



HAL
open science

Contributions aux méthodes de calibration et d'imagerie pour les radio-interféromètres en présence d'interférences

Yassine Mhiri

► **To cite this version:**

Yassine Mhiri. Contributions aux méthodes de calibration et d'imagerie pour les radio-interféromètres en présence d'interférences. Traitement du signal et de l'image [eess.SP]. Université Paris-Saclay, 2023. Français. NNT : 2023UPAST111 . tel-04297295

HAL Id: tel-04297295

<https://theses.hal.science/tel-04297295v1>

Submitted on 21 Nov 2023

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.

Contributions aux méthodes de calibration et d'imagerie pour les radio-interféromètres en présence d'interférences

*Contributions to Calibration and Imaging Methods for
Radio Interferometers in the Presence of Interferences*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°580, Sciences et Technologies de l'Information et de la
Communication (STIC)

Spécialité de doctorat : Sciences du traitement du signal et des images

Graduate School : Sciences de l'Ingénierie et des Systèmes

Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **SATIE (Université Paris-Saclay,
ENS Paris-Saclay, CNRS)**, sous la direction de **Pascal Larzabal**,
Professeur, Université Paris-Saclay, le co-encadrement de **Mohammed-
Nabil El-Korso**, Professeur, Université Paris-Saclay, et le co-encadrement de
Arnaud Breloy, Professeur, CNAM

Thèse soutenue à Paris-Saclay, le 19 octobre 2023, par

Yassine MHIRI

Composition du jury

Membres du jury avec voix délibérative

Guillaume Ginolhac Professeur des universités, Université Savoie Mont-Blanc	Président
Nelly Pustelnik Chargée de recherche, CNRS - ENS de Lyon	Rapporteuse & Examinatrice
Jordi Vilà-Valls Professeur, ISAE Supaero	Rapporteur & Examineur
Barbara Nicolas Directrice de recherche, CNRS - Université de Lyon	Examinatrice
David Mary Professeur des universités, Université de Nice Côte d'Azur	Examineur

Titre : Contributions aux méthodes de calibration et d'imagerie pour les radio-interféromètres en présence d'interférences

Mots clés : Interférométrie, maximum de vraisemblance, calibration, imagerie, interférences de radiofréquence, radioastronomie

Résumé : Les radiotélescopes interféromètres permettent de reconstruire des images des émissions radio provenant de sources célestes. Ces images sont d'un grand intérêt pour les radioastronomes et les astrophysiciens car elles permettent d'obtenir des informations sur les phénomènes physiques en place dans l'univers. Pour cela, une première étape de calibration est nécessaire pour estimer et corriger les effets perturbateurs notamment liés à l'instrument et à la propagation du signal d'intérêt dans l'atmosphère. Des algorithmes d'imagerie sont ensuite mis en œuvre pour reconstruire une image du ciel. Toutefois, la conjonction de l'utilisation croissante de radiofréquences dans de nombreuses activités humaines et la forte sensibilité des antennes rendent nécessaire le développement de traitements robustes à la présence d'interférences (RFI). Plusieurs méthodes permettent de détecter et filtrer ces signaux interférents. Néanmoins, les sources d'interférences de faible et moyenne puissances restent difficiles à détecter avec de telles méthodes. Le but de ce travail est alors de proposer des algorithmes de traitement du signal pour la radioastronomie qui soient robustes à la présence de sources interférentes dans les données. Ce travail sera conduit avec un souci constant de performance statistique et d'efficacité computationnelle en grandes dimensions. Tout d'abord, dans le contexte de la calibration, nous montrons que les interférences peuvent être modélisées par un bruit de rang faible. Cette modélisation nous a permis de développer un al-

gorithme original de calibration multifréquentiel basé sur une variante de l'algorithme Espérance-Maximisation. Grâce à un choix judicieux des données latentes, nous avons obtenu des expressions analytiques à chaque itération permettant de limiter la charge de calcul requise. Les simulations effectuées ont montré la supériorité en précision des solutions proposées par rapport à l'état de l'art en présence de RFI. Nous proposons ensuite des modélisations statistiques des données permettant de prendre en compte l'impact de sources interférentes dans le contexte de l'imagerie. Nous nous appuyons sur la classe des distributions gaussiennes composées. Nous avons ainsi pu proposer un algorithme du maximum de vraisemblance régularisé (type EM) en présence d'interférences. Toutefois, tous les algorithmes proposés pour résoudre l'étape de maximisation de l'EM font appel à une DFT dont la charge de calcul est liée à la coûteuse implémentation d'une DFT sur un vecteur de visibilité de grandes dimensions. Afin d'accélérer cette étape par l'usage d'une FFT, nous avons dû introduire un bruit synthétique pour prendre en compte l'approximation d'une DFT sur une grille non uniforme par une FFT. C'est ainsi que nous avons été amenés à développer un jeu original de variables latentes permettant d'en améliorer l'efficacité computationnelle. Comme attendu, les algorithmes d'imagerie ainsi proposés permettent une reconstruction améliorée par rapport aux algorithmes de l'état de l'art en présence de RFI avec une charge de calcul contrôlée.

Title : Contributions to Calibration and Imaging Methods for Radio Interferometers in the Presence of Interferences.

Keywords : Interferometry, maximum likelihood, calibration, imaging, radiofrequency interference, radioastronomy

Abstract : Radio-interferometers enable the reconstruction of images of radio emissions originating from celestial sources. These images are of great interest to radio astronomers and astrophysicists as they provide informations about the underlying physical phenomena in the universe. To achieve this, an initial calibration step is necessary to estimate and correct perturbing effects, particularly those related to the instrument and the propagation of the signal of interest through the atmosphere. Subsequently, imaging algorithms are implemented to reconstruct a sky image. However, the combination of the increasing use of radio frequencies in various human activities and the high sensitivity of the antennas require the development of robust treatments to address the presence of radio frequency interference (RFI). Several methods exist to detect and filter these interfering signals. Nevertheless, detecting low and moderate power interference sources remains challenging using such methods. The objective of this research is to propose signal processing algorithms for radio astronomy that are resilient to the presence of interfering sources in the data. This work will be conducted with a constant focus on statistical performance and computational efficiency in high-dimensional settings. Firstly, in the context of calibration, we demonstrate that interferences can be modeled as a low-rank noise. This modeling approach has allowed us to develop an innovative

multifrequency calibration algorithm based on a variant of the Expectation-Maximization (EM) algorithm. By carefully selecting a latent data space, we obtained analytical expressions at each iteration to limit the required computational burden. The performed simulations have shown the superior precision of the proposed solutions compared to the state-of-the-art methods in the presence of RFI. Subsequently, we propose a statistical data model that accounts for the impact of interfering sources in the imaging context. To do so, we rely on the class of compound gaussian distributions. Thus, we have been able to propose a regularized maximum likelihood algorithm (EM like) in the presence of interferences. However, all the proposed algorithms for solving the maximization step of the EM algorithm involve a Discrete Fourier Transform (DFT), the computational load of which is linked to the expensive implementation of a DFT on a high-dimensional visibility vector. To accelerate this step by employing a Fast Fourier Transform (FFT), we introduced a synthetic noise to account for the approximation of a non-uniform grid DFT by an FFT. This led us to develop an original set of latent variables to improve its computational efficiency. As expected, the proposed imaging algorithms allow for improved reconstruction compared to state-of-the-art algorithms in the presence of RFI, while maintaining controlled computational complexity.

À ma soeur, Ines,

It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment.

Carl Friedrich Gauss

Remerciements

Cette thèse fut une belle aventure, en partie par la présence et la contribution d'un certain nombre de personnes que je souhaite remercier ici.

Ainsi, je remercie particulièrement l'équipe encadrante qui a su m'accompagner et me guider avec humanité et bienveillance. Merci à Arnaud d'avoir répondu à ce premier mail impromptu dans lequel je demandais un sujet de thèse, en répondant à une annonce de sujet de stage. Merci surtout de m'avoir mis en contact avec Pascal et Nabil, qui ont été dans cette aventure plus que des encadrants. En effet, je vous remercie tous les trois pour votre engagement, votre expertise et votre humanité. Vous m'avez permis de garder une sérénité certaine dans les quelques moments intenses que l'on peut connaître dans une thèse.

Aussi, dans le cadre de cette thèse, j'ai eu l'opportunité de faire plein de rencontres qui m'ont beaucoup apporté, que ce soit sur le plan scientifique ou humain. Merci à Jianhua, mon partenaire de glisse au SATIE, avec qui j'ai pu voyager et découvrir le monde de la recherche. Mais surtout, merci pour ton aide et tes questions qui m'ont permis d'aller plus loin dans ma compréhension et dans mes contributions. Merci à Julien, à Cyril et à l'équipe du Lesia de nous avoir accueillis sur le site du radiotélescope de Nançay. C'était une expérience particulièrement enrichissante. Merci à Oleg Smirnov et à Landman Bester d'avoir cru en mes idées et de m'avoir permis de mettre en pratique mes propositions.

Merci à l'équipe du SATIE qui m'a très bien accueilli, ainsi qu'aux doctorants du labo avec qui j'ai pu apprendre à skier et partager pas mal de cafés. Merci à Lucien, Isabelle, Nawel, Pierre, Nicolas et Nicolas et tous ceux que j'oublie. Tous ces moments de vie partagés ont fait de cette thèse une belle aventure.

Enfin, je me dois aussi de remercier les personnes de l'ombre qui m'ont beaucoup apporté pour que je puisse aller au bout de cette aventure. Merci à mes amis, qui m'ont offert une respiration dans ce marathon qu'est la thèse. À Adrien avec qui j'ai découvert le trail, et sans quoi cette thèse aurait été plus dur à mener. Je remercie mes parents pour leur soutien inconditionnel. Vous avez toujours été pour moi le meilleur exemple de résilience et de détermination. Pour terminer, merci à Lucie, ton soutien et ta présence ont été la meilleure chose que j'aurais pu espérer. Je n'aurais pas pu y arriver sans toi.

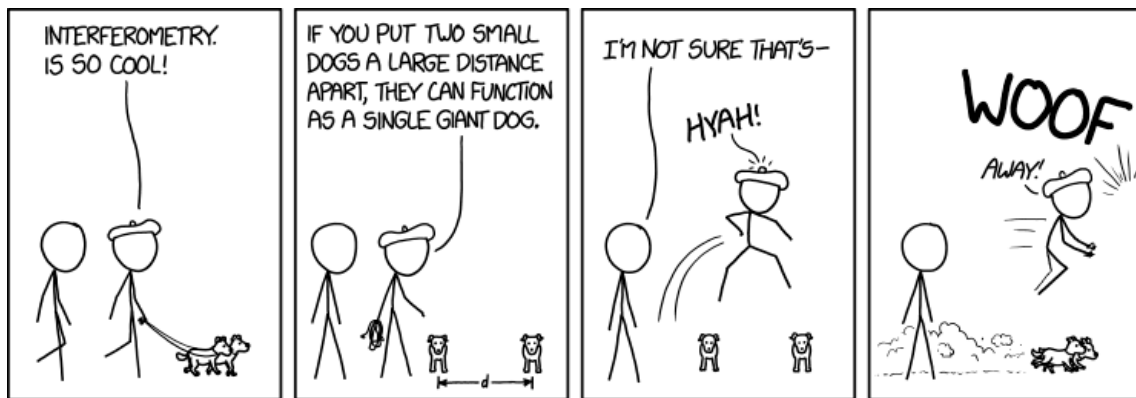


Figure 1 - Interferometry - <https://xkcd.com/1922/>

Table des matières

Remerciements	7
Glossaire	15
1 Introduction	19
1.1 Cadre général	19
1.2 Objectifs de la thèse	23
1.3 Contributions	24
1.4 Valorisations des travaux de recherche	26
1.5 Organisation du document	27
2 État de l'art	29
2.1 Introduction à la radio-interférométrie	29
2.2 Méthodes de calibration en radio-interférométrie	34
2.3 Algorithmes d'imagerie en radio-interférométrie	43
2.4 Méthodes d'atténuation des interférences de radiofréquence	48
2.5 Conclusion et motivations des travaux de thèse	50
3 Calibration multifréquentielle en présence d'interférences	53
3.1 Énonciation du problème de calibration	54
3.2 Modélisation des interférences de radiofréquence	57
3.3 Calibration multifréquentielle par maximum de vraisemblance	59
3.4 Conclusion	67
4 Imagerie interférométrique en présence d'interférences	69
4.1 Modèle de données en présence de RFI	70
4.2 Algorithme EM pour l'imagerie en présence de RFI	71
4.3 Factorisation du modèle direct pour un algorithme EM accéléré	72
4.4 Résultats numériques	75
4.5 EM déroulé : Travaux en cours et résultats préliminaires	79
4.6 Conclusion	83
5 Traitement de données réelles	85
5.1 L'écosystème logiciel en radioastronomie	85
5.2 Intégration dans la librairie Python <i>pfbclean</i>	88
5.3 Imagerie sur données réelles en présence de RFI	92
5.4 Conclusion	95

