte de la dérive ne nous semblant pas primordiale, pour le moment, face aux questions de reconstruction du ciel, elle n'est pas étudiée ici. L'ensemble des résultats présentés utilisent l'arrondi sur les pointages pour le modèle instrument, décrit partie 3.3.2.

Dans un premier temps on étudie, à hyper-paramètres fixés, le gain apporté par la prise en compte d'un modèle instrument et d'une information *a priori* pour le ciel. Il s'agit de la solution des moindres carrés régularisés calculée par optimisation décrite partie 3.2.3.

Dans un deuxième temps on étudie les possibilités de la méthode pour l'estimation des paramètres des lois de probabilité réglant le compromis entre les différentes sources d'information. Les résultats sont également concentrés sur les paramètres pilotant le contenu spatial de l'estimée, c'est-à-dire la précision de la loi gaussienne pour le bruit et des lois *a priori* pour le ciel. Dans ce cas l'estimateur est la moyenne *a posteriori* sur le ciel et les hyper-paramètres et l'algorithme de calcul est l'échantillonneur de Gibbs.

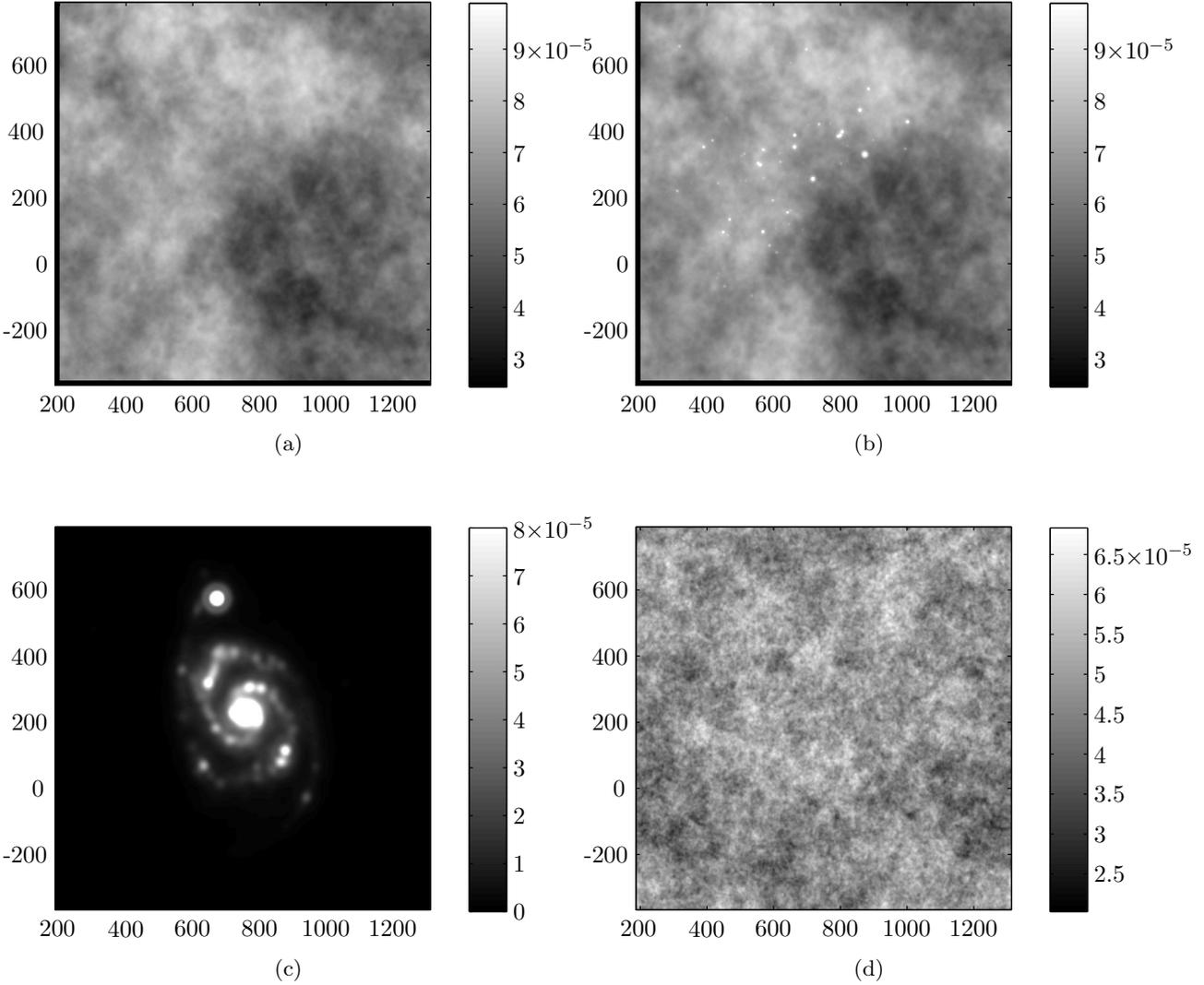
Enfin un troisième aspect porte sur l'estimation de paramètres instruments dans un cadre d'inversion non-supervisé. Cette étude, menée sur un problème de déconvolution plus académique, est décrite sous la forme d'un article présenté dans l'annexe F. La question porte sur l'estimation de paramètres d'une réponse impulsionnelle, comme la largeur d'une gaussienne. Dans ce cas l'estimateur est la moyenne *a posteriori* pour le ciel et les hyper-paramètres, comme précédemment, mais également pour les paramètres instruments. L'algorithme de calcul est l'échantillonneur de Gibbs incluant des étapes de Metropolis-Hastings pour l'échantillonnage des lois *a posteriori* conditionnelles complexes des paramètres instruments.

### 4.1 Conditions de simulation

#### 4.1.1 Ciel étudiés

Les données ont été générées avec le modèle instrument du chapitre 2, il faut donc utiliser un ciel vrai correspondant au modèle. Le pas d'échantillonnage spatial entre coefficients de décomposition du ciel est 2 secondes d'arc. Les coefficients « vrais » utilisés pour représenter le ciel, figure 4.1, sont les images tirées de [CCB<sup>+</sup>06], référence utilisée pour choisir MADMAP comme méthode officielle de traitement haut niveau des données SPIRE.

Tous les ciels considérés sont décomposés sur  $563 \times 579$  coefficients, correspondant à une étendue d'environ 1200 secondes d'arc (20 minutes d'arc) de côté. L'unité des coefficients de décomposition du ciel est en  $\text{W.arcseconde}^{-2}$ . La longueur d'onde n'intervient pas puisqu'elle est intégrée dans le modèle. Pour obtenir la puissance par longueur d'onde il faut donc diviser par la largeur de la bande spectrale. Enfin la valeur du rayonnement est fixée pour avoir une valeur autour de  $5 \times 10^{-5} \text{ W.arcseconde}^{-2}$  correspondant à une valeur typique du milieu interstellaire.



**Figure 4.1** – Ciels vrais utilisés pour l'expérience. La figure 4.1(a) correspond le plus à l'émission étendue. La figure 4.1(b) est identique avec la présence de sources ponctuelles (elle est représentée à la même dynamique que la figure 4.1(a)). La figure 4.1(c) est une galaxie dans une représentation saturée pour faire ressortir l'étendue spatiale. Enfin la figure 4.1(d) est une réalisation de la loi *a priori* pour le ciel  $\phi$ .

Le cas le plus étudié dans ce chapitre est le *cirrus* figure 4.1(a). Il s'agit de l'objet le plus représentatif de l'émission étendue (l'énergie se situe majoritairement dans les basses fréquences). Les autres ciels permettent une première étude de la méthode lorsque le modèle est moins adapté à la réalité terrain.

Le ciel de la figure 4.1(b) est identique au *cirrus* mais avec des sources quasi ponctuelles super-

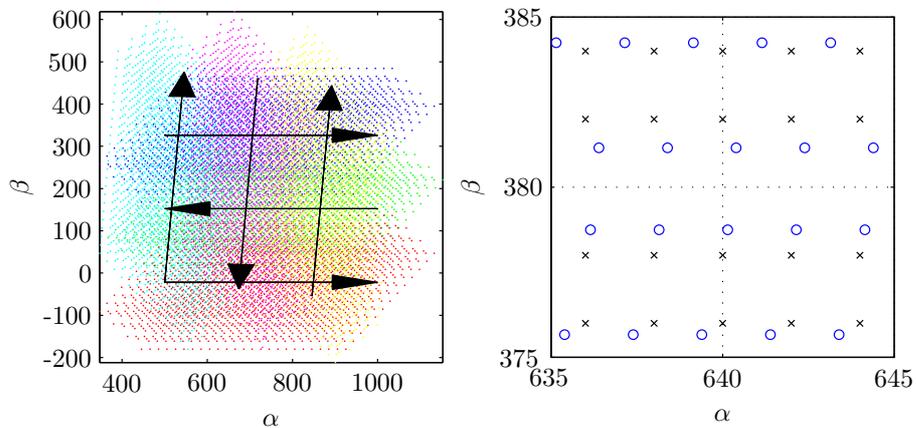
posées. L'amplitude des sources ponctuelles est importante (jusqu'à un facteur 10 par rapport au flux moyen). Le modèle *a priori* pour le ciel ne prenant en compte qu'une information de régularité représente mal ces objets ponctuels. Il s'agit donc d'un cas limite pour la méthode.

Le troisième ciel considéré 4.1(c) correspond à une galaxie. L'objet est relativement régulier spatialement avec cependant une certaine hétérogénéité de la structure spatiale notamment vers les coordonnées (600, 620). Le modèle est donc relativement adapté à ce type de ciel mais la structure spatiale de celui-ci est plus complexe.

Enfin nous avons également considéré une réalisation de la loi *a priori* (gaussien corrélé) figure 4.1(d) pour le ciel continu  $\phi$ . Dans ce cas le modèle *a priori* pour l'inversion correspond parfaitement au modèle utilisé pour la simulation. Ce ciel permet de connaître la valeur du paramètre  $\gamma_{\mathbf{x}}$  ce qui permettra une analyse quantitative de l'estimation des hyper-paramètres.

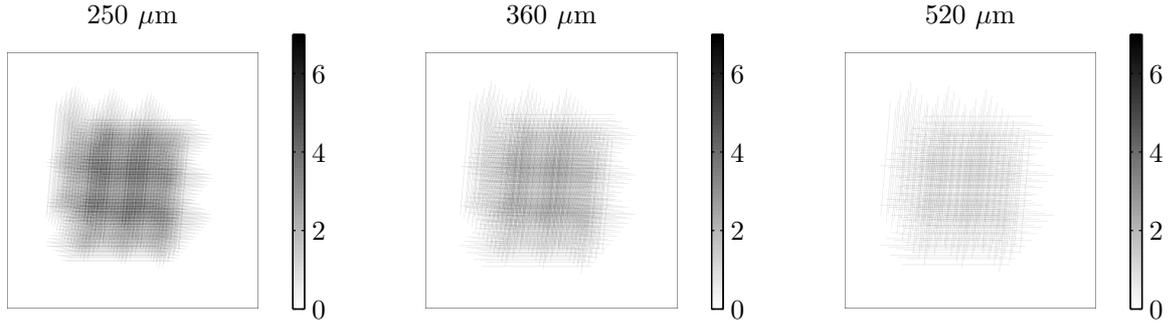
### 4.1.2 Balayage

Le protocole d'observation (*Large map*) est toujours le même avec trois balayages parallèles suivis de trois autres dans une direction de 84,8 degrés. La figure 4.2 montre les pointages effectués ainsi que l'approximation sur la position des bolomètres. La position des bolomètres est arrondie à la position du coefficient de décomposition du ciel le plus proche comme décrit partie 3.3.1, c'est-à-dire 2 secondes d'arc. En fonction du bolomètre l'approximation est plus ou moins importante. Nous avons vu dans la partie 3.3.2 que l'approximation est acceptable, notamment par rapport à l'erreur de mesure du pointage.



**Figure 4.2** – La figure à gauche illustre tous les pointages effectués (en seconde d'arc) avec une couleur différente pour chaque balayage. La figure à droite est un zoom sur une zone particulière. Les pointages réels sont les cercles bleus et les pointages arrondis sont les croix noires.

La figure 4.3 montre les cartes de redondances résultant de l'approximation, correspondant au nombre de fois qu'un pixel est observé. À partir du moment où les pointages sont arrondis à la position des coefficients ceux-ci peuvent être vus plusieurs fois, par des bolomètres différents à des instants différents. On utilise donc l'approximation décrite partie 3.3.2. Dans les cas de la bande spectrale  $250 \mu\text{m}$  la redondance est plus importante que pour  $520 \mu\text{m}$  puisqu'il y a plus de bolomètres et qu'ils sont plus rapprochés.



**Figure 4.3** – Illustration de la redondance pour chaque bande spectrale. Les figures montrent le nombre de fois qu’un coefficient de décomposition est vu pendant toute l’observation. Cette redondance n’a pas de sens sans l’arrondi du pointage.

### 4.1.3 Modèle instrument

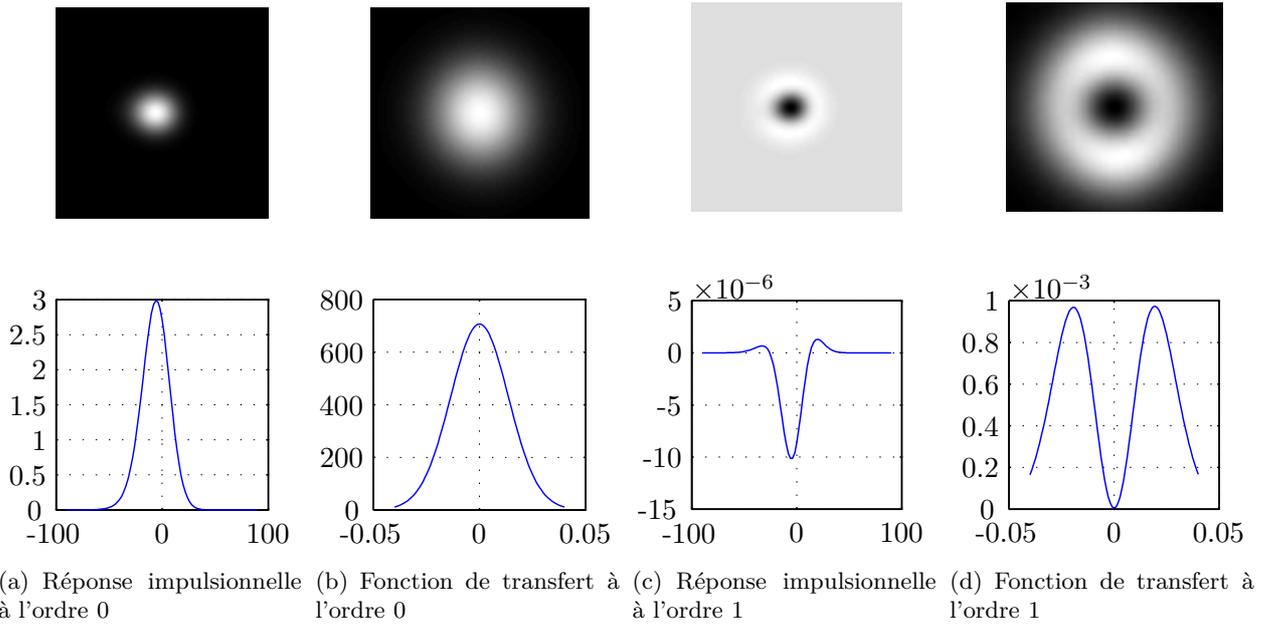
Le modèle instrument utilisé correspond au protocole *Large map* avec une décomposition sur des gaussiennes (partie 2.4.5). L’approximation sur la position des bolomètres, présentée partie 3.3.2, est exploitée. Nous utilisons donc le modèle classiquement rencontré en sur-résolution : le modèle est séparé en une partie convolutive et une partie non-invariante modélisant le pointage ou l’échantillonnage.

**Remarque.** *Le chapitre précédent établit trois modèles d’acquisition. Le premier est la version « exacte » obtenue par application directe des équations du chapitre 2. Il s’agit d’une version lourde numériquement. Le deuxième est obtenu à partir du premier avec une approximation sur les pointages permettant d’extraire l’invariance dans un opérateur de décimation. C’est celui utilisé ici puisqu’il permet une précision suffisante et un calcul rapide. La troisième approximation permet d’obtenir un modèle invariant en considérant que tous les coefficients ont été observés et seulement une seule fois. L’approximation utilisée est potentiellement sévère et la manipulation du modèle est délicate. Cependant les calculs numériques sont potentiellement rapides puisqu’ils peuvent être entièrement menés dans l’espace de Fourier. Ce modèle n’est pas étudié ici.*

La partie convolutive du modèle est calculée par transformée de Fourier rapide. Les valeurs des paramètres du modèle instrument sont les valeurs mesurées en laboratoire (voir annexe D). Les réponses impulsionnelles de la bande spectrale à  $360 \mu\text{m}$  (PMW), sont calculées en discrétisant l’intégrale en longueur d’onde (équation (2.52)) page 69, sont illustrées figure 4.4 pour l’ordre  $k = 0$  et l’ordre  $k = 1$ . La direction de balayage est vers le bas ( $v_\alpha < 0$  et  $v_\beta \approx 0$ ).

Chaque réponse est obtenue par une somme pondérée par  $(\lambda - \lambda_0)^k$  de réponses illustrées figure 2.13(c) page 68 dont la largeur varie puisqu’elle dépend de  $\lambda$ . Concernant l’ordre 0 le résultat ressemble à une gaussienne. Pour l’ordre 1 cependant le résultat ressemble à une différence de gaussiennes, puisque nous avons une pondération par  $(\lambda - \lambda_0)$  ( $\lambda_0$  appartenant à l’intervalle de variation de  $\lambda$ ). La fonction de transfert pour l’ordre 1 présente un caractère passe-bande : les fréquences spatiales autour de la fréquence nulle ne sont pas observées. Les coupes en  $\beta = 0$  mettent en évidence l’effet des bolomètres et de l’électronique de lecture. La réponse est asymétrique et légèrement décalée à cause de la causalité du filtre temporel.

La fonction de transfert modélise la réponse du miroir, l’intégration en longueur d’onde, les cornets, la réponse des bolomètres et de l’électronique de lecture. Tous ces éléments tendent à élargir la réponse par rapport à la figure d’*Airy* modélisant seulement la réponse du miroir. La figure d’*Airy* possède une fréquence de coupure de  $D/\lambda \approx 0,044 \text{ arcsecondes}^{-1}$  à  $360 \mu\text{m}$ . Ici les réponses du modèle ont une atténuation très forte (équivalente à une fréquence de coupure) située



**Figure 4.4** – Réponse avec le protocole *Large map* pour l'ordre 0 et l'ordre 1. Les axes des réponses spatiales sont en arcseconde. Les axes des fonctions de transfert sont en arcseconde<sup>-1</sup>. On observe, en particulier dans les coupes, un décalage dû à la réponse du bolomètre et l'électronique de lecture dans les réponses spatiales.

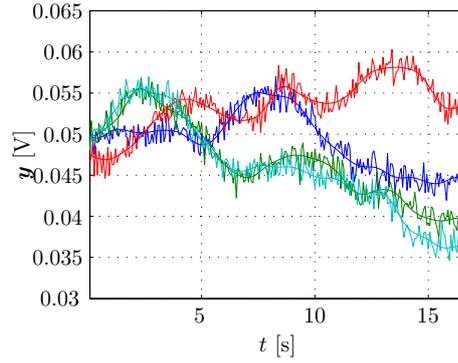
autour de  $0,04 \text{ arcsecondes}^{-1}$  pour l'ordre 0 donc légèrement inférieure à la fréquence de coupure du miroir.

L'espacement entre deux cornets est de  $2\lambda/D$  (partie 2.2.3) soit une fréquence d'échantillonnage de  $D/(2\lambda) \approx 0,022 \text{ arcsecondes}^{-1}$  quatre fois plus faible que la condition de Shannon. L'enjeu de la sur-résolution est notamment de restaurer des fréquences au-delà de cette fréquence d'échantillonnage naturelle du capteur. C'est pourquoi les protocoles d'acquisition effectuent des incréments de pointage de 2 secondes d'arc ce qui revient à une fréquence d'échantillonnage *équivalente* de  $f_e = 0,5 \text{ arcsecondes}^{-1}$ . Cependant, restaurer des fréquences de  $0,25 \text{ arcsecondes}^{-1}$  est probablement très difficile (sans exploiter d'autre information disponible par ailleurs) puisque la fonction de transfert optique (la figure d'*Airy*) atténue fortement toutes les fréquences au-delà de  $0,05 \text{ arcsecondes}^{-1}$ . De plus les possibilités de restauration dépendent beaucoup du niveau de bruit. La question est donc de savoir s'il est possible, et sous quelles conditions, de restaurer des fréquences au-delà de la fréquence d'échantillonnage spatiale naturelle du capteur de  $0,022 \text{ arcseconde}^{-1}$ , et de déterminer empiriquement les limites.

#### 4.1.4 Simulation de données

Tous les résultats présentés ont été obtenus à l'aide de données produites par notre modèle instrument. Le bruit correspond également au modèle, c'est-à-dire gaussien blanc *iid* avec un écart-type de  $\sigma_b = 10^{-3} \text{ V}$ , identique pour tous les bolomètres (divisé par le gain de la réponse instrumentale cela correspond environ à  $1,4 \times 10^{-6} \text{ W.arcseconde}^{-2}$ ). Enfin on ne considère que les résultats pour la bande spectrale  $360 \mu\text{m}$  puisque pour le moment aucune information n'est injectée dans l'inférence justifiant l'utilisation des trois bandes spectrales conjointement. La figure 4.5 montre un exemple de données de quatre bolomètres, générées avec le *cirrus* comme ciel observé.

**Remarque.** *Tous les résultats d'estimation et d'inversion présentés ont été obtenus avec le même*



**Figure 4.5** – Données produites par quatre bolomètres pendant un balayage. La sortie du modèle direct est superposée à la sortie bruitée. Le ciel vrai est le *cirrus*.

*jeu de données (ou la même réalisation de bruit) sauf lorsque cela est nécessaire, par exemple pour le calcul de biais. Les résultats présentés sont des cas nominiaux et plusieurs tests, non présentés, ont été réalisés avec différentes réalisations et niveaux de bruit.*

Les convolutions sont calculées par transformée de Fourier discrète. La gestion des bords est faite ainsi : tout d’abord la taille de la zone observée est calculée puis un bord est ajouté pour permettre un calcul exact de la convolution, par transformée de Fourier, au centre de l’image. Pour la génération des données, le ciel vrai est inscrit dans cette taille d’image figure 4.1.

#### 4.1.5 Algorithmes

L’initialisation de l’algorithme d’optimisation et d’échantillonnage est la co-addition

$$\hat{\mathbf{x}} = \mathbf{P}^t (\mathbf{P}\mathbf{P}^t)^{-1} \mathbf{y}$$

ou la solution

$$\hat{\mathbf{x}} = (\mathbf{P}^t \mathbf{P} + \mu \mathbf{D}^t \mathbf{D})^{-1} \mathbf{P}^t \mathbf{y}$$

c’est-à-dire la co-addition avec un *a priori* de douceur,  $\mathbf{D}$  étant les différences premières. Après une dizaine d’itérations, l’influence de l’initialisation a toujours été visuellement imperceptible. Le paramètre  $\mu$  est déterminé de façon *ad-hoc*.

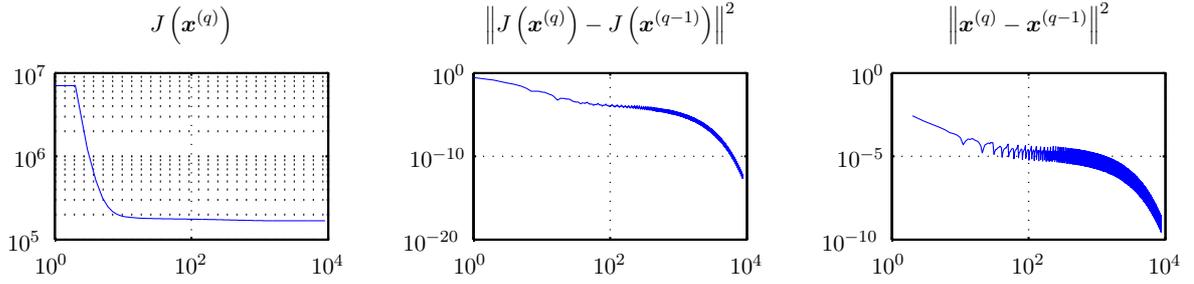
Deux conditions d’arrêt sont utilisées pour stopper l’algorithme d’optimisation :

1. la norme  $\|\mathbf{x}^{(q)} - \mathbf{x}^{(q-1)}\|^2$
2. la décroissance du critère  $J(\mathbf{x}^{(q)}) - J(\mathbf{x}^{(q-1)})$

toutes deux inférieures à  $10^{-6}$ . Une condition sur le nombre d’itérations maximum fixée à 200 est parfois atteinte. L’algorithme d’optimisation adapte le pas automatiquement [BCG<sup>+</sup>06].

Pour déterminer si l’algorithme se comporte bien, la figure 4.6 présente différents éléments typiques de l’évolution de l’algorithme (ces caractéristiques ont toujours été observées). La valeur du critère  $J(\mathbf{x})$  diminue strictement au cours des itérations. Les conditions d’arrêt tendent vers zéro, notamment la différence sur l’objet. L’algorithme a itéré pendant plus de 5000 itérations sans évolution visible de la solution après environ 200 ou 300 itérations quelles que soient les conditions.

Les sauts observés correspondent à une réinitialisation de la direction de descente à la valeur du gradient et non plus la direction conjuguée. Ces évolutions du critère et des conditions d’arrêt ont toujours été observées. En fonction de la taille du problème et des conditions d’arrêt, une



**Figure 4.6** – Évolution du critère et des conditions d’arrêt de l’algorithme en échelle log–log. L’axe horizontal correspond au nombre d’itérations.

optimisation s’effectue en environ 150 itérations calculées en quelques minutes (pour une image de 563 par 579 coefficients, et une bande spectrale).

La condition d’arrêt de la chaîne MCMC est une différence relative (après une période de chauffe de 400 échantillons fixée de manière *ad-hoc*) inférieure à  $e = 10^{-4}$  entre deux moyennes successives

$$\left\| \mathbf{x}_{\text{EAP}}^{(q)} - \mathbf{x}_{\text{EAP}}^{(q-1)} \right\|^2 / \left\| \mathbf{x}_{\text{EAP}}^{(q-1)} \right\|^2 \leq e.$$

Le calcul total nécessite environ quatre jours pour les problèmes considérés. Les étapes nécessitant la plus grande part des ressources de calcul sont dans l’algorithme d’optimisation et plus exactement dans le calcul des sorties modèles  $\mathbf{H}\mathbf{x}$  et des transposés  $\mathbf{H}^t\epsilon$ .

#### 4.1.6 Outils d’évaluation

Pour évaluer la méthode et présenter des résultats, plusieurs outils sont utilisés. Tout d’abord l’image reconstruite et des coupes de l’image sont utilisées pour une évaluation qualitative.

Des spectres de puissances sont également calculés pour visualiser les fréquences restaurées et la limite atteinte. Pour leurs calculs plusieurs étapes sont effectuées. Tout d’abord la partie centrale observée (donc sans les bords) est extraite. Ensuite une apodisation gaussienne sur l’image est appliqués pour s’affranchir des problèmes provoqués par la périodisation. Enfin une moyenne circulaire de la norme au carré de la transformée de Fourier de l’image est calculé. L’apodisation réduit la résolution spectrale ce qui n’est pas un problème pour la fréquence limite que l’on souhaite visualiser. Généralement les figures des spectres contiennent le spectre du vrai, l’estimation par inversion (non-supervisé ou supervisé), la co-addition et l’image uniquement convoluée par la réponse impulsionnelle ( $\mathbf{H}\mathbf{x}^*$  où  $\mathbf{x}^*$  est le vrai ciel). La comparaison entre l’image convoluée et la co-addition permet de mettre en évidence la proportion entre les fréquences atténuées de l’objet et du bruit.

Enfin on utilise également les mesures d’erreur de reconstruction de l’image

$$\ell_2(\gamma_{\mathbf{x}}) = \|\mathbf{x}^* - \hat{\mathbf{x}}(\gamma_{\mathbf{x}})\|^2 / \|\mathbf{x}\|^2 \quad \text{et} \quad \ell_1(\gamma_{\mathbf{x}}) = \|\mathbf{x}^* - \hat{\mathbf{x}}(\gamma_{\mathbf{x}})\|_1 / \|\mathbf{x}\|_1,$$

dont on fait apparaître ici la dépendance au paramètre de régularisation  $\gamma_{\mathbf{x}}$  et où  $\mathbf{x}^*$  est le vrai ciel. Plus exactement l’inversion dépend du rapport  $\gamma_{\mathbf{x}}/\gamma_{\mathbf{b}}$  mais la valeur  $\gamma_{\mathbf{b}}$  a été fixée arbitrairement à la vraie valeur puisque cette connaissance ne suffit pas pour déterminer le rapport. Cette mesure d’erreur permet de déterminer un paramètre optimal lorsque l’on connaît la vérité terrain  $\mathbf{x}^*$ , et de comparer au cas où le rapport  $\gamma_{\mathbf{x}}/\gamma_{\mathbf{b}}$  est estimé.

Dans le corps du texte seront également présentés d’autres outils d’analyse comme les variances *a posteriori* utilisées comme mesure d’incertitude sur la reconstruction ou encore des outils d’analyse de l’estimateur au travers de biais et de variance.

## 4.2 Moindres carrés régularisés

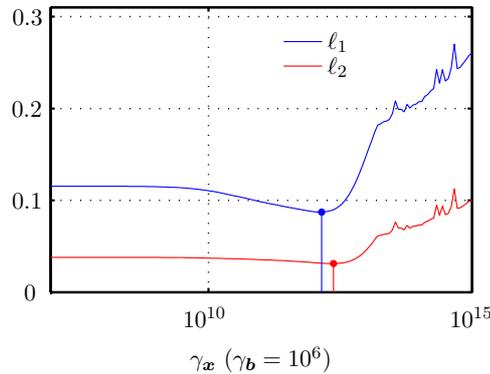
Dans cette partie nous présentons les résultats d'estimation par moindres carrés régularisés à hyper-paramètres fixés. L'estimateur est donc le maximum ou la moyenne *a posteriori* connaissant les hyper-paramètres et les paramètres instruments.

Tout d'abord une étude sur le *cirrus* est menée pour étudier le comportement de la méthode avec un ciel jugé réaliste tout en étant conforme au modèle *a priori* pour l'image. On étudie ensuite la méthode et l'estimateur avec une variation du bruit et du ciel d'entrée. Le biais et la variance de l'estimateur sont également étudiés.

Sauf mention contraire les images présentées sont des images des coefficients  $\mathbf{x}$  et non de la fonction  $\phi$ . La différence pour les cas présentés est en générale minime. On ne considère qu'un seul ordre de développement en  $\lambda$ , une seule régularisation sur les dimensions spatiales de  $\mathbf{x}$  avec un pas d'échantillonnage de 2 secondes d'arc. Il y a un seul niveau de bruit pour tous les bolomètres donc  $\gamma_{\mathbf{x}}$  et  $\gamma_{\mathbf{b}}$  sont des scalaires. Le formalisme présenté dans le chapitre 3 permet d'étendre le problème à l'estimation d'un paramètre de bruit par bolomètre sans difficultés théoriques ou techniques.

### 4.2.1 Restauration du *cirrus*

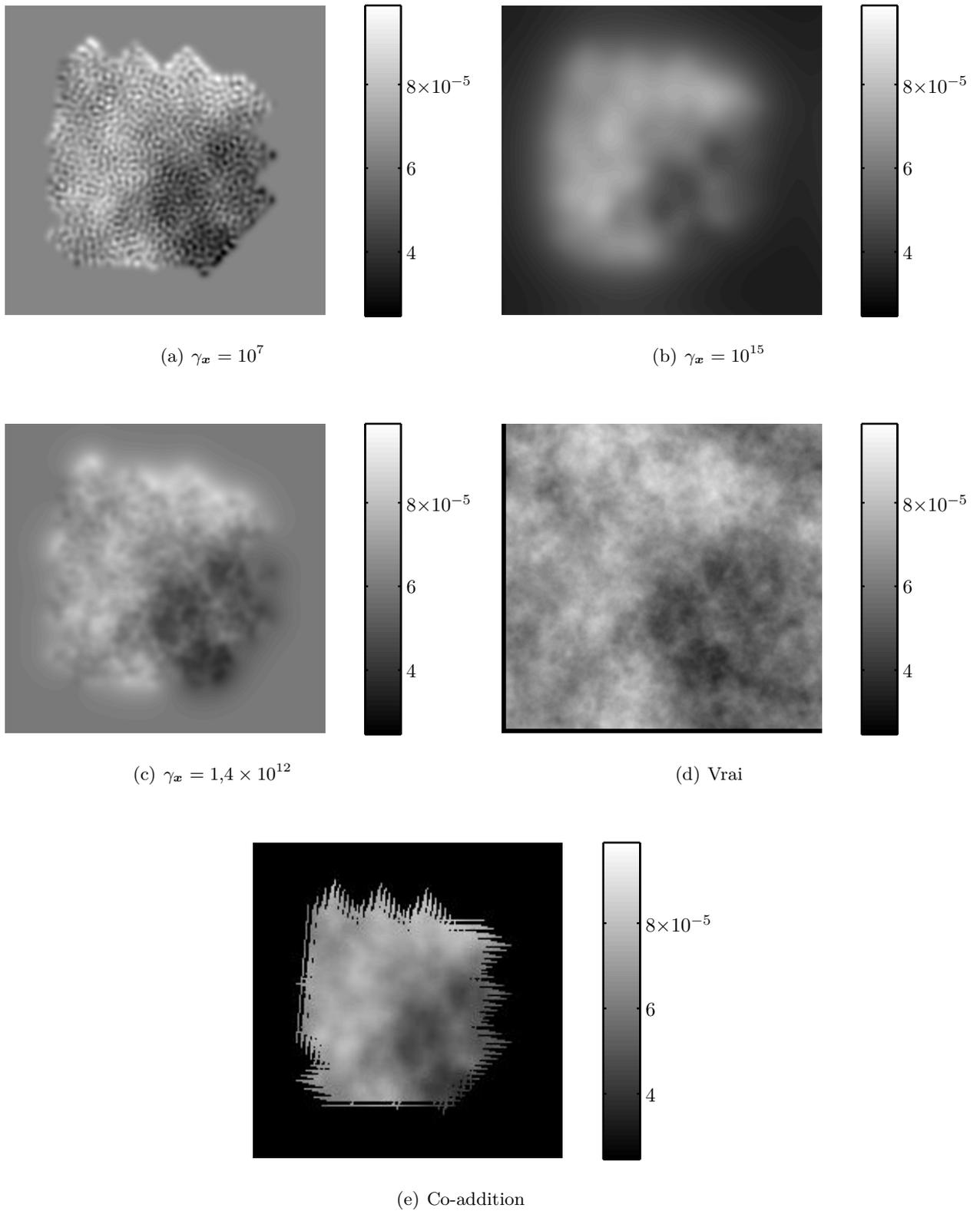
Les résultats présentés ici ont été obtenus avec des données simulées par le modèle direct et le *cirrus* en entrée. Les distances  $\ell$  tracées figure 4.7, permettent de déterminer le rapport optimal  $\gamma_{\mathbf{x}}/\gamma_{\mathbf{b}}$  pour un niveau de bruit  $\gamma_{\mathbf{b}}$  fixé au vrai arbitrairement.



**Figure 4.7** – Distance au ciel vrai en fonction de l'hyper-paramètre en abscisse logarithmique sur 144 points. Le minimum est repéré par un point sur la courbe. Le minimum de la distance  $\ell_1$  est atteint en  $\gamma_{\mathbf{x}} = 1,4 \times 10^{12}$ . Le minimum de la distance  $\ell_2$  est atteint en  $\gamma_{\mathbf{x}} = 2,3 \times 10^{12}$ . La distance absolue est moins sensible aux variations et fournit un paramètre plus petit.

Les figures 4.8(c), 4.9(c) et 4.9(d) correspondent à l'image, la coupe et le spectre obtenus avec le paramètre de régularisation qui minimise la distance  $\ell_1$ . Les images obtenues avec les deux paramètres optimaux sont visuellement très proches aussi bien sur l'image que sur le spectre. Que ce soit sur l'image ou la coupe on voit que des oscillations ou des échelles spatiales sont restaurées jusqu'à une limite de 0,03 arcsecondes<sup>-1</sup>. Cette limite semble correspondre au niveau où le bruit commence à dominer, comme illustré par les spectres figures 4.9(b), 4.9(d) et 4.9(f). Le profil global est retrouvé sauf au bord où aucune donnée n'est disponible.

La figure 4.8(e) présente le résultat de la co-addition, c'est-à-dire  $\hat{\mathbf{x}} = \mathbf{P}^t(\mathbf{P}\mathbf{P}^t)^{-1}\mathbf{y}$  avec une résolution de 6 secondes d'arc pour éviter que certains coefficients ne soient liés à aucune donnée (éviter les « trous »). Les données utilisées ainsi que le pointage sont exactement les mêmes que pour



**Figure 4.8** – Image des coefficients estimés pour différents paramètres de régularisation  $\gamma_x$ . La figure 4.8(a) correspond à l'estimation avec une sous-régularisation. À l'inverse l'image 4.8(b) montre une sur-régularisation. L'image 4.8(c) correspond à la restauration qui minimise la distance  $\ell_1$ . La figure 4.8(d) est le ciel vrai. L'image 4.8(e) est la co-addition des données plus douce que 4.8(c).

l'inversion proposée. En comparaison de la figure 4.8(c), la co-addition est visiblement plus basse-fréquence. La coupe 4.9(c) montre que certaines structures à des fréquences spatiales plus élevées sont visibles. Ce phénomène est également observable sur le spectre figure 4.9(d) : la courbe bleue correspondant à l'image restaurée est au dessus de la courbe noire du spectre de la co-addition.

Les figures 4.8 et 4.9 présentent également les images, coupes et une estimation des spectres de l'image pour trois paramètres de régularisation différents, respectivement  $\gamma_{\mathbf{x}} = 10^7$ ,  $\gamma_{\mathbf{x}} = 1,4 \times 10^{12}$  et  $\gamma_{\mathbf{x}} = 10^{15}$ . Sur la figure 4.8(a) le paramètre de régularisation est trop faible et le ciel est considéré peu régulier. On voit des oscillations importantes visibles également sur la coupe figure 4.9(a). Concernant l'approximation du spectre, des fréquences autour de  $0,03 \text{ arcseconde}^{-1}$  sont amplifiées correspondant au niveau où le bruit commence à dominer les fréquences atténuées venant de l'objet. Si l'algorithme n'avait pas été stoppé les oscillations auraient pu être plus importantes.

A l'inverse les figures 4.8(b), 4.9(e) et 4.9(f) présentent le résultat avec un paramètre de régularisation trop important. Le ciel est supposé très régulier. Les fréquences ne sont pas restaurées. Ce phénomène est bien visible sur le spectre 4.9(f) situé en dessous du spectre, en tirets noirs, de l'image simplement convoluée par la réponse impulsionnelle de la figure 4.4 (c'est-à-dire sans prélèvement pendant les balayages par les bolomètres).

Concernant les effets aux bords de l'observation on constate que la méthode proposée effectue de l'extrapolation grâce à l'*a priori*. Il s'agit d'un phénomène de diffusion d'information vers l'extérieur bien que le manque de données diminue la confiance accordée à l'estimée. La mise en oeuvre par transformée de Fourier implique également une hypothèse de continuité sur les bords de l'image ce qui explique la jonction au niveau des frontières.

La co-addition possède une structure très hachée sur les bords à cause du manque de données liées à certains pixels. La redondance introduite par le balayage permet de composer le faible échantillonnage mais seulement au centre de la zone d'intérêt.

En conclusion, la méthode proposée permet la restauration de fréquences spatiales précédemment atténuées par l'instrument. Le contenu spatial de l'estimée dépend du paramètre de régularisation et un paramètre optimal existe permettant la restauration de structures fines. Le compromis semble correspondre au niveau où les fréquences du bruit dominent les fréquences atténuées de l'objet.