6 Conclusion et perspectives	99
6.1 Conclusion	99
6.2 Perspectives	100
A Calibration multifréquentielle en présence d'interférences	103
B Imagerie interférométrique en présence d'interférences	113
C Systèmes de coordonnées utilisés en radioastronomie	115
D Simulation des visibilités	119

Table des figures

1	Interferometry - https://xkcd.com/1922/	8
1.1	Image composite de la radiogalaxie <i>Hercules A</i> en lumière visible et en onde radio	19
1.2	Antennes du <i>Very Large Array telescope</i>	20
2.1	(a) Position dans le repère terrestre ITRF des 64 antennes du radiotélescope MeerKat [33]. (b) Couverture du plan uv du radiotélescope MeerKat permettant d'échantillonner 2016 visibilitées.	31
2.2	Effets de la rotation de la Terre sur la couverture du plan uv	33
2.3	Boucle de <i>self-calibration</i>	36
2.4	Effets d'une mauvaise calibration	36
2.5	Densité de probabilité d'une gaussienne composée	40
3.1	Évolution du NMSE de Z en fonction du nombre de canaux fréquentiels voisins affectés par les RFI	64
3.2	Évolution du NMSE de Z en fonction de la puissance des RFI	65
3.3	Évolution du NMSE de Z en fonction de la puissance des RFI de forte puissance en présence de RFI additionnelles de faible puissance sur chaque canal fréquentiel.	66
3.4	Évolution du NMSE de Z en fonction de la puissance des RFI pour plusieurs choix du rang de W	66
3.5	Évolution de la log-vraisemblance négative en fonction du nombre d'itérations pour 100 réalisations de Monte Carlo.	67
4.1	Évolution de (a) la SSIM et (b) du SNR des images en fonction la puissance des RFI.	76
4.2	Exemples d'images reconstruites pour un ciel simulé de 25 sources en présence de RFI de puissance -5dB. (a) Image originale, (b) Algorithme EM proposé, (c) Algorithme d'imagerie basé sur un bruit gaussien, (d) <i>dirty image</i>	77
4.3	Images de la radiogalaxie <i>M31</i> reconstruite par différents algorithmes en présence de RFI de puissance -4dB. (a) Image originale, (b) Algorithme EM proposé, (c) Algorithme d'imagerie basé sur un bruit gaussien, (d) <i>dirty image</i>	78
4.4	Temps d'exécution des différents algorithmes d'imagerie en fonction du nombre de visibilitées.	79
4.5	Diagramme de l'architecture neuronale dépliée	81
4.6	Plan uv utilisé pour simuler les jeux de données d'entraînement et de test.	82

4.7	Comparaison des performances de l'algorithme EM pour 100 itérations avec l'architecture neuronale proposée dépliée sur 10 itérations sur (a) le SNR et (b) la SSIM.	82
4.8	(a) Image à reconstruire, (b) Image obtenue en utilisant l'architecture dépliée sur 10 itérations, (c) Image estimée par 10 itérations de l'algorithme EM considérant une t -distribution, (d) Image estimée par 100 itérations de l'algorithme EM considérant une t -distribution	83
4.9	Image reconstruite par l'algorithme déplié sur 10 itérations pour des observations simulées en utilisant les positions des antennes du réseau radio-interférométrique Meerkat. La figure (a) correspond à l'image originale, (b) à la <i>dirty image</i> et (c) à l'image reconstruite par l'algorithme déplié sur 10 itérations.	83
5.1	Hiérarchie du format <i>measurement set</i> utilisé pour les données provenant du LoFAR.	87
5.2	Schéma de l'algorithme EM intégré dans <i>pfbclean</i> . L'étape de maximisation est mise en oeuvre par l'algorithme 5.1.	91
5.3	<i>Dirty image</i> de la source radio SNR G55.7+3.4 obtenue à partir des données du VLA [103].	93
5.4	Impact de la présence de RFI sur la <i>dirty image</i> . Sur chaque image une proportion, $\delta \in \{1\%, 0.5\%, 0.1\%\}$, de visibilité masquées contenant des RFI est ajoutée.	94
5.5	Image obtenue en appliquant l'algorithme 5.1 au <i>measurement set</i> original ($\delta = 0\%$).	96
5.6	Images obtenues par l'algorithme 5.1 ((a) $\delta = 0.1\%$, (b) $\delta = 0.5\%$, (c) $\delta = 1\%$) et par l'algorithme EM proposé ((d) $\delta = 0.1\%$, (e) $\delta = 0.5\%$, (f) $\delta = 1\%$)	96
5.7	Images obtenues par l'algorithme 5.1 pour le <i>measurement set</i> original ($\delta = 0\%$) (a) et pour $\delta = 0.5\%$ (b). L'image obtenue par l'algorithme EM proposé pour $\delta = 0.5\%$ est montrée en (c). Les lignes caractéristiques de la présence de RFI apparaissent en (b). L'image obtenue par l'algorithme EM proposé permet de les éliminer en (c) et obtenir une image plus proche de celle obtenue sur le <i>measurement set</i> original (a).	97
5.8	Images obtenues par l'algorithme 5.1 pour le <i>measurement set</i> original ($\delta = 0\%$) (a) et pour $\delta = 0.1\%$ (b). L'image obtenue par l'algorithme EM proposé pour $\delta = 0.1\%$ est montrée en (c). Les sources faibles (dans la zone entourée en rouge) présentes dans l'image obtenue sur le <i>measurement set</i> original (a), sont atténuées par l'algorithme 5.1 en présence de RFI (b). L'algorithme EM proposé permet de les reconstruire en (c).	97
C.1	Systèmes de coordonnées équatoriales et horaires.	116
C.2	Systèmes de coordonnées horizontales.	117

Liste des tableaux

5.1	Table principale (MAIN) d'un <i>measurement set</i>	86
5.2	Exemples de modules implémentés dans <i>pfb-clean</i>	88
5.3	Bandes spectrales considérées pour l'observation de la source radio <i>SNR</i> <i>G55.7+3.4</i> [103]	93

Glossaire

Symboles mathématiques

\mathbb{R}	Ensemble des réels.
\mathbb{C}	Ensemble des complexes.
\mathbb{R}^+	Ensemble des réels positifs.
j	Unité imaginaire.
$\Re(\cdot)$	Partie réelle.
$\Im(\cdot)$	Partie imaginaire.
$(\cdot)^*$	Opérateur de conjugaison complexe.
$\delta_{i,j}$	Symbole de Kronecker.
\propto	Indique proportionnalité ou égalité à une constante additive près.
$\frac{\partial}{\partial \mathbf{A}}$	Dérivée partielle par rapport à \mathbf{A} .
$\nabla_x f$	Gradient de f par rapport à x .
$\iota_C(\cdot)$	Fonction indicatrice de C .
a	Quantité scalaire.

Opérations sur les vecteurs et matrices

\mathbf{a}	Vecteur colonne.
\mathbf{A}	Matrice.
$[\mathbf{A}]_{i,j}$	Entrée de la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne de la matrice \mathbf{A} .
\mathbf{I}_N	Matrice identité de taille $N \times N$.
$\mathbf{1}_N$	vecteur unitaire de taille N , $\mathbf{1}_N = (1, \dots, 1)$.
$(\cdot)^T$	Opérateur de transposition.
$(\cdot)^H$	Opérateur de transposition Hermitienne.
$\text{vec}(\mathbf{A})$	Vectorisation de \mathbf{A} , les colonnes de \mathbf{A} sont concaténées dans un vecteur.
$\text{vec}^{-1}(\mathbf{a})$	Opération de vectorisation inverse. $\mathbf{A} = \text{vec}^{-1}(\text{vec}(\mathbf{A}))$
$\text{diag}(\mathbf{a})$	Matrice diagonale dont la diagonale contient les éléments du vecteur \mathbf{a} .
$\text{tr}(\mathbf{A})$	Trace de la matrice \mathbf{A} .
$\text{rank}(\mathbf{A})$	Rang de la matrice \mathbf{A} .
$\mathbf{a} \cdot \mathbf{b}$	Produit scalaire entre les vecteurs \mathbf{a} et \mathbf{b} .
\odot	Produit de Hadamard.
$*$	Produit de convolution.
\otimes	Produit de Kronecker.
$\ \cdot\ _p$	Norme ℓ_p .
$ \mathbf{A} $	Déterminant de \mathbf{A} .

Probabilités

$\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Distribution complexe normal de moyenne $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Sigma}$.
$\mathcal{U}(a, b)$	Distribution uniforme de paramètres $a, b \in \mathbb{R}$.
$\mathcal{W}(T, \boldsymbol{\Sigma})$	Distribution de Wishart de moyenne $\boldsymbol{\Sigma}$ et de degré T .
$\mathbb{E}_X [f(X)]$	Espérance de $f(X)$ selon X .
\sim	Équivalence en loi de probabilité.
\xrightarrow{d}	Converge en distribution.
$\mathcal{L}(\boldsymbol{\theta}; \mathbf{y})$	Log-vraisemblance négative de $\boldsymbol{\theta}$ étant donné \mathbf{y} .
$p(\mathbf{x})$	Densité de probabilité de \mathbf{x} .
$p(\mathbf{x} \boldsymbol{\theta})$	Densité de probabilité conditionnelle de \mathbf{x} sachant $\boldsymbol{\theta}$.

Notations

\mathbf{y}	Vecteur de mesures.
\mathbf{v}, \mathbf{V}	vecteur (matrice) de visibilité sans bruits.
\mathbf{R}	Matrice de corrélation.
\mathbf{x}	Image du ciel.
\mathbf{J}	Matrice de Jones.
\mathbf{C}	Matrice de Cohérence.
l, m	Cosinus directeurs.
p, q	Indice d'antennes.
P	Nombre d'antennes du réseau interférométrique.
K	Nombre de visibilité.
N	Nombre de pixels de l'image du ciel.
Q	Nombre de canaux fréquentiels.
D	Nombre de sources.
λ	Longueur d'onde (en m).
f_j	Fréquence (en Hz) du $j^{\text{ème}}$ canal fréquentiel.
Θ	Ensemble des paramètres à estimer.

Acronymes

RFI	RadioFrequency Interference
EM	Espérance-Maximisation
SAGE	Space Alternating Generalized Expectation-maximization
ADMM	Alternate Direction Method of Multiplier
VLA	Very Large Array
LoFAR	LOW Frequency ARray
SKA	Square Kilometer Array
SARAO	South African RADioastronomy Observatory
EHT	Event Horizon Telescope
Jy	Unité Jansky ($\text{W} \cdot \text{m}^{-2} \cdot \text{Hz}^{-1}$).
ITRF	International Terrestrial Reference Frame
ICRF	International Celestial Reference Frame
LST	Local Sidereal Time
FFT	Fast Fourier Transform
DFT	Discrete Fourier Transform
TEC	Total Electron Content
MAP	Maximum a Posteriori
PSF	Point Spread Function
SNR	Signal to Noise Ratio
SSIM	Structuaral SIMilarity
NMSE	Normalized Mean Square Error
CASA	Common Astronomy Software Application
ALMA	Atacama Large Millimeter/submillimeter Array
RATT	Rhodes centre for Radio Astronomy Techniques and Technologies
CPU	Central Processing Unit
ReLU	REctified Linear Unit

1 - Introduction

1.1 Cadre général

La radioastronomie est la branche de l'astronomie qui étudie les phénomènes célestes dans le spectre des ondes radio. Depuis les premières observations radioastronomiques, réalisées dans les années 1930 par l'ingénieur américain Karl Jansky, les techniques d'observation et d'analyse des signaux radio ont été grandement améliorées. Notamment, l'essor des radiotélescopes interféromètres a permis des avancées significatives dans la compréhension de l'univers, en fournissant des images détaillées de structures astronomiques telles que les trous noirs supermassifs au centre des galaxies et les étoiles en formation [1]. La figure 1.1 combine les données acquises dans le visible et en ondes radio pour l'observation de la radiogalaxie Cygnus A.

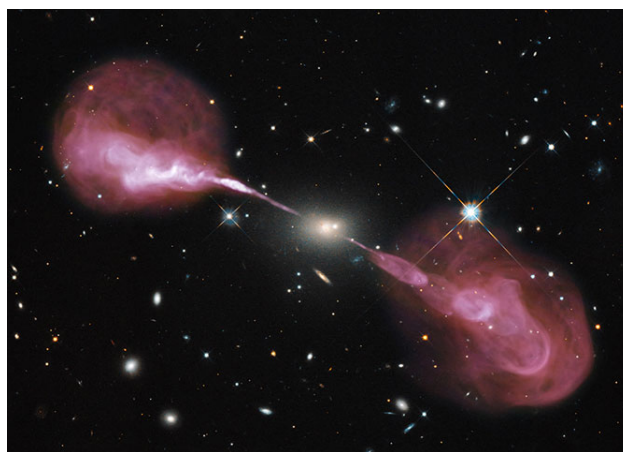


Figure 1.1 – Cette image montre la radiogalaxie *Hercules A*, qui combine les données du télescope spatial *Hubble* en lumière visible et du *Very Large Array telescope* en ondes radio. La galaxie elliptique jaunâtre au centre de l'image apparaît ordinaire en lumière visible. Pourtant, l'observation radio révèle le trou noir supermassif de 2,5 milliards de masses solaires en son centre qui émet des jets puissants de particules relativistes. Ces jets forment des lobes radioélectriques géants qui s'étendent bien au-delà de la structure visible de la galaxie. Crédit : NASA, ESA, S. Baum and C. O'Dea (RIT), R. Perley and W. Cotton (NRAO/AUI/NSF), and the Hubble Heritage Team (STScI/AURA)

Ces instruments se matérialisent par des réseaux d'antennes dont les signaux incidents sont combinés électroniquement pour mesurer des *visibilités* qui correspondent aux corrélations entre les différents capteurs. La figure 1.2 montre les 27 antennes paraboliques du *Very Large Array telescope* (VLA), situé au Nouveau-Mexique. La densité de flux est ensuite reconstruite à partir des visibilités mesurées par des algorithmes de traitement du signal. En comparaison, avec une mesure directe à partir d'une antenne parabolique, les radio-interféromètres permettent d'augmenter la résolution angulaire

des observations et d'améliorer la sensibilité des mesures. Cependant, le nombre de capteurs du réseau interférométrique et leur éloignement spatial ont un impact significatif sur les performances théoriques de l'instrument. Ainsi, les 15 dernières années ont vu la naissance de plusieurs grands radiotélescopes interféromètres apportant la promesse d'observations radioastronomiques de résolutions et sensibilités inédites. Le *Lo-FAR*, pour *Low-Frequency Array*, en service depuis 2010 et encore en construction, est composé de 52 stations, situées pour la plupart au Pays-Bas, mais aussi dans d'autres pays d'Europe pour un total de plus de 8 000 antennes [2]. Le radio-interféromètre européen a permis de lancer une cartographie inédite de l'hémisphère nord de la sphère céleste pour une couverture fréquentielle allant de 120 MHz à 168 MHz [3]. On peut aussi citer l'Event Horizon Telescope (EHT) qui est un ensemble de 11 stations de radioastronomie réparties à travers le monde ayant permis de reconstruire la première image d'un trou noir en 2018 [4]. Enfin, le 5 décembre 2022 a vu le début officiel de la construction du *Square Kilometer Array* (SKA) en Afrique du Sud et en Australie. Il s'agit d'une collaboration internationale ayant vocation à construire un radiotélescope interféromètre de nouvelle génération. Le SKA devrait être composé de plus de 130 000 antennes dans sa première phase de construction pour une résolution angulaire améliorée de plusieurs ordre de grandeur comparée aux instruments existants, apportant la promesse d'une transformation de notre compréhension de l'univers [5].



Figure 1.2 – Le *Very Large Array telescope* (VLA) est un radiotélescope interféromètre, en service depuis 1980, situé au Nouveau-Mexique, aux États-Unis. Il est composé de 27 antennes paraboliques de 25 mètres de diamètre, qui peuvent être disposées selon quatre configurations différentes. Il couvre une gamme de fréquences allant de 74 MHz à 50 GHz.

Néanmoins, l'essor de tels instruments donne naissance à de nouveaux défis technologiques à résoudre afin de pouvoir profiter pleinement de leurs potentiels scientifiques. En ce sens, le développement d'outils de traitement du signal adaptés à cette

nouvelle génération de radiotélescopes interféromètres est crucial à leur bon fonctionnement. Une observation radio-interférométrique est composée des corrélations calculées électroniquement pour chaque couple d'antenne du réseau interférométrique sur un temps d'intégration de l'ordre de quelques secondes [6]. Ainsi, l'observation complète, composée d'un certain nombre d'intégrations sur un temps total pouvant durer plusieurs heures, représente un volume de données important. En conséquence, la production d'une image radio-interférométrique est le fruit d'une chaîne de traitements de données complexes.

Pré-traitement des données : Une première étape de nettoyage des données est nécessaire pour éliminer ou filtrer les différentes sources d'interférences de radio-fréquence, appelés RFI pour *Radio-Frequency Interferences*. Les RFI correspondent principalement aux ondes radio d'origines humaines. Ces dernières ont un impact significatif sur les signaux mesurés du fait de leur forte puissance incidente comparée aux signaux célestes d'intérêts. À titre d'exemple, le futur réseau interférométrique SKA pourra effectuer des mesures allant jusqu'au μJy , où 1Jy correspond à $10^{-26}\text{ W}\cdot\text{m}^{-2}\cdot\text{Hz}^{-1}$ [5]. Le rapport signal sur bruit (SNR) dans le contexte de la radio-interférométrie est alors très peu favorable, de l'ordre de -30dB et pouvant aller jusqu'à -60dB [7]. En conséquence, les antennes des radio-interféromètres sont d'une sensibilité accrue, les rendant vulnérables à tous types d'interférences radio. Ainsi, malgré le choix d'emplacements isolés pour ce genre d'instruments, l'impact des RFI reste important. En particulier, les signaux de télécommunication de tous types sont la principale source de RFI. Ces derniers opèrent sur plusieurs canaux fréquentiels qui se superposent avec ceux d'intérêt à la radioastronomie [8]. La prise en compte de la variabilité fréquentielle et spatiale des RFI constitue un défi pour le développement des grands radiotélescopes interféromètres futurs [8]. En ce sens, plusieurs méthodologies sont appliquées dans le but d'atténuer leur impact sur les observations. Des algorithmes de détection sont notamment implémentés pour éliminer les mesures corrompues. Des opérations de filtrage permettent de prendre en compte efficacement l'impact des RFI de forte puissance. Néanmoins, les interférences de faible puissance restent difficilement détectables et peuvent nécessiter une inspection manuelle des données par les radioastronomes [9]. En effet, la forte puissance des signaux de télécommunication en fait notamment une source d'interférences significatives, même à très longue distance [10].

Calibration : Après un premier pré-traitement des données ayant pour but d'atténuer l'impact des RFI, une calibration des données est nécessaire. La calibration des radiotélescopes interféromètres consiste à prendre en compte l'ensemble des perturbations affectant les émissions radio provenant de sources célestes le long de leur chemin de propagation. Ces perturbations se manifestent par des transformations du front d'onde incident et peuvent être de causes multiples. Les différentes perturbations du front d'onde peuvent être modélisées par des transformations linéaires. Ainsi, la calibration des radio-interféromètres consiste à estimer l'ensemble des paramètres de ces transformations, afin de corriger les mesures. Les différents ef-

fets affectant une onde radio émise par une source céleste se partagent en deux catégories. i) Les effets indépendants de la direction d'observation correspondent principalement aux effets liés à l'instrument. Plus précisément, l'estimation des gains électroniques des antennes, de leur diagramme de rayonnement, ainsi que leur synchronisation temporelle sont nécessaires pour obtenir des mesures exploitables. Cette première étape de calibration s'effectue à partir d'observations sur des sources connues de forte intensité. ii) Dans un second temps, une calibration permettant d'ajuster les mesures aux effets dépendants de la direction est menée. En effet, le long de sa direction de propagation le front d'onde incident subit un certain nombre de transformations modifiant son amplitude, sa phase et sa polarisation. Dans le cas d'observations de sources lointaines, l'impact des sources brillantes de forte puissance proches de la direction d'observation doit être retiré. De plus, la présence d'un milieu ionisant le long du chemin de propagation altère l'onde. Ainsi, la propagation de l'onde dans l'ionosphère a un impact significatif. En effet, l'ionosphère constitue l'ensemble des couches de la haute atmosphère terrestre dans laquelle les ions et électrons ont une densité suffisante pour réfléchir les ondes électromagnétiques. Elle se situe entre 60 km et 1000 km d'altitude. L'onde radio incidente y est en partie réfléchi et subit une réfraction affectant sa phase et sa polarisation. Toutefois, du fait de la non-connaissance des sources présentes dans la direction d'observation, la calibration des mesures n'est pas triviale. Les solutions actuelles sont basées sur des boucles dites de *self-calibration* qui consistent à alterner itérativement entre l'estimation de la densité de flux dans la direction du ciel choisie et l'estimation des paramètres de calibration. Le problème de calibration des radio-interféromètres a été traduit en un problème de maximum de vraisemblance en considérant un bruit additif gaussien [11]. Néanmoins, La présence de bruits additionnels dans les données peut conduire à l'introduction d'un biais dans les solutions de calibration, conduisant à des artefacts dans les images reconstruites [12]. Ainsi, de nombreux travaux ont permis d'améliorer la modélisation des observations radioastronomiques en considérant des distributions robustes pour le bruit additif. Des algorithmes de maximum de vraisemblance ont alors été dérivés pour ces modélisations [13]-[20]. La présence de valeurs aberrantes pour les différentes antennes du réseau peut notamment être modélisée par une t -distribution complexe [15], [21]-[24]. Néanmoins, les approches existantes ne considèrent pas de modélisation statistique des RFI permettant de prendre en compte la variabilité spatiale et fréquentielle de ces dernières de manière indépendante.