#### 4.2.2 Variance *a posteriori* et incertitude

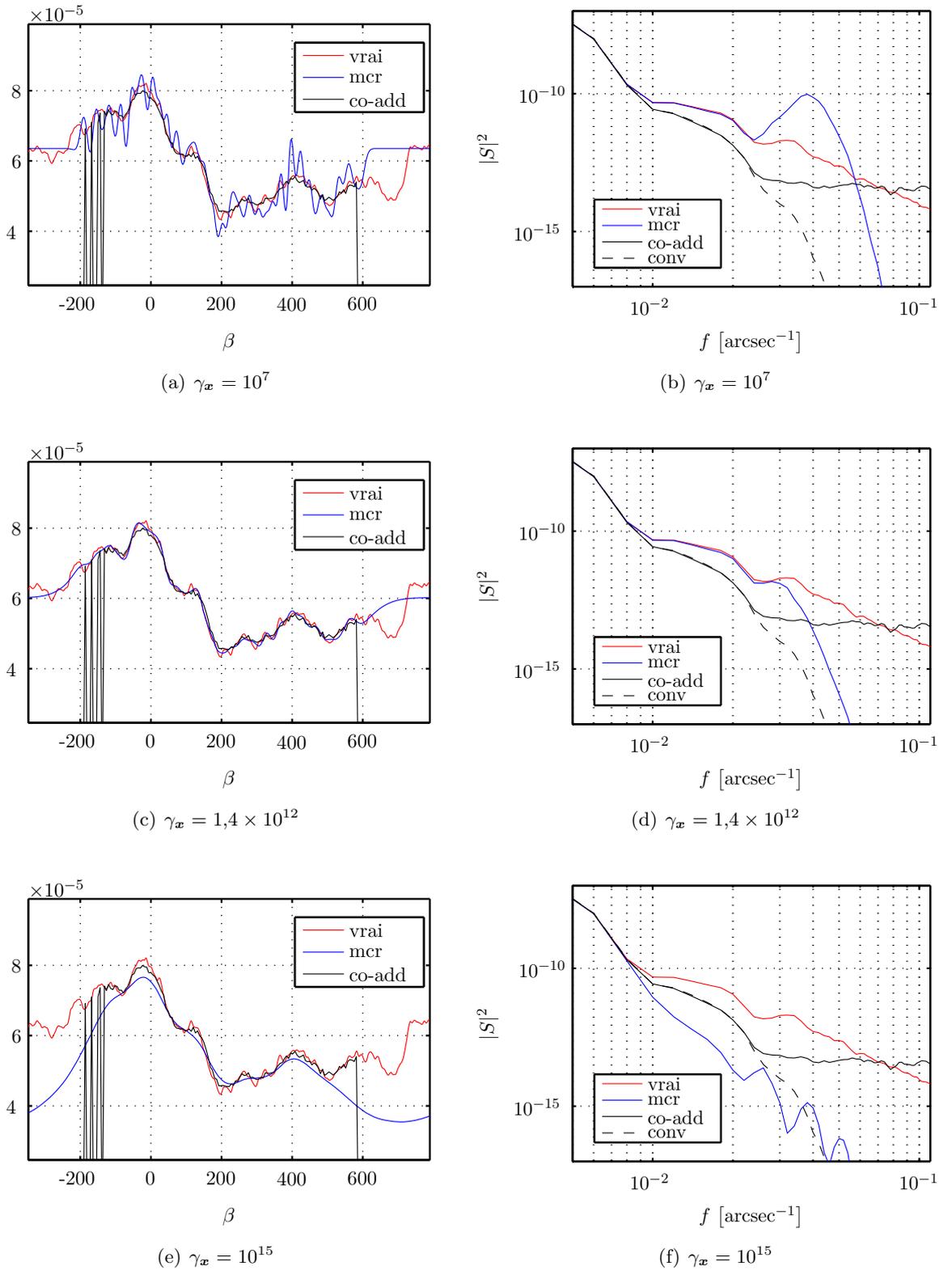
La figure 4.10 page 120 représente l'écart-type de la loi *a posteriori*  $p(\mathbf{x}|\mathbf{y}, \gamma_{\mathbf{x}}, \gamma_{\mathbf{b}})$  de chaque pixel. C'est une mesure de l'incertitude sur l'objet d'intérêt. C'est donc la racine carrée de la diagonale de

$$\mathbb{E} \left[ (\mathbf{x} - \mathbf{m}_{\mathbf{x}}) (\mathbf{x} - \mathbf{m}_{\mathbf{x}})^t \right]$$

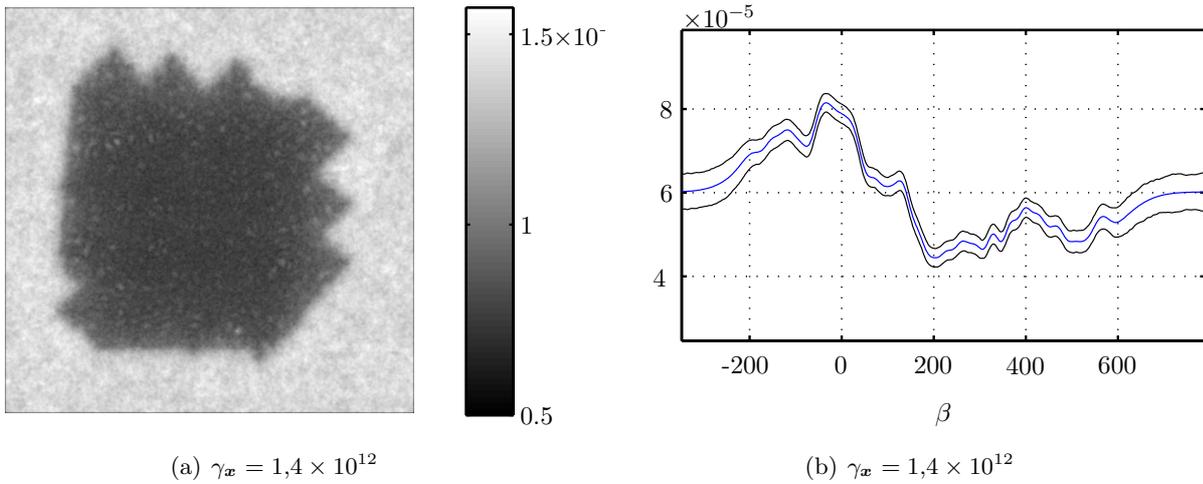
où  $\mathbf{x}$  suit la loi *a posteriori*  $p(\mathbf{x}|\mathbf{y}, \gamma_{\mathbf{x}}, \gamma_{\mathbf{b}})$  de moyenne  $\mathbf{m}_{\mathbf{x}}$  (qui est la solution des moindres carrés régularisés). Cette variance est calculée en générant 800 échantillons de la loi avec l'algorithme décrit partie 3.2.2. Pour générer chaque échantillon, le même jeu de données est utilisé. On perturbe cependant le système pour introduire une fluctuation stochastique permettant d'approcher la variance *a posteriori*. Chaque échantillon produit possède des fluctuations autour de la moyenne pour l'image qui est la solution des moindres carrés régularisés. Ces fluctuations représentent l'incertitude sur la valeur de l'objet.

Visuellement on observe que la variance diminue lorsque des données fournissent une information sur le ciel. La moyenne de l'écart-type au centre de l'image est  $6,8 \times 10^{-7} \text{ W.arcseconde}^{-2}$  pour un ciel d'un niveau moyen autour de  $6 \times 10^{-5} \text{ W.arcseconde}^{-2}$ .

La méthode permet d'extrapoler là où des données sont manquantes ou moins informatives. En



**Figure 4.9** – Coupe, à gauche, et approximation du spectre à moyenne circulaire à droite en échelle log–log, des coefficients restaurés pour différents paramètres de régularisation  $\gamma_x$ . La courbe rouge est le spectre des vrais coefficients et la courbe bleue est le spectre de l’image restaurée. La courbe noire en tirets correspond au spectre de l’image convoluée avec la réponse impulsionnelle sans prélèvement. La courbe noire correspond à la co-addition.



**Figure 4.10** – Écart-type *a posteriori*  $\sigma$  de l’objet calculé avec 800 échantillons pour le meilleur paramètre au sens de la distance  $\ell_1$ . La figure 4.10(a) est une image de l’incertitude visiblement moins importante au centre là où il y a des données. Les extrémités des balayages possèdent une variance plus importante. La figure 4.10(b) est une coupe de l’estimation placée dans un intervalle de  $\pm 3\sigma$ .

examinant l’incertitude on voit nettement que l’écart-type augmente progressivement sur les bords là où il y a de moins en moins de données.

### 4.2.3 Variation du niveau de bruit

Comme on peut le voir sur l’ensemble des spectres, les possibilités de restauration des fréquences sont en partie conditionnées par le niveau du bruit relativement au signal filtré. La figure 4.11(a) illustre le résultat de la méthode avec un niveau de bruit 10 fois plus élevé ( $\sigma_b = 10^{-2}$  V). La figure 4.11(b) correspond quant à elle à un niveau de bruit 10 fois plus faible ( $\sigma_b = 10^{-4}$  V).

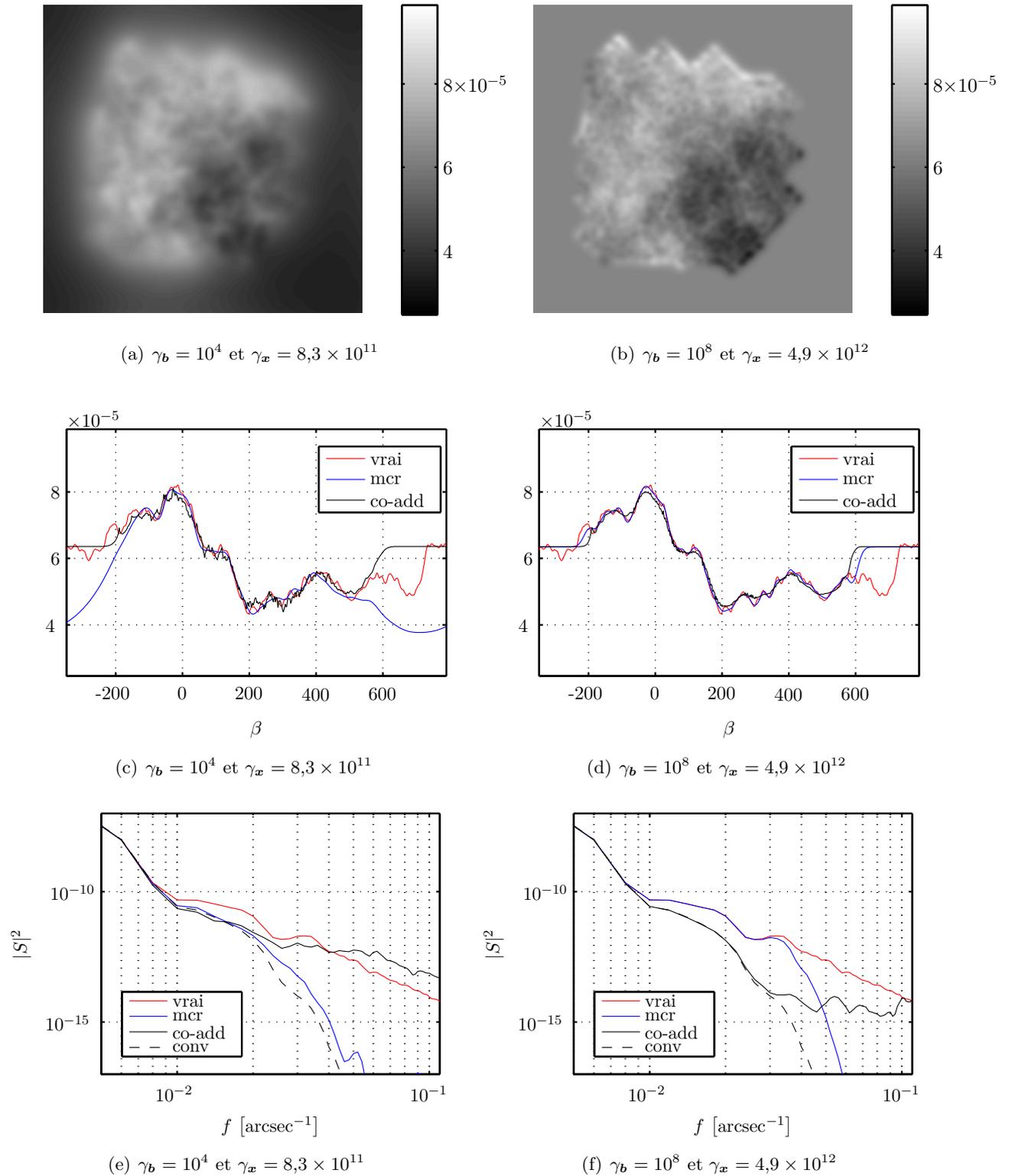
L’estimateur n’est pas linéaire vis-à-vis des paramètres de régularisation. Donc même si le contenu spatial dépend du rapport  $\gamma_x/\gamma_b$  et que l’on connaît la variation de  $\gamma_b$ , cette connaissance ne suffit pas pour régler la méthode puisque nous avons un nouveau jeu de données. Pour cela les distances  $\ell$  ont à nouveau été calculées pour déterminer un paramètre optimal  $\gamma_x$  avec ces jeux de données. Le tableau 4.1 contient les valeurs du paramètre de compromis données-*a priori* et montre bien que celui-ci diminue avec la puissance du bruit.

Bruit « élevé »	Bruit « normal »	Bruit « faible »
$\frac{\gamma_x}{\gamma_b} = \frac{8,3 \times 10^{11}}{10^4} = 8,3 \times 10^7$	$\frac{\gamma_x}{\gamma_b} = \frac{1,4 \times 10^{12}}{10^6} = 1,4 \times 10^6$	$\frac{\gamma_x}{\gamma_b} = \frac{4,9 \times 10^{12}}{10^8} = 4,9 \times 10^4$

**Tableau 4.1** – Paramètre de régularisation optimal au sens de la distance  $\ell_1$  en fonction du niveau de bruit.

Comme attendu, lorsque le niveau de bruit augmente la restauration est moins bonne puisque plus de fréquences ont été dégradées. Dans le spectre figure 4.11(e), l’inversion proposée est au-dessus de la co-addition jusqu’à 0,015 arcsecondes<sup>-1</sup> mais ne parvient pas à atteindre le niveau de l’objet vrai, probablement parce que les fréquences ont été trop dégradées.

A l’inverse lorsque le niveau de bruit diminue, il est possible de restaurer des fréquences un



**Figure 4.11** – Estimation avec deux niveaux de bruit différents. La colonne de gauche correspond à un bruit plus important nécessitant une forte régularisation. À l'inverse la colonne de droite correspond à un bruit faible permettant une régularisation plus faible et une restauration plus importante des hautes fréquences.

peu plus élevées. Il semble visuellement, à la fois dans l'image et la coupe, que des structures fines ont été restaurées. Sur le spectre, la limite atteinte est autour de  $0,35 \text{ arcsecondes}^{-1}$  contre un peu moins de  $0,3 \text{ arcsecondes}^{-1}$  avec un niveau de bruit de  $\sigma_b = 10^{-3}$ . La diminution du niveau de bruit permet une meilleure restauration des fréquences.

#### 4.2.4 Biais et variance de l'estimateur

La figure 4.12 montre une approximation numérique du biais et de la variance de l'estimateur (les expressions explicites nécessitent l'inversion de matrices trop grandes). Le biais  $\mathbb{E}[\hat{\mathbf{x}}] - \mathbf{x}^*$  de l'estimateur  $\hat{\mathbf{x}}$  pour le ciel vrai  $\mathbf{x}^*$  est calculé par moyenne arithmétique des estimateurs calculés avec des jeux de données différents. Chaque jeu de données a été obtenu par simulation à partir du même ciel avec une réalisation de bruit différente. Le calcul est donc

$$\mathbb{E}[\hat{\mathbf{x}}] - \mathbf{x}^* \approx \frac{1}{Q} \sum_{q=0}^{Q-1} \hat{\mathbf{x}}^{(q)} - \mathbf{x}^*.$$

et le calcul a été effectué avec  $Q = 104$  réalisations de bruit. Ce calcul a été effectué pour trois paramètres de régularisation. La variance pour chaque pixel correspond à la diagonale de

$$\mathbb{V}[\hat{\mathbf{x}}] = \mathbb{E} \left[ (\hat{\mathbf{x}} - \mathbb{E}[\hat{\mathbf{x}}]) (\hat{\mathbf{x}} - \mathbb{E}[\hat{\mathbf{x}}])^t \right].$$

La diagonale peut être approchée par

$$\text{diag}[\mathbb{V}[\hat{\mathbf{x}}]] \approx \frac{1}{Q} \sum_{q=0}^{Q-1} \hat{\mathbf{x}}^{(q)} \odot \hat{\mathbf{x}}^{(q)} - \mathbb{E}[\hat{\mathbf{x}}] \odot \mathbb{E}[\hat{\mathbf{x}}]$$

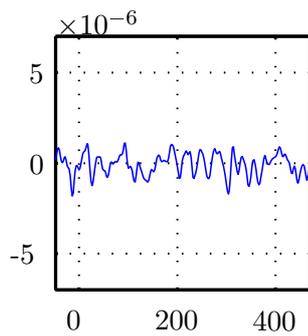
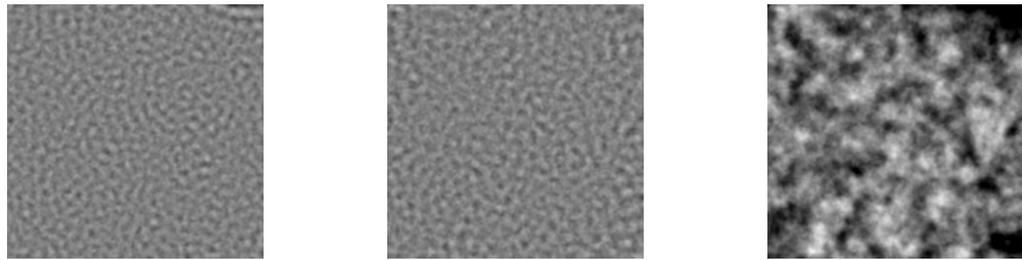
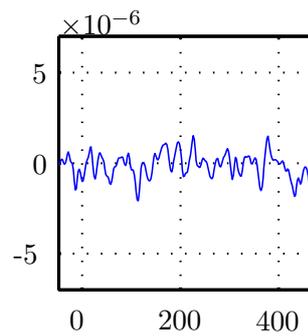
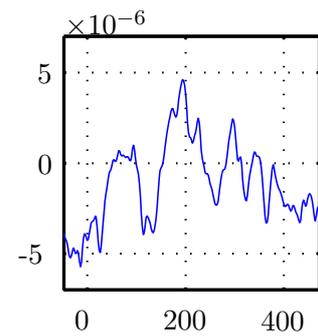
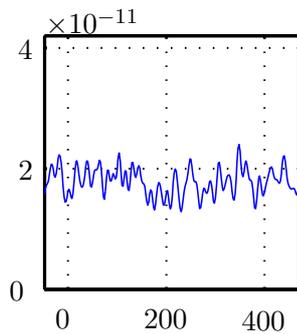
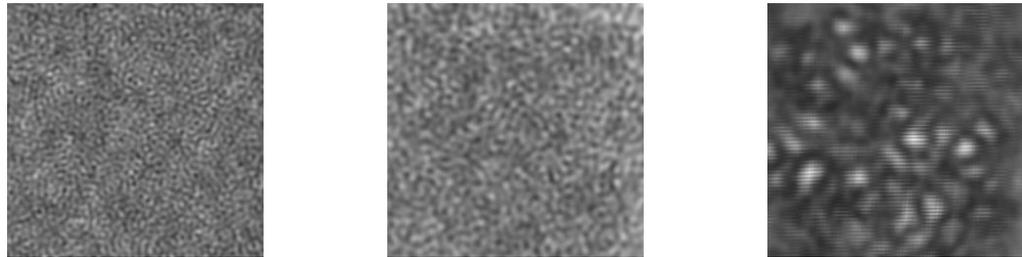
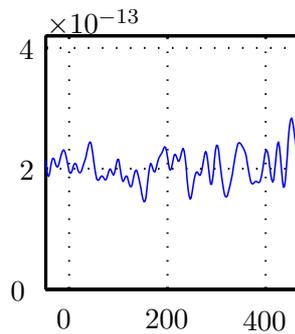
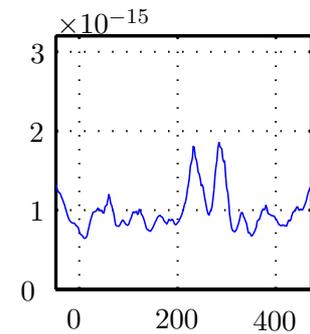
où  $\odot$  correspond au produit terme à terme de deux vecteurs.

Les valeurs de l'estimateur aux bords là où aucune information n'est apportée par les données sont moins intéressantes et gênent la lecture. Elles ne sont donc pas représentées. Les coupes montrent que la moyenne spatiale du biais de l'estimateur est à peu près nulle, ce qui est conforme avec le fait qu'il n'y a pas de pénalisation sur le niveau moyen de l'image mais seulement sur les différences. Le biais possède par contre des fluctuations. Elles correspondent aux structures spatiales, ou aux fréquences, qui ne peuvent pas être reconstruites. En augmentant le paramètre de régularisation  $\gamma_x$ , donc en supposant une plus grande régularité spatiale, des fluctuations de plus grandes échelles sont de plus en plus présentes, voir figure 4.12(c).

A l'inverse, diminuer l'*a priori* de régularité permet de diminuer le biais, ou d'augmenter l'espérance statistique de reconstruire plus de fréquence. Cependant la variance de l'estimateur devient de plus en plus importante. Les fréquences reconstruites venant de l'objet sont noyées dans les fréquences reconstruites provenant du bruit : une petite perturbation sur les données provoque de grandes modifications sur l'estimée. En *moyenne* (au sens de l'espérance statistique lorsque d'autres réalisations des données sont accessibles) le ciel est correctement estimé mais il y a une probabilité élevée pour que le jeu de données existant fournisse une estimation éloignée du vrai. La valeur  $\gamma_x = 1,4 \times 10^{12}$  correspond ainsi à un compromis biais-variance c'est-à-dire au minimiseur de l'erreur quadratique moyenne.

#### 4.2.5 Autres types de ciels

La méthode proposée est dédiée à la reconstruction d'images régulières jusqu'à un certain degré (déterminé par le rapport  $\gamma_x/\gamma_b$ ). Afin d'évaluer la robustesse de la méthode, elle a été évaluée avec deux ciels aux caractéristiques différentes.

(a) Biais pour  $\gamma_x = 10^7$ (b) Biais pour  $\gamma_x = 1,4 \times 10^{12}$ (c) Biais pour  $\gamma_x = 10^{15}$ (d) Variance pour  $\gamma_x = 10^7$ (e) Variance pour  $\gamma_x = 1,4 \times 10^{12}$ (f) Variance pour  $\gamma_x = 10^{15}$ 

**Figure 4.12** – En haut : images et coupes de biais de l'estimateur. En bas : image et coupe de la variance de l'estimateur. Chaque colonne correspond à une valeur donnée du paramètre  $\gamma_x$ . Comme attendu le biais augmente avec l'*a priori* de régularité. À l'inverse la variance diminue avec l'*a priori* de régularité.

- Tout d’abord elle a été évaluée avec le *cirrus* auquel ont été ajoutées des sources ponctuelles. Ces sources introduisent des ruptures franches et le modèle *a priori* est moins adapté.
- Le deuxième ciel étudié est une galaxie. Ce ciel est relativement régulier mais possède une structure plus complexe et une hétérogénéité spatiale. Il constitue un cas proche de la réalité et où le modèle est relativement adapté.

Pour les deux ciels, les distances  $\ell_1$  et  $\ell_2$  ont également été calculées. Les résultats présentés sont ceux qui minimisent ces erreurs de reconstruction.

Le modèle pour le ciel, introduisant une information sur la régularité spatiale ne prend pas bien en compte les phénomènes impulsions présents dans le ciel 4.13(a). Cependant en comparant avec la co-addition, figure 4.13(c), il apparaît que la méthode proposée restaure mieux l’impulsion présente autour de  $\beta = 350$ . De plus des structures sont mieux estimées par exemple autour de  $\beta = 100$ . Cependant, il apparaît un rebond à droite de l’impulsion en  $\beta \approx 350$ , caractéristique d’une déconvolution avec une pénalisation quadratique lorsqu’il y a des ruptures franches.

En ce qui concerne le profil de la galaxie, de nouveau certaines fréquences dégradées par l’instrument sont restaurées autour de  $\beta = 10$  et 250. La partie centrale est également plus résolue qu’avec la co-addition. La mesure de régularité au travers de la dérivée première semble adaptée également dans ce cas qui présente une structure spatiale différente du *cirrus*.

Concernant les deux spectres de puissance, figures 4.13(e) et 4.13(f), la méthode proposée restaure le niveau des fréquences jusqu’à  $0,02\text{--}0,025$  arcsecondes<sup>-1</sup> contrairement à la co-addition. Par rapport au *cirrus*, figure 4.9(d) les performances sont moins bonnes dans le sens où la limite de fréquences restaurées est plus faible qu’avec le *cirrus* seul. Cependant l’ordre de grandeur reste le même.

En conclusion la méthode proposée est relativement robuste par rapport à un ciel d’entrée présentant des caractéristiques mal prises en compte par le modèle *a priori*. Elle permet encore d’obtenir un ciel plus proche de la réalité que ce que permet la co-addition y compris dans un cas moins favorable.

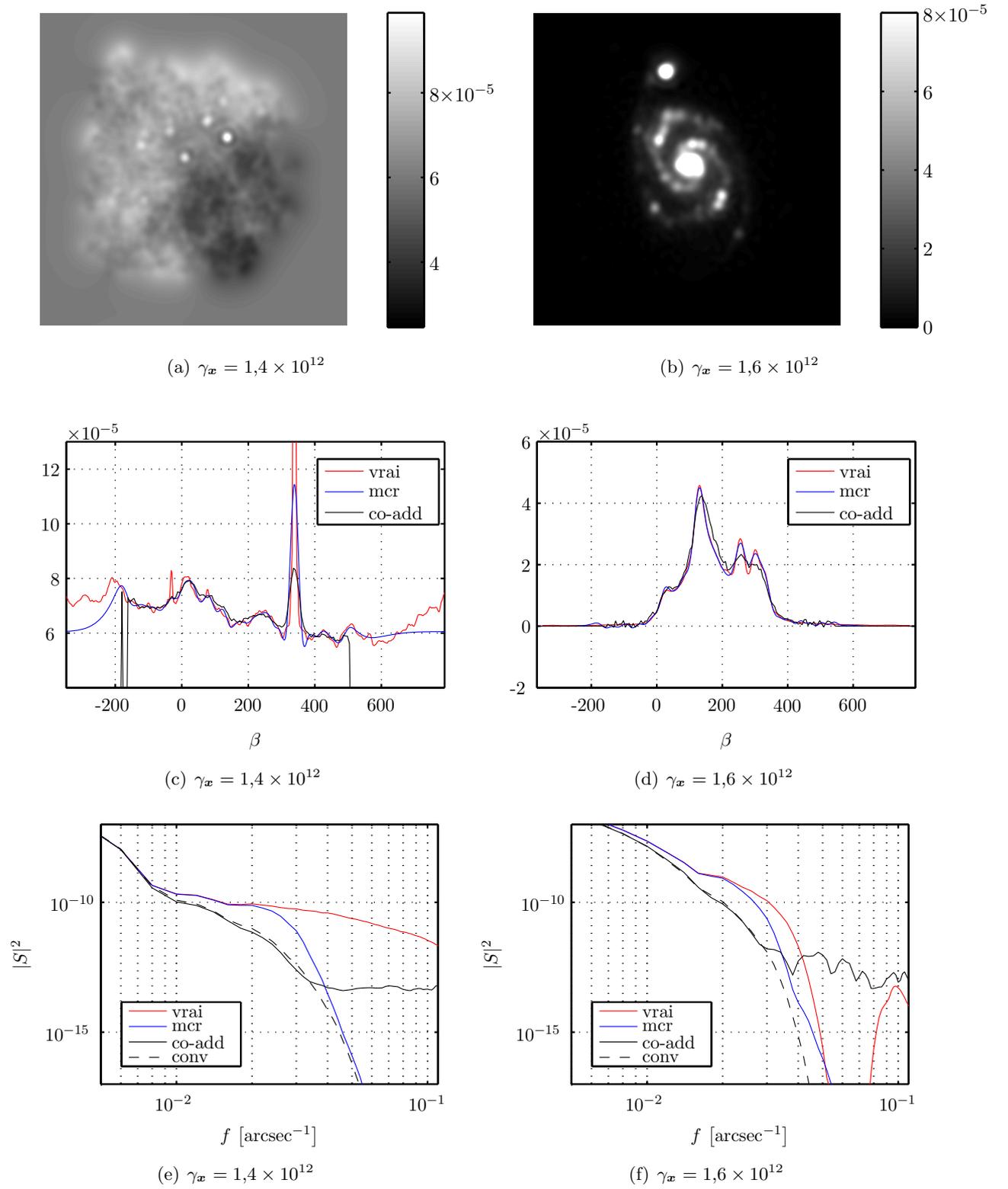
Cependant l’ensemble des résultats présentés ici nécessite de fixer de manière *ad-hoc* les paramètres  $\gamma_{\mathbf{b}}$  et  $\gamma_{\mathbf{x}}$ . Pour mener l’étude les distances  $\ell$  ont été utilisées mais elles sont inutilisables en pratique puisqu’elles nécessitent la connaissance de la vérité terrain. La prochaine partie présente des résultats non-supervisés : les paramètres  $\gamma_{\mathbf{x}}$  et  $\gamma_{\mathbf{b}}$  sont estimés conjointement avec l’image.

### 4.3 Estimation non-supervisée

La partie précédente montre des résultats obtenus avec une méthode où le rapport  $\gamma_{\mathbf{x}}/\gamma_{\mathbf{b}}$  doit être fixé. Nous avons vu que le résultat de l’estimation dépend en grande partie de ce rapport qui n’est jamais connu en pratique. Une deuxième catégorie supplémentaire de paramètres mal connus existe : les paramètres instruments ou du processus d’acquisition. Cette deuxième catégorie se distingue des hyper-paramètres dans le sens où, au moins dans cette application, on connaît avec une incertitude la valeur de ces paramètres.

Cette partie présente des résultats sur l’estimation non-supervisée c’est à dire de l’objet  $\mathbf{x}$  conjointement aux deux paramètres  $\gamma_{\mathbf{x}}$  et  $\gamma_{\mathbf{b}}$ . L’algorithme utilisé est l’échantillonneur de Gibbs, présenté partie 3.2.1, qui successivement échantillonne suivant la loi conditionnelle *a posteriori* de  $\mathbf{x}$ , de  $\gamma_{\mathbf{b}}$  puis de  $\gamma_{\mathbf{x}}$ . L’échantillonnage de  $\mathbf{x}$  est particulier puisqu’il est effectué par optimisation d’un système perturbé décrit partie 3.2.2.

Une étude sur l’estimation de paramètres instruments, avec un modèle convolutif donc invariant cette fois-ci, est présentée dans l’article de l’annexe F. Ce cas présente un modèle plus académique



**Figure 4.13** – Restauration du *cirrus* avec sources ponctuelles, sur la colonne de gauche, et de la galaxie sur la colonne de droite. Les paramètres optimaux ont été déterminés avec la distance  $\ell_1$ .

mais avec l'estimation des paramètres instruments et des hyper-paramètres. L'algorithme utilisé est également un échantillonneur de Gibbs. Les lois *a posteriori* conditionnelles pour les paramètres instruments étant plus complexes, des étapes d'échantillonnage de Metropolis-Hastings sont utilisées.

### 4.3.1 Estimation du ciel

Dans la suite nous présentons les résultats avec trois ciels :

- le *cirrus* étudié en détail
- le *cirrus* avec des sources ponctuelles pour évaluer la robustesse de la méthode
- une réalisation de la loi *a priori* permettant de connaître la vraie valeur du paramètre  $\gamma_{\mathbf{x}}$ .

#### Étude sur le *cirrus*

Nous présentons ici les résultats d'estimation non-supervisée avec le *cirrus*. L'image 4.14(a) correspond à la moyenne *a posteriori*. Elle est visiblement plus haute fréquence que la co-addition représentée figure 4.14(c) avec des structures plus fines. Les structures sont également bien visibles sur la coupe 4.14(b). Concernant l'estimation du spectre 4.14(d) on voit que des fréquences sont restaurées jusqu'à un niveau où le bruit commence à être amplifié. Donc la méthode proposée permet bien de restaurer des fréquences tout en estimant les hyper-paramètres déterminant le contenu spatial de l'objet.

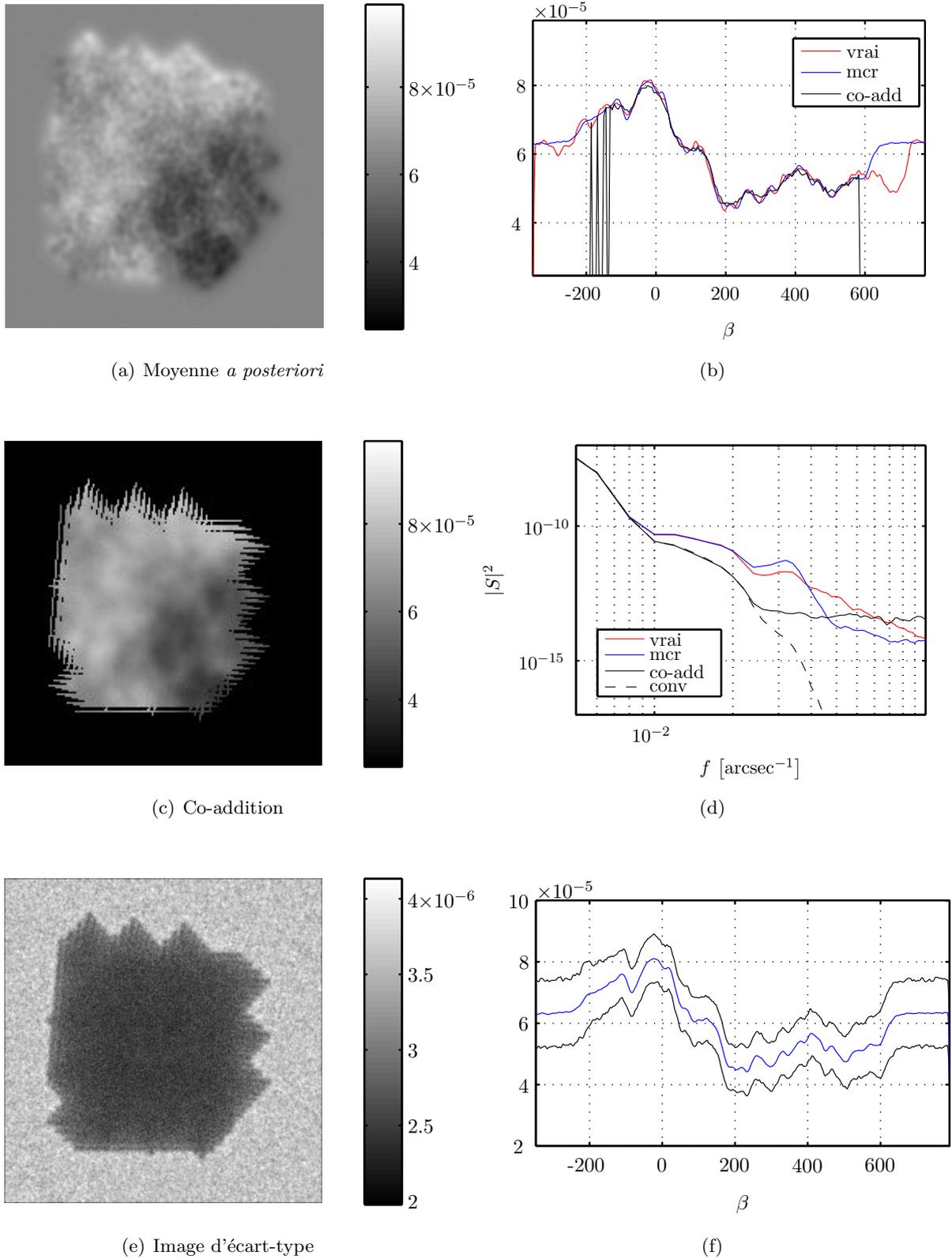
La figure 4.14(e) est une image de l'écart-type *a posteriori* pour chaque pixel obtenu avec les échantillons (environ 600) produits par l'algorithme MCMC. Tout comme pour le cas supervisé l'incertitude est moins importante là où des données fournissent une information. La méthode proposée extrapole le ciel (ou comble le manque de données) dans les zones où peu, où très peu, de données sont acquises. La méthode permet de quantifier l'incertitude sur cette extrapolation de la même façon que dans le cas supervisé. La limite sur les bords observés apparaît clairement comme étant une zone plus incertaine que le centre de l'image tout en étant plus sûre que sur l'extérieur. Notamment on peut observer un aspect de râteau qui s'identifie clairement avec les zones où très peu de données sont disponibles comme sur l'image de co-addition 4.14(c).

La moyenne de l'écart-type au centre de l'image, là où des données fournissent de l'information, est approximativement  $2,55 \times 10^{-6}$  où l'ordre de grandeur du ciel est  $5 \times 10^{-5}$ . Par rapport au cas supervisé, l'incertitude est plus grande, alors que les données sont les mêmes. Une interprétation probable est le fait que les paramètres  $\gamma_{\mathbf{x}}$  et  $\gamma_{\mathbf{b}}$  ne sont pas fixés. On injecte moins d'information dans le problème donc l'incertitude augmente. Il s'agit d'une illustration du fait que l'information présente dans le problème d'inférence dépend de l'état de connaissance de l'utilisateur et non uniquement des données.

En observant en détail le spectre il semble qu'il y ait une légère sous-régularisation. Ce phénomène est visible dans l'image et la coupe. L'image contient des « petites structures » typiques d'un phénomène de Gibbs provenant d'une sous-régularisation. Le spectre semble indiquer que les fréquences au-delà de  $0,02-0,025$  arcseconde<sup>-1</sup> sont plutôt des fréquences du bruit amplifiées.

#### Autres types de ciels

Comme pour le cas supervisé, la méthode est dédiée à la reconstruction d'image régulière. Ici le niveau de régularité est estimé. La méthode a été appliquée, avec les mêmes paramètres de simulation et de réglages, sur deux ciels différents :



**Figure 4.14** – Inversion non-supervisée pour le *cirrus*. La figure 4.14(a) est la moyenne *a posteriori* du ciel et la figure 4.14(b) une coupe. La figure 4.14(e) est une estimation de l'incertitude. La figure 4.14(f) est une coupe de la moyenne *a posteriori* encadrée par  $\pm 3$  fois l'écart-type *a posteriori*. Enfin la figure 4.14(d) est une estimation du spectre de puissance.

- une réalisation de la loi *a priori*, qui permet d’avoir un ciel simulé avec le même modèle utilisé pour l’inversion. De plus on connaît ainsi la vraie valeur du paramètre  $\gamma_{\mathbf{x}}$ , fixé à  $\gamma_{\mathbf{x}}^* = 4 \times 10^{11}$ , ce qui permet une comparaison avec la valeur estimée.
- le *cirrus* avec les sources ponctuelles. Cela permet d’évaluer la robustesse de la méthode et de l’estimation du compromis, lorsque la réalité terrain possède des ruptures franches en contradiction avec le modèle image.

Les figures 4.15(a), 4.15(c) et 4.15(e) présentent l’estimation lorsque le ciel vrai est une réalisation de la loi *a priori*. On constate que la méthode restaure bien des hautes fréquences et la limite de 0,03 arcseconde<sup>-1</sup> est atteinte. L’estimation du paramètre  $\gamma_{\mathbf{x}}$  est étudiée dans la partie suivante.

La figure 4.15 présente l’estimation pour le *cirrus* avec sources ponctuelles. Cette fois-ci le modèle *a priori* est moins adapté pour modéliser la réalité puisqu’il y a des ruptures franches. Cependant comme pour le *cirrus*, étudié dans la partie précédente, des structures spatiales ont visiblement été restaurées. La coupe montre également que le pic vers  $\beta = 450$  est plus résolu qu’avec la co-addition avec cependant des rebonds sur les côtés.

On peut en déduire que la méthode proposée estime correctement l’objet inconnu, ainsi que le compromis données–*a priori*, même si le modèle *a priori* n’est pas adapté à toutes les situations. La méthode ne semble pas sous ou sur-régulariser outre mesure.

### Erreur de reconstruction

Le tableau 4.2 regroupe différentes valeurs d’erreurs de reconstruction  $\ell_1$  et  $\ell_2$  pour l’image. La comparaison porte sur trois cas.

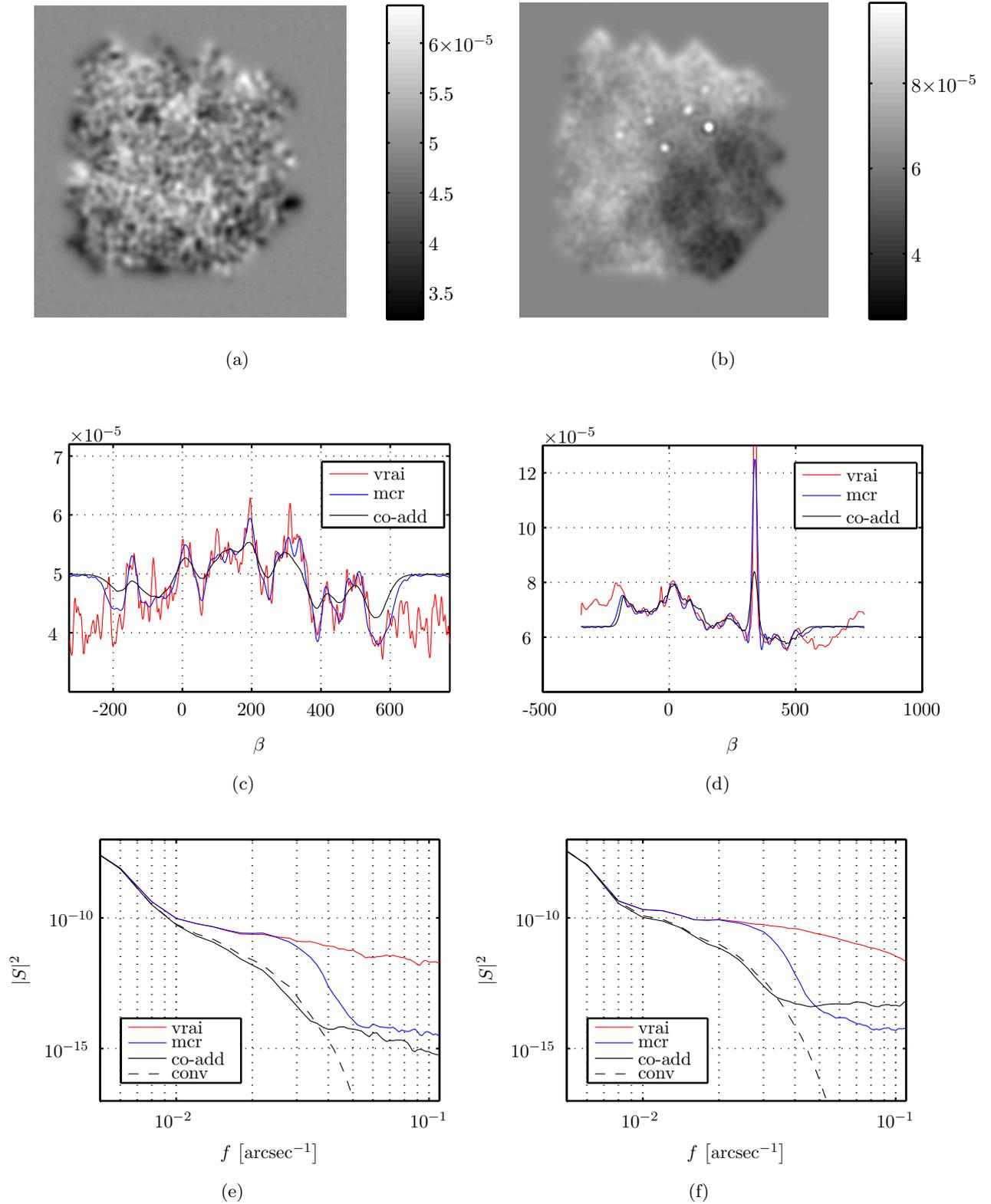
- Tout d’abord la valeur de l’erreur en fixant le paramètre  $\gamma_{\mathbf{x}}$  qui la minimise. Il s’agit donc de la méthode supervisée, étudiée partie 4.2, et ce paramètre est choisi en connaissant la vérité terrain. Il n’est pas possible d’avoir une erreur plus faible.
- Ensuite la valeur de l’erreur de reconstruction avec la méthode non-supervisée, donc sans connaître la vérité terrain.
- Enfin la valeur de l’erreur, avec la méthode supervisée, en fixant le paramètre  $\gamma_{\mathbf{x}}$  à la vraie valeur utilisée pour la simulation. Ce résultat ne peut être obtenu qu’avec la réalisation de la loi *a priori*.

		Réalisation <i>a priori</i>	<i>Cirrus</i>	<i>Cirrus</i> avec sources
Connaissant vérité terrain $\widehat{\gamma}_{\mathbf{x}}(\ell)$	$\ell_1$	6,35%	8,72%	8,82%
	$\ell_2$	0,69%	3,12%	3,19%
Non-supervisé	$\ell_1$	17,7%	9,42%	9,85%
	$\ell_2$	4,7%	3,47%	3,62%
Vraie valeur	$\ell_1$	17,57%	-	-
	$\ell_2$	4,62%	-	-

**Tableau 4.2** – Erreur de reconstruction  $\ell$  de l’image en pourcentage.

Concentrons-nous sur la distance  $\ell_1$ . Premièrement on peut constater que, mis à part la réalisation *a priori* dans le cas non-supervisé, moins le modèle image est adapté, plus l’erreur est importante, dans le cas supervisé et non-supervisé. Dans le cas supervisé on passe de 6,35 % d’erreur pour la réalisation de la loi *a priori* à 8,82% pour le *cirrus* avec des sources ponctuelles.

Bien sûr, par rapport à la méthode supervisée optimale connaissant la vérité terrain, l’erreur augmente lorsque la méthode non-supervisée est utilisée. Sans connaître la vérité terrain on passe, pour le *cirrus*, d’une erreur de 8,71% à 9,42% ce qui est une augmentation relativement faible. Sur la distance  $\ell_2$  on passe également d’une erreur de 3,12% à 3,47%.



**Figure 4.15** – Images, coupes et estimation du spectre à moyenne circulaire en échelle log–log pour la réalisation de la loi *a priori* à gauche et la *cirrus* avec sources ponctuelles à droite.

Un cas particulier et intéressant concerne la réalisation de la loi *a priori*. Tout d'abord on remarque que choisir la vraie valeur  $\gamma_{\mathbf{x}}^*$  pour la méthode supervisée ne donne pas l'erreur de reconstruction la plus faible. Elle est de 17,57% au lieu de 6,35% lorsque le paramètre minimisant l'erreur est choisi. L'erreur avec la méthode non-supervisée est de 17,7% donc très proche de l'erreur supervisée lorsque la vraie valeur est choisie. Ce résultat correspond au fait que la méthode fournit une estimation du paramètre  $\widehat{\gamma}_{\mathbf{x}}$  proche de la vraie valeur, comme nous verrons dans la prochaine partie. Les distances au contraire ne prennent en compte que l'erreur sur l'image. Donc la méthode non-supervisée semble vouloir minimiser l'erreur à la fois sur l'image *et* sur le paramètre.

Il semble cependant qu'il y a un comportement différent en fonction de l'adaptation du modèle *a priori* pour le ciel à la réalité terrain. Lorsque le modèle correspond parfaitement, la méthode fournit une estimation du paramètre  $\gamma_{\mathbf{x}}$  qui tend vers la vraie valeur avec une erreur plus grande sur l'image que dans le cas supervisé. Si le modèle n'est pas adapté alors l'erreur sur l'image est beaucoup plus proche du cas supervisé. Cependant la valeur estimée pour  $\gamma_{\mathbf{x}}$  est tout de même éloignée de la meilleure valeur pour l'erreur sur l'image. Donc lorsque le modèle est moins adapté à l'image, une erreur sur le paramètre a moins d'influence sur l'erreur de reconstruction.

### 4.3.2 Estimation et chaînes des hyper-paramètres

#### Valeur de l'estimation

Le tableau 4.3 regroupe l'estimation des paramètres  $\gamma_{\mathbf{b}}$  et  $\gamma_{\mathbf{x}}$  pour les trois ciels étudiés. Concernant  $\gamma_{\mathbf{b}}$  dans les trois cas la vraie valeur est connue et vaut  $10^6$ . La valeur estimée  $\widehat{\gamma}_{\mathbf{b}}$  dans les trois cas est très proche de la vraie valeur. De plus la vraie valeur est toujours comprise dans l'intervalle  $\pm 3\hat{\sigma}$ , où  $\hat{\sigma}$  est l'écart type *a posteriori* calculé avec les échantillons. Enfin nous avons constaté que la sur ou sous-estimation par rapport à la vraie valeur semble dans une grande mesure dépendre de la réalisation du bruit et non de la méthode.

	<i>Cirrus</i>	<i>Cirrus</i> et sources ponctuelles	Réalisation <i>a priori</i>
$\gamma_{\mathbf{b}}^*$	$10^6$	$10^6$	$10^6$
$\widehat{\gamma}_{\mathbf{b}}(\hat{\sigma})$	$1,009 \times 10^6 (4,07 \times 10^3)$	$1,003 \times 10^6 (4,04 \times 10^3)$	$1,003 \times 10^5 (4,05 \times 10^3)$
$\gamma_{\mathbf{x}}^*$	-	-	$4 \times 10^{11}$
$\widehat{\gamma}_{\mathbf{x}}(\hat{\sigma})$	$2,74 \times 10^{11} (4,62 \times 10^9)$	$2,47 \times 10^{11} (5,99 \times 10^9)$	$3,28 \times 10^{11} (1,07 \times 10^{10})$
$\widehat{\gamma}_{\mathbf{x}}(\ell_1)$	$1,40 \times 10^{12}$	$1,40 \times 10^{12}$	$7,36 \times 10^{11}$
$\widehat{\gamma}_{\mathbf{x}}(\ell_2)$	$3,93 \times 10^{12}$	$3,93 \times 10^{12}$	$8,37 \times 10^{11}$

**Tableau 4.3** – Estimation des hyper-paramètres. Le symbole \* désigne la vraie valeur. Les valeurs  $\widehat{\gamma}_{\mathbf{b}}$  et  $\widehat{\gamma}_{\mathbf{x}}$  sont les estimateurs de la moyenne *a posteriori*. Les distances  $\ell$  sont les distances calculées précédemment lorsque  $\mathbf{x}^*$  est connue.

Concernant l'estimation du paramètre  $\gamma_{\mathbf{x}}$  on constate que la valeur estimée est systématiquement inférieure à celle indiquée par les distances  $\ell_1$  et  $\ell_2$  qui exploite seulement l'erreur sur l'image. Dans le cas de la réalisation de la loi *a priori*, où la vraie valeur  $\gamma_{\mathbf{x}}^*$  est connue, la valeur estimée est plus proche de la vraie valeur que les valeurs fournies par les distances  $\ell$ . Cependant, en ce qui concerne la réalisation de la loi *a priori*, la valeur estimée dépend de la réalisation en étant parfois au-dessous, parfois au-dessus, de la vraie valeur  $4 \times 10^{11}$ .

#### Chaînes des hyper-paramètres

Pour étudier l'évolution de l'algorithme MCMC, la figure 4.16 présente les chaînes et histogrammes, après la période de chauffe, des paramètres  $\gamma_{\mathbf{b}}$  et  $\gamma_{\mathbf{x}}$ , pour les trois ciels étudiés. Concer-

nant la chaîne de  $\gamma_{\mathbf{b}}$  on observe une convergence rapide autour de  $10^6$  qui correspond à la vraie valeur. Dans les trois cas la chaîne se stabilise au-dessus ou au-dessous en fonction de la réalisation du bruit. L'exploration de l'espace est bonne et les échantillons semblent relativement décorrelés.

La chaîne  $\gamma_{\mathbf{x}}$  évolue également autour d'une valeur qui dépend du ciel considéré. La vitesse d'exploration est plus lente et les échantillons sont franchement plus corrélés. Globalement la chaîne a convergé et il ne semble pas qu'une zone à forte probabilité de l'espace n'ait pas été explorée.

En observant l'évolution de la chaîne et de la moyenne empirique courante, qui varie encore, il est possible que la chaîne n'a pas assez exploré l'espace. L'aspect des histogrammes, légèrement bi-modal pour les *cirrus* avec sources ponctuelles et la réalisation de la loi *a priori*, indique qu'il pourrait y avoir plus d'aller et retour de la chaîne autour de la valeur nominale.

L'évolution des chaînes MCMC depuis le premier échantillon est tracée figure 4.17. L'initialisation de l'algorithme est effectuée avec la co-addition pour le ciel. Donc le premier échantillon des paramètres  $\gamma_{\mathbf{b}}$  et  $\gamma_{\mathbf{x}}$  dépend de cette initialisation.

De nouveau on constate une convergence très rapide de la chaîne de  $\gamma_{\mathbf{b}}$  très proche de la vraie valeur en moins d'une dizaine d'échantillons. Ensuite le paramètre évolue autour de la moyenne *a posteriori*.

Pour le paramètre  $\gamma_{\mathbf{x}}$ , l'initialisation étant la co-addition (donc très doux) la chaîne commence à une valeur élevée. Elle décroît de manière presque monotone pour converger autour d'une valeur nominale au bout de 400 échantillons environ. L'algorithme a également été initialisé à d'autres valeurs pour le ciel et le même comportement a été observé (la chaîne pouvant commencer en dessous de la valeur nominale).

### 4.3.3 Conclusion

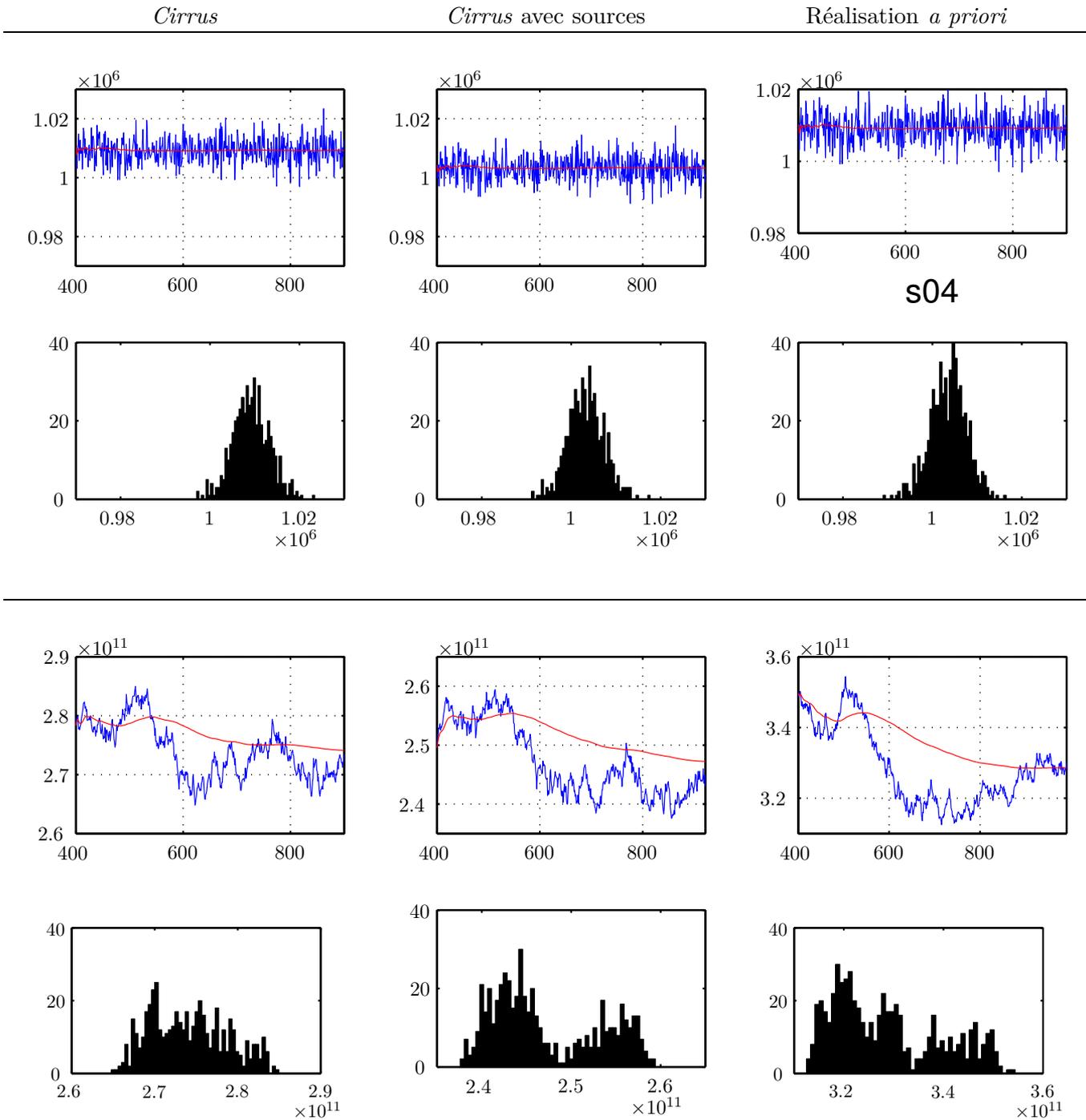
En conclusion l'algorithme non-supervisé, qui estime conjointement l'objet d'intérêt et les paramètres des lois fournit un résultat très proche de ce qui est obtenu lorsque l'on *connaît* la vérité terrain (par une estimation supervisée où le compromis est fixé de manière *ad-hoc*). La valeur estimée des paramètres, étudiée notamment avec la réalisation de la loi *a priori*, est bonne.

Dans le cas du *cirrus* il semble qu'il y ait une légère sous-régularisation. En déterminer l'origine demande d'autres travaux. Une piste pourrait être que la régularisation, en mesurant la régularité par la dérivée première d'une fonction décomposée sur des gaussiennes, n'est pas très adaptée au *cirrus*. En comparant le *cirrus* et la réalisation de la loi *a priori* de  $\phi$ , figure 4.1 page 110, il semble que l'*a priori* laisse plus de hautes fréquences apparaître. Il est également possible de laisser la chaîne MCMC explorer plus l'espace.

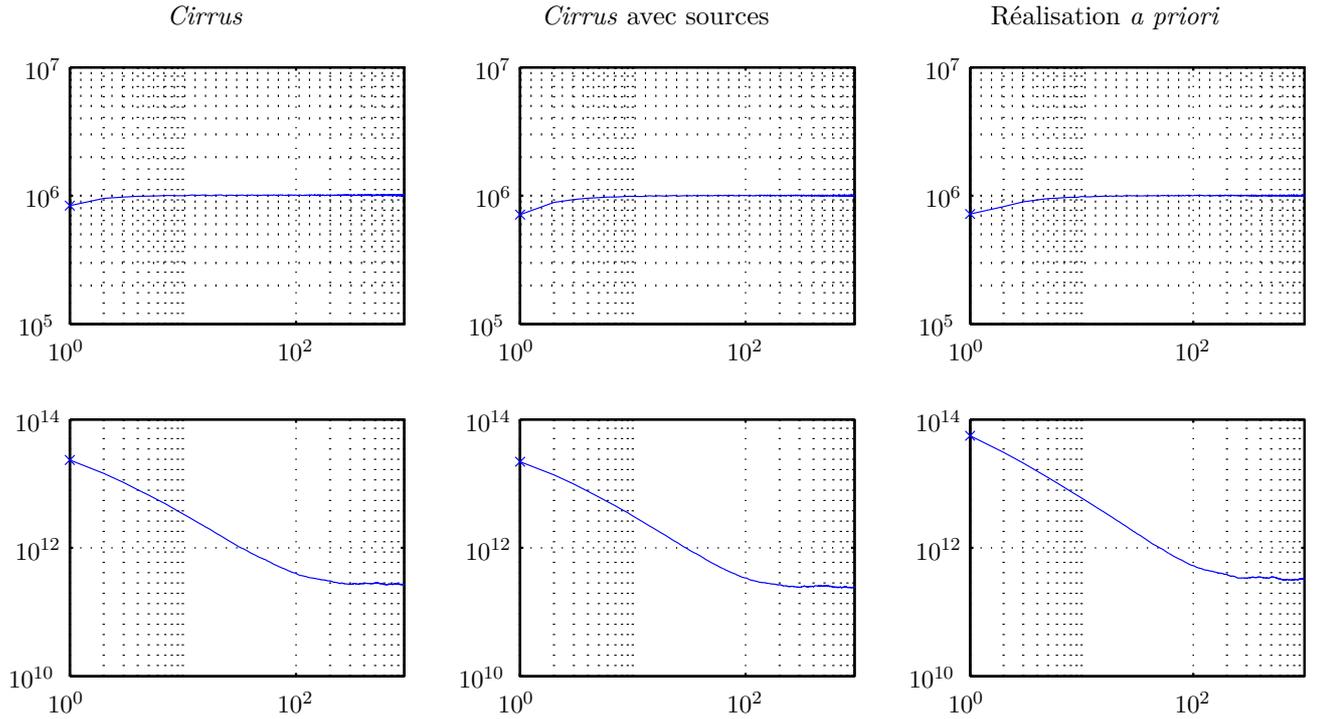
## 4.4 Conclusion

Les résultats présentés sont obtenus à l'aide de simulations. Ils montrent le potentiel de la méthode proposée pour la restauration de fréquences spatiales et de structures spatiales fines dégradées par le processus d'acquisition. De plus l'étude montre la faisabilité de méthodes déterminant automatiquement le compromis entre données et *a priori*.

La première partie, à hyper-paramètres fixés, présente les caractéristiques de l'estimateur des moindres carrés régularisés correspondant à la moyenne *a posteriori*. La solution dépend bien sûr des paramètres de compromis et pour un choix adapté de leur valeur, la solution contient des structures présentes dans l'objet vrai. Ces structures ne peuvent pas être reconstruites par des méthodes, comme la co-addition, ne prenant pas en compte les caractéristiques de l'instrument. Les possibilités



**Figure 4.16** – Chaînes et histogrammes, après l'échantillon 400, des hyper-paramètres  $\gamma_b$ , en haut, et  $\gamma_x$  en bas. La courbe rouge correspond à la moyenne courante. L'axe horizontal des chaînes est l'itération, celui des histogrammes la valeur du paramètre.



**Figure 4.17** – Chaîne complète des hyper-paramètres  $\gamma_{\mathbf{b}}$  en haut et  $\gamma_{\mathbf{x}}$  en bas, en échelle log-log pour observer la période de chauffe.

de sur-résolution, c'est-à-dire de restaurer des fréquences spatiales au-delà de la fréquence d'échantillonnage, dépendent du niveau de bruit. Dans les cas étudiés, on a pu montrer que les fréquences jusqu'à  $0,03 \text{ arcseconde}^{-1}$  peuvent être reconstruites alors que la fréquence d'échantillonnage est de  $0,022 \text{ arcseconde}^{-1}$  et que la fréquence de coupure du miroir est de  $0,044 \text{ arcseconde}^{-1}$ .

Des caractéristiques de l'estimateur ont été mises en évidence à travers la variation du niveau de bruit, du paramètre de régularisation et du ciel d'entrée. Le biais et la variance de l'estimateur ont également été calculés, montrant le comportement classique d'un estimateur bayésien. Le cadre bayésien fournit également une mesure de l'incertitude au travers de la matrice de covariance *a posteriori*. La diagonale de cette matrice permet de calculer une « carte d'incertitude » utilisée pour évaluer la confiance dans la solution proposée.

Dans la deuxième partie un estimateur non-supervisé a été mis en œuvre. L'estimateur est la moyenne *a posteriori* pour les paramètres inconnus comprenant  $\mathbf{x}$  mais également les paramètres  $\gamma_{\mathbf{x}}$  et  $\gamma_{\mathbf{b}}$  qui règlent le compromis entre la fidélité aux données et la fidélité à l'*a priori*. Les résultats indiquent que les valeurs estimées des paramètres de régularisation sont très proches des valeurs optimales obtenues lorsque l'on connaît la vérité terrain. L'image estimée est également très proche de celle obtenue de manière supervisée.

La méthode proposée permet également, au sein du même cadre et d'une façon globale l'estimation de paramètres instruments. Une étude expérimentale de ces aspects est développée, sous la forme d'un article, dans l'annexe F. Cette étude, sur un cas plus standard de déconvolution, montre le potentiel de la méthode pour l'estimation de paramètres instruments mal-connus conjointement avec les autres inconnues. On montre également que les paramètres instruments mal-connus peuvent être estimés conjointement avec les autres paramètres. De plus l'étude montre une corrélation intéressante entre les paramètres de régularisation et les paramètres instruments, ces deux familles expliquant le contenu spectral des données.

En conclusion nous avons montré l'impact de la prise en compte d'un modèle instrument et d'un modèle *a priori* pour le ciel pour la restauration de hautes fréquences spatiales dans les cartes estimées. De plus les résultats montrent également la fiabilité de la méthode pour le réglage automatique, ou l'estimation, des paramètres de compromis entre l'information provenant des données et l'information *a priori* ainsi que la robustesse du modèle *a priori*.

# Chapitre 5

## Conclusions et perspectives

### 5.1 Bilan des travaux

#### 5.1.1 Contributions majeures

Les travaux présentés dans cette thèse concernent l'estimation d'un ciel astrophysique à partir des données issues de l'instrument SPIRE de l'observatoire spatial Herschel. L'estimation est vue sous un angle « problème inverse » et repose essentiellement sur deux points.

**Modèle direct.** L'inversion repose sur un modèle instrument, décrit chapitre 2, établissant le lien entre l'image recherchée et les données produites. Le modèle est construit à partir d'une analyse de la physique de l'instrument. Le modèle prend en compte la convolution par le miroir primaire du télescope, les cornets qui jouent le rôle d'intégrateurs spatiaux, les filtres en longueur d'onde pour sélectionner la bande spectrale et la convolution temporelle introduite par les bolomètres qui sont les éléments sensibles de l'instrument. Le modèle considère le ciel comme un objet défini continûment pour avoir une maîtrise sur les trois dimensions et être au plus proche de la description physique des phénomènes. Les différents protocoles d'observation sont pris en compte.

Le modèle repose également sur une description originale du ciel. En photométrie, la bande spectrale sélectionnée étant considérée « étroite », les méthodes de reconstruction habituelle considèrent généralement que le ciel est monochromatique. Or celui-ci, ainsi que les réponses instrumentales, dépendent de la longueur d'onde. Le modèle proposé utilise un développement de Taylor dans la dimension spectrale  $\lambda$  pour prendre en compte cette information tout en ayant un modèle d'acquisition quasiment explicite.

Le modèle obtenu est linéaire mais non-invariant à cause de l'échantillonnage et du balayage. Par rapport au modèle physique existant implémenté dans le simulateur de Bruce Sibthorpe [SCG09], le modèle proposé se distingue en plusieurs points. Tout d'abord il prend en compte proprement la longueur d'onde. Ensuite le modèle pour les bolomètres est linéaire permettant un calcul plus rapide. Enfin il est conçu pour interpréter les données dans une méthode d'inversion.

Par rapport aux méthodes existantes d'inversion le modèle proposé est plus proche de la réalité. Les méthodes telles que SANEPIC ou MADMAP ne prennent pas en compte la physique de l'instrument mais seulement l'information de pointage. Notre modèle d'acquisition permet donc une meilleure interprétation de l'information présente dans les données.

**Inversion.** L'inversion est basée sur une loi *a posteriori* établie dans un formalisme bayésien usuel chapitre 3. L'utilisation de modèles probabilistes permet de prendre en compte d'autres in-

formations que celles provenant seulement des données. Il est possible ainsi d'inclure une information sur le contenu fréquentiel de la composante basse fréquence provenant de la dérive en température. Le formalisme permet également d'exploiter une mesure plus directe de cette composante fournie par les bolomètres « aveugles ».

Les modèles probabilistes permettent de modifier l'espace des solutions pour le ciel au travers d'une pénalisation de la différence première sur les dimensions spatiales. Il est ainsi possible d'exclure les solutions peu réalistes engendrées par un problème inverse mal conditionné. Les résultats montrent que l'utilisation conjointe d'un modèle instrument et d'un modèle probabiliste pour le ciel permet de restaurer des structures spatiales fines, ce qui est impossible avec un modèle exploitant seulement les informations de pointage. Enfin le formalisme offre un cadre théorique pour l'estimation conjointe de paramètres instruments avec le ciel et la dérive thermique.

L'algorithme de calcul de l'estimateur repose sur un échantillonneur de Gibbs fournissant des échantillons de la loi *a posteriori* globale. La moyenne empirique des échantillons fournit une approximation de la moyenne *a posteriori*. Les différentes étapes d'échantillonnage des lois conditionnelles *a posteriori* se répartissent en trois parties. Les paramètres des différentes puissances ont des lois conditionnelles *a posteriori* gamma dont l'échantillonnage est aisé. D'autres paramètres comme les paramètres de densité spectrale *a priori* ou les paramètres instruments possèdent des lois conditionnelles *a posteriori* non usuelles nécessitant une étape de Metropolis-Hastings.

Enfin l'échantillonnage du ciel, bien qu'il s'agisse d'une simple loi gaussienne multi-variée, est complexe. En effet la matrice de covariance fait intervenir le modèle instrument qui n'est pas invariant, empêchant toute inversion ou factorisation de la covariance. La solution proposée consiste en un algorithme d'optimisation d'un critère perturbé, l'optimum étant un échantillon de la loi cible.

Les résultats expérimentaux de ces deux contributions, présentés dans le chapitre 4, sont très encourageants. La prise en compte d'un modèle instrument et de modèles probabilistes, le tout dans un cadre bayésien unique et cohérent, permet d'estimer un ciel plus proche de la réalité qu'une méthode qui ne prend pas en compte ces informations. De plus nous avons démontré, sur un problème réel et de grande dimension, la possibilité de mettre en oeuvre une méthode « sans paramètres » où le compromis entre fidélité aux données et fidélité à l'*a priori* est réglé automatiquement et non de manière *ad-hoc*. Enfin nous avons également montré que l'estimation conjointe de paramètres instruments mal connus peut être menée, toujours dans le même cadre. L'enjeu est potentiellement important pour des instruments spatiaux dont les caractéristiques peuvent évoluer au cours de leur vie.

### 5.1.2 Autres apports

Nos travaux ont également abouti à d'autres conclusions intéressantes.

- L'analyse de l'instrument a permis de dégager des caractéristiques qui peuvent se retrouver dans d'autres instruments. Le modèle pour le ciel, le miroir et les bolomètres peuvent s'appliquer à d'autres instruments.
- L'étude de l'instrument et des protocoles d'acquisition a permis de dégager des résultats originaux sur le potentiel des méthodes de sur-résolution (ou super-résolution). Nous avons montré comment un protocole d'acquisition permet de produire des données moins repliées dans l'espace de Fourier. L'étude a également permis de poser des conditions d'échantillonnage permettant de déterminer des paramètres optimaux.
- L'analyse du modèle d'acquisition conduit à réduire la complexité et les temps de calcul.

Avec certaines approximations le modèle rentre dans la classe de modèle d'acquisition discret-discret classique en sur-résolution consistant en une convolution suivie d'une décimation. Il est possible dans ce cas d'exploiter la littérature sur le sujet.

- Concernant l'inférence bayésienne, il existe peu de travaux sur des problèmes de sur-résolution avec estimation des hyper-paramètres, encore moins avec une application sur un instrument réel. Une solution permettant une inversion rapide de la covariance par transformée de Fourier permet de fournir une première estimation du ciel mais permet également un algorithme avec estimation des hyper-paramètres en un temps très raisonnable (quelques minutes).

Le temps de calcul nécessaire à l'estimation non-supervisée du ciel observé par SPIRE peut s'avérer décourageant. Cependant, la structure du modèle permet de faire facilement des calculs indépendants en parallèle (par bandes spectrales, par direction de balayage, par bolomètres, *etc.*). L'investissement dans une machine de calcul et une mise en œuvre dans un langage performant permettrait également de réduire le temps de calcul total. Notamment, la « démocratisation » des unités de calcul massivement parallèles comme les processeurs graphiques (GPU) permettra sûrement à certaines méthodes, nécessitant beaucoup de calculs numériques, d'être appliquées sur des problèmes de dimensions importantes rencontrés en pratique. De plus, il convient de rappeler que SPIRE est un instrument unique et qu'il s'inscrit dans un projet de 20 ans ayant coûté plus d'un milliard d'Euros. Quelle que soit la méthode de traitement employée, il est raisonnable qu'elle soit à la hauteur des enjeux et des efforts déjà fournis.

## 5.2 Perspectives

### 5.2.1 Perspectives algorithmiques et calculatoires

Sur les aspects algorithmiques, plusieurs évolutions sont possibles pour améliorer l'exploration de l'espace des configurations et réduire le temps de calcul.

Tout d'abord un grand nombre de calculs peuvent être parallélisés. Les calculs par bande spectrale, par bolomètre, par balayage sont indépendants. De plus le modèle est linéaire donc le traitement par lot de données peut également être fait en parallèle. Ces caractéristiques peuvent réduire le temps de calcul particulièrement sur des problèmes de tailles importantes.

Concernant l'échantillonneur de Gibbs, celui-ci souffre des inconvénients de la corrélation entre échantillons. Cela induit un temps de convergence vers la loi cible de 400 échantillons dans nos expériences donc très long en temps processeur. Un mélange de noyau de transition [Rob96] en utilisant avec une certaine probabilité une étape de Metropolis-Hastings au lieu de l'étape de Gibbs est possible. On bénéficie ainsi de la plus grande variabilité du Metropolis-Hastings tout en ayant toujours une chaîne ergodique convergeant vers la loi cible. On peut également mentionner la sur-relaxation [Mac03, chap. 30]

### 5.2.2 Conditions d'échantillonnages pour la sur-résolution

L'analyse du modèle direct a permis de mettre en évidence comment la redondance introduite par l'acquisition permet d'obtenir des données avec moins de repliement dans l'espace de Fourier. Cette analyse permet d'avoir les équations explicites du spectre et donc d'analyser ses caractéristiques en fonction des paramètres du modèle comme la réponse instrumentale, la vitesse de balayage ou la cadence d'échantillonnage. Il est donc possible de poser des conditions sur ces paramètres garantissant que certaines fréquences ne seront pas (ou peu) dégradées par le repliement.

Pour le moment les conditions sont établies seulement dans certains cas dégénérés avec par

exemple un temps d'observation infini et un champ de vue infini. Il serait intéressant de l'étendre ces conditions à des cas plus réalistes. Une partie de ce travail a été engagé et a fait l'objet d'une publication [ORG09].

### 5.2.3 Bruit impulsionnel

Les observatoires spatiaux souffrent tous des particules à haute énergie qui frappent le détecteur et induisent des comportements parfois difficilement prévisibles. Une recherche approfondie doit être menée mais il semble qu'un premier modèle impulsionnel soit possible. La particule serait déterminée par son amplitude et l'instant d'arrivée où elle frappe un bolomètre. Celui-ci réagit alors à un événement impulsionnel. Cependant le comportement de l'instrument face à ces événements peut être plus complexe avec des effets électriques sur l'ensemble du capteur ou encore une décroissance de la tension plutôt qu'une montée brusque de la température due au dépôt d'énergie.

### 5.2.4 Estimation de paramètres instruments

Certaines caractéristiques de l'instrument et de l'observatoire peuvent évoluer au cours d'une observation ou pendant la durée de vie de la mission. Par exemple la température détermine la valeur de certains paramètres de la réponse des bolomètres.

Pour le moment l'étude de l'estimation des paramètres instruments est menée sur les paramètres d'un modèle convolutif. Il serait intéressant d'étendre au cas du modèle non-invariant de sur-résolution majoritairement étudié dans cette thèse. L'extension à d'autres applications serait également intéressante.

D'un point de vue algorithmique sur les étapes de Metropolis-Hastings, des développements peuvent être faits notamment sur la loi de proposition. On pourrait par exemple évaluer une proposition type marche aléatoire. Il serait également intéressant d'exploiter la structure, ou la forme de la réponse, pour proposer un meilleur échantillon. Des algorithmiques de type Monte-Carlo Hamiltonien ou Langevin, permettant une meilleure exploration de l'espace des configurations, peuvent également constituer des pistes de recherches.

### 5.2.5 Application à d'autres instruments

Dans une perspective à plus long terme il serait possible d'envisager d'appliquer une partie des travaux développés ici à d'autres instruments. D'un point de vue modèle direct, PACS présent dans l'observatoire Herschel est sensiblement identique à SPIRE. De même l'observatoire Planck utilise également des bolomètres.

D'un point de vue méthode d'estimation, les problèmes faisant intervenir l'estimation d'images avec un modèle instrument linéaire proche d'une convolution peuvent exploiter les travaux concernant l'estimation des hyper-paramètres et des paramètres instruments.

### 5.2.6 Sur-résolution myope et non-supervisée

Il serait également intéressant d'approfondir les liens avec les méthodes de super-résolution académique en traitement d'images. Pour le moment nous avons proposé dans ce document une méthode de sur-résolution non-supervisée, sujet de recherche peu étudié pour le moment. L'estimation de paramètres instruments conjointement aux autres paramètres est, à notre connaissance, inexplorée pour le moment et à notre portée.

Enfin il existe des méthodes de sur-résolution avec une pénalisation robuste (norme  $L_1$ ) pour le modèle image, en particulier l’approche proposée par Farsiu *et al.* [FREM04]. Il serait intéressant d’étendre cette approche avec une estimation non-supervisée. L’approche bayésienne apporte certaines réponses [Gio08, CI04, CGLS08].

### 5.2.7 Utilisation conjointe des trois bandes

Les trois bandes spectrales du photomètre de SPIRE sont contiguës dans la dimension spectrale. De plus, on sait que le profil spectral du rayonnement suit la loi des corps noirs sur toute la bande observée par SPIRE. Il y a donc un lien entre les profils spectraux des différentes bandes qu’il serait intéressant d’exploiter. De plus, la position légèrement décalée des bolomètres de chacune des bandes spectrales pourrait probablement améliorer la résolution spatiale de la reconstruction. Il existe au moins trois stratégies possibles.

1. Tout d’abord un lien entre les longueurs d’onde peut être établi en exploitant le modèle proposé dans ce manuscrit. Le développement de Taylor peut se faire sur tout un intervalle recouvrant les trois bandes observées. On peut également envisager d’introduire des contraintes sur la jonction entre les trois bandes.
2. Il est également possible d’introduire une information entre les bandes spectrales au moyen de modèles probabilistes ou de termes de régularisation. En effet, la loi des corps noirs prévoit que sur la bande spectrale totale observée par SPIRE, le profil en échelle log–log est une droite. On peut donc introduire une pénalisation sur l’ordre 1 du développement en  $\lambda$  pour prendre en compte cette information.
3. Une troisième voie consisterait à exploiter directement la loi des corps noirs dans le modèle pour le ciel et ainsi estimer *directement* les paramètres physiques de cette loi. Bien que le potentiel de cette approche soit important, elle conduit à un modèle d’acquisition non linéaire vis-à-vis de ces paramètres physiques ce qui rend plus complexe et difficile l’exploration de la loi et le calcul d’un estimateur. En contre-partie le nombre de paramètres peut être réduit.

Il s’agit, il me semble, d’une des perspectives les plus intéressantes et les plus prometteuses pour la science astrophysique. L’utilisation conjointe de données provenant d’autres instruments et observant d’autres bandes spectrales serait un sujet de recherche encore inexploré.

### 5.2.8 Extension du modèle pour le ciel

Le ciel observé par SPIRE n’est pas composé uniquement d’objets possédant une étendue spatiale. Il y a également la présence d’objets ponctuels comme des sites de formation d’étoiles, des ruptures ou des contours francs. Des travaux pourraient être menés pour la prise en compte de ces objets et adapter le modèle pour le ciel. Plusieurs pistes sont possibles. Tout d’abord la littérature contient de nombreux travaux sur des modèles (ou des pénalisations) permettant de mieux préserver les ruptures franches. On trouve par exemple les potentiels convexes ou non convexes des champs de Markov, mais il est également possible d’envisager des starlettes. Cela pose cependant des problèmes pour l’estimation des paramètres des lois puisque dans ce cas le facteur de normalisation n’est pas connu.