Imagerie : Enfin la synthèse d'image est l'estimation de la densité de flux (en Jy) sur une portion de la voûte céleste. Le théorème de Van-Cittert Zernike montre que les visibilités mesurées sont les coefficients de la transformée de Fourier 2D de l'image du ciel [1]. Plus précisément, pour un couple d'antennes donné, la visibilité correspondante est une mesure du plan de Fourier 2D à des coordonnées continues définies par la ligne de base entre les deux capteurs. Le plan uv désigne alors le plan

de Fourier 2D. Ainsi, la couverture du plan uv est l'ensemble des coordonnées du plan de Fourier échantillonnées par le radio-interféromètre et dépend de la géométrie du réseau. En effet, un nombre plus élevé d'antennes résulte en une meilleure couverture du plan uv . Des antennes spatialement proches, donc à courte ligne de base, permettent de reconstruire les structures étendues des sources célestes en échantillonnant les basses fréquences du plan de Fourier 2D. De la même façon, les antennes à longue ligne de base permettent d'observer les structures fines de la voûte céleste. En conséquence, la meilleure résolution angulaire atteignable par un radiotélescope interféromètre est inversement proportionnelle à sa plus grande ligne de base.

Une première méthodologie de synthèse d'images radio-interférométriques consiste à appliquer une transformée de Fourier 2D inverse aux visibilitées mesurées. L'image résultante, appelée *dirty image*, contient des artefacts liés à la couverture partielle du plan uv . La *dirty image* peut s'exprimer comme la convolution de l'image du ciel avec la fonction d'étalement de point de l'instrument, ou *dirty beam*. C'est ainsi qu'initialement, la reconstruction d'image radio-interférométrique fut définie comme un problème de déconvolution. L'algorithme CLEAN, proposé par Högbom en 1974 [25], est l'une des méthodes couramment utilisées pour résoudre le problème d'imagerie du fait de sa simplicité d'implémentation et de son faible coût computationnel. L'algorithme consiste à modéliser l'image du ciel par une somme de sources ponctuelles, dont les positions et les intensités sont estimées itérativement à partir de la *dirty image*. Un certain nombre d'extensions de l'algorithme CLEAN ont été proposées dans le but d'améliorer les performances de l'algorithme dans le cas de sources non ponctuelles [26], [27]. Néanmoins, cette méthodologie présente des difficultés à reconstruire les structures étendues des sources célestes.

De ce fait, des approches basées sur de l'optimisation convexe ont été développées pour l'imagerie radio-interférométrique. Ces approches consistent à formuler le problème de reconstruction d'images comme un problème d'optimisation. Une régularisation permet alors de favoriser la parcimonie de l'image en adéquation avec la nature des sources célestes observées [28]-[31]. Ces approches permettent de mieux prendre en compte la parcimonie naturelle des images du ciel. Toutefois, les approches de l'état de l'art ne considèrent pas les perturbations affectant les visibilitées, telles que la présence de RFI dans les données.

1.2 Objectifs de la thèse

Les travaux présentés dans cette thèse visent à proposer des méthodologies de traitement statistique du signal permettant de prendre en compte l'impact d'interférences dans le contexte de la calibration et de l'imagerie radio-interférométrique. Pour cela, une modélisation statistique robuste des mesures provenant de radiotélescopes interféromètres est nécessaire. Ainsi, une modélisation statistique compatible avec la réa-

lité physique du système de mesures est proposée. Entre autres, nous considérerons la classe des distributions gaussiennes composées pour modéliser la présence d'interférences de radiofréquence dans les observations.

Néanmoins, le problème d'optimisation associé à la résolution du maximum de vraisemblance dans le cas de ces distributions robustes est non trivial et nécessite une attention particulière. En conséquence, cette thèse vise également à proposer des méthodes algorithmiques pour la résolution du maximum de vraisemblance dans le contexte de données radio-interférométriques. En particulier, les algorithmes proposés sont basés sur des algorithmes de type espérance-maximisation (EM, SAGE).

1.3 Contributions

Modélisation de données radio-interférométrique en présence de RFI

Nous proposons une modélisation statistique des RFI qui a abouti à la conception d'un algorithme de calibration multifréquence de radiotélescopes interféromètres robuste à la présence de RFI dans les données. En effet, l'étape de pré-traitement ne permet pas de détecter la totalité RFI de faible ou moyenne puissance, qui peuvent alors corrompre les données et affecter la qualité des images reconstruites. De plus, ce type de RFI peut être présent sur de nombreux canaux fréquentiels. Ainsi, le modèle proposé est basé sur un bruit de rang faible qui prend en compte la présence de RFI dans les données sur plusieurs canaux fréquentiels. En particulier, le bruit structuré proposé pour les RFI est dérivé du modèle physique d'un tel signal, prenant en compte leur variabilité fréquentielle et spatiale.

Publications associées : [J1], [C1]

Calibration multifréquentielle en présence de RFI

Les paramètres de calibration sont alors estimés par un algorithme du maximum de vraisemblance. Pour cela, nous proposons un algorithme d'Espérance Maximisation sur Espaces Alternés (SAGE) pour lequel nous dérivons originalement deux ensembles de données complètes permettant des expressions explicites pour les mises à jour. Des simulations numériques montrent un gain significatif en performance pour les données corrompues par des RFI en comparaison avec les algorithmes de calibration de l'état de l'art.

Publications associées : [J1], [C1]

Imagerie interférométrique en présence d'interférences

Nous proposons une nouvelle approche pour l'imagerie radio-interférométrique robuste à la présence d'interférences dans les données par un algorithme EM. La principale contribution réside dans l'utilisation de la classe des distributions gaussiennes composées pour modéliser la non gaussianité des données interférométriques. Dans le cas général, l'estimation

du maximum de vraisemblance résultante est intractable. Par conséquent, nous avons proposé un algorithme EM pour la modélisation proposée à partir duquel nous avons dérivé des expressions explicites pour les mises à jour dans le cas d'un bruit t -distribué. Nous avons étendu l'algorithme proposé en exploitant une factorisation de l'opérateur radio-interférométrique direct par le produit d'un opérateur de convolution avec un opérateur de Fourier 2D rapide (*Fast Fourier Transform*, FFT) sur une grille uniforme dans la conception de l'algorithme EM. L'algorithme résultant a une meilleure efficacité computationnelle. Des simulations numériques illustrent les performances des algorithmes proposés.

Publications associées : [J2], [C2]

Apprentissage informé pour l'imagerie en présence de RFI

Dans le cas où le modèle statistique proposé est trop complexe ou mal connu pour permettre une estimation par maximum de vraisemblance, nous proposons une approche d'apprentissage informé pour l'imagerie radio-interférométrique robuste à la présence d'interférences dans les données. Pour cela, nous proposons une architecture neuronale dépliée à partir de l'algorithme EM proposé dans la section 4.2. Des résultats préliminaires montrent l'intérêt de cette approche dans le contexte de l'imagerie radio-interférométrique.

Publications associées : [C3], [E1]

Imagerie multifréquentielle parcimonieuse en présence de RFI

Dans le cadre d'une collaboration avec des membres de l'Observatoire de Radioastronomie d'Afrique du Sud (SARAO), l'algorithme que nous avons proposé dans [J2] a été intégré à une librairie Python dédiée à l'imagerie de données radio-interférométriques de grande dimension. Le critère de vraisemblance est régularisé simultanément dans les domaines spectraux et d'ondelette, en adéquation la nature des sources imagées. Pour cela, nous avons considéré pour l'étape de maximisation un algorithme de type *forward-backward*, pour lequel un préconditionneur basé sur le calcul de la Hessienne permet une convergence accélérée.

Publication associée : [J2]

Traitement de données réelles

Nous illustrons les performances de l'algorithme proposé et de l'implémentation mise en oeuvre sur un jeu de données réelles provenant du *Very Large Array telescope* (VLA). Nous montrons notamment le bénéfice de la modélisation du bruit additif proposé et les performances de l'algorithme EM sur l'impact des interférences.

Publication associée : [E2]

1.4 Valorisations des travaux de recherche

1.4.1 Publications de revues

- [J1] **Mhiri, Y.**, El Korso, M. N., Breloy, A., Larzabal, P., « Multifrequency array calibration in presence of radio frequency interferences, » *Signal Processing*, t. 199, p. 108-113, 2022.
- [J2] **Mhiri, Y.**, El Korso, M. N., Breloy, A., Larzabal, P., « Regularized maximum likelihood estimation for radio interferometric imaging in the presence of radiofrequency interferences, » *Signal Processing*, Soumis en juin 2023.

1.4.2 Publications dans des conférences

- [C1] **Mhiri, Y.**, El Korso, M. N., Breloy, A., Larzabal, P., « Algorithme SAGE pour la calibration robuste de radio-interféromètres en présence d'interférences, » in *GRETSI*, 2022.
- [C2] **Mhiri, Y.**, El Korso, M. N., Breloy, A., Larzabal, P., « A Robust EM Algorithm for Radio Interferometric Imaging in The Presence of Outliers, » in *2022 IEEE Workshop on Signal Processing Systems (SiPS)*, IEEE, 2022, p. 1-5.
- [C3] **Mhiri, Y.**, El Korso, M. N., Breloy, A., Larzabal, P., « Imagerie radio-interférométrique robuste par dépliement neuronal, » in *GRETSI*, 2023.
- [C4] **Mhiri, Y.**, El Korso, M. N., Bacharach, L., Breloy, A., Larzabal, P., « Expectation-Maximization Based Direction of Arrival Estimation Under a Mixture of Noise, » in *2021 29th European Signal Processing Conference (EUSIPCO)*, IEEE, 2021, p. 1890-1893.
- [C5] Wang, J., Bacharach, L., **Mhiri, Y.**, El Korso, M. N., Larzabal, P., « Borne de Barankin vs Borne de Cramer-Rao pour la géométrie de réseaux d'antennes en radioastronomie, » in *GRETSI*, 2023.

1.4.3 Travaux en cours

- [E1] **Mhiri, Y.**, El Korso, M. N., Breloy, A., Larzabal, P., « Unrolled EM algorithm for radio-interferometric imaging in the presence of interference, » *en cours de rédaction*,
- [E2] **Mhiri, Y.**, Bester, L., Smirnov, O., « t -distribution based imaging in radiointerferometry : theory and validation on MeerKat data, » *en cours de rédaction*,

1.4.4 Séminaires

- [S1] 9 fév. 2023 : ***A Robust algorithm for radio interferometric imaging***, Séminaire S3, L2S Centrale-Supélec.
- [S2] 23 jan. 2023 : ***Traitement statistique du signal pour les grands radiotélescopes interféromètres***, Séminaire SATIE, ENS Paris-Saclay.

1.4.5 Intégrations logicielles

- [L1] ***pfb-clean***, Intégration de l’algorithme d’imagerie robuste à la présence d’interférences, proposé dans la librairie Python *pfb-clean* en collaboration avec le *Rhodes University Centre for Radio Astronomy Techniques and Technologies*. Adresse : <https://github.com/ratt-ru/pfb-clean>.

1.5 Organisation du document

Le chapitre 2 permet de positionner et motiver les travaux présentés dans cette thèse par rapport aux méthodes de l’état de l’art. Après une introduction aux principes de la radio-interférométrie, nous faisons état des méthodes de calibration et d’imagerie de la littérature, notamment concernant l’impact de RFI dans les observations. Les méthodes d’atténuation des RFI et les problématiques qui y sont associées sont ensuite présentées.

Le chapitre 3 présente une modélisation statistique de données interférométriques multifréquentielles permettant de prendre en compte la présence de RFI ainsi que la variabilité de leur impact sur les différents canaux fréquentiels. Ainsi, l’estimation des paramètres de calibration est effectuée par maximum de vraisemblance. Pour cela, nous introduisons un algorithme SAGE, pour lequel nous dérivons deux jeux de données complètes originaux permettant des mises à jour explicites des paramètres. Nous illustrons les performances de l’algorithme proposé sur des données corrompues simulées.