Une autre piste pourrait être une approche bi-modèle. Le ciel serait alors formé de deux composantes. Une composante douce pour laquelle un *a priori* quadratique corrélé serait adapté et une composante ponctuelle. Dans ce cas des *a priori* séparables et robustes ont été utilisés [GC05, RZ09]. D’autres approches basées sur de la déconvolution impulsionnelle seraient également possibles [Maz05].



# Annexes



# Annexe A

## Physique des bolomètres

Il existe toute une littérature sur les bolomètres et leur modélisation. Le modèle présenté ici s'appuie en grande partie sur les travaux de Bruce Sibthorpe [Sib07] qui a travaillé sur le simulateur officiel<sup>1</sup> de SPIRE. Ses travaux reposent également largement sur le modèle thermique de bolomètre de Sudiwala *et al.* [SGW02] présenté ici. Ce modèle est lui-même une amélioration des modèles thermiques classiquement utilisés pour modéliser les bolomètres [Ban05, Cam02, Tor92]. Cette partie présente tout d'abord les principes de base des bolomètres. On établit ensuite la réponse du bolomètre à l'aide d'un bilan thermodynamique que nous linéariserons.

### A.1 Principes

Les bolomètres sont constitués de deux parties au rôle distinct. Premièrement l'absorbant modifie sa température en fonction de la puissance incidente. La dissipation de cette puissance dans l'environnement, notée  $P_f$  pour puissance de fuite, se fait au travers d'un lien thermique caractérisé par une conductivité  $\kappa$ . Deuxièmement la lecture de la température du bolomètre se fait à l'aide d'une thermistance. La mesure de la résistance, illustrée par la figure A.1, s'effectue à l'aide d'une polarisation par un courant  $I_p$ . Ce courant est obtenu avec un générateur de tension  $V_p$  en série et une résistance de charge  $R_c \gg R(t)$ . Un étage d'adaptation d'impédance constitué de transistors est placé aux bornes de la résistance pour la mesure.

La polarisation induit nécessairement un dégagement de puissance par effet Joule. Cette puissance  $P_j$  est absorbée par le bolomètre ce qui augmente sa température.

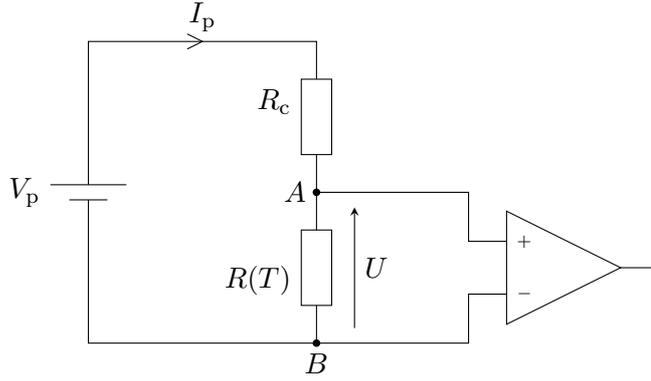
### A.2 Bilan de puissance

Pour connaître la réponse du système à une puissance incidente  $P(t)$  nous allons déterminer l'évolution de la température au cours du temps à l'aide d'un bilan de puissance. Les puissances mises en jeux sont :

- la puissance de rayonnement  $P(t)$  en entrée,
- la puissance dissipée par effet Joule  $P_j(t)$ ,
- la puissance de fuite  $P_f(t)$  diffusée dans l'environnement à travers le lien thermique en sortie.

---

<sup>1</sup>[http://www.astro.cardiff.ac.uk/groups/instrumentation/projects/spire/Public\\_SPIRESimulator.htm](http://www.astro.cardiff.ac.uk/groups/instrumentation/projects/spire/Public_SPIRESimulator.htm)



**Figure A.1** – Schéma de lecture des bolomètres. Afin de mesurer la valeur de la résistance  $R(T)$  on injecte un courant de polarisation  $I_p$  pour obtenir une différence de tensions. Ce courant est généré à l'aide d'un générateur en tension et d'une résistance de charge  $R_c \gg R(T)$ .

Enfin le comportement dans le temps du bolomètre est caractérisée par une capacité calorifique  $C_0 T^b$ . Ainsi, le bilan de puissance s'écrit sous la forme d'une équation différentielle

$$C_0 T^b \frac{dT(t)}{dt} = P(t) + P_J(t) - P_f(t). \quad (\text{A.1})$$

Dans la perspective du modèle instrument, la puissance  $P$  est donnée puisqu'il s'agit de la grandeur d'entrée. Il reste à expliciter les deux puissances  $P_f$  et  $P_J$ .

### A.2.1 Puissance par effet Joule $P_J$

La puissance dissipée par effet Joule a pour origine la polarisation de la résistance. Avec l'approximation  $R(T) \ll R_c$ , elle s'écrit simplement

$$P_J(t) = R(T(t)) I_p^2(t). \quad (\text{A.2})$$

On considère le bolomètre dit « idéal » avec pour expression de la résistance en fonction de la température

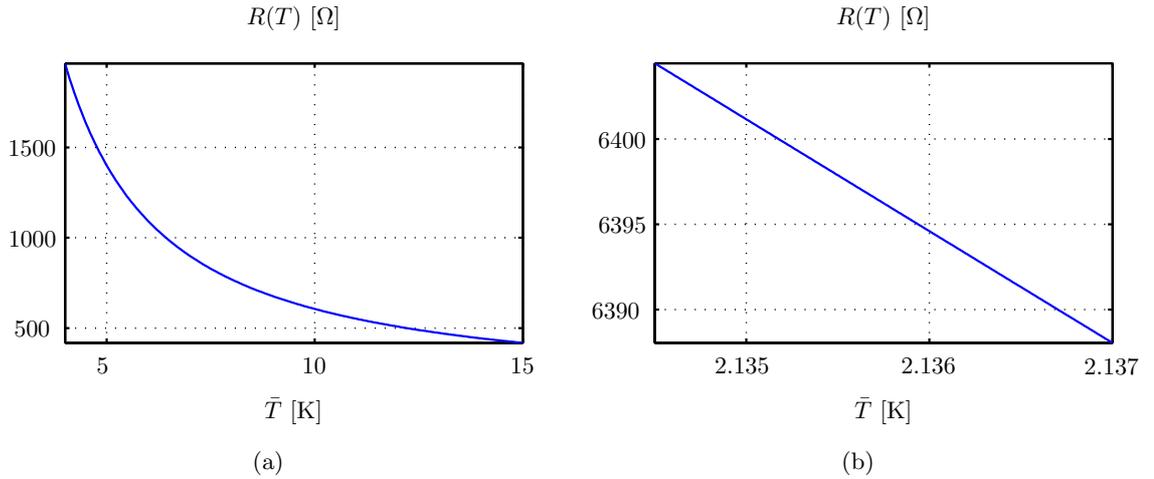
$$R(T) = R_g \exp \left[ \left( \frac{T_g}{T} \right)^n \right]. \quad (\text{A.3})$$

Les grandeurs  $R_g$ ,  $T_g$  et  $n$  sont des paramètres caractéristiques du bolomètre et de sa conception. Elles sont mesurées expérimentalement et leur valeur sont reportées dans le tableau D page 167. L'évolution de la résistance en fonction de la température est tracée sur la figure A.2. On remarque une décroissance de la résistance en fonction de la température. La figure A.2(b) correspond aux valeurs prises par la résistance dans une plage de variations nominales de SPIRE.

### A.2.2 Puissance de fuite $P_f$

Dans le modèle de Sudiwala *et al.* [SGW02] la puissance de fuite s'écrit

$$P_f = \frac{\int_{T_0}^T \kappa(\tau) d\tau}{\int_0^L 1/S(x) dx}$$



**Figure A.2** – Évolution de la résistance en fonction de la température. La figure A.2(a) illustre le régime non linéaire de la thermistance. La courbe A.2(b) correspond aux valeurs prises par la thermistance pour de faibles températures typiques obtenues dans SPIRE. Cette courbe correspond au fonctionnement de la résistance décrit par l'équation (A.3). Les valeurs numériques viennent du tableau D page 167.

où  $T_0$  est la température du cryostat,  $T$  la température du bolomètre,  $S(x)$  la section efficace entre le bolomètre et le lien thermique et  $L$  la taille du lien comme illustré figure 2.7. Ce modèle permet de prendre en compte la variation de température le long du lien thermique. La conductivité thermique  $\kappa(T)$  s'exprime comme

$$\kappa(T) = K_0 \left( \frac{T}{T_0} \right)^\beta$$

avec  $K_0 = \kappa(T_0)$  la conductivité thermique. Si on fait l'hypothèse que la section efficace est uniforme  $S(x) = S$  la puissance de fuite devient

$$\begin{aligned} P_f(T) &= \frac{K_0 S}{L T_0^\beta} \int_{T_0}^T \tau^\beta d\tau \\ &= \frac{G_0}{T_0^\beta (\beta + 1)} \left( T^{\beta+1} - T_0^{\beta+1} \right) \end{aligned} \quad (\text{A.4})$$

avec  $G_0 = K_0 S / L$ . C'est une fonction en puissance de la température  $T$ . A moins d'une variation de  $T_0$ , la puissance est nécessairement positive (avec  $T > T_0$ ). En effet le bolomètre absorbe de l'énergie et la dissipation s'effectue nécessairement vers le milieu le plus froid c'est-à-dire le cryostat.

### A.2.3 Équation du bolomètre

En injectant les expressions (A.2) et (A.4) dans le bilan de puissance (A.1), on obtient « l'équation du bolomètre »

$$C_0 T^b \frac{dT}{dt} - \frac{R(T) V_p^2}{R_c^2} + \frac{G_0}{T_0^\beta (\beta + 1)} \left( T^{\beta+1} - T_0^{\beta+1} \right) = P(t). \quad (\text{A.5})$$

Tout le travail consiste donc à trouver  $T(t)$  connaissant  $P(t)$  c'est à dire la réponse  $h_T$  telle que

$$P(t) = h_T(T(t))$$

À ma connaissance, l'état de l'art ne contient pas de résolution explicite de cette équation différentielle. Cependant, en linéarisant cette équation différentielle, on peut obtenir une réponse explicite qui modélise au premier ordre le fonctionnement du bolomètre. C'est l'objet de la partie suivante.

### A.3 Linéarisation

Pour simplifier l'équation (A.5) nous faisons tout d'abord deux approximations :

- la résistance de charge  $R_c$  est grande devant la résistance du bolomètre  $R_c \gg R(T)$ . Le courant de polarisation  $I_p$  est par conséquent constant dans le temps puisqu'il ne dépend plus de la température

$$\frac{V_p}{R_c + R(T)} = \frac{V_p}{R_c} = I_p = \text{cst.}$$

- On néglige l'évolution de la capacité thermique du bolomètre en fonction de la température

$$C_0 T^b = C \text{ en } [\text{JK}^{-1}].$$

L'équation différentielle (A.5) se réécrit

$$C \frac{dT}{dt} - I_p^2 R_g \exp \left[ \left( \frac{T_g}{T} \right)^n \right] + \frac{G_0}{T_0^\beta (\beta + 1)} (T^{\beta+1} - T_0^{\beta+1}) = P$$

Par souci de clarté on réécrit l'équation précédente comme

$$C \frac{dT}{dt} + \phi(T) = P \tag{A.6}$$

avec

$$\phi(T) = \frac{G_0}{T_0^\beta (\beta + 1)} (T^{\beta+1} - T_0^{\beta+1}) - I_p^2 R_g \exp \left[ \left( \frac{T_g}{T} \right)^n \right], \tag{A.7}$$

où  $T$  et  $P$  sont deux fonctions  $L_2$  du temps  $t$ , définies de  $\mathbb{R}$  dans  $\mathbb{R}$ . Dans ce cas  $\phi$  est une fonctionnelle  $\phi : L_2 \rightarrow L_2$  de la température  $T$ . Il n'existe pas pour le moment dans la littérature de solution explicite à cette équation différentielle. Avec certaines approximations il est cependant possible de linéariser cette équation à l'aide d'un développement au premier ordre et d'obtenir ainsi une réponse  $h_T$  explicite.

- Tout d'abord on considère que les signaux d'entrée et de sortie se décomposent en une partie constante dans le temps et variable dans le temps soit

$$\begin{aligned} P(t) &= \bar{P} + \tilde{P}(t) \\ T(t) &= \bar{T} + \tilde{T}(t) \end{aligned}$$

et on suppose que les parties variables sont intégrables car d'énergie finie. L'équation différentielle se réécrit ainsi

$$C \frac{d\tilde{T}}{dt} + \phi(\bar{T} + \tilde{T}) = \bar{P} + \tilde{P}.$$

- Ensuite, les variations de la température autour d'une température nominale sont considérées petites. Cette approximation permet d'approcher la fonction  $\phi$  par son développement au premier ordre

$$\phi(\bar{T} + \tilde{T}) \approx \phi(\bar{T}) + \left. \frac{d\phi}{dT} \right|_{\bar{T}} \tilde{T}.$$

L'équation différentielle se réécrit

$$C \frac{d\tilde{T}}{dt} + \phi(\bar{T}) + \left. \frac{d\phi}{dT} \right|_{\bar{T}} \tilde{T} = \bar{P} + \tilde{P}.$$

Il est possible maintenant de décrire la réponse du bolomètre comme la somme de deux réponses : celle à un signal constant  $\bar{P}$  et celle à un signal variable  $\tilde{P}$ . Nous allons établir chacune d'elle.

## A.4 Réponse à un signal constant

Comme les fonctions  $\tilde{P}$  et  $\tilde{T}$  sont intégrables alors

$$\lim_{\substack{t_0 \rightarrow -\infty \\ t_1 \rightarrow +\infty}} \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \tilde{P}(t) dt = \lim_{\substack{t_0 \rightarrow -\infty \\ t_1 \rightarrow +\infty}} \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \tilde{T}(t) dt = 0$$

et on a

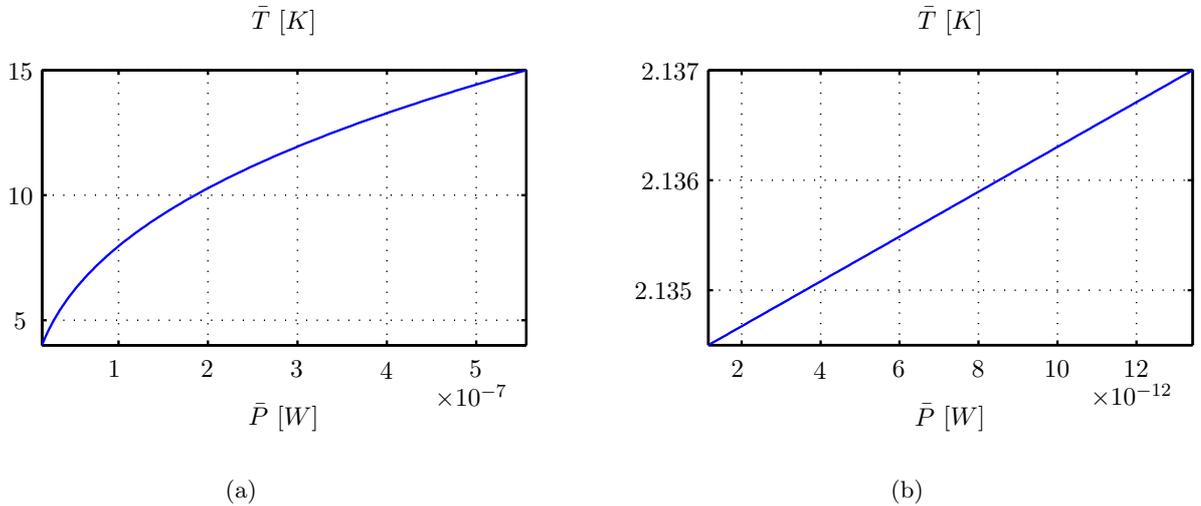
$$\phi(\bar{T}) = \bar{P}. \quad (\text{A.8})$$

La réponse à un signal constant du bolomètre, appelée point de fonctionnement et utilisée pour définir l'état à l'équilibre du bolomètre, est

$$\bar{T} = \phi^{-1}(\bar{P}). \quad (\text{A.9})$$

Cette fonction  $\phi^{-1}$ , sous réserve qu'elle existe, n'est pas connue. On détermine généralement sa valeur prise en  $\bar{P}$  par un calcul numérique approché (recherche exhaustive, optimisation de critères *etc.*) [SG06].

À l'aide des paramètres reportés dans le tableau D page 167 on peut tracer l'évolution de la température en fonction de la puissance incidente. On obtient les courbes tracées sur la figure A.3. La figure A.3(a) illustre la réponse du bolomètre à un signal constant, c'est-à-dire le point de



**Figure A.3** – Régime permanent du bolomètre. Sur la figure A.3(a) on observe un fonctionnement non linéaire de la température  $\bar{T}$  en fonction de la puissance incidente  $\bar{P}$ . Par contre sur la figure A.3(b) on observe un régime presque linéaire pour une puissance incidente très faible de l'ordre de de la dizaine de pico Watts.

fonctionnement, sur un grande plage de puissance incidente. La réponse apparaît comme étant non-linéaire. La figure A.3(b) illustre la même réponse mais pour une puissance incidente très faible de l'ordre du rayonnement de fond  $\approx 1,5$  pW. On observe alors un régime de la température qui est une fonction quasiment linéairement de la puissance incidente.

## A.5 Réponse à un signal variable

D'après (A.8) on sait

$$C \frac{d\tilde{T}}{dt} + \phi' \tilde{T} = \tilde{P} \quad (\text{A.10})$$

avec

$$\phi' = \left. \frac{d\phi}{dT} \right|_{\bar{T}}$$

une constante qui détermine en partie la vitesse de réaction, au premier ordre, du bolomètre. Elle s'exprime à l'aide des paramètres du modèle physique et s'écrit en dérivant l'équation (A.7)

$$\begin{aligned} \phi' &= \frac{G_0}{T_0^\beta} \bar{T}^\beta - \frac{nT_g^n}{\bar{T}^{n+1}} R_g \exp \left[ \left( \frac{T_g}{\bar{T}} \right)^n \right] I_p^2 \\ &= G_d - \nu \bar{R} I_p^2 \\ &= K \end{aligned}$$

avec  $G_d = G_0 \bar{T}^\beta / T_0^\beta$  et  $\nu = nT_g^n / \bar{T}^{n+1}$ .

L'équation différentielle (A.10) apparaît alors simplement comme une équation du premier ordre à coefficients constants et avec second membre

$$\frac{d\tilde{T}}{dt} + \frac{K}{C} \tilde{T} = \frac{\tilde{P}}{C}.$$

La solution est donc la convolution entre  $\tilde{P}$  et la réponse impulsionnelle  $\tilde{h}_T(t)$

$$\tilde{T}(t) = \int_{t'} \tilde{P}(t') \tilde{h}_T(t - t') dt' \quad (\text{A.11})$$

où

$$\tilde{h}_T(t) = S \exp \left[ -\frac{t}{\tau_e} \right] \quad (\text{A.12})$$

avec le gain  $S = 1/C$  et la constante de temps  $\tau_e = C / (G_d - \alpha \bar{R} I_p^2)$ . Cette réponse est tracée figure A.4 et correspond à un filtre passe-bas du premier ordre.

En conclusion, la réponse temporelle du bolomètre en température a été modélisée comme la somme de deux réponses : une réponse dite constante  $\phi^{-1}(P)$ , qui réagit à la partie constante du signal incident  $P$  avec

$$\phi(T) = I_p^2 R_g \exp \left[ \left( \frac{T_g}{T} \right)^n \right] + \frac{G_0}{T_0^\beta (\beta + 1)} (T^{\beta+1} - T_0^{\beta+1})$$

et une réponse dite variable  $\tilde{h}_T(t)$

$$\tilde{h}_T(t) = S \exp \left[ -\frac{t}{\tau_e} \right]$$

qui est une réponse impulsionnelle et qui ne réagit qu'aux variations du signal. Nous avons donc comme évolution de la température au cours du temps à la sortie du bolomètre

$$T(t) = \phi^{-1}(\bar{P}) + \int_{t'} \tilde{P}(t') \tilde{h}_T(t - t') dt'. \quad (\text{A.13})$$

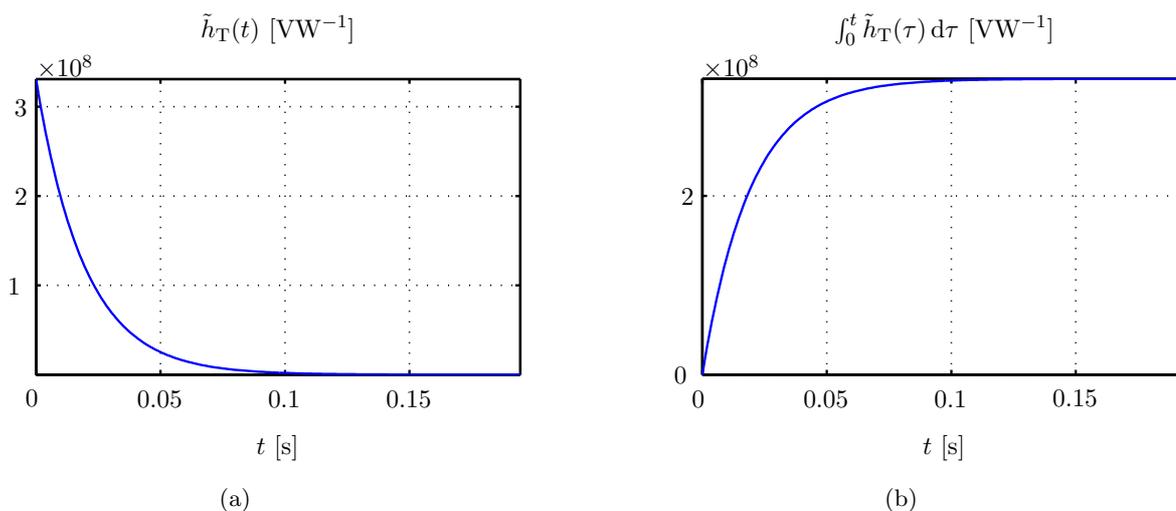


Figure A.4 – Réponse impulsionnelle A.4(a) et indicielle A.4(b) du bolomètre.

## A.6 Réponse en tension

Maintenant que nous connaissons la température du système par l'équation (A.13), on peut écrire la tension  $U(T) = I_p R(T)$  à l'aide de (A.3) page 144. L'évolution de la tension en fonction de la puissance incidente est

$$\begin{aligned} U(P) &= I_p R(\bar{T} + \tilde{T}) \\ &= I_p R(\phi^{-1}(\bar{P}) + \tilde{P} \star \tilde{h}_T) \end{aligned}$$

qui est non-linéaire puisque l'expression de la résistance fait intervenir une exponentielle. Cependant, l'évolution de la résistance a déjà été linéarisée par les calculs précédents pour déterminer la puissance dissipée par effet Joule en considérant de faibles variations de la température. La réponse de la résistance en fonction de  $T$  se modélise comme

$$\begin{aligned} R(T) &= R(\bar{T}) + \left. \frac{dR}{dT} \right|_{\bar{T}} \tilde{T} \\ &= \bar{R} + \nu \bar{R} \tilde{T} \end{aligned}$$

avec  $\nu = nT_g^n / \bar{T}^{n+1}$ . Finalement la réponse en tension s'écrit

$$\begin{aligned} U(P) &= \bar{R} I_p + \nu \bar{R} I_p \tilde{T} \\ &= R(\bar{T}) I_p + \nu R(\bar{T}) I_p \tilde{T} \end{aligned}$$

avec  $\bar{T} = \phi^{-1}(\bar{P})$  et  $\tilde{T} = \tilde{P} \star \tilde{h}_T$ .

En conclusion la réponse du système de bolomètre a été modélisé en linéarisant l'équation du bilan thermique (A.1) autour d'un point de fonctionnement défini par la partie constante de la puissance incidente. La valeur de la tension de sortie est fonction du flux lumineux provenant des cornets. La réponse du système obtenue est la somme de deux réponses :

- une réponse à un signal constant et qui délivre un signal constant
- une réponse à un signal variable, la sortie s'obtenant par convolution.

On obtient à la sortie du système pour un bolomètre en particulier

$$\begin{aligned} U(i,j,t) &= \bar{R}I_p + \alpha\bar{R}I_p\tilde{T}(i,j,t) \\ &= \bar{R}I_p + \int_t \tilde{P}(i,j,t)h_b(t' - t)dt. \end{aligned}$$

On suppose que les parties constantes et variables de  $P$  sont obtenues à partir des parties constantes  $\bar{\phi}_c$  et variables  $\tilde{\phi}_c$  du signal sortant des cornets. Ainsi

$$U(i,j,t) = \bar{R}I_p + \int_t \int_{(\alpha,\beta) \in \mathcal{S}} \int_{\lambda \in \mathbb{R}_+} \tilde{\phi}_c(i,j,\alpha,\beta,\lambda,t)h_b(t' - t)dtd\alpha d\beta d\lambda \quad (\text{A.14})$$

avec

$$h_b(t) = S' \exp\left[-\frac{t}{\tau_e}\right] \quad (\text{A.15})$$

où

$$\left\{ \begin{array}{l} S' = \nu\bar{R}I_p/C \\ \tau_e = C/(G_d - \nu\bar{R}I_p^2) \\ \bar{R} = R(\bar{T}) \\ \nu = nT_g^n/\bar{T}^{n+1} \\ \bar{T} = \phi^{-1}(\bar{P}) \\ G_d = G_0\bar{T}^\beta T_0^\beta \\ \bar{P} = \int \bar{\phi}_c d\alpha d\beta d\lambda. \end{array} \right. \quad (\text{A.16})$$

sont les paramètres physiques de la réponse.

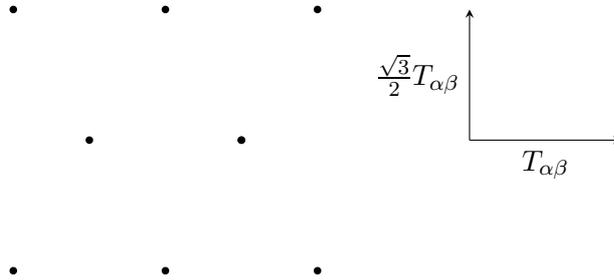
## Annexe B

# Compléments sur le modèle direct

### B.1 Grille hexagonale

Un peigne de Dirac hexagonal, illustré figure B.1, s'exprime comme une somme de deux peignes de Dirac sur grille rectangulaire en quinconce. L'opération d'échantillonnage sur une grille hexagonale s'exprime

$$p(\alpha, \beta) = \sum_{l, m} \delta \left( \alpha - \left( \frac{1}{4} (1 - (-1)^m) + l \right) T_{\alpha\beta}, \beta - \frac{\sqrt{3} T_{\alpha\beta}}{2} m \right).$$



**Figure B.1** – Grille d'échantillonnage hexagonale.

Dans le cas d'une rotation du capteur d'un angle  $\theta$  nous avons

$$\begin{pmatrix} \alpha_{lm} \\ \beta_{lm} \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \left( \frac{1}{4} (1 - (-1)^m) + l \right) T_{\alpha\beta} \\ \frac{\sqrt{3} T_{\alpha\beta}}{2} m \end{pmatrix}$$

soit

$$\begin{aligned} \alpha_{lm} &= \left( \frac{1}{4} (1 - (-1)^m) + l \right) T_{\alpha\beta} \cos \theta - \frac{\sqrt{3} T_{\alpha\beta} \sin \theta}{2} m \\ \beta_{lm} &= \left( \frac{1}{4} (1 - (-1)^m) + l \right) T_{\alpha\beta} \sin \theta + \frac{\sqrt{3} T_{\alpha\beta} \cos \theta}{2} m. \end{aligned}$$

La transformée de Fourier de ce peigne sans rotation est également un peigne mais de période inversée et tournée d'un quart de tour

$$\dot{P}(f_\alpha, f_\beta) = \sum_{l, m} \delta \left( f_\alpha - \frac{l}{T_{\alpha\beta}}, f_\beta - \left( \frac{1}{4} (1 - (-1)^l) + m \right) \frac{2}{\sqrt{3} T_{\alpha\beta}} \right),$$

et avec la rotation, identique dans le domaine de Fourier, nous avons comme grille d'échantillonnage dans le plan fréquence

$$\begin{pmatrix} f_{e\alpha} \\ f_{e\beta} \end{pmatrix} = \begin{pmatrix} \frac{1}{T_{\alpha\beta}} \cos \theta - \left( \frac{1}{4} (1 - (-1)^l) + m \right) \frac{2}{\sqrt{3}T_{\alpha\beta}} \sin \theta \\ \frac{1}{T_{\alpha\beta}} \sin \theta + \left( \frac{1}{4} (1 - (-1)^l) + m \right) \frac{2}{\sqrt{3}T_{\alpha\beta}} \cos \theta \end{pmatrix}.$$

## B.2 Décomposition sur des gaussiennes et protocole *Large map*

Cette partie présente le calcul explicite de l'équation (2.47) page 66

$$f(nT_e) = \frac{1}{2\pi} \frac{S'}{\Sigma_\alpha \Sigma_\beta} \exp \left[ \frac{nT_e}{\tau_e} \right] \int_{nT_e}^{+\infty} \exp \left[ -\frac{1}{2} \frac{(v_\alpha t + o_\alpha)^2}{\Sigma_\alpha^2} \right] \exp \left[ -\frac{1}{2} \frac{(v_\beta t + o_\beta)^2}{\Sigma_\beta^2} \right] \exp \left[ \frac{-t}{\tau_e} \right] dt \quad (\text{B.1})$$

avec

$$o_\alpha = c_\alpha + \alpha^{ij} - \alpha_{lm} \quad \text{et} \quad o_\beta = c_\beta + \beta^{ij} - \beta_{lm}.$$

### B.2.1 Calcul de l'argument

Nous allons ici exprimer la forme quadratique de l'argument de l'exponentielle

$$-\frac{1}{2} \frac{\tau_e \Sigma_\beta^2 (v_\alpha t + o_\alpha)^2 + \tau_e \Sigma_\alpha^2 (v_\beta t + o_\beta)^2 + 2\Sigma_\alpha^2 \Sigma_\beta^2 t}{\Sigma_\alpha^2 \Sigma_\beta^2 \tau_e} = -\frac{1}{2} \frac{n(t)}{\Sigma_\alpha^2 \Sigma_\beta^2 \tau_e}.$$

On développe et factorise le numérateur  $n(t)$  soit

$$\begin{aligned} n(t) &= \tau_e \Sigma_\beta^2 (v_\alpha^2 t^2 + 2v_\alpha o_\alpha t + o_\alpha^2) + \tau_e \Sigma_\alpha^2 (v_\beta^2 t^2 + 2v_\beta o_\beta t + o_\beta^2) + 2\Sigma_\alpha^2 \Sigma_\beta^2 t \\ &= \tau_e (\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2) t^2 + 2 [\tau_e \Sigma_\beta^2 v_\alpha o_\alpha + \tau_e \Sigma_\alpha^2 v_\beta o_\beta + \Sigma_\alpha^2 \Sigma_\beta^2] t + \tau_e (\Sigma_\beta^2 o_\alpha^2 + \Sigma_\alpha^2 o_\beta^2) \\ &= \Sigma^{-2} \left( t^2 + 2 \frac{\tau_e \Sigma_\beta^2 v_\alpha o_\alpha + \tau_e \Sigma_\alpha^2 v_\beta o_\beta + \Sigma_\alpha^2 \Sigma_\beta^2}{\tau_e (\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2)} t + \frac{\tau_e (\Sigma_\beta^2 o_\alpha^2 + \Sigma_\alpha^2 o_\beta^2)}{\tau_e (\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2)} \right) \\ &= \frac{t^2 + 2at + b + (a^2 - a^2)}{\Sigma^2} \\ &= \frac{(t+a)^2 + b - a^2}{\Sigma^2} \end{aligned}$$

où

$$\begin{aligned} a &= \Sigma^2 (\tau_e \Sigma_\beta^2 v_\alpha o_\alpha + \tau_e \Sigma_\alpha^2 v_\beta o_\beta + \Sigma_\alpha^2 \Sigma_\beta^2) \\ b &= \tau_e \Sigma^2 (\Sigma_\beta^2 o_\alpha^2 + \Sigma_\alpha^2 o_\beta^2) \end{aligned}$$

et

$$\Sigma^2 = \frac{1}{\tau_e (\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2)}.$$

### B.2.2 Report dans l'intégrale (2.47)

Donc en reportant cette forme quadratique en  $t$  dans l'intégrale on obtient

$$\begin{aligned}
 f(nT_e) &= \frac{1}{2\pi} \frac{S'}{\Sigma_\alpha \Sigma_\beta} \exp\left[\frac{nT_e}{\tau_e}\right] \int_{nT_e}^{+\infty} \exp\left[-\frac{1}{2} \frac{(t+a)^2 + b - a^2}{\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e}\right] dt \\
 &= \frac{1}{2\pi} \frac{S'}{\Sigma_\alpha \Sigma_\beta} \exp\left[\frac{nT_e}{\tau_e}\right] \exp\left[-\frac{1}{2} \frac{b - a^2}{\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e}\right] \int_{nT_e}^{+\infty} \exp\left[-\frac{1}{2} \frac{(t+a)^2}{\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e}\right] dt \\
 &= \frac{1}{2\pi} S' \frac{\sqrt{\pi \tau_e} \Sigma}{\sqrt{2}} \exp\left[\frac{nT_e}{\tau_e} - \frac{1}{2} \frac{b - a^2}{\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e}\right] \left(1 - \operatorname{erf}\left(\frac{nT_e + a}{\sqrt{2\tau_e} \Sigma \Sigma_\alpha \Sigma_\beta}\right)\right) \quad (\text{B.2})
 \end{aligned}$$

où

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\theta^2} d\theta.$$

L'expression de la convolution temporelle dans le cas du balayage *Large map* à vitesse constante et lorsque le ciel est décomposé sur des gaussiennes est bien explicite. Cette expression peut être simplifiée en faisant apparaître la fonction  $\operatorname{erfc}x = \exp(x^2)\operatorname{erfc}(x)$  où  $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$ . Pour cela il faut travailler l'argument de l'exponentielle pour faire ressortir l'argument de la fonction  $\operatorname{erf}$ .

### B.2.3 Argument de l'exponentielle

Pour éclaircir les écritures temporairement on pose  $S = \Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2$ . Le facteur qui intervient dans la fonction  $\exp$  est

$$\begin{aligned}
 \frac{nT_e}{\tau_e} - \frac{b - a^2}{2\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e} &= \frac{nT_e}{\tau_e} - \frac{b}{2S\tau_e} + \frac{a^2}{2S\tau_e} + \left(\frac{n^2 T_e^2}{2S\tau_e} - \frac{n^2 T_e^2}{2S\tau_e}\right) + \left(\frac{2nT_e a}{2S\tau_e} - \frac{2nT_e a}{2S\tau_e}\right) \\
 &= \frac{nT_e}{\tau_e} - \frac{b}{2S\tau_e} - \frac{2nT_e a}{2S\tau_e} - \frac{n^2 T_e^2}{2S\tau_e} + \left(\frac{nT_e + a}{\sqrt{2S\tau_e}}\right)^2
 \end{aligned}$$

Donc en injectant cette expression dans (B.2) la fonction  $\operatorname{erfc}x$  apparaît

$$\begin{aligned}
 \exp\left[\frac{nT_e}{\tau_e} - \frac{b - a^2}{2S\tau_e}\right] \left(1 - \operatorname{erf}\left(\frac{nT_e + a}{\sqrt{2S\tau_e}}\right)\right) &= \\
 \exp\left[\frac{nT_e}{\tau_e} - \frac{b}{2S\tau_e} - \frac{2nT_e a}{2S\tau_e} - \frac{n^2 T_e^2}{2S\tau_e}\right] \operatorname{erfc}x\left(\frac{nT_e + a}{\sqrt{2S\tau_e}}\right).
 \end{aligned}$$

Les valeurs respectives de  $S$ ,  $a$  et  $b$  peuvent être remplacées et beaucoup de choses vont se simplifier. Tout d'abord l'argument de l'exponentielle s'écrit

$$\begin{aligned}
 \frac{nT_e}{\tau_e} - \frac{\tau_e \Sigma^2 (\Sigma_\beta^2 o_\alpha^2 + \Sigma_\alpha^2 o_\beta^2)}{2\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e} - \frac{2nT_e \Sigma^2 (\tau_e \Sigma_\beta^2 v_\alpha o_\alpha + \tau_e \Sigma_\alpha^2 v_\beta o_\beta + \Sigma_\alpha^2 \Sigma_\beta^2)}{2\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e} - \frac{n^2 T_e^2}{2S\tau_e} &= \\
 \frac{nT_e}{\tau_e} - \frac{o_\alpha^2}{2\Sigma_\alpha^2} - \frac{o_\beta^2}{2\Sigma_\beta^2} - \frac{2nT_e v_\alpha o_\alpha}{2\Sigma_\alpha^2} - \frac{2nT_e v_\beta o_\beta}{2\Sigma_\beta^2} - \frac{nT_e}{\tau_e} - \frac{n^2 T_e^2}{2\Sigma_\alpha^2 \Sigma_\beta^2 \tau_e}.
 \end{aligned}$$

Les termes  $nT_e/\tau_e$  se simplifient. Ensuite on devine une identité remarquable quadratique mais pour cela il faut remplacer l'expression de  $\Sigma^2$  soit

$$\frac{n^2 T_e^2}{2\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e} = \frac{n^2 T_e^2 \tau_e (\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2)}{2\Sigma_\alpha^2 \Sigma_\beta^2 \tau_e} = \frac{n^2 T_e^2 v_\alpha^2}{2\Sigma_\alpha^2} + \frac{n^2 T_e^2 v_\beta^2}{2\Sigma_\beta^2}.$$

Donc finalement l'argument de l'exponentielle dans (B.2) s'écrit

$$\begin{aligned}
 -\frac{o_\alpha^2}{2\Sigma_\alpha^2} - \frac{o_\beta^2}{2\Sigma_\beta^2} - \frac{2nT_e v_\alpha o_\alpha}{2\Sigma_\alpha^2} - \frac{2nT_e v_\beta o_\beta}{2\Sigma_\beta^2} - \frac{n^2 T_e^2 v_\alpha^2}{2\Sigma_\alpha^2} - \frac{n^2 T_e^2 v_\beta^2}{2\Sigma_\beta^2} = \\
 -\frac{(o_\alpha + nT_e v_\alpha)^2}{2\Sigma_\alpha^2} - \frac{(o_\beta + nT_e v_\beta)^2}{2\Sigma_\beta^2} \quad (\text{B.3})
 \end{aligned}$$

qui est exactement l'argument d'une gaussienne bi-variée puisqu'il s'agit de l'argument de la fonction exponentielle. On trouve de nouveau les mêmes « écarts-types »  $\Sigma_\alpha$  et  $\Sigma_\beta$ . Cependant la réponse de l'optique gaussiennes qui est décalée en  $o_\alpha$  et  $o_\beta$  est décalée maintenant de  $nT_e v$ , c'est à dire le décalage de pointage entre deux échantillons temporels successifs. L'expression est donc logique puisqu'il s'agit de la réponse optique décalée par le mouvement du télescope.

#### B.2.4 Argument de la fonction erfcx

Il y a cependant un terme supplémentaire pour connaître la réponse globale, venant de la fonction erfcx, qui correspond à l'influence du bolomètre. L'argument de la fonction erfcx est

$$\begin{aligned}
 \frac{nT_e + a}{\sqrt{2\tau_e \Sigma \Sigma_\alpha \Sigma_\beta}} &= \frac{nT_e + \Sigma^2 (\tau_e \Sigma_\beta^2 v_\alpha o_\alpha + \tau_e \Sigma_\alpha^2 v_\beta o_\beta + \Sigma_\alpha^2 \Sigma_\beta^2)}{\sqrt{2\tau_e \Sigma \Sigma_\alpha \Sigma_\beta}} \\
 &= \frac{nT_e}{\sqrt{2\tau_e \Sigma \Sigma_\alpha \Sigma_\beta}} + \Sigma \frac{\tau_e \Sigma_\beta^2 v_\alpha o_\alpha + \tau_e \Sigma_\alpha^2 v_\beta o_\beta + \Sigma_\alpha^2 \Sigma_\beta^2}{\sqrt{2\tau_e \Sigma_\alpha \Sigma_\beta}} \\
 &= \frac{nT_e}{\sqrt{2\tau_e \Sigma \Sigma_\alpha \Sigma_\beta}} + \frac{\tau_e \Sigma_\beta^2 v_\alpha o_\alpha + \tau_e \Sigma_\alpha^2 v_\beta o_\beta + \Sigma_\alpha^2 \Sigma_\beta^2}{\sqrt{2\tau_e \Sigma_\alpha \Sigma_\beta} \sqrt{\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2}} \\
 &= \frac{nT_e \Sigma_v}{\sqrt{2\Sigma_\alpha \Sigma_\beta}} + \frac{\Sigma_\beta v_\alpha o_\alpha}{\sqrt{2\Sigma_\alpha \Sigma_v}} + \frac{\Sigma_\alpha v_\beta o_\beta}{\sqrt{2\Sigma_\beta \Sigma_v}} + \frac{\Sigma_\alpha \Sigma_\beta}{\sqrt{2\tau_e \Sigma_v}} \\
 &= \frac{\Sigma_\beta v_\alpha (o_\alpha + nT_e v_\alpha)}{\sqrt{2\Sigma_\alpha \Sigma_v}} + \frac{\Sigma_\alpha v_\beta (o_\beta + nT_e v_\beta)}{\sqrt{2\Sigma_\beta \Sigma_v}} + \frac{\Sigma_\alpha \Sigma_\beta}{\sqrt{2\tau_e \Sigma_v}} \quad (\text{B.4})
 \end{aligned}$$

où

$$\Sigma_v = \sqrt{\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2}$$

l'intérêt ici est que les mêmes facteurs interviennent dans l'argument de l'exponentielle. Pour connaître la réponse globale nous devons rassembler le tout. En injectant les expressions des arguments (B.3) et (B.4), nous obtenons

$$\begin{aligned}
 f(nT_e) &= \frac{1}{2\pi} S' \frac{\sqrt{\pi \tau_e \Sigma}}{\sqrt{2}} \exp \left[ \frac{nT_e}{\tau_e} - \frac{1}{2} \frac{b - a^2}{\Sigma^2 \Sigma_\alpha^2 \Sigma_\beta^2 \tau_e} \right] \left( 1 - \operatorname{erf} \left( \frac{nT_e + a}{\sqrt{2\tau_e \Sigma \Sigma_\alpha \Sigma_\beta}} \right) \right) \\
 &= \frac{S'}{2\sqrt{2\pi} \Sigma_v} \exp \left[ -\frac{(o_\alpha + nT_e v_\alpha)^2}{2\Sigma_\alpha^2} - \frac{(o_\beta + nT_e v_\beta)^2}{2\Sigma_\beta^2} \right] \\
 &\quad \operatorname{erfcx} \left( \frac{\Sigma_\beta v_\alpha (o_\alpha + nT_e v_\alpha)}{\sqrt{2\Sigma_\alpha \Sigma_v}} + \frac{\Sigma_\alpha v_\beta (o_\beta + nT_e v_\beta)}{\sqrt{2\Sigma_\beta \Sigma_v}} + \frac{\Sigma_\alpha \Sigma_\beta}{\sqrt{2\tau_e \Sigma_v}} \right)
 \end{aligned}$$

avec

$$\begin{aligned}
 \Sigma_{\alpha,\beta}^2 &= \sigma_{\alpha,\beta}^2 + \sigma_m^2 + \sigma_c^2 & \text{et} & \quad \Sigma_v = \sqrt{\Sigma_\beta^2 v_\alpha^2 + \Sigma_\alpha^2 v_\beta^2}, \\
 o_\alpha &= c_\alpha + \alpha^{ij} - \alpha_{lm} & \text{et} & \quad o_\beta = c_\beta + \beta^{ij} - \beta_{lm}.
 \end{aligned}$$

La grandeur physique de l'argument de l'exponentielle

$$\frac{([\mathbf{m}] + [\mathbf{s}] \left[\frac{\mathbf{m}}{\mathbf{s}}\right])^2}{[\mathbf{m}]^2} = \frac{[\mathbf{m}]^2}{[\mathbf{m}]^2}$$

est bien sans dimension ainsi que celui de la fonction  $\operatorname{erfc}x$

$$\frac{[\mathbf{m}] \left[\frac{\mathbf{m}}{\mathbf{s}}\right] ([\mathbf{m}] + [\mathbf{s}] \left[\frac{\mathbf{m}}{\mathbf{s}}\right])}{[\mathbf{m}] \left([\mathbf{m}]^2 \left[\frac{\mathbf{m}}{\mathbf{s}}\right]^2\right)^{1/2}} + \frac{[\mathbf{m}]^2}{[\mathbf{s}] \left([\mathbf{m}]^2 \left[\frac{\mathbf{m}}{\mathbf{s}}\right]^2\right)^{1/2}} = \frac{[\mathbf{m}]^3 [\mathbf{s}]^{-1}}{[\mathbf{m}]^3 [\mathbf{s}]^{-1}} + \frac{[\mathbf{m}]^2}{[\mathbf{m}]^2}.$$

### B.3 Calculs explicites pour les protocoles *Point source* et *Small map*

Cette partie présente le calcul de

$$f(nT_e) = \int_t g(p_\alpha(t) + \alpha^{ij} - \alpha_{lm}) g(p_\beta(t) + \beta^{ij} - \beta_{lm}) h_b(t - nT_e) dt$$

où  $g$  est soit une fonction gaussienne soit une fonction  $\operatorname{esinc}$  dans le cas de SPIRE. Dans le cas des protocoles *Point source* et *Small map* les pointages sont

$$p_\alpha(t) = \sum_{q=1}^Q c_\alpha^q \mathbb{1}_{[t_q \ t_{q+1}[}(t) \quad \text{et} \quad p_\beta(t) = \sum_{q=1}^Q c_\beta^q \mathbb{1}_{[t_q \ t_{q+1}[}(t)$$

où  $q$  est la  $q^{\text{e}}$  position pointée. Donc

$$g(p_\alpha(t) + \alpha^{ij} - \alpha_{lm}) = \sum_q g(\alpha_q + \alpha^{ij} - \alpha_{lm}) \mathbb{1}_{[t_q \ t_{q+1}[}(t)$$

et par conséquent

$$\begin{aligned} f(nT_e) &= \sum_q g(\alpha_q + \alpha^{ij} - \alpha_{lm}) g(\beta_q + \beta^{ij} - \beta_{lm}) \int_t \mathbb{1}_{[t_q \ t_{q+1}[}(t) h_b(t - nT_e) dt \\ &= \sum_q g(\alpha_q + \alpha^{ij} - \alpha_{lm}) g(\beta_q + \beta^{ij} - \beta_{lm}) \\ &\quad \int_{\mathbb{R}} S' \exp\left[-\frac{t - nT_e}{\tau_e}\right] \mathbb{1}_{[t_q \ t_{q+1}[}(t) \mathbb{1}_{[nT_e \ +\infty[}(t) dt. \end{aligned} \quad (\text{B.5})$$

La causalité du filtre passe-bas des bolomètres amène à considérer plusieurs cas

$$\begin{aligned} \mathbb{1}_{[t_q \ t_{q+1}[}(t) \mathbb{1}_{[nT_e \ +\infty[} &= 0 & \text{si} \quad t_{q+1} \leq nT_e \\ &= \mathbb{1}_{[nT_e \ t_{q+1}[} & \text{si} \quad t_q \leq nT_e \\ &= \mathbb{1}_{[t_q \ t_{q+1}[} & \text{si} \quad nT_e \leq t_q. \end{aligned}$$

Finalement l'intégration dans le temps donne

$$\begin{aligned} \int_{\mathbb{R}} \exp\left[-\frac{t - nT_e}{\tau_e}\right] \mathbb{1}_{[t_q \ t_{q+1}[}(t) \mathbb{1}_{[nT_e \ +\infty[}(t) dt &= \exp\left[\frac{nT_e}{\tau_e}\right] \int_{\max(t_q, nT_e)}^{t_{q+1}} \exp\left[-\frac{t}{\tau_e}\right] dt \\ &= \tau_e \left( \exp\left[\frac{nT_e - \max(t_q, nT_e)}{\tau_e}\right] - \exp\left[\frac{nT_e - t_{q+1}}{\tau_e}\right] \right) \end{aligned}$$

ou 0 si  $t_{q+1} \leq nT_e$ . Donc en reportant l'expression dans (B.5) nous avons

$$\begin{aligned} f(nT_e) &= \tau_e S' \sum_q g(\alpha_q + \alpha^{ij} - \alpha_{lm}) g(\beta_q + \beta^{ij} - \beta_{lm}) \\ &\quad \left( \exp\left[\frac{nT_e - \max(t_q, nT_e)}{\tau_e}\right] - \exp\left[\frac{nT_e - t_{q+1}}{\tau_e}\right] \right). \end{aligned}$$

## B.4 Autres modèles d'acquisition

Dans cette annexe nous mettons les équations obtenues pour des modèles d'acquisition dans d'autre cas que ceux décrits dans le paragraphe 2.4.5. Il s'agit plutôt d'une énumération puisque l'essentiel des calculs se trouve dans les parties B.2 et B.3 ainsi que dans le corps du texte partie 2.4.2.

**Large map et sinus cardinaux.** Dans ce cas le modèle s'exprime

$$y^{lm}(nT_e) = K' T_\alpha^{ij} T_\beta^{ij} \sum_{k=0}^K \sum_{ij} \sum_q c^k(i,j) \int_\lambda h_f(\lambda) (\lambda - \lambda_0)^k \int_t \text{esinc}_{T_\alpha^{ij}, \sigma_{mc}}(o_\alpha + nT_e v_\alpha) \text{esinc}_{T_\beta^{ij}, \sigma_{mc}}(o_\beta + nT_e v_\beta) \tilde{h}_T(t - nT_e) dt d\lambda$$

et il nécessite la discrétisation de deux intégrales.

**Large map et réponse du bolomètre négligée.** Les modèles sont bien sûr plus simples puisque nous n'avons pas à calculer une intégrale supplémentaire. Dans ce cas avec une décomposition du ciel sur des gaussiennes nous avons

$$y^{lm}(nT_e) = \frac{K'}{\sqrt{2\pi}\Sigma_\alpha\Sigma_\beta} \sum_{kij} c^k(i,j) \int_\lambda h_f(\lambda) (\lambda - \lambda_0)^k \exp\left[-\frac{(o_\alpha + nT_e v_\alpha)^2}{2\Sigma_\alpha^2} - \frac{(o_\beta + nT_e v_\beta)^2}{2\Sigma_\beta^2}\right] d\lambda$$

avec une décomposition sur des sinus cardinaux

$$y^{lm}(nT_e) = K' T_\alpha^{ij} T_\beta^{ij} \sum_{k=0}^K \sum_{ij} c^k(i,j) \int_\lambda h_f(\lambda) (\lambda - \lambda_0)^k \text{esinc}_{T_\alpha^{ij}, \sigma_{mc}}(o_\alpha + nT_e v_\alpha) \text{esinc}_{T_\beta^{ij}, \sigma_{mc}}(o_\beta + nT_e v_\beta) d\lambda.$$

**Small map et Point source et réponse du bolomètre négligée.** De même les modèles sont plus simples et nous avons

$$y^{lm}(nT_e) = \frac{K'}{2\pi\Sigma_\alpha\Sigma_\beta} \sum_{k=0}^K \sum_{ij} \sum_q c^k(i,j) \int_\lambda h_f(\lambda) (\lambda - \lambda_0)^k \exp\left[-\frac{1}{2} \frac{(\alpha_q + \alpha^{ij} - \alpha_{lm})^2}{\Sigma_\alpha^2}\right] \exp\left[-\frac{1}{2} \frac{(\beta_q + \beta^{ij} - \beta_{lm})^2}{\Sigma_\beta^2}\right] d\lambda$$

si le ciel est décomposé sur des gaussiennes et

$$y^{lm}(nT_e) = K' T_\alpha^{ij} T_\beta^{ij} \sum_{k=0}^K \sum_{ij} \sum_q c^k(i,j) \int_\lambda h_f(\lambda) (\lambda - \lambda_0)^k \text{esinc}_{T_\alpha^{ij}, \sigma_{mc}}(\alpha_q - \alpha^{ij} - \alpha_{lm}) \text{esinc}_{T_\beta^{ij}, \sigma_{mc}}(\beta_q - \beta^{ij} - \beta_{lm}) d\lambda$$

si le ciel est décomposé sur des sinus cardinaux.

## Annexe C

# Calculs d'hivers rigoureux

### C.1 Convolution de deux gaussiennes

L'intégration du produit de deux gaussiennes quelconques, avec  $(a, b, m_1, \sigma_1, m_2, \sigma_2) \in \mathbb{R}^7$ ,

$$\int_a^b \exp \left[ -\frac{1}{2} \frac{(x - m_1)^2}{\sigma_1^2} \right] \exp \left[ -\frac{1}{2} \frac{(x - m_2)^2}{\sigma_2^2} \right] dx = \int_a^b \exp [p(x)] dx$$

a pour résultats les deux équations (C.3) et (C.4) et s'obtiennent avec les développements suivants. En développant le polynôme on obtient

$$\begin{aligned} p(x) &= -\frac{1}{2} \left( x^2 \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) - 2x \left( \frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2} \right) \right) - \frac{1}{2} \left( \frac{m_1^2}{\sigma_1^2} + \frac{m_2^2}{\sigma_2^2} \right) \\ &= -\frac{1}{2} \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \left( x^2 - 2 \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} \left( \frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2} \right) x \right) - \frac{1}{2} \left( \frac{m_1^2}{\sigma_1^2} + \frac{m_2^2}{\sigma_2^2} \right) \\ &= -\frac{1}{2} \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \left( x - \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} \left( \frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2} \right) \right)^2 \end{aligned} \quad (\text{C.1})$$

$$- \frac{1}{2} \left( \frac{m_1^2}{\sigma_1^2} + \frac{m_2^2}{\sigma_2^2} \right) + \frac{1}{2} \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \left[ \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} \left( \frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2} \right) \right]^2 \quad (\text{C.2})$$

Le premier facteur (C.1) reste dans l'intégrale puisqu'il dépend de  $x$ . Le deuxième facteur (C.2) sort de l'intégrale. On note pour la suite

$$\left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \Sigma^2$$

et

$$\left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} \left( \frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2} \right) = \frac{\sigma_2^2 m_1 + \sigma_1^2 m_2}{\sigma_1^2 + \sigma_2^2} = \mu$$

Le facteur (C.1) se calcule en utilisant la fonction erreur  $\text{erf}(a) = \frac{2}{\sqrt{\pi}} \int_0^a e^{-t^2} dt$  et le changement de variable

$$\begin{aligned} t = \frac{x - \mu}{\Sigma\sqrt{2}} \Bigg|_{x=0} &\rightarrow t = \frac{\mu}{\Sigma\sqrt{2}} \\ dt = \frac{dx}{\Sigma\sqrt{2}} \Bigg|_{x=a} &\rightarrow t = \frac{a - \mu}{\Sigma\sqrt{2}}. \end{aligned}$$

On obtient alors

$$\begin{aligned} \int_a^b \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\Sigma^2}\right) dx &= \int_{-\infty}^b \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\Sigma^2}\right) dx - \int_{-\infty}^a \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\Sigma^2}\right) dx \\ &= \frac{\Sigma\sqrt{\pi}}{\sqrt{2}} \left( \operatorname{erf} \frac{b-\mu}{\Sigma\sqrt{2}} - \operatorname{erf} \frac{a-\mu}{\Sigma\sqrt{2}} \right) \end{aligned}$$

Le deuxième facteur (C.2) s'obtient en développant et factorisant soit

$$\begin{aligned} -\frac{1}{2} \left( \frac{m_1^2}{\sigma_1^2} + \frac{m_2^2}{\sigma_2^2} \right) + \frac{1}{2} \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \left[ \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} \left( \frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2} \right) \right]^2 \\ = -\frac{1}{2} \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1} \left[ \left( \frac{m_1^2}{\sigma_1^2} + \frac{m_2^2}{\sigma_2^2} \right) \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) - \left( \frac{m_1}{\sigma_1^2} + \frac{m_2}{\sigma_2^2} \right)^2 \right] \\ = -\frac{1}{2} \frac{\sigma_1^2 \sigma_2^2}{1\sigma_1^2 + \sigma_2^2} \left[ \frac{1}{\sigma_1^2 \sigma_2^2} (1m_1^2 - 2m_1 m_2 + m_2^2) \right] = -\frac{1}{2} \frac{(m_1 - m_2)^2}{\sigma_1^2 + \sigma_2^2}. \end{aligned}$$

Donc pour tout  $a$  et  $b \in \mathbb{R}^2$  nous avons

$$\boxed{\int_a^b \exp\left(-\frac{1}{2} \frac{(x-m_1)^2}{\sigma_1^2}\right) \exp\left(-\frac{1}{2} \frac{(x-m_2)^2}{\sigma_2^2}\right) dx = \frac{\Sigma\sqrt{\pi}}{\sqrt{2}} \exp\left(-\frac{1}{2} \frac{(m_1 - m_2)^2}{\sigma_1^2 + \sigma_2^2}\right) \left[ \operatorname{erf} \frac{b-\mu}{\Sigma\sqrt{2}} - \operatorname{erf} \frac{a-\mu}{\Sigma\sqrt{2}} \right].} \quad (\text{C.3})$$

En particulier lorsque les bornes tendent vers l'infini

$$\begin{aligned} \lim_{x \rightarrow -\infty} \operatorname{erf}(x) &= -1 \\ \lim_{x \rightarrow +\infty} \operatorname{erf}(x) &= 1 \end{aligned}$$

l'intégration est la convolution de deux gaussiennes

$$\boxed{\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \frac{(x-m_1)^2}{\sigma_1^2}\right) \exp\left(-\frac{1}{2} \frac{(x-m_2)^2}{\sigma_2^2}\right) dx = \sqrt{2\pi}\Sigma \exp\left(-\frac{1}{2} \frac{(m_1 - m_2)^2}{\sigma_1^2 + \sigma_2^2}\right).} \quad (\text{C.4})$$

## C.2 Convolution d'une gaussienne et d'un sinus cardinal

On souhaite connaître la convolution entre une gaussienne et un sinus cardinal

$$y(c-m) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma_g} \exp\left[-\frac{1}{2} \frac{(x-m)^2}{\sigma_g^2}\right] T\operatorname{sinc}(T(x-c)) dx \quad (\text{C.5})$$

dont le résultat est directement disponible équation (C.9) page 162. Il est possible d'exprimer explicitement cette intégrale à l'aide de fonctions spéciales. On pose

$$\begin{aligned} u = T(x-c) & \left| \begin{array}{l} x = -\infty \rightarrow u = -\infty \\ x = +\infty \rightarrow u = +\infty \end{array} \right. \\ du = T dx & \end{aligned}$$

et par changement de variable  $x = u/T + c$

$$y(c - m) = \frac{1}{\sqrt{2\pi}\sigma_g} \int_{\mathbb{R}} \exp \left[ -\frac{1}{2} \frac{(u + m_c)^2}{\sigma^2} \right] \operatorname{sinc}(u) du$$

avec  $m_c = cT - mT$  et  $\sigma = T\sigma_g$ . Alors en développant le terme au carré de la gaussienne on peut exprimer cette intégrale ainsi

$$y(c - l) = \frac{1}{\sqrt{2\pi}\sigma_g} \exp \left[ -\frac{1}{2} \frac{m_c^2}{\sigma^2} \right] \int_{\mathbb{R}} \exp \left[ -\beta u^2 - \gamma u \right] \sin(u) \frac{du}{u}$$

où

$$\beta = \frac{1}{2} \frac{1}{\sigma^2} \quad \text{et} \quad \gamma = \frac{1}{2} \frac{2m_c}{\sigma^2}.$$

Intuitivement le résultat de la convolution entre une gaussienne et un sinus cardinal ressemblera à un sinus cardinal dont les lobes seront atténués. En analysant cette expression on devine que la partie gaussienne en  $m_c^2$  formera le lobe principal et que le résultat de l'intégrale compensera sa décroissance exponentielle pour former les lobes secondaires. Toute la difficulté réside justement dans le calcul de cette intégrale. En fait il est possible d'avancer en exploitant l'expression Eq. 3.953.1 de [GR80] p. 496

$$\int_0^{\infty} x^{\mu-1} e^{-\gamma x - \beta x^2} \sin(ax) dx = \frac{-i}{2(2\beta)^{\mu/2}} \exp \left[ \frac{\gamma^2 - a^2}{8\beta} \right] \Gamma(\mu) \left\{ \exp \left[ -\frac{ia\gamma}{4\beta} \right] D_{-\mu} \left( \frac{\gamma - ia}{\sqrt{2\beta}} \right) - \exp \left[ \frac{ia\gamma}{4\beta} \right] D_{-\mu} \left( \frac{\gamma + ia}{\sqrt{2\beta}} \right) \right\}. \quad (\text{C.6})$$

si les parties réelles  $\Re[\mu] > -1$ ,  $\Re[\beta] > 0$  et  $a > 0$ , avec  $\Gamma$  la fonction gamma et  $D$  une fonction parabolique cylindrique ([GR80] p. 1064). Cette expression paraît compliquée mais il s'agit d'un cas particulier des valeurs  $\mu, \gamma, \beta$  et  $a$ , et de plus certains facteurs vont se simplifier puisque notre intégrale est sur  $\mathbb{R}$  au lieu de  $\mathbb{R}^+$ . Tout d'abord développons notre intégrale en deux parties justement

$$\begin{aligned} I &= \int_{\mathbb{R}} \exp \left[ -\beta u^2 - \gamma u \right] \sin(u) \frac{du}{u} \\ &= \int_0^{+\infty} \exp \left[ -\beta u^2 - \gamma u \right] \sin(u) \frac{du}{u} + \int_{-\infty}^0 \exp \left[ -\beta u^2 - \gamma u \right] \sin(u) \frac{du}{u} \\ &= \int_0^{+\infty} \exp \left[ -\beta u^2 - \gamma u \right] \sin(u) \frac{du}{u} - \int_0^{+\infty} \exp \left[ -\beta u^2 - \gamma u \right] \sin(u) \frac{du}{u}. \end{aligned}$$

En posant  $u = -v$  et comme  $\sin(-x) = -\sin(x)$ , on obtient

$$I = \int_0^{+\infty} \exp \left[ -\beta u^2 - \gamma u \right] \sin(u) \frac{du}{u} + \int_0^{+\infty} \exp \left[ -\beta v^2 - \gamma' v \right] \sin(u) \frac{dv}{v} \quad (\text{C.7})$$

où  $\gamma' = -\gamma$ . A partir de là on peut utiliser (C.6). Or nous avons le cas particulier  $a = 1$ ,  $\mu = 0$  et  $\beta \in \mathbb{R}_+^*$  vérifie bien la condition, alors

$$\int_0^{\infty} e^{-\gamma x - \beta x^2} \sin(x) \frac{dx}{x} = \frac{-i}{2} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \left\{ \exp \left[ -\frac{i\gamma}{4\beta} \right] D_0 \left( \frac{\gamma - i}{\sqrt{2\beta}} \right) - \exp \left[ \frac{i\gamma}{4\beta} \right] D_0 \left( \frac{\gamma + i}{\sqrt{2\beta}} \right) \right\}.$$

En exploitant ce résultat dans (C.7) on obtient

$$\begin{aligned} I &= \frac{-i}{2} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \left\{ \exp \left[ -\frac{i\gamma}{4\beta} \right] D_0 \left( \frac{\gamma - i}{\sqrt{2\beta}} \right) - \exp \left[ \frac{i\gamma}{4\beta} \right] D_0 \left( \frac{\gamma + i}{\sqrt{2\beta}} \right) \right\} + \\ &\quad \frac{-i}{2} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \left\{ \exp \left[ \frac{i\gamma}{4\beta} \right] D_0 \left( \frac{-\gamma - i}{\sqrt{2\beta}} \right) - \exp \left[ -\frac{i\gamma}{4\beta} \right] D_0 \left( \frac{-\gamma + i}{\sqrt{2\beta}} \right) \right\}. \end{aligned}$$

L'expression paraît lourde et complexe mais certains facteurs peuvent être factorisés et d'autres facteurs interviennent plusieurs fois sous plusieurs formes. Pour éclaircir, sachant que  $(\beta, \gamma) \in \mathbb{R}^2$ , on pose

$$\begin{aligned} z &= \frac{\gamma + i}{\sqrt{2\beta}} & z^* &= \frac{\gamma - i}{\sqrt{2\beta}} \\ -z &= \frac{-\gamma - i}{\sqrt{2\beta}} & -z^* &= \frac{-\gamma + i}{\sqrt{2\beta}}. \end{aligned}$$

Après factorisation, l'intégrale s'exprime

$$I = \frac{-i}{2} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \left\{ \exp \left[ -\frac{i\gamma}{4\beta} \right] \left( D_0(z^*) - D_0(-z^*) \right) - \exp \left[ \frac{i\gamma}{4\beta} \right] \left( D_0(z) - D_0(-z) \right) \right\}. \quad (\text{C.8})$$

Les choses sont déjà plus claires. On aimerait bien par ailleurs pouvoir exploiter des propriétés de la fonction parabolique cylindrique  $D_p$  pour simplifier un peu plus cette expression. Or, justement d'après Eq. 9.240 de [GR80] p. 1064

$$D_p(z) = 2^{p/2} e^{-z^2/4} \left\{ \frac{\sqrt{\pi}}{\Gamma\left(\frac{1-p}{2}\right)} \Phi\left(-\frac{p}{2}, \frac{1}{2}; \frac{z^2}{2}\right) - \frac{\sqrt{2\pi}z}{\Gamma\left(\frac{-p}{2}\right)} \Phi\left(\frac{1-p}{2}, \frac{3}{2}; \frac{z^2}{2}\right) \right\}$$

où  $\Phi$  est la fonction hyper-géométrique *cf.* Eq. 9.210.1 de [GR80] p. 1058. Nous concernant nous avons  $p = 0$  et  $\Phi(0, a; b) = 1$  pour tout  $(a, b)$ . De plus la fonction gamma  $\Gamma(0) = 1$  *cf.* Eq. 9.310.2 de [GR80] p. 933. En exploitant cette expression pour  $D_0$  nous avons

$$D_0(z) = e^{-z^2/4} \left\{ \frac{\sqrt{\pi}}{\Gamma(1/2)} - \sqrt{2\pi}z \Phi\left(\frac{1}{2}, \frac{3}{2}; \frac{z^2}{2}\right) \right\}.$$

L'avantage de cette expression est que d'une fonction qui dépend de  $z$  nous passons à une expression qui dépend de  $z^2$ . On en déduit les autres facteurs qui interviennent dans (C.8)

$$D_0(z^*) = e^{-(z^*)^2/4} \left\{ \frac{\sqrt{\pi}}{\Gamma(1/2)} - \sqrt{2\pi}z^* \Phi\left(\frac{1}{2}, \frac{3}{2}; \frac{(z^*)^2}{2}\right) \right\},$$

$$D_0(-z) = e^{-z^2/4} \left\{ \frac{\sqrt{\pi}}{\Gamma(1/2)} + \sqrt{2\pi}z \Phi\left(\frac{1}{2}, \frac{3}{2}; \frac{z^2}{2}\right) \right\}$$

et

$$D_0(-z^*) = e^{-(z^*)^2/4} \left\{ \frac{\sqrt{\pi}}{\Gamma(1/2)} + \sqrt{2\pi}z^* \Phi\left(\frac{1}{2}, \frac{3}{2}; \frac{(z^*)^2}{2}\right) \right\}.$$

Or dans (C.8) ils interviennent sous forme de différence entre eux soit

$$D_0(z^*) - D_0(-z^*) = e^{-(z^*)^2/4} \left\{ -2\sqrt{2\pi}z^* \Phi\left(\frac{1}{2}, \frac{3}{2}; \frac{(z^*)^2}{2}\right) \right\},$$

et

$$D_0(z) - D_0(-z) = e^{-z^2/4} \left\{ -2\sqrt{2\pi}z \Phi\left(\frac{1}{2}, \frac{3}{2}; \frac{z^2}{2}\right) \right\}.$$

et en reportant ces expressions dans (C.8) on obtient

$$\begin{aligned} I &= \frac{-i}{2} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \left\{ -\exp \left[ \frac{-i\gamma}{4\beta} \right] e^{-(z^*)^2/4} 2\sqrt{2\pi}z^* \Phi\left(\frac{1}{2}, \frac{3}{2}; \frac{(z^*)^2}{2}\right) - \right. \\ &\quad \left. - \exp \left[ \frac{i\gamma}{4\beta} \right] e^{-z^2/4} 2\sqrt{2\pi}z \Phi\left(\frac{1}{2}, \frac{3}{2}; \frac{z^2}{2}\right) \right\}. \end{aligned}$$

En examinant cette expression on voit qu'entre les deux termes, l'un exploite  $z$  et l'autre  $z^*$ . Donc en posant

$$Z = \frac{i\gamma}{4\beta}$$

en exploitant les propriétés

$$\begin{aligned} (z^*)^n &= (z^n)^* \\ \exp [z^*] &= \exp [z]^* \end{aligned}$$

et

$$\begin{aligned} \Phi(a,b; z^*) &= 1 + \frac{a}{b}z^* + \frac{a(a+1)}{b(b+1)} \frac{(z^*)^2}{2!} + \dots \\ &= \Phi(a,b; z)^* \end{aligned}$$

on obtient

$$I = -i\sqrt{2\pi} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \left\{ \exp [Z] e^{-z^2/4} z \Phi \left( \frac{1}{2}, \frac{3}{2}; \frac{z^2}{2} \right) - \exp [Z]^* (e^{-z^2/4})^* z^* \Phi \left( \frac{1}{2}, \frac{3}{2}; \frac{z^2}{2} \right)^* \right\}.$$

Donc nous avons la différence entre un complexe et son conjugué, soit en posant

$$K = z \exp [Z] \exp [-z^2/4] \Phi \left( \frac{1}{2}, \frac{3}{2}; \frac{z^2}{2} \right),$$

on a

$$\begin{aligned} I &= -i\sqrt{2\pi} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \{K - K^*\} \\ &= -i\sqrt{2\pi} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] 2i\Im[K] \\ &= 2\sqrt{2\pi} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \Im[K] \\ &= 2\sqrt{2\pi} \exp \left[ \frac{\gamma^2 - 1}{8\beta} \right] \Im \left[ \exp \left[ \frac{i\gamma}{4\beta} \right] \exp \left[ -\frac{z^2}{4} \right] z \Phi \left( \frac{1}{2}, \frac{3}{2}; \frac{z^2}{2} \right) \right] \end{aligned}$$

qui est bien un nombre réel après un passage dans le domaine complexe (la convolution de deux fonctions qui prennent leur valeur dans  $\mathbb{R}$  est également dans  $\mathbb{R}$ ). On peut réunir les arguments des exponentielles soit

$$\begin{aligned} \frac{\gamma^2 - 1}{8\beta} + \frac{i\gamma}{4\beta} - \frac{z^2}{4} &= \frac{\gamma^2 - 1}{8\beta} + \frac{2i\gamma}{8\beta} - \frac{1}{4} \frac{\gamma^2 + 2i\gamma - 1}{2\beta} \\ &= \frac{\gamma^2 + 2i\gamma - 1}{8\beta} - \frac{\gamma^2 + 2i\gamma - 1}{8\beta} \\ &= 0. \end{aligned}$$

Donc l'expression de l'intégrale est encore plus simple

$$I = 2\sqrt{2\pi} \Im \left[ z \Phi \left( \frac{1}{2}, \frac{3}{2}; \frac{z^2}{2} \right) \right].$$

Il nous reste maintenant à exprimer cette fonction avec les paramètres définis au départ soit

$$\begin{aligned} z &= \frac{\gamma + i}{\sqrt{2\beta}} = \frac{m_c/\sigma^2 + i}{\sqrt{1/\sigma^2}} \\ &= \frac{c - m}{\sigma_g} + iT\sigma_g \end{aligned}$$

et donc

$$I = 2\sqrt{2\pi}\Im \left[ \left( \frac{c - m}{\sigma_g} + iT\sigma_g \right) \Phi \left( \frac{1}{2}, \frac{3}{2}; \frac{1}{2} \left( \frac{c - m}{\sigma_g} + iT\sigma_g \right)^2 \right) \right].$$

Finalement nous pouvons enfin exprimer la convolution entre une gaussienne et un sinus cardinal soit

$$\boxed{\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma_g} \exp \left[ -\frac{1}{2} \frac{(x - m)^2}{\sigma_g^2} \right] \operatorname{sinc}(T(x - c)) \, dx = \frac{2}{\sigma_g} \exp \left[ -\frac{1}{2} \frac{(c - m)^2}{\sigma_g^2} \right] \Im \left[ \left( \frac{c - m}{\sigma_g} + iT\sigma_g \right) \Phi \left( \frac{1}{2}, \frac{3}{2}; \frac{1}{2} \left( \frac{c - m}{\sigma_g} + iT\sigma_g \right)^2 \right) \right]} \quad (\text{C.9})$$

avec  $\Im$  la partie imaginaire et  $\Phi$  la fonction hypergéométrique dégénérée voir Eq. 9.210.1 de [GR80] p. 1058. disponible entre autres dans la toolbox `symbolic` de Matlab<sup>TM</sup>. A notre connaissance c'est la première fois que le résultat explicite de cette convolution est mis ainsi en évidence.

### C.3 Lemme d'inversion de matrices partitionnées

Le lemme d'inversion de matrices partitionnées est notamment utilisé pour le calcul du hessien dans l'annexe suivante. Soit une matrice inversible

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix}$$

alors

$$\mathbf{P}^{-1} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

où

$$\begin{aligned} \mathbf{A}_{11} &= \left( \mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21} \right)^{-1} \\ \mathbf{A}_{22} &= \left( \mathbf{P}_{22} - \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12} \right)^{-1} \\ \mathbf{A}_{12} &= -\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{A}_{22} \\ \mathbf{A}_{21} &= -\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\mathbf{A}_{11}. \end{aligned}$$

### C.4 Calculs du filtre de Wiener–Hunt et inverse du hessien

Cette partie décrit le calcul du filtre de Wiener–Hunt de l'équation (3.50) page 107 par approximation diagonale. Nous allons séparer le calcul en deux en calculant d'abord la rétro-projection

$D^t H_c^t \Sigma^t \tilde{\mathbf{y}}$  puis l'inverse du hessien. Pour simplifier, en reprenant huit balayages et deux dans chaque direction on a

$$\begin{aligned}
 D^t H_c^t \Sigma^t \tilde{\mathbf{y}} &= \begin{pmatrix} I \\ I \\ \vdots \\ I \end{pmatrix}^t \begin{pmatrix} H_0^1 & & & & \\ & H_1^1 & & & \\ & & H_0^2 & & \\ & & & H_1^2 & \\ & & & & \ddots \\ & & & & & H_1^4 \end{pmatrix}^t \begin{pmatrix} I & I & & & & & & \\ & & I & I & & & & \\ & & & & I & I & & \\ & & & & & & I & I \\ & & & & & & & I & I \end{pmatrix}^t \tilde{\mathbf{y}} \\
 &= \begin{pmatrix} H_0^{1t} & & & & \\ & H_1^{1t} & & & \\ & & H_0^{2t} & & \dots \\ & & & H_1^{2t} & \dots \\ & & & & \dots & H_1^{4t} \end{pmatrix} \begin{pmatrix} I \\ I \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{y}}^1 \\ \tilde{\mathbf{y}}^2 \\ \tilde{\mathbf{y}}^3 \\ \tilde{\mathbf{y}}^4 \end{pmatrix} \\
 &= \begin{pmatrix} H_0^{1t} & H_0^{2t} & H_0^{3t} & H_0^{4t} \\ H_1^{1t} & H_1^{2t} & H_1^{3t} & H_1^{4t} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{y}}^1 \\ \tilde{\mathbf{y}}^2 \\ \tilde{\mathbf{y}}^3 \\ \tilde{\mathbf{y}}^4 \end{pmatrix} \\
 &= \sum_{v=1}^4 \begin{pmatrix} H_0^{vt} \tilde{\mathbf{y}}^v \\ H_1^{vt} \tilde{\mathbf{y}}^v \end{pmatrix}
 \end{aligned}$$

Le calcul numérique est donc extrêmement simple et consiste à appliquer, par ordre de développement de Taylor, la somme des transposés des convolutions sur chaque « image » obtenue par retroprojection des données acquises avec cette direction de balayage.

Pour calculer l'inverse du hessien

$$\left( D^t H_c^t \Sigma^t \Sigma H_c D + \Sigma_x^{-1} \right)^{-1}$$

procédons par étape. Tout d'abord

$$\Sigma^t \Sigma = \begin{pmatrix} I \\ I \end{pmatrix} \begin{pmatrix} I & I & & & & & & \\ & & I & I & & & & \\ & & & & I & I & & \\ & & & & & & I & I \\ & & & & & & & I & I \\ & & & & & & & & I & I \end{pmatrix} = \begin{pmatrix} I & I & & & & & & \\ I & I & & & & & & \\ & & I & I & & & & \\ & & & & I & I & & \\ & & & & & & I & I \\ & & & & & & & I & I \\ & & & & & & & & I & I \\ & & & & & & & & & I & I \end{pmatrix}.$$



et l'inverse du hessien étant obtenu avec le lemme d'inversion de matrices partitionnées (voir annexe C.3) soit

$$\begin{aligned}
 \mathbf{A}_{11} &= \left( \sum_v |\Lambda_0^v|^2 + |\Lambda_{D_0}|^2 + \sum_v \Lambda_0^{v\dagger} \Lambda_1^v \left( \sum_v |\Lambda_1^v|^2 + |\Lambda_{D_1}|^2 \right)^{-1} \sum_v \Lambda_1^{v\dagger} \Lambda_0^v \right)^{-1} \\
 \mathbf{A}_{22} &= \left( \sum_v |\Lambda_1^v|^2 + |\Lambda_{D_1}|^2 + \sum_v \Lambda_1^{v\dagger} \Lambda_0^v \left( \sum_v |\Lambda_0^v|^2 + |\Lambda_{D_0}|^2 \right)^{-1} \sum_v \Lambda_0^{v\dagger} \Lambda_1^v \right)^{-1} \\
 \mathbf{A}_{12} &= - \left( \sum_v |\Lambda_0^v|^2 + |\Lambda_{D_0}|^2 \right)^{-1} \left( \sum_v \Lambda_0^{v\dagger} \Lambda_1^v \right) \mathbf{A}_{22} \\
 \mathbf{A}_{21} &= - \left( \sum_v |\Lambda_1^v|^2 + |\Lambda_{D_1}|^2 \right)^{-1} \left( \sum_v \Lambda_1^{v\dagger} \Lambda_0^v \right) \mathbf{A}_{11}.
 \end{aligned}$$

Chaque partie est diagonale et donc très facile à calculer et inverser. La généralisation à un nombre d'ordre de développement de Taylor est immédiate. Nous avons tout d'abord une augmentation du nombre de lignes des matrices, une par ordre. Enfin le hessien aura sur la diagonale la norme de la réponse pour chaque ordre et les termes croisés entre les ordres deux à deux dans les autres parties.



## Annexe D

# Valeur numérique des paramètres du modèle

Cette annexe présente la valeur connue à la publication de cette thèse des paramètres instruments [Gri06]. Ce sont en particulier ces valeurs qui ont été employées pour les simulations du chapitre 4.