Le chapitre 4 présente deux algorithmes d’imagerie radio-interférométrique par maximum de vraisemblance régularisée pour des observations en présence de RFI. Ainsi, un modèle de données permettant de prendre en compte la présence d’interférences est dérivé. Un algorithme EM est alors exprimé pour résoudre le problème de maximum de vraisemblance associé. Nous proposons ensuite une extension de l’algorithme proposé en considérant un second jeu de données complètes en considérant comme information manquante des visibilités virtuelles. Cela résulte en une étape de maximisation pouvant être résolue efficacement en tirant partie de l’algorithme de FFT. Des simulations numériques illustrent les performances des algorithmes proposés par rapport aux méthodes de l’état de l’art. Enfin, nous présentons une architecture neuronale dépliée à partir de l’algorithme EM proposé (Algorithme 4.1). L’architecture déroulée proposée est composée d’un réseau neuronal récurrent qui remplace l’étape d’espérance. L’architecture est entraînée sur des données simulées affectées par des valeurs aberrantes qui sont tirées d’un mélange de bruit gaussien composé. Des résultats préliminaires montrent l’intérêt de cette approche dans le contexte de l’imagerie radio-interférométrique.

Le chapitre 5 présente l’intégration de l’algorithme d’imagerie précédemment proposé dans le cadre du développement de la librairie Python *Preconditioned forward-backward clean algorithm (pfb-clean)*, en collaboration avec Landman Bester de l’Observatoire Sud-Africain de Radio Astronomie (SRAO). Après avoir présenté l’écosystème logiciel en radioastronomie, ainsi que des considérations pratiques pour la prise en main de données réelles, nous exposons les résultats obtenus pour une des méthodes

d'imagerie proposées. Nous présentons d'abord un algorithme *forward-backward* pré-conditionné pour l'imagerie radio-interférométrique qui sera utilisé pour l'étape de maximisation de l'algorithme EM proposé ainsi que la librairie Python *pfbclean* dédiée au traitement de données radio-interférométriques de grande dimension. Nous présentons enfin les résultats obtenus en appliquant la méthode proposée sur des données réelles provenant du *Very Large Array telescope* permettant d'illustrer son efficacité en présence d'interférences.

Enfin, **le chapitre 6** conclut ce manuscrit et présente les perspectives et travaux en cours.

2 - État de l'art

2.1 Introduction à la radio-interférométrie

Les sources cosmiques d'intérêts observées en radioastronomie sont de natures diverses et permettent d'obtenir des informations sur les phénomènes physiques en jeu dans l'univers. Parmi les phénomènes observables dans le domaine des ondes radio, on distingue les rayonnements à spectres continus des rayonnements sur raies spectrales. Dans le cas de spectres continus, on parle alors de sources radio qui peuvent être observées sur une large gamme de fréquences. Ces dernières sont notamment générées par des émissions synchrotrons, propres aux rayonnements issus d'électrons de hautes énergies en mouvement dans un champ magnétique, mais peuvent aussi correspondre à un rayonnement thermique de corps noir [1]. Les rayonnements sur raies spectrales sont quant à eux émis sur des bandes spectrales étroites. Ce sont principalement des rayonnements générés par des transitions atomiques ou moléculaires. La raie de 21 cm de l'hydrogène neutre est un exemple de raie spectrale étudiée en radioastronomie [1].

La puissance d'un signal provenant d'une source cosmique est exprimée par sa densité de flux, mesurée en $\text{W}\cdot\text{m}^{-2}\cdot\text{Hz}^{-1}$. Du fait de la faible puissance des signaux cosmiques, les radioastronomes utilisent une unité de mesure spécifique, le Jansky (Jy) où $1 \text{ Jy} = 10^{-26} \text{ W}\cdot\text{m}^{-2}\cdot\text{Hz}^{-1}$. Lorsque l'on cherche à produire une image d'une source radio, on mesure en réalité la densité de flux par angle solide. On parle alors d'*intensité* qui se mesure donc en Jy/sr (Jansky par stéradian) [1].

Un radiotélescope interféromètre est un réseau d'antennes. Les signaux mesurés par les différentes antennes sont ensuite combinés électroniquement pour produire une estimée des corrélations pour chaque couple d'antennes. Concrètement, le signal radiofréquence incident est divisé sur plusieurs sous-bandes de quelques centaines de kHz puis converti en signal complexe sur bande de base. Les corrélations sont ensuite calculées électroniquement pour chaque couple d'antennes sur chaque canal fréquentiel. Pour un radio-interféromètre composé de P antennes, la corrélation mesurée sur un canal fréquentiel entre 2 antennes d'indice p et q s'écrit alors,

$$v_{pq} = \frac{1}{T} \sum_{k=1}^T e_p(t_k) e_q(t_k)^*, \quad (2.1)$$

où $e_p(t_k)$ et $e_q(t_k)$ représentent les $k^{\text{ème}}$ échantillons de l'enveloppe complexe des signaux mesurés aux bornes des $p^{\text{ème}}$ et $q^{\text{ème}}$ antennes. Le nombre d'échantillons temporels utilisés dans le calcul des corrélations, T , est choisi de sorte à assurer la stationnarité des signaux observés et correspond à un temps d'intégration pouvant varier de quelques millisecondes à quelques dizaines de secondes [1]. La corrélation mesurée, v_{pq} , est appelée *visibilité*. Ainsi, le vecteur de mesure associé au radio-interféromètre s'écrit comme la concaténation de l'ensemble des visibilités mesurées pour l'ensemble

des couples d'antennes pouvant être formés par le réseau,

$$\mathbf{v} = \begin{bmatrix} v_{12} \\ v_{13} \\ \vdots \\ v_{(P-1)P} \end{bmatrix}. \quad (2.2)$$

De la même façon, l'ensemble des corrélations mesurées par le radiotélescope peut être exprimé à travers la matrice de corrélation des mesures,

$$\mathbf{R} = \frac{1}{T} \sum_{k=1}^T \mathbf{e}(t_k) \mathbf{e}(t_k)^H, \quad (2.3)$$

$$v_{pq} = [\mathbf{R}]_{p,q},$$

où $\mathbf{e}(t) = [e_1(t), \dots, e_P(t)]$ est le vecteur contenant l'ensemble des signaux mesurés par le réseau d'antennes à l'instant t . Pour un réseau interféromètre composé de P antennes, le nombre de couples d'antennes et donc de visibilité mesurées est

$$K = \frac{P(P-1)}{2}. \quad (2.4)$$

On notera que les auto-corrélations qui correspondent aux coefficients diagonaux de la matrice de corrélation, \mathbf{R} , ne sont pas prises en compte du fait du fort impact du bruit sur ces coefficients [32]. La corrélation mesurée entre les signaux incidents pour un couple d'antennes p et q s'exprime par la cohérence mutuelle du champ électrique, E , mesurée en deux points de l'espace, \mathbf{r}_p et \mathbf{r}_q , qui correspondent aux positions des antennes p et q [1],

$$v_{pq} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E(t, \mathbf{r}_p) E^*(t, \mathbf{r}_q) dt. \quad (2.5)$$

Les sources radios observées par les radiotélescopes peuvent être considérées comme incohérentes, quasi monochromatiques et se trouvant dans le champ lointain des antennes. Dans ce cas, l'onde incidente peut-être considérée plane [1]. Ainsi, les délais de propagation induits par la géométrie de réseau se matérialisent par un déphasage du signal reçu, conduisant à la formulation suivante du modèle de propagation,

$$v_{pq} = \int_{\mathcal{S}^2} \frac{\mathcal{A}(l, m) \mathcal{I}(l, m)}{n} e^{-j2\pi(lu_{pq} + mv_{pq} + nw_{pq})} dl dm, \quad (2.6)$$

où l et m , $n = \sqrt{1 - l^2 - m^2}$ sont des cosinus directeurs et \mathcal{S}^2 est la sphère unité. La réponse des antennes du réseau est notée $\mathcal{A}(l, m)$, $\mathcal{I}(l, m)$ est la distribution d'intensité de la source observée et u_{pq} , v_{pq} , w_{pq} , sont les coordonnées du vecteur ligne de base $\mathbf{b}_{pq} = \mathbf{r}_p - \mathbf{r}_q \in \mathbb{R}^3$. Ils s'expriment en unité de longueur d'onde, λ , ainsi,

$$u_{pq} = \frac{\mathbf{b}_{pq} \cdot \mathbf{e}_u}{\lambda}, \quad v_{pq} = \frac{\mathbf{b}_{pq} \cdot \mathbf{e}_v}{\lambda}, \quad w_{pq} = \frac{\mathbf{b}_{pq} \cdot \mathbf{e}_w}{\lambda}, \quad (2.7)$$

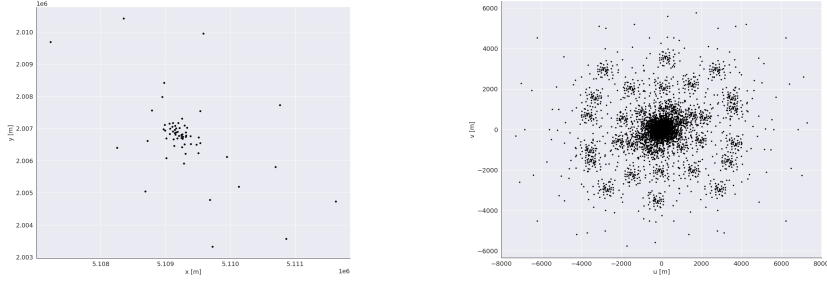


Figure 2.1 – (a) Position dans le repère terrestre ITRF des 64 antennes du radiotélescope MeerKat [33]. (b) Couverture du plan uv du radiotélescope MeerKat permettant d'échantillonner 2016 visibilités.

où (e_u, e_v, e_w) est une base de \mathbb{R}^3 , appelée plan uv , tel que e_w soit colinéaire à la direction d'observation. La figure 2.1 montre les positions des antennes et le plan uv associé au radiotélescope Meerkat pour une observation au zénith [33].

Le théorème de Van-Cittert Zernike relie les visibilités mesurées par un radio-interféromètre à la transformée de Fourier 2D de la distribution d'intensité de la source observée sous les hypothèses suivantes :

- Le champ de vue considéré est étroit de sorte à pouvoir supposer $l, m \ll 1$ et donc $n = \sqrt{1 - l^2 - m^2} \approx 1$.
- Les antennes du réseau interféromètre se trouvent sur le même plan. Ainsi, pour tout couple d'antennes (p, q) , le vecteur ligne de base $\mathbf{b}_{pq} = \mathbf{r}_p - \mathbf{r}_q$ est orthogonal à la direction de propagation du champ électrique et la composante w du plan uv est alors nulle.