### Paramètre du modèle de l'optique

Primaire [m]	Secondaire [m]	Focale [m]	$c$ [".m <sup>-1</sup> ](où $\sigma_m^2 + \sigma_c^2 = c^2 \lambda^2$ )
3,285	0,56	8,68	$6,857 \times 10^4$

### Paramètres du modèle physique des bolomètres

$V_p$ [V]	$R_c$ [M $\Omega$ ]	$G_0$ [pWK <sup>-1</sup> ]	$T_0$ [K]	$\beta$	$R_g$ [ $\Omega$ ]	$T_g$ [K]	$n$
0,015	16	65	0,2	1,7	80	41	0,5

### Paramètres du modèle linéaire des bolomètres

$\nu$	$\bar{R}$	$V_p$ [mV]	$R_c$ [M $\Omega$ ]	$G_d$ [pWK <sup>-1</sup> ]	$C$ [pJK <sup>-1</sup> ]	$S'$ [VW <sup>-1</sup> ]	$\tau_e$ [ms]	$f_c$ [Hz]
-	-	3	$\approx 16$	$\approx 65$	$\approx 0.6$	$3,31 \times 10^8$	19,45	51,4



# Annexe E

## Algorithmes

Cette annexe présente l'algorithme établi partie 3.2 en langage pseudo-code. Il correspond à l'algorithme utilisé pour produire les résultats de la partie 4. La plupart de ces algorithmes ont été mis en œuvre sur la plateforme MATLAB<sup>TM</sup> version 2009a en environ 4200 lignes dont 2000 de commentaires. Le code source sera, à terme, disponible librement.

Plusieurs hypothèses de modélisation sont considérées pour obtenir ces algorithmes.

- Le ciel est décomposé dans l'espace sur des fonctions disposées régulièrement.
- La position des bolomètres est arrondie pour permettre une mise en œuvre rapide du modèle direct par convolution puis échantillonnage.
- Les matrices  $\mathbf{H}_c$ ,  $\mathbf{D}$  et  $\mathbf{\Sigma}_c$  sont des matrices de convolution ou composées de matrices de convolution.
- Contrairement aux écritures mathématiques, les symboles en capitale grasse représentent non pas des matrices mais les noyaux de convolution, c'est à dire la première ligne de la matrice.
- L'opération de convolution précédemment notée  $\mathbf{H}_c \mathbf{x}$  se note  $\mathbf{H}_c \star \mathbf{x}$  et le produit terme à terme  $\mathbf{H}_c \odot \mathbf{x}$ , sous-entendu que les tailles des tableaux correspondent.
- Les éléments  $\mathbf{P}$  et  $\mathbf{I}$  sont requis. Pour  $\mathbf{P}$  il s'agit de l'information de pointage liant les coefficients observés aux données considérées valides. Cette information est obtenue par arrondis des pointages fournis par l'instrument. Pour  $\mathbf{I}$  il s'agit des informations sur les données valides à utiliser pour le terme d'adéquation aux données pour  $\mathbf{c}$  (pour éliminer le bord des balayages par exemple). A noter que dans les algorithmes présentés ici, on conserve toutes les données des bolomètres aveugles mais une extension de  $\mathbf{I}$  aux bolomètres aveugles est possible.

Certains abus de notations peu importants sont présents pour alléger les écritures et faciliter la compréhension globale. L'ensemble à été décomposé en blocs plus ou moins élémentaires.

Le premier algorithme 1 page 171 présente le calcul d'une sortie modèle, c'est à dire la reproduction des données connaissant le ciel à l'entrée  $\mathbf{x}$  et le signal commun basse-fréquence  $\mathbf{c}$  vue par tous les bolomètres, lorsque les autres inconnues sont fixés. La convolution de la ligne 1.5 est de loin le calcul prenant le plus de temps processeur dans l'échantillonneur de Gibbs (plus de 90% du temps dans les résultats présentés partie 4). Le deuxième algorithme 2 correspond au transposé pour les coefficients  $\mathbf{x}$ , soit une mis en œuvre de

$$\sum_{b=0}^{B-1} \gamma_{b_b} \mathbf{H}_b^t (\mathbf{H}_b \mathbf{x} + \mathbf{T}_b \mathbf{c} - \tilde{\mathbf{y}}_b)$$

voir équation (3.44) page 98.

Les algorithmes 3 et 4 page 172 sont une présentation de l'algorithme d'optimisation, de la partie 3.2.3, pour  $\mathbf{x}$  et  $\mathbf{c}$  par descente de gradient à hyper-paramètres fixés. Dans cette version

l'optimisation est faite sur  $\boldsymbol{x}$  et  $\boldsymbol{c}$  indépendamment. Les deux peuvent être menées en même temps ou encore de manière alternée. Cette version est intéressante dans une perspective d'échantillonnage. Les lignes 3.11 et 4.11 sont des étapes de minimisation en ligne pour obtenir un pas permettant une descente du critère importante. En pratique le logiciel GPAC [BCG<sup>+</sup>06] a été utilisé.

Les deux algorithmes 5 et 6 page 173 englobent les algorithmes d'optimisation pour effectuer l'échantillonnage des lois conditionnelles *a posteriori* de  $\boldsymbol{x}$  et  $\boldsymbol{c}$ . Ils consistent principalement à perturber le système avant l'optimisation proprement dite.

Les deux algorithmes 7 et 8 page 174 sont des mises en œuvre d'un échantillonneur de Metropolis-Hastings pour les paramètres instrument et les paramètres de la corrélation *a priori* pour  $\boldsymbol{c}$  respectivement.

Enfin l'algorithme 9 page 175 décrit l'échantillonneur de Gibbs permettant l'estimation des paramètres et de l'objet. Chaque ligne d'échantillonnage d'un paramètre considéré inconnu peut être ignorée et dans ce cas la méthode considère le paramètre connu et fixé. A noter que la même initialisation est utilisée pour l'échantillonnage de  $\boldsymbol{x}$  et  $\boldsymbol{c}$  par optimisation. L'utilisation des échantillons précédents est possible. On peut remarquer que ce qui est nécessaire maintenant se résume à l'information de pointage, la régularisation et les données. Le reste est estimé.

---

**Algorithm 1** Calcul d'une sortie modèle pour le protocole *Large map*. Les paramètres sont les coefficients du ciel  $\mathbf{x}$ , la composante basse-fréquence  $\mathbf{c}$ , la partie convolutive du modèle  $\mathbf{H}_c$ , la matrice de pointage  $\mathbf{P}$  et les périodes par balayage  $\mathbf{I}$  d'observation de  $\mathbf{c}$ .

---

```

1: function DIRECTMODEL( $\mathbf{x}, \mathbf{c}, \mathbf{H}_c, \mathbf{P}, \mathbf{I}$ )
2:    $\mathbf{x}_c \leftarrow 0$ 

3:   for  $k \in \{0, \dots, K-1\}$  do
4:     for  $v \in \{0, 1, 2, 3\}$  do
5:        $\mathbf{x}_c[v] \leftarrow \mathbf{x}_c[v] + \mathbf{x}[k] \star \mathbf{H}_c[k, v]$  ▷ Somme des ordres de Taylor
6:     end for
7:   end for

8:   for  $s \in \{0, \dots, S\}$  do ▷ Pour chaque balayage
9:      $v \leftarrow \text{IDENTIFYSPEED}(s)$  ▷ Identification de la vitesse  $v$  associée au balayage  $s$ 
10:     $\mathbf{y}[s] \leftarrow \mathbf{P}[s]\mathbf{x}[v]$  ▷ Extraction et arrangement des coefficients pointés
11:     $\mathbf{y}[s] \leftarrow \mathbf{y}[s] + \mathbf{I}[s]\mathbf{c}$  ▷ Ajout de la composante basse-fréquence
12:  end for

13:  return  $\mathbf{y}$ 
14: end function

```

---

**Algorithm 2** Calcul du transposé du modèle pour le protocole *Large map*. Les paramètres sont l'erreur de reproduction des données  $\mathbf{e}$ , la partie convolutive du modèle  $\mathbf{H}_c$ , la matrice de pointage  $\mathbf{P}$  et les périodes par balayage  $\mathbf{I}$  d'observation de  $\mathbf{c}$ .

---

```

1: function TRANSPPOSEMODEL( $\mathbf{e}, \mathbf{c}, \mathbf{H}_c, \mathbf{P}, \mathbf{I}$ )
2:    $\mathbf{x}_c \leftarrow 0$ 

3:   for  $s \in \{0, \dots, S\}$  do ▷ Pour chaque balayage
4:      $v \leftarrow \text{IDENTIFYSPEED}(s)$ 
5:      $\mathbf{x}_c[v] \leftarrow \mathbf{x}_c[v] + \mathbf{P}^t[s](\mathbf{e}[s] - \mathbf{I}[s]\mathbf{c})$  ▷ Retroprojection des données sur la grille du ciel
6:   end for

7:   for  $v \in \{0, 1, 2, 3\}$  do
8:     for  $k \in \{0, \dots, K-1\}$  do
9:        $\mathbf{x}[k] \leftarrow \mathbf{x}[k] + \mathbf{x}_c[v] \star \mathbf{H}_c^t[k, v]$  ▷ Somme des vitesses pour chaque ordre
10:    end for
11:  end for

12:  return  $\mathbf{x}$ 
13: end function

```

---

**Algorithm 3** Algorithme d'optimisation pour  $\mathbf{x}$ .

---

```

1: function MINIMIZEX( $\mathbf{y}, \mathbf{x}^{(0)}, \mathbf{c}, \mathbf{H}_c, \mathbf{P}, \mathbf{I}, \mathbf{D}, \gamma_b, \gamma_x, \mathbf{m}_x$ )
2:    $q \leftarrow 0$ 
3:    $\mathbf{g}_x^{(0)} \leftarrow \mathbf{0}$ 
4:   repeat

5:      $\mathbf{e} \leftarrow \text{DIRECTMODEL}(\mathbf{x}^{(q)}, \mathbf{c}, \mathbf{H}_c, \mathbf{P}, \mathbf{I}) - \mathbf{I}\mathbf{y}$ 
6:      $\mathbf{g}_x \leftarrow 2\gamma_b \text{TRANSPOSEMODEL}(\mathbf{e}, \mathbf{c}, \mathbf{H}_c, \mathbf{P}, \mathbf{I})$  ▷ Adequation aux données
7:     for  $k \in \{0, \dots, K\}$  do ▷ Adequation à l'a priori
8:        $\mathbf{g}_x[k] \leftarrow \mathbf{g}_x[k] + 2\gamma_x[k]\mathbf{D}[k] \star (\mathbf{x}[k] - \mathbf{m}_x[k])$ 
9:     end for

10:     $\mathbf{g}_x^{(q+1)} \leftarrow \text{FINDDIRECTION}(\mathbf{g}_x^{(q-1)}, \mathbf{g}_x^{(q)})$  ▷ Gradient conjugué par exemple
11:     $\eta \leftarrow \text{FINDSTEP}(\mathbf{x}^{(q)}, \mathbf{c}^{(q)}, \mathbf{d}^{(q+1)})$ 

12:     $\mathbf{x}^{(q+1)} \leftarrow \mathbf{x}^{(q)} - \eta\mathbf{g}_x^{(q+1)}$ 
13:     $q \leftarrow q + 1$ 
14:  until Some criterion is meet
15:  return  $\mathbf{x}^{(q+1)}$ 
16: end function

```

---

**Algorithm 4** Algorithme d'optimisation pour  $\mathbf{c}$ .

---

```

Require:  $\mathbf{o} = \mathbf{P}\mathbf{H}_c\mathbf{x}$  ▷ Sortie modèle sans  $\mathbf{c}$ 
1: function MINIMIZEC( $\mathbf{y}, \mathbf{o}, \mathbf{c}^{(0)}, \mathbf{I}, \Sigma_c, \gamma_b, \gamma_c, \mathbf{m}_c$ )
2:    $q \leftarrow 0$ 
3:    $\mathbf{g}_c^{(0)} \leftarrow \mathbf{0}$ 
4:   repeat

5:     for  $s \in \{0, \dots, S\}$  do ▷ Pour chaque balayage
6:        $\mathbf{e}[s] \leftarrow \mathbf{o}[s] + \mathbf{I}[s]\mathbf{c} - \mathbf{I}[s]\mathbf{y}$  ▷ Adéquation aux données
7:     end for
8:      $\mathbf{g}_c \leftarrow 2\gamma_b \left( \mathbf{I}^t \mathbf{e} + \|\mathbf{c}^{(q)} - \mathbf{y}[B]\|^2 + \|\mathbf{c}^{(q)} - \mathbf{y}[B+1]\|^2 \right)$  ▷ Avec les bolomètres aveugles

9:      $\mathbf{g}_c \leftarrow \mathbf{g}_c + 2\gamma_c \Sigma_c^{-1}(\mathbf{c} - \mathbf{m}_c)$  ▷ Adéquation à l'a priori

10:     $\mathbf{g}_c^{(q+1)} \leftarrow \text{FINDDIRECTION}(\mathbf{g}_c^{(q-1)}, \mathbf{g}_c^{(q)})$  ▷ Gradient conjugué par exemple
11:     $\eta \leftarrow \text{FINDSTEP}(\mathbf{c}^{(q)}, \mathbf{g}_c^{(q+1)})$ 

12:     $\mathbf{c}^{(q+1)} \leftarrow \mathbf{c}^{(q)} - \eta\mathbf{g}_c^{(q+1)}$ 
13:     $q \leftarrow q + 1$ 
14:  until Some criterion is meet
15:  return  $\mathbf{c}^{(q+1)}$ 
16: end function

```

---

---

---

**Algorithm 5** Échantillonnage du ciel  $\mathbf{x}$  par optimisation.

---

```
1: function SKYRND( $\mathbf{y}, \mathbf{x}^{(0)}, \mathbf{c}, \mathbf{H}_c, \mathbf{P}, \mathbf{I}, \mathbf{D}, \gamma_b, \gamma_x$ )
2:   for  $b \in \{0, \dots, B + 1\}$  do ▷ Perturbation des données pour chaque bolomètre
3:      $\tilde{\mathbf{y}}[b] \leftarrow \mathbf{y}[b] + \gamma_b[b]^{-1/2} \text{RANDN}(\text{size}(\mathbf{y}[b]))$ 
4:   end for

5:   for  $k \in \{0, \dots, K\}$  do ▷ Perturbation du ciel pour chaque ordre
6:      $\mathbf{m}_x[k] \leftarrow \gamma_x[k]^{-1/2} \mathbf{D}[k]^{-1/2} \star \text{RANDN}(\text{size}(\mathbf{x}[k]))$  ▷  $\mathbf{D}$  ne doit pas contenir de zeros
7:   end for

8:    $\mathbf{x} \leftarrow \text{MINIMIZEX}(\tilde{\mathbf{y}}, \mathbf{x}^{(0)}, \mathbf{c}, \mathbf{H}_c, \mathbf{P}, \mathbf{I}, \mathbf{D}, \gamma_b, \gamma_x, \mathbf{m}_x)$  ▷ Échantillonnage par optimisation
9:   return  $\mathbf{x}$ 
10: end function
```

---

---

---

**Algorithm 6** Échantillonnage de  $\mathbf{c}$  par optimisation.

---

**Require:**  $\mathbf{o} = \mathbf{P}\mathbf{H}_c\mathbf{x}$

```
1: function COMRND( $\mathbf{y}, \mathbf{o}, \mathbf{c}^{(0)}, \mathbf{I}, \Sigma_c, \gamma_b, \gamma_c$ )
2:   for  $b \in \{0, \dots, B + 1\}$  do ▷ Perturbation des données pour chaque bolomètre
3:      $\tilde{\mathbf{y}}[b] \leftarrow \mathbf{y}[b] + \gamma_b[b]^{-1/2} \text{RANDN}(\text{size}(\mathbf{y}[b]))$ 
4:   end for

5:    $\mathbf{m}_c \leftarrow \gamma_c^{-1/2} \Sigma_c^{1/2} \star \text{RANDN}(\text{size}(\mathbf{c}))$ 

6:    $\mathbf{c} \leftarrow \text{MINIMIZEC}(\tilde{\mathbf{y}}, \mathbf{o}, \mathbf{c}^{(0)}, \mathbf{I}, \Sigma_c, \gamma_b, \gamma_c, \mathbf{m}_c)$  ▷ Échantillonnage par optimisation
7:   return  $\mathbf{c}$ 
8: end function
```

---

**Algorithm 7** Échantillonnage de  $\theta_i$  par Metropolis-Hastings.**Require:**  $P, I$ 

- 1: **function** MODELRND( $\theta_i, \mathbf{x}, \mathbf{c}, \mathbf{y}, \gamma_b$ )
- 2:      $\mathbf{H}_c \leftarrow \text{MODEL}(\theta_i)$  ▷ Voir équation (2.55) page 70
- 3:      $\mathbf{m}_y \leftarrow \text{DIRECTMODEL}(\mathbf{x}, \mathbf{c}, \mathbf{H}_c, P, I)$
  
- 4:      $\theta_i^p \sim p(\theta_i)$  ▷ Proposition
- 5:      $\mathbf{H}_c^p \leftarrow \text{MODEL}(\theta_i^{(p)})$
- 6:      $\mathbf{m}_y^p \leftarrow \text{DIRECTMODEL}(\mathbf{x}, \mathbf{c}, \mathbf{H}_c^p, P, I)$
  
- 7:      $P \leftarrow \exp \left[ \sum_{b=0}^{B-1} -\gamma_{b_b} \|\mathbf{y}_b - \mathbf{m}_{y_b}^p\|^2/2 + \sum_{b=0}^{B-1} -\gamma_{b_b} \|\mathbf{y}_b - \mathbf{m}_{y_b}\|^2/2 \right]$
  
- 8:      $u \sim \mathcal{U}_{[0, 1]}(u)$
- 9:     **if**  $u < \min\{1, P\}$  **then** ▷ Acceptation / conservation
- 10:         **return**  $\theta_i^p$
- 11:     **else**
- 12:         **return**  $\theta_i$
- 13:     **end if**
- 14: **end function**

**Algorithm 8** Échantillonnage de  $f_c$  et  $d$  par Metropolis-Hastings.

- 1: **function** CORRELRND( $f_c, d, \mathbf{c}, \gamma_c$ )
- 2:      $\Sigma_c \leftarrow \text{CORREL}(f_c, d)$  ▷ Voir équation (3.6) page 76
  
- 3:      $f_c^p \sim p(f_c)$  ▷ Proposition
- 4:      $d^p \sim p(d)$
- 5:      $\Sigma_c^p \leftarrow \text{CORREL}(f_c^{(q+1)}, d^{(q+1)})$
  
- 6:      $P \leftarrow \det[\Sigma_c^p]^{-1/2} \det[\Sigma_c^{(q)}]^{1/2} \exp \left[ -\frac{1}{2} \gamma_c \mathbf{c}^t \left( (\Sigma_c^p)^{-1} - \Sigma_c^{-1} \right) \mathbf{c} \right]$
  
- 7:      $u \sim \mathcal{U}_{[0, 1]}(u)$
- 8:     **if**  $u < \min\{1, P\}$  **then** ▷ Acceptation / conservation
- 9:         **return**  $[f_c, d]^p$
- 10:     **else**
- 11:         **return**  $[f_c, d]$
- 12:     **end if**
- 13: **end function**

---

**Algorithm 9** Échantillonneur de Gibbs. L'algorithme nécessite ici une initialisation pour  $\mathbf{x}$  et  $\mathbf{c}$ . Ces initialisations sont utilisées pour l'algorithme d'optimisation. On peut remarquer que ce qui est nécessaire maintenant se résume à l'information de pointage, la régularisation et les données. Le reste est estimé.

---

**Require:**  $\mathbf{P}, \mathbf{I}, \mathbf{D}, \mathbf{x}_0, \mathbf{c}_0$

```

1: procedure GIBBSAMPLER( $\mathbf{y}, \mathbf{x}^{(0)}, \mathbf{c}^{(0)} \mathbf{P}, \mathbf{I}, \mathbf{D}$ )
2:    $q \leftarrow 0$ 
3:   INITIALIZE( $\boldsymbol{\theta}_i^{(q)}, f_c^{(q)}, d^{(q)}$ )
4:    $\mathbf{H}_c \leftarrow \text{MODEL}(\boldsymbol{\theta}_i^{(q)})$  ▷ Voir équation (2.55) page 70
5:    $\mathbf{x}^{(q)} \leftarrow \mathbf{x}_0$ 
6:    $\mathbf{c}^{(q)} \leftarrow \mathbf{c}_0$ 
7:    $\boldsymbol{\Sigma}_c \leftarrow \text{CORREL}(f_c^{(q)}, d^{(q)})$  ▷ Voir équation (3.6) page 76
8:   repeat
9:     for  $k \in \{0, \dots, K\}$  do
10:       $\gamma_{\mathbf{x}}^{(q)}[k] \leftarrow \text{GAMRND}(\alpha_{\mathbf{x}} + (IJ - 1)/2, (\beta_{\mathbf{x}}^{-1} + \mathbf{x}^{(q)}[k] \odot (\mathbf{D}[k] \star \mathbf{x}^{(q)}[k]) / 2)^{-1})$ 
11:    end for
12:     $\gamma_{\mathbf{c}}^{(q)} \leftarrow \text{GAMRND}(\alpha_{\mathbf{c}} + N/2, (\beta_{\mathbf{c}}^{-1} + \mathbf{c}^{(q)} \odot (\boldsymbol{\Sigma}_{\mathbf{c}}^{-1} \star \mathbf{c}^{(q)}))^{-1})$ 
13:     $\mathbf{o} \leftarrow \text{DIRECTMODEL}(\mathbf{x}^{(q)}, \mathbf{c}^{(q)}, \mathbf{H}_c, \mathbf{P}, \mathbf{I})$ 
14:    for  $b \in \{0, \dots, B - 1\}$  do
15:       $\gamma_b^{(q)}[b] \leftarrow \text{GAMRND}(\alpha_b + N/2, (\beta_b^{-1} + \|\mathbf{y}[b] - \mathbf{o}[b]\|^2)^{-1})$ 
16:    end for
17:    for  $b \in \{B, B + 1\}$  do ▷ Bolomètres aveugles
18:       $\gamma_b^{(q)}[b] \leftarrow \text{GAMRND}(\alpha_b + N/2, (\beta_b^{-1} + \|\mathbf{y}[b] - \mathbf{c}^{(q)}\|^2)^{-1})$ 
19:    end for
20:     $\mathbf{x}^{(q)} \leftarrow \text{SKYRND}(\mathbf{y}, \mathbf{x}_0, \mathbf{c}^{(q)}, \mathbf{H}_c, \mathbf{P}, \mathbf{I}, \mathbf{D}, \gamma_b^{(q)}, \gamma_{\mathbf{x}}^{(q)})$ 
21:     $\mathbf{o} \leftarrow \mathbf{P} \mathbf{H}_c \mathbf{x}^{(q)}$ 
22:     $\mathbf{c}^{(q)} \leftarrow \text{COMRND}(\mathbf{y}, \mathbf{o}, \mathbf{c}_0, \mathbf{I}, \boldsymbol{\Sigma}_c, \gamma_b^{(q)}, \gamma_{\mathbf{c}}^{(q)})$ 
23:     $\boldsymbol{\theta}_i^{(q+1)} \leftarrow \text{MODEL RND}(\boldsymbol{\theta}_i^{(q)}, \mathbf{x}^{(q)}, \mathbf{c}^{(q)}, \mathbf{y}, \gamma_b^{(q)})$ 
24:     $\mathbf{H}_c \leftarrow \text{MODEL}(\boldsymbol{\theta}_i^{(q+1)})$ 
25:     $[f_c, d]^{(q+1)} \leftarrow \text{CORREL RND}(f_c^{(q)}, d^{(q)}, \mathbf{c}, \gamma_{\mathbf{c}})$ 
26:     $\boldsymbol{\Sigma}_c \leftarrow \text{CORREL}(f_c^{(q+1)}, d^{(q+1)})$ 
27:     $\bar{\mathbf{x}}^{(q)} \leftarrow \frac{1}{q+1} \sum_{i=0}^q \mathbf{x}^{(i)}$  ▷ Moyenne empirique
28:     $q \leftarrow q + 1$ 
29:    until Some criterion is meet ▷ Par exemple  $\|\bar{\mathbf{x}}^{(q)} - \bar{\mathbf{x}}^{(q-1)}\| / |\bar{\mathbf{x}}^{(q)}| \geq \epsilon$ 
30: end procedure

```

---



## Annexe F

# Déconvolution myope non-supervisée

Cette annexe est un projet de publication. Les travaux présentés concernent la déconvolution d'image non-supervisée myope (avec estimation des hyper-paramètres et des paramètres instruments) avec une approche globale. Le modèle cette fois est totalement invariant (convolutif). La réponse impulsionnelle est une gaussienne pilotée par trois paramètres considérés mal-connus.

La démarche est faite dans le formalisme bayésien. L'estimation repose sur une loi *a posteriori* étendue sur l'image, les paramètres des lois et les paramètres instruments. La loi est explorée à l'aide d'un échantillonneur de Gibbs permettant d'approcher la loi *a posteriori*. En particulier on calcule la moyenne *a posteriori* comme estimateur des paramètres.

La vraisemblance et l'*a priori* étant quadratiques tous les calculs nécessaires à l'algorithme peuvent être menés dans l'espace de Fourier. Le calcul est donc très rapide.

# Unsupervised and Myopic Deconvolution Using a Full Bayesian Approach

F. Orieux, J.-F. Giovannelli and T. Rodet

**Abstract**—This paper proposes a Bayesian approach for unsupervised and myopic image deconvolution. The parameters of the instrument response and the parameters of the probabilistic model (or regularization parameters in the energetic point of view) are unknown and estimated with the image. With a Bayesian approach, the solution is found by inferring on a global *a posteriori* law for unknown object and parameters. Work is done on the image prior law to facilitate posterior law manipulation. Study of the posterior law degeneracy is also addressed. The estimate is chosen in the sense of the posterior mean, numerically calculated by means of a Monte-Carlo Markov chain algorithm. The computation is efficiently done in Fourier space and the practicability of the method is shown on simulated examples. Results show an image estimate with restoration of high-frequencies as well as correct estimations for the parameters. Our method provides a solution to the problem of hyperparameters and instrument parameters estimation in deconvolution with a global and coherent approach.

**Index Terms**—Hyperparameter estimation, myopic estimation, unsupervised estimation, Bayesian methods, quadratic penalization, Monte-Carlo Markov chain.

## I. INTRODUCTION

**D**ECONVOLUTION is an active research field [1–3]. Examples of application are medical imaging, astronomy, nondestructive testing and more generally imagery problems. In these applications, the degradation introduced by the instrument limits the data interpretation while the need in resolution can be important. For example this is particularly critical for long-wavelength astronomy (*e.g.*, [4]). In addition, the cost of instruments can be high and investments in appropriate data processing methods are decisive. Even for poor performance and cheap instruments, restoration method can be used to bypass their limitations. In other word well adapted numerical methods to restore images have a significant impact.

The deconvolution problem is ill-posed and a well-known solution relies on the introduction of information in addition to the data and the acquisition model [5, 6]. Since information is added, the methods are specific to a given class of images. For these reasons, this paper is dedicated to the restoration of smooth images encountered in application such as [4, 7].

In the Bayesian approach [1], information is introduced by means of probabilistic models [8]. Once the models for data and object are defined, the next step is the construction of the *a posteriori* law *i.e.*, given the measured data. The solution is then defined as a representative point of the law and the two most classical are (1) the maximum, which leads to an

optimisation problem, and (2) the mean, which leads to an integration problem. Anyway, the resulting estimate depends on two sets of variables in addition to the data.

- Firstly, the solution depends on the parameters of the probability laws named hyperparameters (mean, variance, parameters of correlation matrix,...). These parameters define the shape of the laws but also the compromise, like the regularisation parameter, between the information provided by the *a priori* and the information provided by the data. In real experiments, their values are unknown and in this case the problems is named unsupervised.
- Secondly, the estimate naturally depends on the instrument response model (the point spread function or PSF). In case of an unknown response parametrized by only a few parameters the problem is generally named myopic. In the non-parametric case the problem is named blind. A book about these problems has been recently published [3]. The main difficulty here is an ambiguity problem between the transfer function and the image spectrum, since the data spectrum is the product of both of them. In order to resolve the ambiguity problem prior information must be added [3, 9].

For both families of parameters, two approaches exist. In the first one, the parameter values are empirically tuned or estimated in a preliminary step (with Maximum Likelihood [7] or calibration [10] for example), then the values are used in a second image estimation step. In the second one the parameters are jointly estimated with the object [2, 9].

In Bayesian context, Likas *et al.* [9] address the blind deconvolution problem. To resolve the ambiguity problem, a smoothness feature with Gaussian law is introduced for the instrument model. They consider a non-parametric response which provides a Gaussian conditional law for the instrument model given the other parameters. The inference is done with a variational algorithm [11] approximating the true posterior law in the sense of Kullback-Leibler divergence with a separable law between the object and the instrument response. Molina *et al.* [2] have extended this approach with a hierarchical Bayesian model to estimate the hyperparameters. For the myopic problem, Jalobeanu *et al.* [12] address the case where the instrument response is a symmetric Gaussian. The width parameter and the noise variance are estimated in a preliminary step with Maximum-Likelihood.

In the present paper we address the myopic and unsupervised deconvolution problem. We propose a new method that jointly estimates the image, the prior law parameters and the instrument parameters. This is done in a coherent and global way with an extended *a posteriori* law for all unknown

F. Orieux, J.-F. Giovannelli and T. Rodet are with the Laboratoire des Signaux et Systèmes (CNRS – SUPELEC – Univ. Paris-Sud 11), SUPELEC, Plateau de Moulon, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France. E-mail: orieux, giova, rodet@lss.supelec.fr.

variables. The posterior law is obtained with the Bayes rules, based on *a priori* laws: Gaussian for the image and noise, gamma or Jeffreys for the hyperparameters and uniform for the instrument parameters.

We have worked on the image prior law to facilitate joint law manipulation such as integration, conditioning or hyperparameter estimation. The degeneracy of the *a posteriori* law in some limit cases is also studied. Moreover, the myopic case is addressed with a physical study providing a parametric expression for the response (with parameters such as width or orientation).

The estimate is chosen as the mean of the posterior law and is computed using Monte-Carlo simulation. To obtain samples, Monte-Carlo Markov chain (MCMC) algorithms permit to simulate the *a posteriori* law despite of its complexity. For more information on MCMC methods, the reader may consult reference books such as [13].

This paper is presented in the following manner. Sec. II presents the notation and states the problem. The three following sections describe our methodology. Firstly the Bayesian probabilistic model for the parameters is described in details in Sec. III. Then a proper posterior law is established in Sec. IV. To compute the estimate, the MCMC algorithm is described in Sec. V and numerical results are shown in Sec. VI. Finally, Sec. VII is devoted to the conclusion and the perspectives.

## II. NOTATIONS AND CONVOLUTION MODEL

We consider  $N$  pixels real square images represented in lexicographic order by vector  $\mathbf{x} \in \mathbb{R}^N$ , with generic elements  $x_n$ . The forward model is written

$$\mathbf{y} = \mathbf{H}_w \mathbf{x} + \mathbf{n} \quad (1)$$

where  $\mathbf{y} \in \mathbb{R}^N$  are the data,  $\mathbf{H}_w$  a convolution matrix parametrized by  $w$ ,  $\mathbf{x}$  the image of interest and  $\mathbf{n}$  the model errors. The parameters, such as width or orientation, of the parametric model (*e.g.*, Gaussian) are collected in vector  $w \in \mathbb{R}^P$ . The matrix  $\mathbf{H}_w$  is modeled as block-circulant circulant-block (BCCB) for computational efficiency of the convolution in Fourier space.

The diagonalization of  $\mathbf{H}_w$  is written  $\Lambda_H = \mathbf{F} \mathbf{H}_w \mathbf{F}^\dagger$  where  $\Lambda_H$  is a diagonal matrix [14],  $\mathbf{F}$  the unitary Fourier matrix and  $\dagger$  the transpose conjugate symbol. The convolution, in Fourier space, is written

$$\hat{\mathbf{y}} = \Lambda_H \hat{\mathbf{x}} + \hat{\mathbf{n}} \quad (2)$$

where  $\hat{\mathbf{x}} = \mathbf{F} \mathbf{x}$ ,  $\hat{\mathbf{y}} = \mathbf{F} \mathbf{y}$  and  $\hat{\mathbf{n}} = \mathbf{F} \mathbf{n}$  are the 2D discrete Fourier transform (DFT-2D) of image, data and noise, respectively. The notation  $\hat{\mathbf{x}}_* \in \mathbb{R}^{N-1}$  is the vector without its first coefficient, the null frequency.

Since  $\Lambda_H$  is diagonal, the convolution is computed with a term-wise product in Fourier space. There is a strict equivalence between a description in spatial and Fourier spaces. Consequently, for coherent description and computational efficiency, all development are done in Fourier space.

## III. BAYESIAN PROBABILISTIC MODEL

This section presents the prior law for each set of parameters. In order to account for smoothness, the image law introduces penalization of high-frequency through a difference operator on the pixel. Conjugate law for the hyperparameters and uniform law for the instrument parameters are considered.

Special attention is paid to the image law. The correlation matrix of the image law may be singular and, in this case, leads to a degenerated prior law<sup>1</sup>. We present a parametrization helpful to study the problem. Thanks to this parametrization, Sec. IV studies the degeneracy of the posterior in relation with the prior law.

### A. Image prior law

The probability law for the image is a toroidal Gaussian field  $p(\mathbf{x}|\gamma) \sim \mathcal{N}(0, \mathbf{R})$  parametrized by vector  $\gamma$ . In the Fourier space, the probability law is also Gaussian (see Appendix A) and writes

$$p(\hat{\mathbf{x}}|\gamma) = (2\pi)^{-N/2} \det[\Lambda_{\mathbf{R}}^{-1}]^{1/2} \exp \left[ -\frac{1}{2} \hat{\mathbf{x}}^\dagger \Lambda_{\mathbf{R}}^{-1} \hat{\mathbf{x}} \right] \quad (3)$$

where the diagonal covariance matrix  $\Lambda_{\mathbf{R}} = \mathbf{F} \mathbf{R} \mathbf{F}^\dagger$  depends on the parameters  $\gamma$ . This paper focus on smooth images, thus on positive correlation between pixels introduced by penalization of the high-frequencies. It is obtained with a circulant difference operator  $\mathbf{D}$  and its diagonalization  $\Lambda_{\mathbf{D}} = \mathbf{F} \mathbf{D} \mathbf{F}^\dagger$ . The operator is build with a high-pass filter, the Laplacian for example. Then with an inverse variance, named precision and noted  $\gamma_1 > 0$ , the precision matrix  $\Lambda_{\mathbf{R}}^{-1}$  writes

$$\gamma_1 \Lambda_{\mathbf{D}}^\dagger \Lambda_{\mathbf{D}} = \text{diag} \left( 0, \gamma_1 |\hat{d}_1|^2, \dots, \gamma_1 |\hat{d}_{N-1}|^2 \right) \quad (4)$$

where  $\text{diag}$  builds a diagonal matrix from a vector and  $\hat{d}_n$  is the  $n$ -th DFT-2D coefficient of  $\mathbf{D}$ .

With this first parametrization this matrix is not invertible and its first eigenvalue is null. No information is introduced about the average of the image *i.e.*, the null frequency  $\hat{x}_0$ . This expression for  $\Lambda_{\mathbf{R}}^{-1}$  leads to a degenerated prior law for  $\mathbf{x}$  with a null determinant  $\det[\gamma_1 \Lambda_{\mathbf{D}}^\dagger \Lambda_{\mathbf{D}}] = 0$ . Unfortunately, a degenerated prior law results in a degenerated joint law and possibly a degenerated posterior law.

To manage this difficulty, several approaches have been proposed. Some authors [2] still use this degenerated prior law for the posterior law and the inference. By removing the first coordinate with  $\Lambda_{\mathbf{D}*}$ , it is possible to use approximately the same prior law for the Fourier coefficients except the null frequency  $\hat{x}_0$  [15]. In this case, no information is introduced about this coordinate and this leads to a law written

$$p(\hat{\mathbf{x}}_*|\gamma_1) = (2\pi)^{-(N-1)/2} \prod_{n=1}^{N-1} |\hat{d}_n|^2 \gamma_1^{(N-1)/2} \exp \left[ -\frac{\gamma_1}{2} \|\Lambda_{\mathbf{D}*} \hat{\mathbf{x}}_*\|^2 \right] \quad (5)$$

where  $\Lambda_{\mathbf{D}*}$  is  $\Lambda_{\mathbf{D}}$  without it first line and column. This *a priori* is not degenerated because it is defined without  $\hat{x}_0$ . Consequently, the Bayes rule is not applied to this coordinate

<sup>1</sup>A  $p(\mathbf{x})$  law is degenerated if  $p(\mathbf{x}) = 0$  for all  $\mathbf{x}$ . The law is improper if  $\int p(\mathbf{x}) d\mathbf{x} = +\infty$ . The law is proper if  $\int p(\mathbf{x}) d\mathbf{x} = 1$ .

since there is no prior law for it. In the Bayesian point of view, the null frequency is a known parameter, as data, and the *a posteriori* law writes

$$p(\hat{\mathbf{x}}_*, \gamma, \mathbf{w} | \hat{\mathbf{y}}, \hat{\mathbf{x}}_0) = \frac{p(\hat{\mathbf{y}}, \hat{\mathbf{x}}_*, \gamma, \mathbf{w} | \hat{\mathbf{x}}_0)}{p(\hat{\mathbf{y}})}. \quad (6)$$

Another standard possibility to manage this difficulty relies on the energy of  $\mathbf{x}$ . The additional term  $\gamma_0 \mathbf{I}$ , with  $\gamma_0 > 0$  as parameter in the precision matrix [16, 17], introduces information on all frequencies including  $\hat{\mathbf{x}}_0$ . The precision matrix writes

$$\begin{aligned} \Lambda_{\mathbf{R}}^{-1} &= \gamma_0 \mathbf{I} + \gamma_1 \Lambda_{\mathbf{D}}^\dagger \Lambda_{\mathbf{D}} \\ &= \text{diag} \left( \gamma_0, \gamma_0 + \gamma_1 |\hat{d}_1|^2, \dots, \gamma_0 + \gamma_1 |\hat{d}_{N-1}|^2 \right) \end{aligned} \quad (7)$$

with a determinant

$$\det[\Lambda_{\mathbf{R}}^{-1}] = \prod_{n=0}^{N-1} \left( \gamma_0 + \gamma_1 |\hat{d}_n|^2 \right). \quad (8)$$

The obtained Gaussian *a priori* law is not degenerated and undoubtedly leads a proper *a posteriori* law. As a consequence the non-null frequencies  $\hat{\mathbf{x}}_*$  are controlled by two parameters  $\gamma_0$  and  $\gamma_1$

$$p(\hat{\mathbf{x}} | \gamma_0, \gamma_1) = p(\hat{\mathbf{x}}_0 | \gamma_0) p(\hat{\mathbf{x}}_* | \gamma_0, \gamma_1). \quad (9)$$

Moreover, the determinant (8) is not separable between  $\gamma_0$  and  $\gamma_1$ . Consequently, the conditional posterior law for these parameters is not a classical law and future development will be more difficult.

Our approach to manage the degeneracy difficulty relies on the addition of a term for the null frequency only

$$\begin{aligned} \gamma_0 \Lambda_{\mathbb{1}}^\dagger \Lambda_{\mathbb{1}} &= \text{diag} \left( \gamma_0 |\hat{\mathbb{1}}_1|^2, \gamma_0 |\hat{\mathbb{1}}_2|^2, \dots, \gamma_0 |\hat{\mathbb{1}}_{N-1}|^2 \right) \\ &= \text{diag} (\gamma_0, 0, \dots, 0) \end{aligned} \quad (10)$$

where  $\Lambda_{\mathbb{1}} = \mathbf{F} \mathbb{1} \mathbf{F}^\dagger$  is a matrix with only one non-null coefficient at null frequency. The average operator  $\mathbb{1}$  is a matrix of  $1/N$  elements with  $\hat{\mathbb{1}}_n$  the DFT-2D coefficients and  $\hat{\mathbb{1}}_0 = 1$ . By adding (4) with (10) we obtain the precision matrix

$$\begin{aligned} \Lambda_{\mathbf{R}}^{-1} &= \text{diag} \left( \gamma_0, \gamma_1 |\hat{d}_1|^2, \dots, \gamma_1 |\hat{d}_{N-1}|^2 \right) \\ &= \gamma_0 \Lambda_{\mathbb{1}}^\dagger \Lambda_{\mathbb{1}} + \gamma_1 \Lambda_{\mathbf{D}}^\dagger \Lambda_{\mathbf{D}}. \end{aligned} \quad (11)$$

The determinant has a separable expression

$$\begin{aligned} \det[\Lambda_{\mathbf{R}}^{-1}] &= \prod_{n=0}^{N-1} \left( \gamma_0 |\hat{\mathbb{1}}_n|^2 + \gamma_1 |\hat{d}_n|^2 \right) \\ &= \gamma_0 \gamma_1^{N-1} \prod_{n=1}^{N-1} |\hat{d}_n|^2. \end{aligned} \quad (12)$$

The precisions parameter have been factorized. In addition, the two terms control different frequencies:  $\gamma_0$  the average and  $\gamma_1$  the smoothness. Setting  $\gamma_0$  or  $\gamma_1$  to 0 leads again to a

degenerated prior law. With the correlation expression in (11), we have the non-degenerated explicit prior law for the image

$$\begin{aligned} p(\hat{\mathbf{x}} | \gamma_0, \gamma_1) &= p(\hat{\mathbf{x}}_0 | \gamma_0) p(\hat{\mathbf{x}}_* | \gamma_1) \\ &= (2\pi)^{-N/2} \prod_{n=1}^{N-1} |\hat{d}_n| \gamma_0^{1/2} \gamma_1^{(N-1)/2} \\ &\quad \exp \left[ -\frac{\gamma_0}{2} \|\hat{\mathbf{x}}_0\|^2 - \frac{\gamma_1}{2} \|\Lambda_{\mathbf{D}} \hat{\mathbf{x}}_*\|^2 \right]. \end{aligned} \quad (13)$$

Finally, we have the *a priori* law for the image (13) above. The next step is to explicit the *a priori* law for the noise and the other parameters, including the law parameters  $\gamma$  and the instrument parameters  $\mathbf{w}$ .

### B. Noise and data laws

The noise is modeled as white zero-mean Gaussian with unknown precision parameter  $\gamma_n$

$$p(\hat{\mathbf{n}} | \gamma_n) = (2\pi)^{-N/2} \gamma_n^{N/2} \exp \left[ -\frac{\gamma_n}{2} \|\hat{\mathbf{n}}\|^2 \right]. \quad (14)$$

Consequently the likelihood of the parameters given the data writes

$$p(\hat{\mathbf{y}} | \hat{\mathbf{x}}, \gamma_n, \mathbf{w}) = (2\pi)^{-N/2} \gamma_n^{N/2} \exp \left[ -\frac{\gamma_n}{2} \|\hat{\mathbf{y}} - \Lambda_{\mathbf{H}} \hat{\mathbf{x}}\|^2 \right] \quad (15)$$

with a spatial equivalent description

$$p(\mathbf{y} | \mathbf{x}, \gamma_n, \mathbf{w}) = (2\pi)^{-N/2} \gamma_n^{N/2} \exp \left[ -\frac{\gamma_n}{2} \|\mathbf{y} - \mathbf{H}_w \mathbf{x}\|^2 \right]. \quad (16)$$

It depends, of course, on the image  $\hat{\mathbf{x}}$ , on the noise parameter  $\gamma_n$  and instrument parameters  $\mathbf{w}$  embedded in  $\Lambda_{\mathbf{H}}$ , and naturally involves a least square data fitting  $\|\mathbf{y} - \mathbf{H}_w \mathbf{x}\|^2 = \|\hat{\mathbf{y}} - \Lambda_{\mathbf{H}} \hat{\mathbf{x}}\|^2$ .

### C. Hyperparameters law

A classical choice for hyperparameter prior law is conjugate law [18] with computational efficiency justification. With conjugate prior, the conditional posterior law for the hyperparameters is in the same family than its prior law.

The three parameters  $\gamma_0$ ,  $\gamma_1$  and  $\gamma_n$  are precisions of Gaussian laws (13) and (15). A conjugate law for this parameters is the Gamma law parametrized by two values  $(\alpha_i, \beta_i)$ , see Appendix B

$$p(\gamma_i) = \frac{1}{\beta_i^{\alpha_i} \Gamma(\alpha_i)} \gamma_i^{\alpha_i-1} \exp(-\gamma_i/\beta_i) \quad (17)$$

with  $i = 0, 1$  or  $n$ .

In addition to computational efficiency, the goal is also to use non-informative *a priori* law. With specific parameter values, one obtains two improper non-informative prior<sup>2</sup>: the Jeffreys's law  $p(\gamma) = 1/\gamma$  and the uniform law  $p(\gamma) = \mathcal{U}_{[0, +\infty)}(\gamma)$  with  $(0, +\infty)$  and  $(1, +\infty)$ , respectively. Jeffreys's law is a classical law for the precisions and is considered as non-informative [19]. This law is also invariant to transformations of the form  $\eta = \gamma^n$  [19, 20]. For this reasons development are done with the Jeffreys's law.

<sup>2</sup>Alternatively, a non-flat Gamma law can be considered to introduce more information.

#### D. Instrument parameters law

For the instrument parameters  $\mathbf{w}$ , a physical study or the datasheet provides a nominal value with uncertainty in a given interval  $[\mathbf{a} \ \mathbf{b}]$ . The "Principle of Insufficient Reason" [19] leads to a uniform prior on this interval

$$p(\mathbf{w}) = \mathcal{U}_{[\mathbf{a} \ \mathbf{b}]}(\mathbf{w}) = \frac{1}{M} \mathbb{1}_{[\mathbf{a} \ \mathbf{b}]}(\mathbf{w}) \quad (18)$$

with  $\mathbb{1}_{[\mathbf{a} \ \mathbf{b}]}(\mathbf{w}) = 1$ , if  $\mathbf{w} \in [\mathbf{a} \ \mathbf{b}]$ , 0 elsewhere and the normalization constant  $M = \prod_{i=1}^P (b_i - a_i)$ . However, within the proposed framework, the choice is not limited and other laws, such as Gaussian, are possible.

#### IV. PROPER POSTERIOR LAW

At this point the law of each parameter is available: the image, the hyperparameters and the instrument parameters. Thus, the joint law for all the parameters is built by multiplying the likelihood (15) and the *a priori* laws (13), (17) and (18)

$$p(\hat{\mathbf{x}}, \gamma_n, \gamma_0, \gamma_1, \mathbf{w}, \hat{\mathbf{y}}) = p(\hat{\mathbf{y}}|\hat{\mathbf{x}}, \gamma_n, \mathbf{w}) p(\hat{\mathbf{x}}|\gamma_0, \gamma_1) p(\gamma_n) p(\gamma_0) p(\gamma_1) p(\mathbf{w}) \quad (19)$$

and explicitly

$$p(\hat{\mathbf{x}}, \gamma_n, \gamma_0, \gamma_1, \mathbf{w}, \hat{\mathbf{y}}) = \frac{(2\pi)^{-N} \prod_{n=1}^{N-1} |\hat{d}_n|}{M \beta_n^{\alpha_n} \Gamma(\alpha_n) \beta_0^{\alpha_0} \Gamma(\alpha_0) \beta_1^{\alpha_1} \Gamma(\alpha_1)} \gamma_n^{\alpha_n + N/2 - 1} \gamma_0^{\alpha_0 - 1/2} \gamma_1^{\alpha_1 + (N-1)/2 - 1} \exp \left[ -\frac{\gamma_n}{\beta_n} - \frac{\gamma_0}{\beta_0} - \frac{\gamma_1}{\beta_1} \right] \mathbb{1}_{[\mathbf{a} \ \mathbf{b}]}(\mathbf{w}) \exp \left[ -\frac{\gamma_n}{2} \|\hat{\mathbf{y}} - \mathbf{\Lambda}_H \hat{\mathbf{x}}\|^2 - \frac{\gamma_0}{2} \|\hat{\mathbf{x}}_0\|^2 - \frac{\gamma_1}{2} \|\mathbf{\Lambda}_D \hat{\mathbf{x}}\|^2 \right]. \quad (20)$$

Thanks to the Bayes rule, the *a posteriori* law reads

$$p(\hat{\mathbf{x}}, \gamma_n, \gamma_0, \gamma_1, \mathbf{w}|\hat{\mathbf{y}}) = \frac{p(\hat{\mathbf{x}}, \gamma_n, \gamma_0, \gamma_1, \mathbf{w}, \hat{\mathbf{y}})}{p(\hat{\mathbf{y}})} \quad (21)$$

where  $p(\hat{\mathbf{y}})$  is a normalization constant

$$p(\hat{\mathbf{y}}) = \int p(\hat{\mathbf{y}}, \hat{\mathbf{x}}, \gamma, \mathbf{w}) d\hat{\mathbf{x}} d\gamma d\mathbf{w}. \quad (22)$$

As described before, setting  $\gamma_0 = 0$  leads to a degenerated prior and joint laws. However, system generally preserves null frequency and only prior information on the smoothness is available. In this specific case  $\gamma_0$  can be considered as a nuisance parameter.

In Bayesian framework, a solution to get rid of the nuisance parameters is to integrate them out in the *a posteriori* law. Thanks to our parametrization Sec. III-A, the integration of  $\gamma_0$  is the integration of a Gamma law. Application of Appendix B-B on  $\gamma_0$  in the *a posteriori* law (21) provides

$$p(\hat{\mathbf{x}}, \gamma_n, \gamma_1, \mathbf{w}|\hat{\mathbf{y}}) = \frac{p(\hat{\mathbf{x}}_0) p(\hat{\mathbf{y}}, \hat{\mathbf{x}}_*, \gamma_n, \gamma_1, \mathbf{w}|\hat{\mathbf{x}}_0)}{\int p(\hat{\mathbf{x}}_0) p(\hat{\mathbf{y}}, \hat{\mathbf{x}}_*, \gamma_n, \gamma_1, \mathbf{w}|\hat{\mathbf{x}}_0) d\gamma_n d\gamma_1 d\mathbf{w} d\hat{\mathbf{x}}_* d\hat{\mathbf{x}}_0} \quad (23)$$

with

$$p(\hat{\mathbf{x}}_0) = \int p(\hat{\mathbf{x}}_0|\gamma_0) p(\gamma_0) d\gamma_0 = \left( 1 + \frac{\beta_0 \hat{\mathbf{x}}_0^2}{2} \right)^{-\alpha_0 - 1/2}. \quad (24)$$

Now the parameter is integrated, the parameter  $\alpha_0$  and  $\beta_0$  are set to remove the null frequency penalization. Since we have  $\alpha_0 > 0$  and  $\beta_0 > 0$  we get  $(1 + \beta_0 \hat{\mathbf{x}}_0^2/2)^{-\alpha_0 - 1/2} \leq 1$  and the joint law is majored

$$\left( 1 + \frac{\beta_0 \hat{\mathbf{x}}_0^2}{2} \right)^{-\alpha_0 - 1/2} p(\hat{\mathbf{y}}, \hat{\mathbf{x}}_*, \gamma_n, \gamma_1, \mathbf{w}|\hat{\mathbf{x}}_0) \leq p(\hat{\mathbf{y}}, \hat{\mathbf{x}}_*, \gamma_n, \gamma_1, \mathbf{w}|\hat{\mathbf{x}}_0). \quad (25)$$

Consequently, by the dominated convergence theorem [21], the limit of the law with  $\alpha_0 \rightarrow 1$  and  $\beta_0 \rightarrow 0$  can be placed under the integral sign in the denominator. Then the null-frequency penalization  $p(\hat{\mathbf{x}}_0)$  in the numerator and denominator are removed. It is equivalent to integrate the  $\gamma_0$  parameter under a Dirac (see appendix B). The equation is simplified and the integration with respect to  $\hat{\mathbf{x}}_0$  in the denominator (22)

$$\int_{\mathbb{R}} p(\hat{\mathbf{y}}|\hat{\mathbf{x}}, \gamma_n, \mathbf{w}) p(\hat{\mathbf{x}}_*|\gamma_1) p(\gamma_1, \gamma_n, \mathbf{w}) d\hat{\mathbf{x}}_0 \propto \int_{\mathbb{R}} p(\hat{\mathbf{y}}_0|\hat{\mathbf{x}}_0, \gamma_n, \mathbf{w}) d\hat{\mathbf{x}}_0 \propto \int_{\mathbb{R}} \exp \left[ -\frac{\gamma_n}{2} (\hat{\mathbf{y}}_0 - \hat{h}_0 \hat{\mathbf{x}}_0)^2 \right] d\hat{\mathbf{x}}_0 \quad (26)$$

converges if and only if  $\hat{h}_0 \neq 0$ : the null frequency is observed. If this condition is met, (23) with  $\beta_0 = 0$  and  $\alpha_0 = 1$  is a proper posterior law for the image, the precision parameters and the instrument parameters. In other word, if the average is observed, the degeneracy of the *a priori* law is not transmitted to the *a posteriori* law.

Then, the obtained *a posteriori* law writes

$$p(\hat{\mathbf{x}}, \gamma_n, \gamma_1, \mathbf{w}|\hat{\mathbf{y}}) \propto \gamma_n^{\alpha_n + N/2 - 1} \gamma_1^{\alpha_1 + (N-1)/2 - 1} \mathbb{1}_{[\mathbf{a} \ \mathbf{b}]}(\mathbf{w}) \exp \left[ -\frac{\gamma_n}{2} \|\hat{\mathbf{y}} - \mathbf{\Lambda}_H \hat{\mathbf{x}}\|^2 - \frac{\gamma_1}{2} \|\mathbf{\Lambda}_D \hat{\mathbf{x}}\|^2 \right] \exp \left[ -\frac{\gamma_n}{\beta_n} - \frac{\gamma_1}{\beta_1} \right]. \quad (27)$$

Finally, inference is done on this law (27). If the null frequency is not observed, or information must be added, the previous (21) can be used.

#### V. POSTERIOR MEAN ESTIMATOR AND LAW EXPLORATION

This section presents the algorithm to explore the posterior law (21) or (27) and to compute an estimate of the parameters. For this purpose Monte Carlo Markov chain is used to provides samples. Firstly, the obtained samples are used to compute different moments of the law. Afterwards, they are also used to compute approximation of the marginal laws with histograms. This two representations are helpful to analyse the *a posteriori* law, the structure of the available information and the incertitude. For example, in the present paper, histograms are used to illustrate the ambiguity in myopic problem.

Here the samples of the *a posteriori* law are obtained by a Gibbs sampling algorithm [13, 22, 23]. The Gibbs algorithm consists in sampling, iteratively, a conditional posterior law of a set of parameters given all the others parameters obtained at previous iteration. Typically, the sampled laws are the law of  $\hat{\mathbf{x}}$ ,  $\gamma_i$  and  $\mathbf{w}$ . After a burn-in time, the complete set of the samples are under the joint *a posteriori* law. The three next sections present each sampling step.

### A. Sampling the image

The conditional posterior law of the image is a Gaussian law

$$\hat{\mathbf{x}}^{(k+1)} \sim p\left(\hat{\mathbf{x}}|\hat{\mathbf{y}}, \gamma_n^{(k)}, \gamma_0^{(k)}, \gamma_1^{(k)}, \mathbf{w}^{(k)}\right) \quad (28)$$

$$\sim \mathcal{N}\left(\boldsymbol{\mu}^{(k+1)}, \boldsymbol{\Sigma}^{(k+1)}\right). \quad (29)$$

The covariance matrix is diagonal and writes

$$\boldsymbol{\Sigma}^{(k+1)} = \gamma_n^{(k)} |\boldsymbol{\Lambda}_H^{(k)}|^2 + \gamma_0^{(k)} |\boldsymbol{\Lambda}_I|^2 + \gamma_1^{(k)} |\boldsymbol{\Lambda}_D|^2 \quad (30)$$

and the mean

$$\boldsymbol{\mu}^{(k+1)} = \gamma_n^{(k)} \left(\boldsymbol{\Sigma}^{(k+1)}\right)^{-1} \boldsymbol{\Lambda}_H^* \hat{\mathbf{y}}. \quad (31)$$

where  $*$  is the conjugate symbol. The vector  $\boldsymbol{\mu}^{(k+1)}$  is the regularized least square solution at current iteration (or the Wiener-Hunt filter). Clearly, if the null-frequency is not observed  $\hat{h}_0 = 0$  and if  $\gamma_0 = 0$ , the covariance matrix  $\boldsymbol{\Sigma}$  is not invertible and the estimate is not defined as described Sec. IV.

Finally, since the matrix is diagonal, the sample  $\hat{\mathbf{x}}^{(k+1)}$  is obtained by a term-wise product of a complex white Gaussian noise with the standard deviation matrix  $\left(\boldsymbol{\Sigma}^{(k+1)}\right)^{-1/2}$  followed by the addition of the mean  $\boldsymbol{\mu}^{(k+1)}$  also computed with term-wise products (31). Consequently, the sampling of the image is effective even with high-dimensional object.

### B. Sampling precision parameters

The conditional posterior laws of the precisions are Gamma corresponding to their prior law with parameters updated by the likelihood

$$\gamma_i^{(k+1)} \sim p\left(\gamma_i|\hat{\mathbf{y}}, \hat{\mathbf{x}}^{(k+1)}, \mathbf{w}^{(k)}\right) \quad (32)$$

$$\sim \mathcal{G}\left(\gamma_i|\alpha_i^{(k+1)}, \beta_i^{(k+1)}\right). \quad (33)$$

For  $\gamma_n, \gamma_0$  and  $\gamma_1$  the parameters law are at iteration  $k+1$ , respectively,

$$\alpha_n = \alpha_n + N/2 \quad \text{and} \quad \beta_n^{-1} = \left(\beta_n^{-1} + \frac{\|\hat{\mathbf{y}} - \boldsymbol{\Lambda}_H^{(k)} \hat{\mathbf{x}}^{(k+1)}\|^2}{2}\right) \quad (34)$$

$$\alpha_0 = \alpha_0 + 1/2 \quad \text{and} \quad \beta_0^{-1} = \left(\beta_0^{-1} + \frac{\|\hat{\mathbf{x}}_0^{(k+1)}\|^2}{2}\right) \quad (35)$$

$$\alpha_1 = \alpha_1 + \frac{N-1}{2} \quad \text{and} \quad \beta_1^{-1} = \left(\beta_1^{-1} + \frac{\|\boldsymbol{\Lambda}_D \hat{\mathbf{x}}^{(k+1)}\|^2}{2}\right). \quad (36)$$

In the case of Jeffreys's prior, the parameters are

$$\alpha_n = N/2 \quad \text{and} \quad \beta_n = 2/\|\hat{\mathbf{y}} - \boldsymbol{\Lambda}_H^{(k)} \hat{\mathbf{x}}^{(k+1)}\|^2, \quad (37)$$

$$\alpha_0 = 1/2 \quad \text{and} \quad \beta_0 = 2/\|\hat{\mathbf{x}}_0^{(k+1)}\|^2, \quad (38)$$

$$\alpha_1 = (N-1)/2 \quad \text{and} \quad \beta_1 = 2/\|\boldsymbol{\Lambda}_D \hat{\mathbf{x}}^{(k+1)}\|^2. \quad (39)$$

**Remark 1** — *If the a posteriori law (27) without  $\gamma_0$  is considered, there is no need to sample this parameter ((35) and (38) are not useful) and  $\gamma_0^{(k)} = 0$  in (30).*

### C. Sample instrument parameters

The conditional law for instrument parameters writes

$$\mathbf{w}^{(k+1)} \sim p\left(\mathbf{w}|\hat{\mathbf{y}}, \hat{\mathbf{x}}^{(k+1)}, \gamma_n^{(k+1)}\right) \quad (40)$$

$$\propto \exp\left[-\frac{\gamma_n^{(k+1)}}{2} \|\hat{\mathbf{y}} - \boldsymbol{\Lambda}_{H, \mathbf{w}} \hat{\mathbf{x}}^{(k+1)}\|^2\right] \quad (41)$$

where parameters  $\mathbf{w}$  are embedded in the instrument response  $\boldsymbol{\Lambda}_H$ . This law is not standard and intricate, so no algorithm exists for direct sampling. The Metropolis-Hastings (M.-H.) method is used to bypass this difficulty. In M.-H. algorithm, a sample  $\mathbf{w}_p$  is proposed and accepted with a certain probability. This probability depends on the ratio between the proposed likelihood and the likelihood of the current value  $\mathbf{w}^{(k)}$ . In practice, in the independent form described in appendix C, with prior law as proposition law, it is divided in several steps.

1) PROPOSITION: Sample a proposition

$$\mathbf{w}_p \sim p(\mathbf{w}) = \mathcal{U}_{[\mathbf{a} \ \mathbf{b}]}(\mathbf{w}). \quad (42)$$

2) PROBABILITY OF ACCEPTATION: Calculate the criterion

$$J\left(\mathbf{w}^{(k)}, \mathbf{w}_p\right) = \frac{\gamma_n^{(k+1)}}{2} \left(\|\hat{\mathbf{y}} - \boldsymbol{\Lambda}_{H, \mathbf{w}^{(k)}} \hat{\mathbf{x}}^{(k+1)}\|^2 - \|\hat{\mathbf{y}} - \boldsymbol{\Lambda}_{H, \mathbf{w}_p} \hat{\mathbf{x}}^{(k+1)}\|^2\right). \quad (43)$$

3) ACCEPTATION/REJECT: Sample  $t \sim \mathcal{U}_{[0 \ 1]}$  and takes

$$\mathbf{w}^{(k+1)} = \begin{cases} \mathbf{w}_p & \text{if } \log t < \min\{J, 0\} \\ \mathbf{w}^{(k)} & \text{otherwise.} \end{cases} \quad (44)$$

### D. Empirical mean

The sampling of  $\hat{\mathbf{x}}$ ,  $\gamma$  and  $\mathbf{w}$  are repeated iteratively until the law has been sufficiently explored. These samples  $[\hat{\mathbf{x}}^{(k)}, \gamma^{(k)}, \mathbf{w}^{(k)}]$  follow the global *a posteriori* law of (21). By the large numbers law, the estimate, defined as the posterior mean, is approximated with

$$\hat{\mathbf{x}} = \mathbf{F}^\dagger \mathbb{E}[\hat{\mathbf{x}}] \approx \mathbf{F}^\dagger \left[\frac{1}{K} \sum_{k=0}^{K-1} \hat{\mathbf{x}}^{(k)}\right]. \quad (45)$$

As described by (45), to obtain an estimate of the image in spatial space, all the computation are achieved recursively in Fourier space with a single IFFT at the end. An implementation example in pseudo code is described Fig. 1.

```

1: Initialisation of  $[\hat{\mathbf{x}}^{(0)}, \gamma^{(0)}, \mathbf{w}^{(0)}, k = 0]$ 
2: repeat
   % Sample of  $\hat{\mathbf{x}}$ 
3:  $\Sigma \leftarrow \gamma_n^{(k)} |\Lambda_H|^2 + \gamma_0^{(k)} |\Lambda_{\mathbb{1}}|^2 + \gamma_1^{(k)} |\Lambda_D|^2$ 
4:  $\mu \leftarrow \gamma_n^{(k)} \Sigma^{-1} \Lambda_H^* \hat{\mathbf{y}}$ 
5:  $\hat{\mathbf{x}}^{(k)} \leftarrow \mu + \Sigma^{-1/2} \cdot \text{randn}$ 
   % Sample of  $\gamma$ 
6:  $\gamma_n^{(k)} \leftarrow \text{gamrnd}(\alpha_n, \beta_n)$ 
7:  $\gamma_1^{(k)} \leftarrow \text{gamrnd}(\alpha_1, \beta_1)$ 
8:  $\gamma_0^{(k)} \leftarrow \text{gamrnd}(\alpha_0, \beta_0)$ 
   % Sample of  $\mathbf{w}$ 
9:  $\mathbf{w}_p \leftarrow \text{rand} * (\mathbf{a} - \mathbf{b}) + \mathbf{a}$ 
10:  $J \leftarrow \gamma_n (|\hat{\mathbf{y}} - \Lambda_H \hat{\mathbf{x}}^{(k)}|^2 - |\hat{\mathbf{y}} - \Lambda_{H, \mathbf{w}_p} \hat{\mathbf{x}}^{(k)}|^2) / 2$ 
11: if  $\log(\text{rand}) < \min\{J, 0\}$  then
12:    $\mathbf{w}^{(k)} \leftarrow \mathbf{w}_p$ 
13:    $\Lambda_H \leftarrow \Lambda_{H, \mathbf{w}_p}$ 
14: else
15:    $\mathbf{w}^{(k)} \leftarrow \mathbf{w}^{(k-1)}$ 
16: end if
   % Empirical mean
17:  $k \leftarrow k + 1$ 
18:  $\hat{\mathbf{x}}^{(k)} \leftarrow \sum_i \hat{\mathbf{x}}^{(i)} / k$ 
19: until  $|\bar{\mathbf{x}}^{(k)} - \bar{\mathbf{x}}^{(k-1)}| / |\bar{\mathbf{x}}^{(k)}| \geq \text{criterion}$ 

```

Fig. 1. The algorithm in pseudo code. The `randn`, `rand` and `gamrnd` are the name of sampling functions for white complex Gaussian, uniform and gamma laws, respectively.

## VI. DECONVOLUTION RESULTS

This section presents results obtained with the proposed approach. In order to evaluate completely the method, true values of all parameters  $\mathbf{x}$ ,  $\mathbf{w}$ ,  $\gamma_n$  but also  $\gamma_1, \gamma_0$  are needed. To achieve this, we study an entirely simulated case: the true image and noise are simulated with their respective prior laws (13) and (14) with known values of  $\gamma_0, \gamma_1$  and  $\gamma_n$ . With this protocol, all experimental conditions are controlled and the estimation is evaluated entirely.

The method has also been applied in different conditions (lower signal to noise ratio, broader instrument response, different and realistic images, ...) and showed similar characteristics. However, in case of realistic images, since the true values of the hyperparameters are not known, the evaluation cannot be complete.

### A. Practical experimental conditions

Concretely, the image is generated with a complex white Gaussian noise multiplied by the *a priori* standard deviation matrix  $\Sigma = (\gamma_0 \Lambda_{\mathbb{1}}^\dagger \Lambda_{\mathbb{1}} + \gamma_1 \Lambda_D^\dagger \Lambda_D)^{-1/2}$ , see (11). The image is square with  $128 \times 128$  pixels. The parameter values are  $\gamma_0 = 1$  and  $\gamma_1 = 2$ . The matrix  $\Lambda_D$  is obtained with the FFT-2D of the Laplacian  $[0 \ 1 \ 0; 1 \ -4 \ 1; 0 \ 1 \ 0] / 8$ .

These parameters provide the image illustrated Fig. 2(a). It is an image with smooth features similar to a cloud. Pixel have numerical values between  $-100$  and  $150$ . The profile line 68 shows fluctuations around a value of  $-40$ .

The *a priori* law of the hyperparameters are set to the non-informative Jeffreys's law by fixing the parameters  $(\alpha_i, \beta_i)$  to

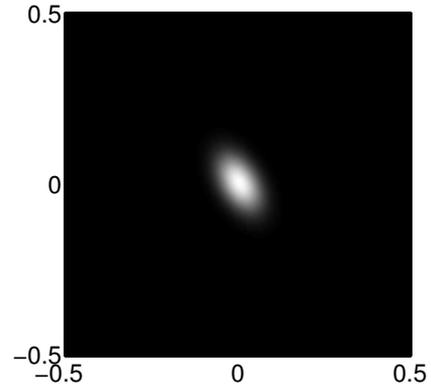


Fig. 3. Illustration of the instrument response  $\Lambda_H$  with  $w_\alpha = 20$ ,  $w_\beta = 7$  and  $\varphi = \pi/3$ .

( $0, +\infty$ ) see Sec. III-C. In addition, the instrument response  $\Lambda_H$  is obtained by sampling a normalized Gaussian PSF written in Fourier space

$$\begin{aligned} \hat{h}(\nu_\alpha, \nu_\beta) = \exp \left( -2\pi^2 \left( \nu_\alpha^2 (w_\alpha \cos^2 \varphi + w_\beta \sin^2 \varphi) \right. \right. \\ \left. \left. + \nu_\beta^2 (w_\alpha \sin^2 \varphi + w_\beta \cos^2 \varphi) \right. \right. \\ \left. \left. + 2\nu_\alpha \nu_\beta \sin \varphi \cos \varphi (w_\alpha - w_\beta) \right) \right) \quad (46) \end{aligned}$$

with frequencies  $(\nu_\alpha, \nu_\beta) \in [-0.5; 0.5]^2$ . This low-pass filter, illustrated Fig. 3, is controlled by three parameters:

- two width parameters  $w_\alpha$  and  $w_\beta$  set to 20 and 7, respectively. Their *a priori* laws are  $p(w_\alpha) = \mathcal{U}_{[19 \ 21]}$  and  $p(w_\beta) = \mathcal{U}_{[6 \ 8]}$ . This corresponds to uncertainty of approximately  $\pm 5\%$  and  $15\%$  around the nominal value (see Sec III-D).
- a rotation parameter  $\varphi$  set to  $\pi/3$  with an *a priori* law  $p(\varphi) = \mathcal{U}_{[\pi/4 \ \pi/2]}$ . The corresponding uncertainty is approximately  $\pm 50\%$  around the nominal value.

Then, the convolution is computed in Fourier space. The data illustrated Fig. 2(b) are obtained by adding a white Gaussian noise with precision  $\gamma_n = 0.5$ . The data are of course smoother than the true image. The fluctuations of the signal are less visible and corrupted by the noise. The average is correctly observed (the first DFT-2D coefficient of  $\mathbf{H}_w$  is  $\hat{h}_0 = 1$ ) so the parameter  $\gamma_0$  is considered to be a nuisance parameter. Consequently it is integrated out under a Dirac (see Sec. IV). It is equivalent to fix its value to 0 in the algorithm Fig. 1 line 4.

Finally, the method is evaluated on two different situations:

- 1) the unsupervised non-myopic case. The parameters  $\mathbf{w}$  are known and fixed in the algorithm. Consequently, there is no Metropolis-Hastings step (Sec. V-C). To see sufficient law exploration, the algorithm is run until the difference between two successive empirical means is less than  $10^{-3}$ . In this case 921 samples, obtained in approximately 12 seconds<sup>3</sup>, are necessary

<sup>3</sup>All the computation have been done on a processor at 2.66 GHz with Matlab. Gamma laws are simulated with the `gamrnd` functions.

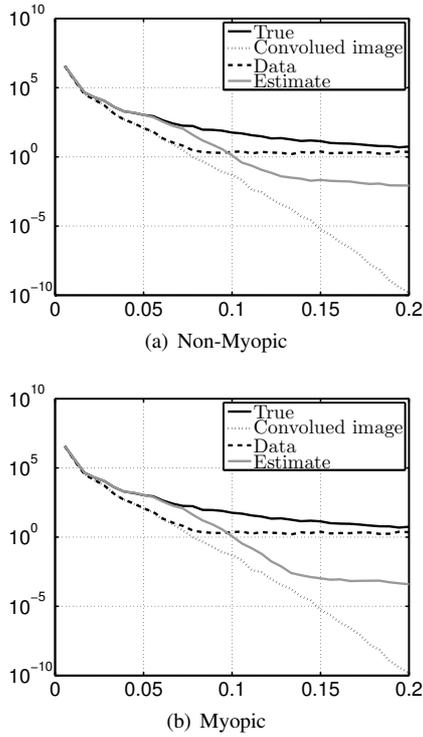


Fig. 4. Circular mean of the power spectral density (PSD) of the image, the convolved image, the data (filtered image corrupted by noise) and the estimates, in reduced frequency with y-axis in logarithmic scale.

- 2) the unsupervised and myopic case. All the parameters are estimated. To see sufficient law exploration, the algorithm is run until the difference between two successive empirical means is less than  $5 \cdot 10^{-5}$ . In this case 18 715 samples are needed. They are obtained in approximately 7 minutes.

**Remark 2** — *The algorithm has been run until 1 000 000 samples, in both cases, without perceptible qualitative changes.*

### B. Estimation results

The two results for the image are illustrated Figs. 2(c) and 2(d) for the non-myopic and the myopic cases, respectively. The image is restored, more details are visible and the profiles are closer to the true image than data. More precisely, the profiles around pixel 20 of the 68-th line Fig. 2 shows the retrieval of the curve dynamic. Between pixels 70 and 110, fluctuations not visible in data are restored. There is no visual differences between the non-myopic and myopic cases. The circular mean of the power spectral density of the objects are illustrated Figs. 4(a) and 4(b) in non-myopic and myopic cases, respectively. The spectrum of the true image is retrieve, in both cases, up to the frequency  $f \approx 0.075$  limits where the noise start to be dominant. Effectively, after this frequency the information about the image is almost lost.

In order to quantify, the distances

$$\epsilon = \|\hat{\mathbf{x}} - \mathbf{x}\|^2/N \quad (47)$$

	Data	Non-myopic	Myopic
Error ( $\epsilon$ )	28.27	8.80	8.85
$\hat{\sigma}$ of $\mathbf{x}$ law	-	3.16	3.25

TABLE I  
ERROR  $\epsilon$  AND AVERAGED STANDARD DEVIATION  $\hat{\sigma}$  OF THE POSTERIOR IMAGE LAW.

between the true image and the estimates are computed and reported in Tab. I. The error values for the non-myopic and the myopic cases are almost the same, with a slightly lower value for the non-myopic case as expected. This difference between the both cases is coherent with the intuition since in non-myopic case the true instrument parameters values are known: more information is injected. Both are lower than the data error  $\|\mathbf{y} - \mathbf{x}\|^2/N \approx 28$ .

The worst case is when in non-myopic case the instrument parameters are fixed to the lowest or biggest limit interval of their respective prior. In these cases the obtained errors  $\epsilon$  are between 13 and 20. So even in the less favorable case of myopic deconvolution, the estimation is close to the non-myopic result when true instrument parameter value are used (the most favorable case).

Concerning the other parameters, their estimates are close to the true values and are reported in Tab. II. The  $\gamma_n$  estimation is very close with  $\hat{\gamma}_n = 0.49$  instead of 0.5 in the two cases. The error for the instrument parameters are 0.39%, 2.74% and 1.68% for  $w_\alpha$ ,  $w_\beta$  and  $\varphi$ , respectively. The value of  $\gamma_1$  is underestimated in the two cases with approximately 1.7 instead of 2. All the true values fall in the  $\hat{\mu} \pm 3\hat{\sigma}$  interval.

Anyway, an automatic balance between the two sources of information (*a priori* and data), as described in the introduction, is found with an estimated image closer to the true image than data.

### C. A posteriori law characteristics

This section describes the *a posteriori* law characteristics with histograms and variances of the parameters. The samples histograms, Figs. 5 and 6, provide an approximation of the marginal posterior law for each parameter. The empirical variances permits to quantify the uncertainty. Tabs. I and II report the variance for the image and parameters law, respectively.

1) *Hyperparameters characteristics*: The histograms of  $\gamma_n$  and  $\gamma_1$ , Fig. 5, are concentrated around their mean in both non-myopic and myopic case. The variance of  $\gamma_n$  is lower than the  $\gamma_1$  one. Effectively there is a degradation, by the convolution, of the information about  $\gamma_1$  present in the image which is an input of the instrument model. This is not the case for  $\gamma_n$ , which is directly observed in the output, resulting in a lower variance for  $\gamma_n$ .

A second observation is the larger variance for  $\gamma_1$  in the myopic case than in the non-myopic case. An explanation is the addition of information when the instrument parameters are fixed to the true values. The same explanation applies to the larger variance of the image  $\mathbf{x}$ , see Tab. I, in the myopic case.

		$\hat{\gamma}_n \pm \hat{\sigma}$	$\hat{\gamma}_1 \pm \hat{\sigma}$	$\hat{w}_\alpha \pm \hat{\sigma}$	$\hat{w}_\beta \pm \hat{\sigma}$	$\hat{\varphi} \pm \hat{\sigma}$
	<b>True value</b>	<b>0.5</b>	<b>2</b>	<b>20</b>	<b>7</b>	<b>1.05 (<math>\pi/3</math>)</b>
<b>Non-myopic</b>	Estimate	0.49 $\pm$ 0.0056	1.78 $\pm$ 0.14	-	-	-
<b>Myopic</b>	Estimate	0.49 $\pm$ 0.0056	1.65 $\pm$ 0.15	20.07 $\pm$ 0.53	7.19 $\pm$ 0.38	1.03 $\pm$ 0.04
	Acceptation rate	-	-	14.50 %	9.44 %	2.14 %

TABLE II  
QUANTITATIVE EVALUATION: TRUE AND ESTIMATED VALUES.

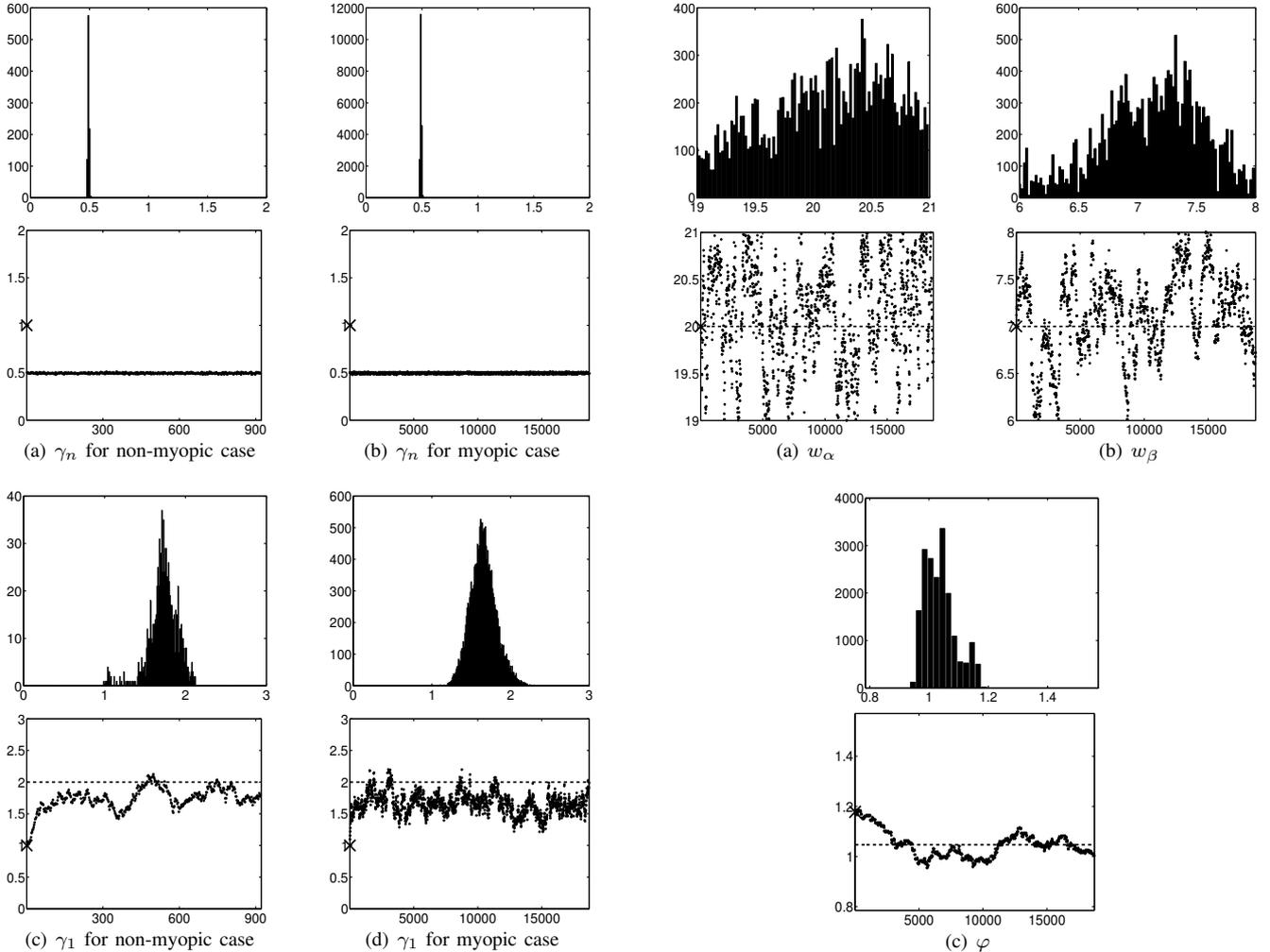


Fig. 5. Histograms and chains for the non-myopic on Figs. 5(a) and 5(c) and the myopic case on Figs. 5(b) and 5(d) for  $\gamma_n$  and  $\gamma_1$ , respectively. Histograms are above the chain. The symbol  $\times$  represents the initialisation value. The dashed line corresponds to the true value.