L'équation 2.5 peut alors être réécrite comme suit,

$$v_{pq} = \int_{S^2} \mathcal{A}(l, m) \mathcal{I}(l, m) e^{-j2\pi(lu_p + mv_q)} dl dm. \quad (2.8)$$

L'équation (2.8) est une transformée de Fourier 2D, ainsi, par abus de langage, le plan uv désigne alors le plan de Fourier à 2 dimensions généré par les vecteurs unitaires (e_u, e_v) . Un radio-interféromètre permet donc d'échantillonner des coefficients de Fourier de l'image du ciel observé, en fonction de la géométrie du réseau d'antennes. Dans le but de reconstruire une image du ciel, l'équation (2.8) est discrétisée sur une grille uniforme des directions du ciel. On note alors \mathbf{x} l'image du ciel composée de N pixels,

$$\mathbf{x} = [\mathcal{A}(l_0, m_0) \mathcal{I}(l_0, m_0), \dots, \mathcal{A}(l_N, m_N) \mathcal{I}(l_N, m_N)]^T. \quad (2.9)$$

Le théorème de Van-Cittert Zernike peut alors être réécrit sous la forme d'un modèle linéaire où le vecteur de mesure \mathbf{v} est relié à l'image du ciel \mathbf{x} par une transformée de Fourier discrète 2D (DFT),

$$\mathbf{v}_{pq} = \sum_{n=0}^N \mathbf{x}[n] e^{-j2\pi(l_n u_{pq} + m_n v_{pq})}, \quad (2.10)$$

$$\mathbf{v} = \mathbf{H} \mathbf{x}, \quad (2.11)$$

où $[\mathbf{H}]_{pq,n} = e^{-j2\pi(l_n u_{pq} + m_n v_{pq})}$. Le vecteur de mesure \mathbf{v} peut alors s'exprimer par le produit de Hadamard, noté \odot , de la transformée de Fourier de l'image du ciel $\tilde{\mathbf{x}}$ avec une fonction d'échantillonnage \mathcal{M} ,

$$\mathbf{v} = \mathcal{M} \odot \tilde{\mathbf{x}}. \quad (2.12)$$

La *dirty image* est l'image obtenue en appliquant une transformée de Fourier inverse discrète (DFT^{-1}) sur les visibilitées mesurées par le radio-interféromètre et peut notamment s'exprimer comme un produit de convolution,

$$\mathbf{x}_{\text{dirty}} = DFT^{-1}\{\mathbf{v}\} = \mathbf{H}^H \mathbf{v}, \quad (2.13)$$

$$= DFT^{-1}\{\mathcal{M} \odot \tilde{\mathbf{x}}\} = \mathbf{x} * \mathcal{B}, \quad (2.14)$$

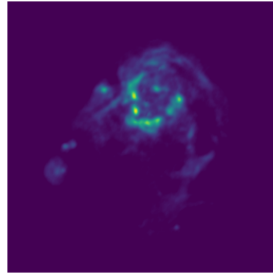
où $\mathcal{B} = DFT^{-1}\{\mathcal{M}\}$ est appelé le *dirty beam* et représente la fonction d'étalement de point (*Point Spread Function*, PSF) du réseau interféromètre.

En conséquence, plus la couverture du plan uv est dense, plus l'information spectrale de l'image du ciel est riche. Toutefois, la couverture du plan uv est limitée par la géométrie du réseau d'antennes. En effet, la plus haute fréquence spatiale échantillonnée par le réseau est déterminée par la plus longue ligne de base du réseau et la plus basse fréquence spatiale échantillonnée par la plus courte ligne de base du réseau. Dans le but d'augmenter la couverture du plan uv , les radioastronomes ont alors recours à de la *synthèse d'image par rotation de la Terre* (*Earth synthesis imaging*). Il s'agit de tirer partie de la stationnarité des sources observées pour effectuer des observations sur de longues durées. En effet, la projection des vecteurs de lignes de bases sur le plan uv est modifiée au cours du temps du fait de la rotation de la Terre sur elle-même. En conséquence, la couverture du plan uv est augmentée en considérant la concaténation des visibilitées mesurées à différents instants. Ce procédé est illustré sur la figure 2.2 par la couverture du plan uv du *Very Large Array* (VLA) pour une observation simulé de la radiogalaxie *M31*.

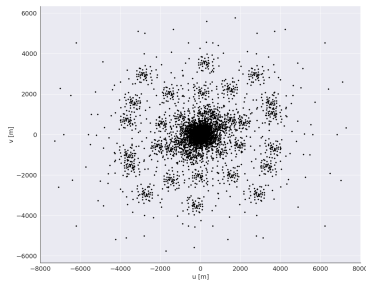
Cela résulte en une forte augmentation du volume de données à traiter. Notamment, le calcul de la DFT par un produit matriciel est coûteux en mémoire et en temps de calcul. Ces considérations ont conduit à approximer le calcul de la DFT par une interpolation linéaire suivie d'une FFT. Cette opération, appelée *gridding*, revient à factoriser l'opérateur de DFT ainsi,

$$\mathbf{v} = \mathbf{H}\mathbf{x} \approx \mathbf{S}\mathbf{F}\mathbf{x}, \quad (2.15)$$

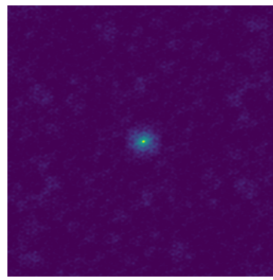
où \mathbf{S} est un opérateur d'interpolation et \mathbf{F} est l'opérateur de FFT. Les visibilitées mesurées sont donc interpolées sur une grille régulière du plan uv sur laquelle une FFT peut être appliquée. L'interpolation est effectuée par une convolution des visibilitées mesurées avec un noyau judicieusement choisi [34]. Cette technique permet alors de réduire drastiquement le coût en mémoire et en temps de calcul. De plus, l'erreur d'approximation de cette factorisation a été fortement réduite au cours des dernières décennies par le choix de noyaux de convolution adaptés et par l'utilisation de méthodes d'interpolation plus précises [34], [35].



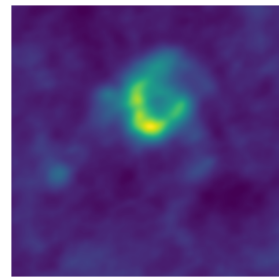
(a) Image haute résolution de la radiogalaxie M31.



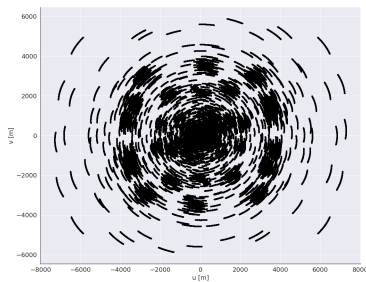
(b) plan uv instantané.



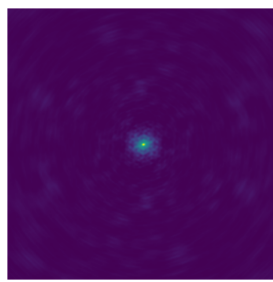
(c) PSF instantané.



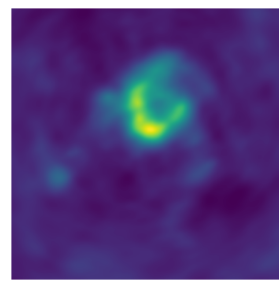
(d) *Dirty image* instantané.



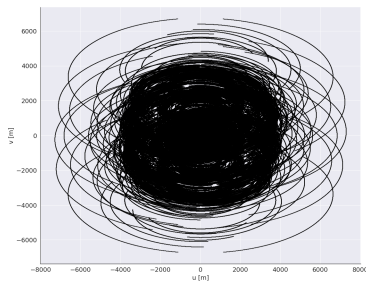
(e) plan uv après 1 heure.



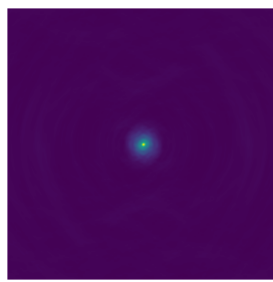
(f) PSF après 1 heure.



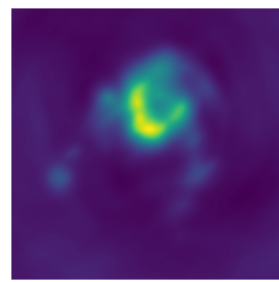
(g) *Dirty image* après 1 heure.



(h) plan uv après 8 heures.



(i) PSF après 8 heures.



(j) *Dirty image* après 8 heures.

Figure 2.2 – Effet de la rotation de la Terre sur la couverture du plan uv et son impact sur la *dirty image*. (a) Image haute résolution de la radiogalaxie M31. (b) à (j) montrent le plan uv , la PSF et la *dirty image* pour plusieurs temps d'observation.

Jusqu'à présent nous avons considéré que les signaux reçus par les antennes étaient des signaux scalaires. En réalité, les antennes des radiotélescopes interféromètres sont à polarisation duale. Ainsi, les réponses à deux polarisations orthogonales sont mesurées, ce qui résulte en des signaux complexes de dimension deux,

$$\mathbf{e}_p(t) = \begin{bmatrix} e_{a_p}(t) \\ e_{b_p}(t) \end{bmatrix}, \quad (2.16)$$

où $p \in \{1, 2, \dots, P\}$ est l'indice de l'antenne du réseau et a et b sont les indices des deux polarisations orthogonales (verticale et horizontale par exemple).

Les paramètres de Stokes sont quatre scalaires, I , Q , U et V , qui permettent de décrire l'état de polarisation d'une onde électromagnétique. Le paramètre I correspond à l'intensité totale de la source et ne dépend pas de sa polarisation. Les paramètres Q et U décrivent la proportion de polarisation linéaire de la source et le paramètre V décrit la proportion de polarisation circulaire. Pour une source radio monochromatique d'enveloppe complexe notée \mathbf{s} , on définit alors sa matrice de cohérence en fonction de ses paramètres de Stokes, qui dans le cas d'antennes à polarisations verticale et horizontale s'écrit comme suit,

$$\mathbf{C} = \mathbb{E} [\mathbf{s}\mathbf{s}^H] = \begin{bmatrix} I + Q & U + jV \\ U - jV & I - Q \end{bmatrix}. \quad (2.17)$$

En conséquence, les visibilité mesurées pour chaque couple d'antennes sont des matrices complexes de dimension deux,

$$\mathbf{V}_{pq} = \mathbb{E} [\mathbf{e}_p(t)\mathbf{e}_q^H(t)] = \begin{bmatrix} \mathbb{E} [e_{p,x}(t)e_{q,x}^*(t)] & \mathbb{E} [e_{p,x}(t)e_{q,y}^*(t)] \\ \mathbb{E} [e_{p,y}(t)e_{q,x}^*(t)] & \mathbb{E} [e_{p,y}(t)e_{q,y}^*(t)] \end{bmatrix}. \quad (2.18)$$

Le théorème de Van Cittert-Zernike défini en (2.8) peut-être généralisé à des signaux polarisés en considérant les matrices de cohérence pour chaque direction du ciel [36],

$$\mathbf{V}_{pq} = \int_{S^2} \mathbf{C}(l, m) e^{-2j\pi(lu_{pq} + mv_{pq})} dl dm. \quad (2.19)$$

Ainsi, les relations définies plus haut se retrouvent alors en considérant la *visibilité en intensité*,

$$v_{pq} = \frac{1}{2}([\mathbf{V}_{pq}]_{11} + [\mathbf{V}_{pq}]_{22}). \quad (2.20)$$

Les sources cosmiques observées sont en général considérées non polarisées et l'image reconstruite est alors une image en intensité.

2.2 Méthodes de calibration en radio-interférométrie

2.2.1 Contexte

Le long de son chemin de propagation, une onde électromagnétique subit un certain nombre d'effets perturbatifs affectant son amplitude, sa phase et son état de polarisation. Dans ce contexte, la calibration des radiotélescopes est une étape cruciale dans le but de reconstruire des images de sources célestes.

Parmi les effets perturbatifs affectant l'onde incidente, on fera notamment la distinction entre les effets instrumentaux et les effets environnementaux. Les effets instrumentaux correspondent à l'ensemble des gains affectant le signal incident le long de la chaîne de traitement des données. En effet, les antennes des radio-interféromètres sont définies par leur diagramme de rayonnement qui décrit la réponse de l'antenne pour chaque direction du ciel. Une calibration précise de ce diagramme de rayonnement est alors nécessaire pour prendre en compte la présence de lobes secondaires dans la réponse des antennes, afin de limiter l'influence de sources voisines [36]-[38]. Par ailleurs, la variation des conditions de température et de pression affecte les propriétés des antennes et induit des variations des gains, pour lesquels une calibration est également nécessaire.

La troposphère et l'ionosphère sont des couches spécifiques de l'atmosphère terrestre qui constituent la principale source de perturbations environnementales [39]. En effet, la forte quantité d'électrons ionisés dans ces milieux induit un certain nombre d'effets perturbatifs sur le front d'onde. La réfraction de l'onde incidente conduit à des déphasages et des délais de propagation qui doivent être pris en compte dans le traitement des données. La rotation de Faraday est la conséquence d'indices de réfraction différents pour les deux composantes de polarisation circulaire de l'onde incidente, conduisant à une rotation de l'état de polarisation de l'onde. L'ensemble de ces effets perturbatifs sont dépendant de la fréquence d'observation et de la quantité d'électrons ionisés dans l'atmosphère, mesuré en TEC (*Total Electron Content*). Une calibration des effets ionosphériques est alors nécessaire sur des intervalles réguliers du fait de la forte variabilité du TEC au cours du temps [39], [40].

La calibration des radiotélescopes consiste alors à estimer et corriger les effets affectant le front d'onde incident le long de son chemin de propagation. Pour cela, une première approche consiste à estimer les différents effets instrumentaux et environnementaux à partir de l'observation de sources de références brillantes. Cependant, les directions d'observation des sources observées ne coïncident pas nécessairement avec les directions des sources de calibration. Ainsi, du fait de la variabilité en fonction de la direction d'observation des effets environnementaux, notamment liés aux passages de l'onde dans l'atmosphère, des stratégies de calibration plus élaborées ont été développées. Les boucles de *self-calibration*, dont le principe est présenté en figure 2.3, consistent à affiner les solutions de calibration obtenues à partir de sources de références. Pour cela, une calibration est effectuée de manière itérative à partir de l'estimé courant du modèle de ciel pour la direction d'observation. L'étape de calibration est primordiale dans la chaîne de traitement d'un radiotélescope interféromètre. En effet, comme illustré sur la figure 2.4, une mauvaise calibration conduira inévitablement à une image reconstruite fortement dégradée.

2.2.2 Le formalisme de Jones

Les principaux phénomènes perturbatifs affectant l'onde incidente peuvent être représentés comme des transformations linéaires de l'enveloppe complexe de l'onde in-

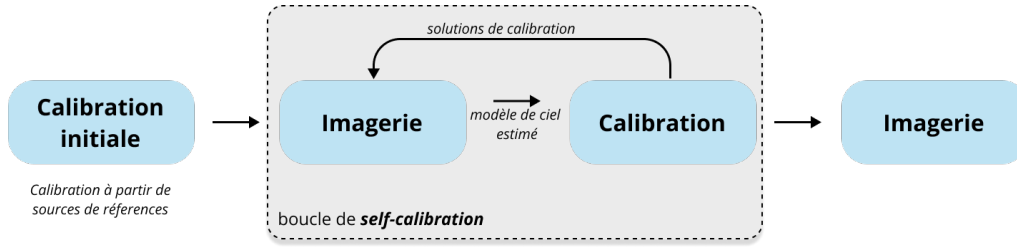
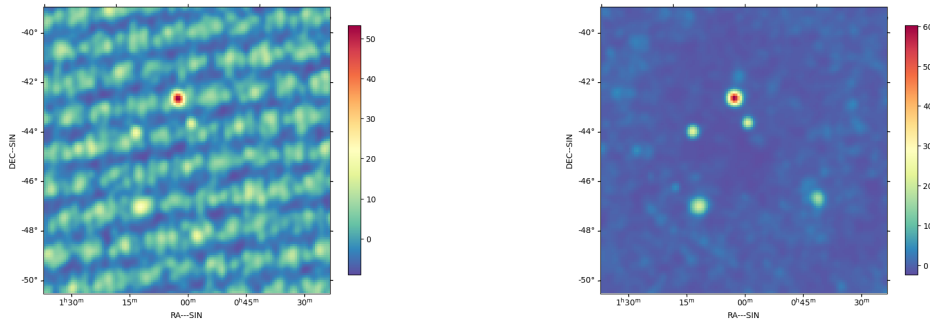


Figure 2.3 – Boucle de *self-calibration*



(a) *dirty image* reconstruite à partir de visibilités non calibrées.

(b) *dirty image* reconstruite à partir de visibilités calibrées.

Figure 2.4 – Effets d’une mauvaise calibration

cidente. Cela a conduit au développement du *formalisme de Jones* qui permet de modéliser l’ensemble des perturbations affectant le front d’onde le long de son chemin de propagation par des transformations linéaires paramétrées.

Considérons le cas d’une source ponctuelle monochromatique ayant une matrice de cohérence C observée par un radiotélescope interféromètre composé de P antennes. Les antennes étant à polarisation duale, le *formalisme de Jones* permet d’écrire le signal incident aux bornes d’une antenne p comme suit,

$$\mathbf{e}_p(t) = \mathbf{J}_p \mathbf{s}(t), \quad (2.21)$$

où \mathbf{J}_p est la matrice de Jones associée à l’antenne p pour la direction d’observation correspondant à la source observée. L’intérêt du formalisme de Jones réside dans la factorisation de la matrice de Jones, \mathbf{J}_p , en un produit de κ sous-matrices de Jones structurées, $\mathbf{J}_p^{[i]}$, chacune associée à un effet perturbatif spécifique,

$$\mathbf{J}_p = \mathbf{J}_p^1 \mathbf{J}_p^2 \dots \mathbf{J}_p^\kappa. \quad (2.22)$$

On distingue ainsi des structures de Jones associées aux différents phénomènes pouvant affecter le front d’onde. Un déphasage affectant les deux polarisations de manière identique est modélisé par une matrice scalaire de la forme,

$$\mathbf{K}(\phi) = e^{i\phi} \mathbf{I}_2, \quad (2.23)$$

où ϕ est le déphasage affectant les deux polarisations. Les gains électroniques, et plus généralement tous gains affectant les deux polarisations de manière indépendante, sont matérialisés par une matrice diagonale de la forme,

$$\mathbf{G} = \begin{bmatrix} g_a & 0 \\ 0 & g_b \end{bmatrix}, \quad (2.24)$$

où g_a et g_b sont les gains associés à chaque polarisation.

Les matrices de Jones contenant des éléments off-diagonaux non nuls modélisent des effets affectant la polarisation de l'onde incidente. En particulier, la rotation de Faraday, conséquence de la propagation de l'onde incidente dans l'ionosphère, est modélisée par une matrice de rotation de la forme,

$$\mathbf{F}(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}, \quad (2.25)$$

où θ est l'angle de rotation de la polarisation. Par ailleurs, une matrice 2×2 quelconque permettra de modéliser une transformation quelconque du front d'onde.

Le formalisme de Jones permet de considérer deux types de calibration. Une calibration dite *structurée* consiste à modéliser les effets perturbatifs de manière indépendante par des matrices de Jones structurées et à en estimer les paramètres. Dans ce cas, le nombre de paramètres à estimer est fortement réduit du fait de la structure imposée sur les matrices de Jones. Toutefois, cette approche nécessite de connaître les différentes sources de perturbations affectant le front d'onde, ainsi que la structure des matrices de perturbations correspondantes. Des erreurs de modélisation conduisent alors à des biais systématiques sur les paramètres de calibration estimés.

Une calibration dite *non-structurée* consiste à modéliser les effets perturbatifs par des matrices de Jones générique de type (2.21) et à en estimer les coefficients. Cela conduit à une modélisation des données plus générale ne nécessitant pas de modélisation précise des effets perturbatifs mais demandant l'estimation d'un plus grand nombre de paramètres. Par ailleurs, il est possible de combiner les deux approches, notamment en raffinant l'estimation obtenue initialement par une calibration non-structurée. Une calibration structurée permet alors de prendre en compte des effets perturbatifs spécifiques [13]. Dans ces travaux, nous considérons une forme non structurée des matrices de Jones.

2.2.3 Modèle de données

Dans le contexte de la calibration des radio-interféromètres, le ciel est supposé composé de D sources connues, dites de *calibration*, indépendantes entre elles. Le modèle sur les signaux observés aux bornes de la $p^{\text{ème}}$ antenne s'écrit alors pour un canal fréquentiel,

$$e_p(t) = \sum_{i=1}^D \mathbf{J}_{i,p} \mathbf{s}_i(t), \quad (2.26)$$

où $\mathbf{J}_{i,p}$ est la matrice de Jones associée à la $i^{\text{ème}}$ antenne pour la direction d'observation correspondant à la $i^{\text{ème}}$ source d'enveloppe complexe $s_i(t)$. Les sources étant indépendantes entre elles, la matrice de visibilité pour un couple d'antennes d'indice (p, q) s'exprime alors,

$$\mathbf{V}_{pq} = \mathbb{E} [\mathbf{e}_p \mathbf{e}_q^H] = \sum_{i=1}^D \mathbf{J}_{i,p} \mathbf{C}_i \mathbf{J}_{i,q}^H, \quad (2.27)$$

où $\mathbf{C}_i = \mathbb{E} [s_i s_i^H]$ est la matrice de cohérence de la $i^{\text{ème}}$ source. Un bruit additif est introduit dans le modèle pour tenir compte des erreurs de mesures et des sources de bruits non modélisées, conduisant au modèle suivant,

$$\mathbf{Y}_{pq} = \mathbf{V}_{pq} + \mathbf{N}_{pq}, \quad (2.28)$$

où \mathbf{Y}_{pq} est la matrice de visibilité bruitées et \mathbf{N}_{pq} la matrice de bruit associée au couple d'antennes (p, q) . Il est courant dans la littérature de considérer une vectorisation du modèle défini en (2.28),

$$\mathbf{y}_{pq} = \sum_{i=1}^D \mathbf{v}_{pq,i}(\mathbf{J}) + \mathbf{n}_{pq}, \quad (2.29)$$

$$\mathbf{v}_{pq,i}(\mathbf{J}) = (\mathbf{J}_{i,q}^* \otimes \mathbf{J}_{i,p}) \mathbf{c}_i, \quad (2.30)$$

où $\mathbf{c}_i = \text{vec}(\mathbf{C}_i)$ et \otimes indique le produit de kronecker. Le vecteur de mesure se définit alors comme la concaténation des visibilité mesurées, \mathbf{y}_{pq} pour tous les couples d'antennes du réseau interféromètre,

$$\mathbf{y} = [\mathbf{y}_{12}^T, \mathbf{y}_{13}^T, \dots, \mathbf{y}_{1P}^T, \mathbf{y}_{23}^T, \dots, \mathbf{y}_{(P-1)P}^T]^T \in \mathbb{C}^{4P(P-1) \times 1}. \quad (2.31)$$

2.2.4 Calibration par maximum de vraisemblance

En présupposant un modèle statistique sur les observations, l'estimation des matrices de Jones peut être effectuée par maximum de vraisemblance. Ainsi, dans le cas d'un bruit additif gaussien, le modèle de données considéré est le suivant,

$$\mathbf{y} = \sum_{i=1}^D \mathbf{v}_i(\mathbf{J}_i) + \mathbf{n}, \quad (2.32)$$

où $\mathbf{v}_i(\mathbf{J}) = [\mathbf{v}_{12,i}(\mathbf{J})^T, \dots, \mathbf{v}_{1N,i}(\mathbf{J})^T, \dots, \mathbf{v}_{N(N-1),i}(\mathbf{J})^T]^T$ est la concaténation des visibilité provenant de la $i^{\text{ème}}$ source de calibration et $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$. Cela conduit à la log-vraisemblance suivante,

$$\mathcal{L}(\mathbf{J}; \mathbf{y}) = -\log \sigma^2 - \frac{1}{\sigma^2} \|\mathbf{y} - \sum_{i=1}^D \mathbf{v}_i(\mathbf{J}_i)\|_2^2. \quad (2.33)$$

Le problème de maximum de vraisemblance associé peut être formulé comme un problème de moindres carrés,

$$\hat{\mathbf{J}} = \underset{\mathbf{J}}{\text{argmin}} \|\mathbf{y} - \sum_{i=1}^D \mathbf{v}_i(\mathbf{J}_i)\|_2^2. \quad (2.34)$$

Le problème d'optimisation résultant est de nature non linéaire et est notamment résolu par des algorithmes de type Gauss-Newton ou Levenberg-Marquardt [41]. Ces approches furent améliorées par l'introduction d'algorithmes de types Espérance-Maximisation (EM). Nous présentons notamment l'algorithme EM proposé dans [11] pour la calibration de radio-interféromètres ayant permis d'améliorer la précision et la vitesse de convergence des algorithmes de calibration dans le cas d'observations monofréquentielles pour un bruit additif gaussien.