Fig. 6. Histogram and chain for the instrument parameters  $w_\alpha$  on Fig. 6(a),  $w_\beta$  on Fig. 6(b) and  $\varphi$  on Fig. 6(c). Histograms are above the chain. The symbol  $\times$  corresponds to the initialisation value. The dashed line corresponds to the true value. The x-axis limits for the histograms are limits of *a priori* law.

2) *Instrument parameters characteristics*: The histogram of the three instrument parameters are different. The histograms of  $w_\alpha$  and  $w_\beta$ , Figs. 6(a) and 6(b), are not so concentrated around the true value than the hyperparameter histograms. Their variances are quite large with regards to the interval of the *a priori* law. The histogram of the parameter  $\varphi$ , Fig. 6(c), has the smallest variance. An interpretation is the larger influence of this parameter on the data, or data are much more informative about its value.

3) *Myopic ambiguity*: Finally, a correlation between parameters  $(\gamma_1, w_\beta)$  and  $(\gamma_1, w_\alpha)$  is visible on their joint histograms Fig. 7, with a larger correlation for  $(\gamma_1, w_\beta)$ . This is an

illustration of the ambiguity in the myopic deconvolution problem. It can be interpreted in the following manner: the parameters  $\gamma_1$  and  $w$  both explain the spectral content of data. Effectively, an increase of  $w_\beta$  results in a decrease of the cutoff frequency: the data should be smoother. To explain the spectral content of measured data, the spectrum of the image must contain more high frequency *i.e.*, smaller  $\gamma_1$ . This is also observed on the histogram illustrated Fig. 7(a).

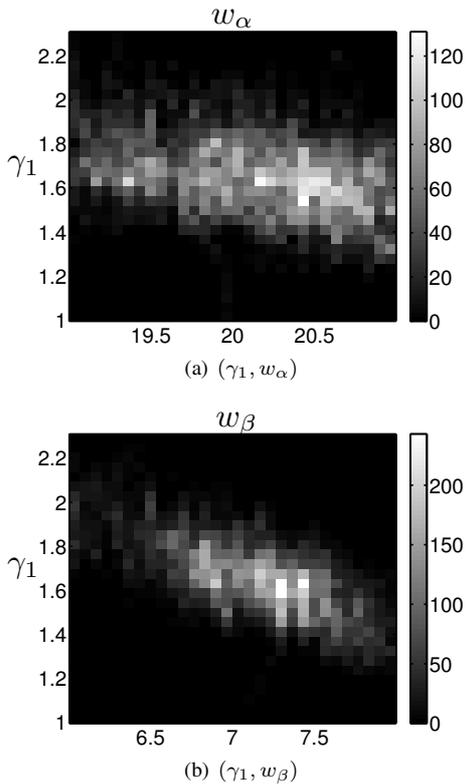


Fig. 7. Joint histogram of the couple  $(\gamma_1, w_\alpha)$  and  $(\gamma_1, w_\beta)$  on Figs. 7(a) and 7(b) respectively.

#### D. MCMC algorithm characteristics

Globally, the chains, Figs. 5 and 6, have a Markov feature (correlated) and explore the parameter space. They have a burn-in time followed by a stationary state. This characteristic has always been seen whatever the initialization. For fixed experimental conditions, the stationary state of multiple runs was always around the same value with the burn-in time. With different initialization, the only visible change is on the length of the burn-in time.

More precisely, the chain of  $\gamma_n$  explores the space and samples are concentrated in a small interval. The burn-in time is very short (less than 10 samples) and its evolution looks independent of the other parameters. The chain of  $\gamma_1$  has a larger exploration. Its burn-in time is approximately 200 samples and is the same for the myopic and non-myopic cases.

About the instrument parameters, the behaviour is different for  $(w_\alpha, w_\beta)$  and  $\varphi$ . The chain of the two width parameters has a very good exploration with quasi-instantaneous burn-in period. Conversely, the chain of  $\varphi$  is more concentrated and its burn-in time is approximately 4 000 samples of which 80 are accepted. Acceptation rates are reported in Tab. II: they are quite small, especially for the angle parameter. Conversely, the *a priori* interval for this parameter is larger, leading to a rejection probability in the Metropolis-Hastings algorithm larger too.

#### VII. CONCLUSION AND PERSPECTIVES

This paper presents a new global and coherent method for unsupervised myopic deconvolution. It is build within

a Bayesian framework and a proper extended *a posteriori* law for the image, the hyperparameters and the instrument parameters. The estimate, defined as the posterior mean, is computed by means of an MCMC algorithm in less than a few minutes. The results show that the deconvolved image is closer to the true image than the data and show restored high-frequencies. In addition, a parametric instrument response and an automatic balance between data information and *a priori* information are jointly estimated with the image.

The paper focuses on a linear invariant model often encountered in astronomy, medical imaging, nondestructive testing and especially around optic problems. Linear models without invariance can also be used to address other applications like spectrometry [4]. The loss of invariance property makes the calculation in Fourier space impossible but the methodology is still the same.

Gaussian law, related to  $L_2$  penalization, is known for excessive high-frequency penalization of the restored object. The use of convex  $L_2 - L_1$  penalization [24, 25] such Huber function or non convex  $L_2 - L_0$  penalization [26] can overcome this limitation. In these cases a difficulty is the no-longer known partition function [1, 7, 27] that depends on hyperparameters, except for a special recent field [28]. For the noise, Gaussian law limits the robustness to outliers. The use of robust law can help to bypass the limitation. Again, the partition function and its dependency on hyperparameters are no longer known. Consequently, the estimation of the hyperparameters is more difficult.

Finally, estimation of parameters of correlation matrix (cut-off frequency, attenuation coefficients,...) is possible within the same methodological framework. This could be done for the correlation matrix of the object or the noise. As for the instrument response parameters, the approach will be an extended *a posteriori* law and a Metropolis-Hastings step for the sampling.

#### VIII. ACKNOWLEDGMENT

The authors would like to thank professor A. Abergel in IAS laboratory at Université Paris-Sud 11, France, for his fruitful discussions and constructive suggestions.

#### APPENDIX A LAW IN FOURIER SPACE

We are interested in the law of  $\hat{x} = \mathbf{F}x$ , the FFT of  $x$ , when  $x \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . It is a known result: since  $x$  is Gaussian and the transform is linear,  $\hat{x}$  is Gaussian too. The law is then determined by its two first moments:

- The mean is

$$\mathbb{E}[\hat{x}] = \mathbf{F}\mathbb{E}[x] = \mathbf{F}\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}. \quad (48)$$

- The covariance matrix is

$$\begin{aligned} \mathbb{V}[\hat{x}] &= \mathbb{E}[(\hat{x} - \hat{\boldsymbol{\mu}})(\hat{x} - \hat{\boldsymbol{\mu}})^\dagger] \\ &= \mathbf{F}\mathbb{E}[(x - \boldsymbol{\mu})(x - \boldsymbol{\mu})^\dagger]\mathbf{F}^\dagger \\ &= \mathbb{E}[(\mathbf{F}x - \mathbf{F}\boldsymbol{\mu})(\mathbf{F}x - \mathbf{F}\boldsymbol{\mu})^\dagger] \\ &= \mathbf{F}\boldsymbol{\Sigma}\mathbf{F}^\dagger. \end{aligned} \quad (49)$$

	$\alpha$	$\beta$
Jeffreys	0	$+\infty$
Uniform	1	$+\infty$
Dirac	-	0

TABLE III  
SPECIFIC LAWS OBTAINED BY LIMIT OF THE GAMMA LAW

Moreover, if the covariance matrix  $\Sigma$  is circular it writes

$$\mathbb{V}[\hat{\mathbf{x}}] = \mathbf{F}\Sigma\mathbf{F}^\dagger = \mathbf{F}\mathbf{F}^\dagger\Lambda_\Sigma\mathbf{F}\mathbf{F}^\dagger = \Lambda_\Sigma. \quad (50)$$

## APPENDIX B THE GAMMA PROBABILITY DENSITY

### A. Definition

The Gamma probability density for  $\gamma > 0$ , with known parameter  $\alpha > 0$  and  $\beta > 0$ , is written

$$\mathcal{G}(\gamma|\alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \gamma^{\alpha-1} \exp(-\gamma/\beta). \quad (51)$$

This law has the following properties:

- The mean is  $\mathbb{E}_{\mathcal{G}}[\gamma] = \alpha\beta$
- The variance is  $\mathbb{V}_{\mathcal{G}}[\gamma] = \alpha\beta^2$
- The mode is  $m = \beta(\alpha - 1)$  only and only if  $\alpha > 1$

For some limits of the parameters, we obtain the laws in Tab. III.

### B. Marginalisation

First we consider a  $p(\mathbf{x}|\gamma)$   $N$  multivariate Gaussian law with a known precision matrix  $\gamma\Lambda^t\mathbf{A}$ ,  $\gamma \geq 0$ , and zero mean. So we consider the conjugate probability law for  $\gamma$  as a Gamma law with parameter  $(\alpha, \beta)$ . The marginalization of the joint law is known [29]

$$\begin{aligned} p(\mathbf{x}) &= \int_{\mathbb{R}_+} p(\mathbf{x}|\gamma)p(\gamma) d\gamma \\ &= \frac{\beta^{N/2} \det[\mathbf{A}]\Gamma(\alpha + N/2)}{(2\pi)^{N/2}\Gamma(\alpha)} \left(1 + \frac{\beta\|\mathbf{A}\mathbf{x}\|^2}{2}\right)^{-\alpha-N/2} \\ &= \mathcal{ST}(\mathbf{x}) \end{aligned} \quad (52)$$

which is a  $N$  multivariate  $t$ -Student law of  $2\alpha$  degrees of freedom with a  $\beta\Lambda^t\mathbf{A}$  precision matrix.

## APPENDIX C THE METROPOLIS-HASTINGS ALGORITHM

The Metropolis-Hastings algorithm provides samples of a targeted law  $p(\mathbf{w}|\hat{\mathbf{y}}, \hat{\mathbf{x}}^{(k+1)}, \hat{\gamma}_n^{(k+1)}) = f(\mathbf{w})$  that cannot be directly sampled but can be evaluated. With a law  $q(\mathbf{w}_p|\mathbf{w}^{(t)})$  named "instrument law", samples are obtained with the following iterative steps.

- 1) Sample a proposition  $\mathbf{w}_p \sim q(\mathbf{w}_p|\mathbf{w}^{(t)})$ .
- 2) Compute the probability

$$\rho = \min \left\{ \frac{f(\mathbf{w}_p) q(\mathbf{w}^{(t)}|\mathbf{w}_p)}{f(\mathbf{w}^{(t)}) q(\mathbf{w}_p|\mathbf{w}^{(t)})}, 1 \right\}. \quad (53)$$

3) Take

$$\mathbf{w}^{(t+1)} = \begin{cases} \mathbf{w}_p & \text{with } \rho \text{ probability} \\ \mathbf{w}^{(t)} & \text{with } 1 - \rho \text{ probability} \end{cases}. \quad (54)$$

At convergence, the samples are under the  $f(\mathbf{w})$  law [13, 22]. When

$$q(\mathbf{w}_p|\mathbf{w}^{(t)}) = q(\mathbf{w}_p) \quad (55)$$

the algorithm is named independent Metropolis-Hastings. In addition, if the instrument law is uniform, the acceptance probability gets simpler in

$$\rho = \min \left\{ \frac{f(\mathbf{w}_p)}{f(\mathbf{w}^{(t)})}, 1 \right\}. \quad (56)$$

## REFERENCES

- [1] J. Idier, Ed., *Bayesian Approach to Inverse Problems*. London: ISTE Ltd and John Wiley & Sons Inc., 2008.
- [2] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. Image Processing*, vol. 15, no. 12, pp. 3715–3727, December 2006.
- [3] P. Campisi and K. Egiazarian, Eds., *Blind Image Deconvolution*. CRC Press, 2007.
- [4] T. Rodet, F. Orieux, J.-F. Giovannelli, and A. Abergel, "Data inversion for over-resolved spectral imaging in astronomy," *IEEE J. of Sel. Topics in Signal Proc.*, vol. 2, no. 5, pp. 802–811, October 2008.
- [5] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*. Washington, DC: Winston, 1977.
- [6] S. Twomey, "On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature," *J. Assoc. Comp. Mach.*, vol. 10, pp. 97–101, 1962.
- [7] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia, "Hyperparameter estimation for satellite image restoration by a MCMC maximum likelihood method," *Pattern Recognition*, vol. 35, no. 2, pp. 341–352, 2002.
- [8] G. Demoment, "Image reconstruction and restoration: Overview of common estimation structure and problems," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-37, no. 12, pp. 2024–2036, December 1989.
- [9] A. C. Likas and N. P. Galatsanos, "A variational approach for Bayesian blind image deconvolution," *IEEE Trans. Image Processing*, vol. 52, no. 8, pp. 2222–2233, August 2004.
- [10] M. Cannon, "Blind deconvolution of spatially invariant image blurs with phase," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 24, no. 1, pp. 58–63, 1976.
- [11] M. J. Beal, "Variational algorithms for approximate bayesian inference," PHD, University of Cambridge, UK, 2003.
- [12] A. Jalobeanu, L. Blanc-Feraud, and J. Zerubia, "Estimation of blur and noise parameters in remote sensing," in *Proc. IEEE ICASSP*, vol. 4, 2002, pp. 3580–3583.
- [13] C. P. Robert and G. Casella, *Monte-Carlo Statistical Methods*, ser. Springer Texts in Statistics. New York, NY: Springer, 2000.
- [14] B. R. Hunt, "A matrix theory proof of the discrete convolution theorem," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 285–288, 1971.
- [15] K. Mardia, J. Kent, and J. Bibby, *Multivariate Analysis*. San Diego : Academic Press, 1992, ch. 2, pp. 36–43.
- [16] R. Molina, A. K. Katsaggelos, and J. Mateos, "Bayesian and regularization methods for hyperparameter estimation in image restoration," *IEEE Trans. Image Processing*, vol. 8, no. 2, pp. 231–246, February 1999.
- [17] C. A. Bouman and K. D. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Processing*, vol. 2, no. 3, pp. 296–310, July 1993.
- [18] D. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [19] R. E. Kass and L. Wasserman, "The selection of prior distributions by formal rules," *J. Amer. Statist. Assoc.*, vol. 91, pp. 1343–1370, 1996.
- [20] E. T. Jaynes, *Probability Theory: The Logic of Science*, G. L. Bretthorst, Ed. Cambridge University Press, 2003.
- [21] S. Lang, *Real and functional analysis*. Springer, 1993.
- [22] P. Brémaud, *Markov Chains. Gibbs fields, Monte Carlo Simulation, and Queues*, ser. Texts in Applied Mathematics 31. New York, NY: Springer, 1999.

- [23] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, November 1984.
- [24] H. R. Künsch, "Robust priors for smoothing and image restoration," *Ann. Inst. Stat. Math.*, vol. 46, no. 1, pp. 1–19, 1994.
- [25] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Processing*, vol. 6, no. 2, pp. 298–311, February 1997.
- [26] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Trans. Image Processing*, vol. 4, no. 7, pp. 932–946, July 1995.
- [27] X. Descombes, R. Morris, J. Zerubia, and M. Berthod, "Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood," *IEEE Trans. Image Processing*, vol. 8, no. 7, pp. 954–963, 1999.
- [28] J.-F. Giovannelli, "Unsupervised bayesian convex deconvolution based on a field with an explicit partition function," *IEEE Trans. Image Processing*, vol. 17, no. 1, pp. 16–26, January 2008.
- [29] G. E. P. Box and G. C. Tiao, *Bayesian inference in statistical analysis*. Addison-Wesley publishing, 1972.

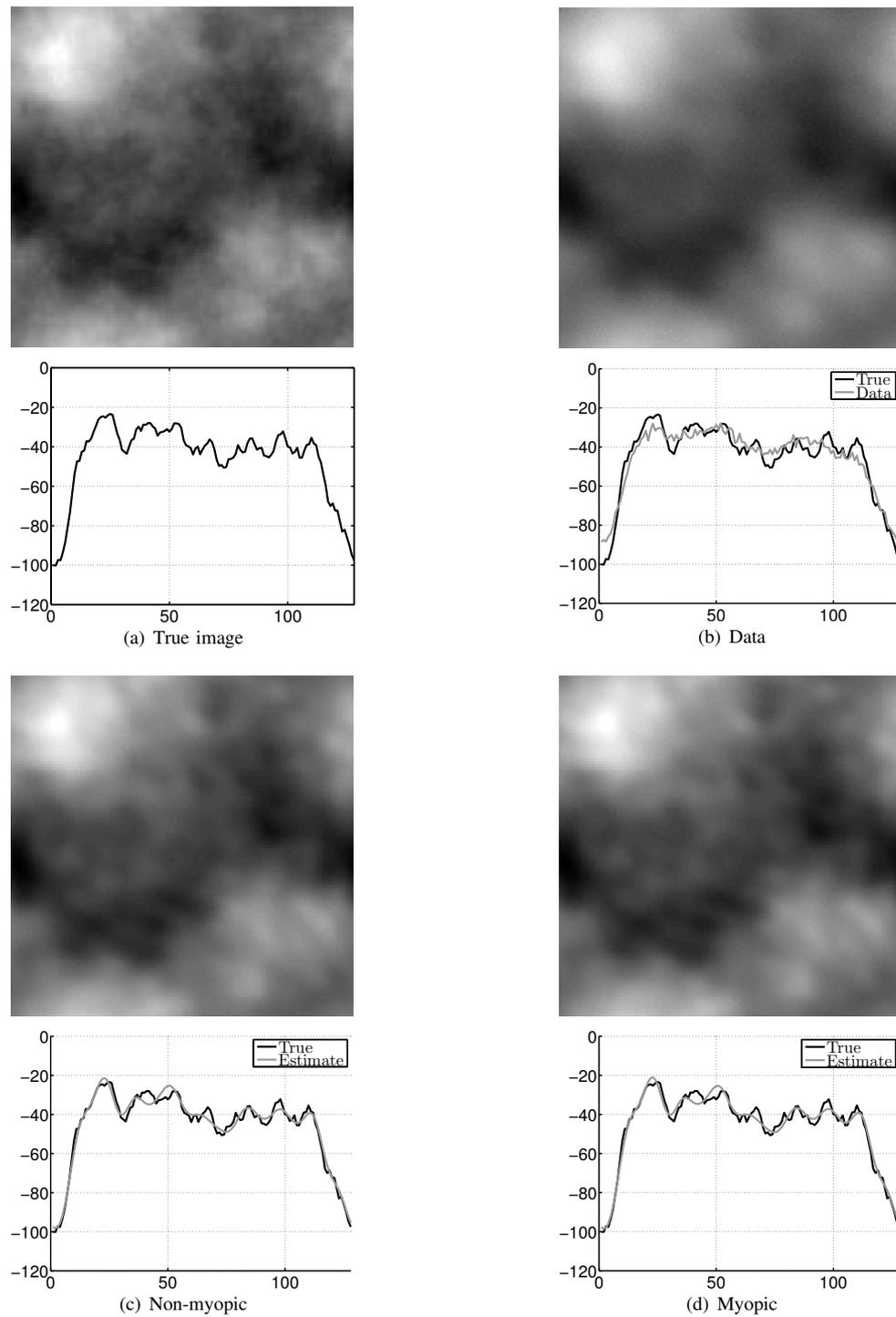


Fig. 2. The figure 2(a) represents a sample of the *a priori* law of the object. The image is 128 pixels aside with  $\gamma_0 = 1$  and  $\gamma_1 = 2$ . Fig. 2(b) is the data computed with the instrument response. Figs. 2(c) and 2(d) are the estimate with fixed instrument parameters and the myopic estimate, respectively. Profiles correspond to the 68-th line.



# Références

- [AIG06] M. ALLAIN, J. IDIER et Y. GOUSSARD : On global and local convergence of half-quadratic algorithms. *Image Processing, IEEE Transactions on*, 15(5):1130–1142, mai 2006.
- [AMD08] H. AYASSO et A. MOHAMMAD-DJAFARI : Variational bayes with Gauss-Markov-Potts prior models for joint image restoration and segmentation. *In Int. Conf. on Computer Vision and Applications, Visapp Proceedings, Funchal, Madaira, Portugal*, 2008.
- [Ban05] D. BANCELIN : Simulation de la chaine de détection de l’instrument Planck HFI. Rapport technique, Collège de France, juillet 2005. Rapport de licence, sous la direction de Michel Piat.
- [BCG<sup>+</sup>06] S. BRETTE, H. CARFANTAN, J.-F. GIOVANNELLI, T. MARTIN, J.-F. BERCHER, C. HEINRICH, J. IDIER et C. SOUSSEN : Gradient à pas adaptatif avec corrections, une mise en oeuvre matlab : GPAC.m. Logiciel déposé, Déclaration d’invention, DI-0258-01, Logibox APP 21961, IDDN.FR.001.070033.000.S.P.2006.000.10800, GPI – L2S, France, 2006.
- [Bea03] M. J. BEAL : *Variational Algorithms for approximate Bayesian inference*. Thèse de doctorat, University of Cambridge, UK, 2003.
- [BMI99] A. BLANC, L. MUGNIER et J. IDIER : Identifiabilité et estimation myope d’aberrations optiques par diversité de phase. *In Actes 17<sup>e</sup> coll. GRETSI*, pages 63–66, Vannes, septembre 1999.
- [BMK09] S. D. BABACAN, R. MOLINA et A. K. KATSAGGELOS : Variational Bayesian Blind Deconvolution Using a Total Variation Prior. *IEEE Trans. Image Processing*, 18(1): 12–26, janvier 2009.
- [BR96] M. J. BLACK et A. RANGARAJAN : On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *IJCV*, 19(1):57–91, 1996.
- [BZ87] A. BLAKE et A. ZISSERMAN : *Visual reconstruction*. The MIT Press, Cambridge, MA, USA, 1987.
- [Cam02] P. CAMUS : *Principe des bolomètres*, juin 2002. Détection de rayonnements à très basse température, 5<sup>e</sup> école thématique.
- [CCB<sup>+</sup>06] D. CLEMENTS, P. CHANIAL, G. BENDO, K. XU, B. SCHULZ, T. WASKETT, B. SIBTHORPE et G. LAURENT : SPIRE mapmaking algorithm review report. Rapport technique, Astrophysics group at Imperial College London, 2006.

- [CE07] P. CAMPISI et K. EGIAZARIAN, éditeurs. *Blind Image Deconvolution*. CRC Press, 2007.
- [CGB<sup>+</sup>03] G. CHATTOPADHYAY, J. G. GLENN, J. J. BOCK, B. K. ROWND, Ma. CALDWELL et M. J. GRIFFIN : Feed horn coupled bolometer arrays for spire — design, simulation, and measurements. *IEEE Trans. Microwave Theory Tech.*, 51(10):2139–2146, octobre 2003.
- [CGLS08] G. CHANTAS, N. GALATSANOS, A. LIKAS et M. SAUNDERS : Variational Bayesian image restoration based on a product of  $t$ -distributions image prior. *IEEE Trans. Image Processing*, 17(10):1795–1805, octobre 2008.
- [CGW07] G. K. CHANTAS, N. P. GALATSANOS et N. A WOODS : Super-resolution based on fast registration and maximum *a posteriori* reconstruction. *IEEE Transactions on Image Processing*, 16(7):1821–1830, juillet 2007.
- [CI93] F. CHAMPAGNAT et J. IDIER : Un nouvel algorithme de déconvolution impulsionnelle avec prise en compte de saturation. In *Actes 14<sup>e</sup> coll. GRETSI*, pages 559–562, Juan-les-Pins, septembre 1993.
- [CI04] F. CHAMPAGNAT et J. IDIER : A connection between half-quadratic criteria and EM algorithm. *IEEE Signal Processing Letters*, 11(9):709–712, septembre 2004.
- [CLBK09] F. CHAMPAGNAT, G. LE BESNERAIS et C. KULCSÁR : Statistical performance modeling for superresolution : a discrete data-continuous reconstruction framework. *Journal of the Optical Society of America A*, 26(7):1730–1746, 2009.
- [Dem] G. DEMOMENT : Modélisation des incertitudes, inférence logique, et traitement des données expérimentales. Laboratoire des signaux et systèmes, Groupe des Problèmes Inverses, Université Paris-Sud 11.
- [Dem89] G. DEMOMENT : Image reconstruction and restoration : Overview of common estimation structure and problems. *IEEE Trans. Acoust. Speech, Signal Processing*, ASSP-37(12):2024–2036, décembre 1989.
- [Dev85] L. DEVROYE : *Non-Uniform Random Variate Generation*. Springer Verlag, New York, NY, USA, 1985.
- [DMZB99] X. DESCOMBES, R. MORRIS, J. ZERUBIA et M. BERTHOD : Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood. *IEEE Trans. Image Processing*, 8(7):954–963, 1999.
- [EF97] M. ELAD et A. FEUER : Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *IEEE Trans. Image Processing*, 6(12):1646–1658, décembre 1997.
- [EF99] M. ELAD et A. FEUER : Superresolution restoration of an image sequence : Adaptive filtering approach. *IEEE Trans. Image Processing*, 8(3):387–395, mars 1999.
- [ESA07] ESA. *SPIRE Observers' Manual*, septembre 2007.
- [FD93] N. FORTIER et Y. DEMOMENT, G. GOUSSARD : GCV and ML methods of determining parameters in image restoration by regularization : Fast computation in the spatial domain and experimental comparison. *J. Visual Comm. Image Repres.*, 4(2):157–170, juin 1993.

- [Fer07] M. FERLET : Spire calibration files : Photometer beam profiles. Rapport technique, Université de Wales Cardiff, août 2007.
- [FHMP88] C. L. FALES, F. O. HUCK, J. A. MCCORMICK et S. K. PARK : Wiener restoration of sampled image data : an end-to-end analysis. *J. Opt. Soc. Amer. (A)*, 5(3):300–314, mars 1988.
- [FREM04] S. FARSIU, M.D. ROBINSON, M. ELAD et P. MILANFAR : Fast and robust multiframe super resolution. *Image Processing, IEEE Transactions on*, 13(10):1327–1344, octobre 2004.
- [Fér06] O. FÉRON : *Champs de Markov cachés pour les problèmes inverses. Application à la fusion de données et à la reconstruction d'images en tomographie micro-onde*. Thèse de Doctorat, Université de Paris-Sud, Orsay, novembre 2006.
- [GBG02] M. J. GRIFFIN, J. J. BOCK et W. K. GEAR : Relative performance of filled and feedhorn-coupled focal-plane architectures. *Applied Optics*, 41(31):6543–6554, novembre 2002.
- [GC05] J.-F. GIOVANNELLI et A. COULAIS : Positive deconvolution for superimposed extended source and point sources. *Astron. Astrophys.*, 439:401–412, 2005.
- [GG84] S. GEMAN et D. GEMAN : Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6(6):721–741, novembre 1984.
- [Gio08] J.-F. GIOVANNELLI : Unsupervised Bayesian convex deconvolution based on a field with an explicit partition function. *IEEE Trans. Image Processing*, 17(1):16–26, janvier 2008.
- [Goo72] J. W. GOODMAN : *Introduction à l'optique de Fourier et à l'holographie*. Masson, Paris, 1972.
- [GR80] I. S. GRADSHTEYN et I. M. RYZHIK : *Table of integrals, series, and products*. Academic Press, Inc., 4-th édition, 1980.
- [GR92] D. GEMAN et G. REYNOLDS : Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(3):367–383, mars 1992.
- [Gri06] M. J. GRIFFIN : Revised photometer sensitivity model, décembre 2006. Working version after sensitivity review meeting.
- [Gri07] M. J. GRIFFIN : The SPIRE analogue signal chain and photometer detector data processing pipeline. Rapport technique, Université de Wales Cardiff, avril 2007.
- [GY95] D. GEMAN et C. YANG : Nonlinear image recovery with half-quadratic regularization. *IEEE Trans. Image Processing*, 4(7):932–946, juillet 1995.
- [Had01] J. HADAMARD : Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton Univ. Bull.*, 13, 1901.
- [Han92] P.C. HANSEN : Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.*, 34:561–580, 1992.
- [Hun71] B. R. HUNT : A matrix theory proof of the discrete convolution theorem. *IEEE Trans. Automat. Contr.*, AC-19:285–288, 1971.

- [Hun72] B. R. HUNT : Deconvolution of linear systems by constrained regression and its relationship to the Wiener theory. *IEEE Trans. Automat. Contr.*, AC-17:703–705, 1972.
- [Idi01a] J. IDIER, éditeur. *Approche bayésienne pour les problèmes inverses*. Traité IC2, Série traitement du signal et de l'image, Hermès, Paris, 2001.
- [Idi01b] J. IDIER : Convex half-quadratic criteria and interacting auxiliary variables for image restoration. *IEEE Trans. Image Processing*, 10(7):1001–1009, juillet 2001.
- [Idi08] J. IDIER, éditeur. *Bayesian Approach to Inverse Problems*. ISTE Ltd and John Wiley & Sons Inc., London, 2008.
- [Jay03] E. T. JAYNES : *Probability Theory : The Logic of Science*. Cambridge University Press, 2003.
- [JBF02] A. JALOBEANU et J. BLANC-FÉRAUD, L. ZERUBIA : Estimation of blur and noise parameters in remote sensing. *In ICASSP*, volume 4, pages 3580–3583, 2002.
- [JBFZ02] A. JALOBEANU, L. BLANC-FÉRAUD et J. ZERUBIA : Hyperparameter estimation for satellite image restoration by a MCMC maximum likelihood method. *Pattern Recognition*, 35(2):341–352, 2002.
- [KMR97] J. KALIFA, S. MALLAT et B. ROUGÉ : Restauration d'images par paquets d'ondelettes. *In Actes 16<sup>e</sup> coll. GRETSI*, pages 1339–1342, Grenoble, septembre 1997.
- [KW96] R. E. KASS et L. WASSERMAN : The selection of prior distributions by formal rules. *JASAS*, 91:1343–1370, 1996.
- [Mac03] D. MACKAY : *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [Maz05] V. MAZET : *Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies*. Thèse de doctorat, Université Henri Poincaré, Nancy 1, 2005.
- [MIB05] V. MAZET, J. IDIER et D. BRIE : Déconvolution impulsionnelle positive myope. *In Actes 20<sup>e</sup> coll. GRETSI*, Louvain-la-Neuve, Belgique, septembre 2005.
- [MKB92] K.V. MARDIA, J.T. KENT et J.M. BIBBY : *Multivariate Analysis*, chapitre 2, pages 36–43. San Diego : Academic Press, 1992.
- [MKM99] R. MOLINA, A. K. KATSAGGELOS et J. MATEOS : Bayesian and regularization methods for hyperparameter estimation in image restoration. *IEEE Trans. Image Processing*, 8(2):231–246, février 1999.
- [MMK06] R. MOLINA, J. MATEOS et A. K. KATSAGGELOS : Blind deconvolution using a variational approach to parameter, image, and blur estimation. *IEEE Trans. Image Processing*, 15(12):3715–3727, décembre 2006.
- [MT93] S.P. MEYN et R.L. TWEEDIE : *Markov chains and stochastic stability*. Springer-Verlag, London, 1993.
- [Nas76] M. Z. NASHED : *Generalized Inverses and Applications*. Academic, New York, 1976.
- [NW00] J. NOCEDAL et S. J. WRIGHT : *Numerical Optimization*. Series in Operations Research. Springer Verlag, New York, 2000.

- [ORG09] F. ORIEUX, T. RODET et J.-F. GIOVANNELLI : Super-resolution with continuous scan shift. *In Proc. of IEEE Int. Conf. on Image Processing (ICIP 2009)*, Cairo, Egypt, Novembre 2009. Accepté à ICIP 2009.
- [PAB<sup>+</sup>08] G. PATANCHON, P. A. R. ADE, J. J. BOCK, E. L. CHAPIN, M. J. DEVLIN, S. DICKER, M. GRIFFIN, J. O. GUNDERSEN, M. HALPERN, P. C. HARGRAVE, D. H. HUGHES, J. KLEIN, G. MARSDEN, P. G. MARTIN, P. MAUSKOPF, C. B. NETTERFIELD, L. OLMI, E. PASCALE, M. REX, D. SCOTT, C. SEMISCH, M. D. P. TRUCH, C. TUCKER, G. S. TUCKER, M. P. VIERO et D. V. WIEBE : Sanepic : A mapmaking method for time stream data from large arrays. *The Astrophysical Journal*, 681(1):708–725, 2008.
- [Pol71] E. POLAK : *Computational methods in optimization*. Academic Press, New York, NY, USA, 1971.
- [PPK03] S. C. PARK, M. K. PARK et M. G. KANG : Super-resolution image reconstruction : a technical overview. *IEEE Trans. Signal Processing Mag.*, pages 21–36, mai 2003.
- [PSCD03] G. PATANCHON, H. SNOUSSI, J.-F. CARDOSO et J. DELABROUILLE : Component separation for cosmic microwave background data : a blind approach based on spectral diversity. *In PSIP*, Grenoble, janvier 2003.
- [RC00] C. ROBERT et G. CASELLA : *Monte-Carlo Statistical Methods*. Springer Texts in Statistics. Springer, New York, NY, USA, 2000.
- [RCLBG06] G. ROCHEFORT, F. CHAMPAGNAT, G. LE BESNERAIS et J.-F. GIOVANNELLI : An improved observation model for super-resolution under affine motion. *IEEE Trans. Image Processing*, 15(11):3325–3337, novembre 2006.
- [Ric72] W. H. RICHARDSON : Bayesian-based iterative method of image restoration. *J. Opt. Soc. Amer. (A)*, 62:55–59, janvier 1972.
- [Rob92] C. ROBERT : *L’analyse statistique bayésienne*. Economica, Paris, 1992.
- [Rob96] C. ROBERT : *Méthodes de Monte-Carlo par chaînes de Markov*. Economica, Paris, 1996.
- [Rob05] C. ROBERT : *Le choix bayésien*. Springer Editions, 2005.
- [Rod] T. RODET : Note sur la régularisation en pénalisant la norme de la fonction et sa dérivée. Rapport interne.
- [RZ09] T. RODET et Y. ZHENG : Approche bayésienne variationnelle : application à la déconvolution conjointe d’une source ponctuelle dans une source étendue. *In Actes 22<sup>e</sup> coll. GRETSI*, volume 1, septembre 2009.
- [SBL<sup>+</sup>08] B. SCHULZ, J. J. BOCK, N. LU, H. T. NGUYEN, C. K. XU, L. ZHANG, C. D. DOWELL, M. J. GRIFFIN, G. T. LAURENT, T. L. LIM et B. M. SWINYARD : Noise performance of the Herschel-SPIRE bolometers during instrument ground tests. *In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7020 de *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, août 2008.
- [SCG09] B. SIBTHORPE, P. CHANIAL et M. J. GRIFFIN : SPS : A software simulator for the Herschel-SPIRE photometer. *ArXiv e-prints*, juin 2009.

- [SG06] B. SIBTHORPE et M. J. GRIFFIN : Spire photometer simulator. Rapport technique, Université de Wales Cardiff, avril 2006.
- [SGW02] R. V. SUDIWALA, M. J. GRIFFIN et A. L. WOODCRAFT : Thermal modelling and characterisation of semiconductor bolometers. *International Journal of Infrared and Millimeter Waves*, 23(4):545 – 573, avril 2002.
- [Sib07] B. SIBTHORPE : *Characterisation and Optimisation of the Herschel-SPIRE Imaging Photometer Through Simulations*. Thèse de doctorat, Cardiff University, 2007.
- [SS96] R. R. SCHULTZ et R. L. STEVENSON : Extraction of high-resolution frames from video sequences. *Image Processing, IEEE Transactions on*, 5(6):996–1011, juin 1996.
- [TA76] A. TIKHONOV et V. ARSÉNINE : *Méthodes de résolution de problèmes mal posés*. Éditions MIR, Moscou, Russie, 1976.
- [Tik63] A. TIKHONOV : Regularization of incorrectly posed problems. *Soviet. Math. Dokl.*, 4:1624–1627, 1963.
- [Tor92] J.-P. TORRE : *Les bolomètres pour la photométrie*, septembre 1992. Détection de rayonnements aux très basses températures, 2e École d'Automne.
- [Two63] S. TWOMEY : On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature. *Journal of the Association for Computing Machinery*, 10:97–101, 1963.
- [VSVV07] P. VANDEWALLE, L. SBAIZ, J. VANDEWALLE et M. VETTERLI : Super-resolution from unregistrered and totally aliased signals using subspace methods. *IEEE Trans. Signal Processing*, 55(7):3687–703, juillet 2007.
- [YBK<sup>+</sup>96] G. Z. YANG, P. BURGER, P. J. KILNER, S. P. KARWATOSKI et D. M. FIRMIN : Dynamic range extension of cine velocity measurements using motion-registered spatiotemporal phase unwrapping. *J. Magn. Reson. Imaging*, 6:495–502, 1996.



## Résumé

Les travaux concernent le traitement de données pour l'imagerie sur-résolue avec une application en astronomie. On s'intéresse en particulier aux données issues de l'instrument SPIRE de l'observatoire spatial Herschel de l'ESA, dédié aux infrarouges lointains. Les problèmes soulevés sont principalement : la réponse de l'optique qui atténue les fréquences spatiales, le sous-échantillonnage ainsi que la présence d'une dérive thermique. L'approche proposée est l'inversion de données, c'est à dire la prise en compte du processus d'acquisition en plus d'information *a priori*, pour estimer le ciel d'intérêt.

La première partie des travaux concerne la modélisation du processus d'acquisition des données. Le modèle est composé d'une optique, de filtres en longueur d'onde, du capteur sensible à la température à base de bolomètres ainsi que du protocole de pointage. Le modèle obtenu, linéaire mais non invariant à cause de l'échantillonnage, est étudié. Des propriétés intéressantes, notamment dans une perspective de traitement de données, sont dégagées en particulier en lien avec la sur-résolution. L'analyse du modèle permet également de faire ressortir des propriétés utiles pour un algorithme de calcul.

La deuxième partie des travaux repose sur une démarche d'inférence inscrite dans le formalisme bayésien usuel. Toute l'information ne passant pas à travers l'instrument ou étant dégradée, le problème inverse est mal-conditionné. La méthode employée, en plus de formaliser une information de régularité spatiale sur le ciel permettant de lever le problème de conditionnement, propose l'estimation des paramètres des lois réglant le compromis entre les différentes sources d'information (hyper-paramètres). De plus, l'approche proposée permet l'estimation de paramètres instruments ainsi que l'estimation d'une dérive thermique lente affectant l'ensemble du capteur conjointement aux autres paramètres. L'ensemble de l'information utilisée pour résoudre le problème est formalisé au travers d'une loi *a posteriori* jointe pour l'ensemble des inconnues. L'estimateur choisi est la moyenne *a posteriori* calculée par un algorithme MCMC. Une étude expérimentale démontre la capacité de la méthode à restaurer de hautes fréquences spatiales. L'étude montre également le potentiel de l'approche pour l'estimation des hyper-paramètres et des paramètres instruments.

## Abstract

The work involves data processing for super-resolution imaging with an application in astronomy. We are particularly interested in data from the SPIRE instrument of the spatial observatory Herschel of ESA, dedicated to far-infrared. The problems are mainly : the convolution of the spatial optical response, the sub-sampling and the presence of a thermal drift. The proposed approach is the inversion of data, ie taking into account the acquisition process as well as prior information to estimate the sky of interest.

The first part of the work concerns the modelisation of the data acquisition process. The model consists of the mirror and feed-horns, the wavelength filter, the temperature sensitive sensor based on bolometers and scanning protocol. The obtained model, linear but not invariant due to sampling, is studied. Interesting properties, including perspective on data processing, are obtained especially in connection with the super-resolution capacity. The analysis also show useful properties for efficient algorithm.

The second part of the work relies on a inference framework based on the usual Bayesian formalism. Since information is degraded, the inverse problem is ill-conditioned. The used method offers parameter estimation laws governing the balance between different information sources (hyper-parameters), in addition to the formalization of spatial regularity. Moreover, the proposed approach allows instrument parameters estimation and estimation of a slow thermal drift affecting all the sensor in conjunction with all the other parameters. All the information used to solve the problem is formalized through a posterior law. The estimator is chosen as the posterior mean calculated by means of an MCMC algorithm. An experimental study demonstrates the capacity of the approach to restore high spatial frequencies. The study also shows the potential of the approach for estimating hyper-parameters and instrument parameters.