Pour cela, un espace de données complètes est défini comme étant la concaténation des visibilité correspondant à chaque source de calibration,

$$\boldsymbol{\xi} = [\mathbf{w}_1^T, \dots, \mathbf{w}_D^T]^T, \quad (2.35)$$

où $\mathbf{w}_i = \mathbf{v}_i(\mathbf{J}_i) + \mathbf{n}_i$, et \mathbf{n}_i correspond à la contribution de la $i^{\text{ème}}$ source au bruit global. Le vecteur de mesure observé s'exprime alors en fonction des données complètes comme suit,

$$\mathbf{y} = \sum_{i=1}^D \mathbf{w}_i, \quad \mathbf{n} = \sum_{i=1}^D \mathbf{n}_i. \quad (2.36)$$

Les différentes contributions au bruit, \mathbf{n}_i sont supposées indépendantes et distribuées selon une gaussienne de moyenne nulle et de covariance définie par,

$$\mathbb{E} [\mathbf{n}_i \mathbf{n}_j^H] = \beta_i \delta_{ij} \sigma^2 \mathbf{I}, \quad \beta_i \in [0, 1], \quad i \in \llbracket 1, D \rrbracket, \quad j \in \llbracket 1, D \rrbracket, \quad \sum_{i=1}^D \beta_i = 1. \quad (2.37)$$

Ainsi, à partir de cette définition des données complètes, l'estimation du maximum de vraisemblance est menée par l'algorithme EM suivant :

$$\begin{aligned} \text{Étape-E:} \quad \hat{\mathbf{w}}_i^{(m)} &= \mathbb{E}_{\mathbf{w}_i | \mathbf{y}, \mathbf{J}_i} [\mathbf{w}_i]. \\ \text{Étape-M:} \quad \hat{\mathbf{J}}_i^{(m)} &= \underset{\mathbf{J}_i}{\operatorname{argmin}} \|\hat{\mathbf{w}}_i^{(m)} - \mathbf{v}_i(\mathbf{J}_i)\|_2^2. \end{aligned} \quad (2.38)$$

Le problème d'optimisation résultant reste non linéaire et est résolu par un algorithme de descente de gradient de type Levenberg-Marquardt [41]. L'utilisation d'algorithmes EM dans la calibration de radio-interféromètres en présence de bruits gaussiens a permis de réduire la charge de calculs nécessaire et d'améliorer la précision et la vitesse de convergence des algorithmes de calibration de l'état de l'art [11].

2.2.5 Calibration en présence de bruits non gaussiens

L'hypothèse de gaussiannité du bruit additif dans le contexte de la calibration des radiotélescopes fut remise en question dans des travaux récents de la littérature [13], [21], [42]. En effet, il a été montré que les visibilité mesurées, issues de l'observation des sources de calibration, contiennent des valeurs aberrantes de multiples natures. L'impact de sources voisines de faible puissance ou de sources brillantes, apparaissant au travers des lobes secondaires des antennes, constitue une source de perturbations

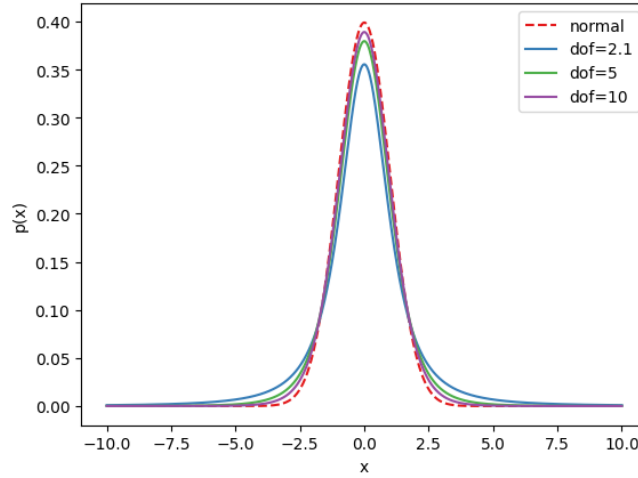


Figure 2.5 – Densité de probabilité d’une distribution gaussienne et d’une t -distribution pour plusieurs valeurs du degré de liberté. La queue de la distribution dans le cas des gaussiennes composées est plus élevée, affectant donc des probabilités plus importantes à la présence de valeurs aberrantes dans les données.

ne pouvant être modélisée par une simple distribution gaussienne. Aussi, RFI, qui correspondent aux émissions d’ondes radio d’origine humaine, ne peuvent pas être filtrées ou détectées dans le cas d’interférences de puissance du même ordre de grandeur que les sources de calibration [21].

Ces considérations ont conduit à l’utilisation de distributions issues de la classe des gaussiennes composées pour la modélisation du bruit additif [13], [21]. Ces dernières sont définies par le produit de deux composantes : la *speckle* qui est un bruit gaussien de moyenne nulle, et la *texture* qui est un réel aléatoire positif. Le bruit associé au vecteur de visibilités du couple d’antennes (p, q) s’écrit alors,

$$\mathbf{n}_{pq} = \sqrt{\tau_{pq}} \mathbf{t}_{pq}, \quad (2.39)$$

où $\mathbf{t} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ et $\tau \sim p(\tau)$ dénotent respectivement la *speckle* et la *texture*. La distribution de la texture peut alors être interprétée comme étant la distribution a priori de la variance de la composante de *speckle*. L’utilisation de telles distributions est notamment motivée par leur capacité à modéliser des distributions à queues lourdes. En effet, dans [13], il est montré que la présence de sources voisines de faible puissance dans les mesures conduit à élever la queue de la distribution du bruit. Dans [21], la t -distribution, qui constitue un cas particulier des gaussiennes composées, est utilisée notamment pour atténuer l’impact de RFI de faible puissance dans les données.

La résolution du problème de maximum de vraisemblance associé dans le cas de ces distributions conduit généralement à des problèmes d’optimisation intractables. Il a été proposé, dans une première approche, de recourir à un algorithme EM pour l’estimation des paramètres de calibration et du bruit de manière jointe dans le cas d’un bruit modélisé par une t -distribution [21], [42]. Pour cela, le paramètre de *texture*

est incorporé dans le jeu de données complètes,

$$\boldsymbol{\xi} = [\tau_{12}, \dots, \tau_{P(P-1)}, \mathbf{y}]. \quad (2.40)$$

L'étape-E consiste alors à calculer l'espérance des *textures* sachant le vecteur de mesure et l'estimé courant des matrices de Jones à la $m^{\text{ème}}$ itération,

$$\hat{\tau}_{pq}^{(m)} = \mathbb{E}_{\tau_{pq} | \mathbf{y}, \hat{\mathbf{J}}^{(m)}} [\tau_{pq}^{-1}]. \quad (2.41)$$

Dans le cas de la t -distribution de degré de liberté ν , (2.41) peut se calculer de manière explicite [43], conduisant à l'étape-E suivante,

$$\hat{\tau}_{pq}^{(m)} = \frac{\nu + 1}{\nu + \|\mathbf{y}_{pq} - \mathbf{v}_{pq}(\hat{\mathbf{J}}^{(m)})\|_2^2}. \quad (2.42)$$

L'étape-M de l'algorithme EM consiste alors à résoudre un problème de moindres carrés pondérées utilisant l'espérance des textures précédemment évaluées,

$$\hat{\mathbf{J}}^{(m)} = \underset{\mathbf{J}}{\operatorname{argmin}} \sum_{pq} \hat{\tau}_{pq}^{(m)} \|\mathbf{y}_{pq} - \mathbf{v}_{pq}(\mathbf{J}_i)\|_2^2. \quad (2.43)$$

Une seconde approche propose d'effectuer l'estimation des paramètres du bruit (*texture* et covariance du *speckle*) et des matrices de Jones indépendamment de manière itérative [13]. Pour cela, des estimateurs de Maximum à Posteriori (MAP) ont été dérivés pour plusieurs distributions de *texture* (Gamma, Inverse Gamma, Inverse gaussienne, Exponentielle, Cauchy). L'estimation des matrices de Jones se fait ensuite par un algorithme EM considérant le jeu de données complètes précédemment présenté.

L'application de ces méthodes de calibration sur des données réelles a permis de montrer la pertinence des distributions gaussiennes composées dans la modélisation des perturbations additives en radio-interférométrie [13], [21].

2.2.6 Calibration multifréquentielle

Nous avons considéré jusqu'à présent une calibration sur un canal fréquentiel. Les observations radio-interférométriques sont en réalité effectuées sur plusieurs canaux fréquentiels et une calibration multifréquentielle est alors pertinente. En effet, les perturbations affectant le front d'onde ainsi que les contributions des bruits ont une dépendance fréquentielle qui ne peut être négligée. À titre d'exemple, les déphasages et délais de propagation causés par la réfraction de l'onde incidente dans l'ionosphère sont fortement liés à la fréquence d'observation.

Le vecteur de mesure dans le cas d'observations multifréquentielles se matérialise par la concaténation des vecteurs de mesures correspondant à chaque canal fréquentiel,

$$\mathbf{y} = [\mathbf{y}^{[1]T}, \dots, \mathbf{y}^{[Q]T}]^T \in \mathbb{C}^{4P(P-1)Q \times 1}, \quad (2.44)$$

où Q est le nombre de canaux fréquentiels. Il est alors possible d'effectuer une calibration multifréquentielle en introduisant une dépendance fréquentielle aux matrices de

Jones. Ainsi, la visibilité sans bruit pour le $j^{\text{ème}}$ canal fréquentiel s'écrit,

$$\mathbf{V}_{pq}^{[j]} = \sum_{i=1}^D \mathbf{J}_{i,p}^{[j]} \mathbf{C}_i^{[j]} \mathbf{J}_{i,q}^{[j]H}, \quad (2.45)$$

où $\mathbf{J}_{i,p}^{[j]}$ est la matrice de Jones associée à la $p^{\text{ème}}$ antenne pour la direction d'observation correspondant à la $i^{\text{ème}}$ source et pour le $j^{\text{ème}}$ canal fréquentiel et $\mathbf{C}_i^{[j]}$ est la matrice de cohérence de la $i^{\text{ème}}$ source pour le $j^{\text{ème}}$ canal fréquentiel.

Toutefois, cette approche conduit à un nombre de paramètres à estimer trop important relativement au nombre de visibilités observées. Il est alors courant de modéliser les variations spectrales des paramètres de calibration par des polynômes de degré faible [13], [14], [44],

$$\begin{aligned} \mathbf{J}_{i,p}^{[j]} &= \sum_{c=1}^{\zeta} \mathbf{Z}_{i,p,c} b_c(f_j), \\ b_c(f_j) &= \left(\frac{f_j - f_0}{f_0} \right)^{c-1}, \end{aligned} \quad (2.46)$$

où f_0 est la fréquence centrale des observations, ζ est le degré du polynôme et $\mathbf{Z}_{i,p,c}$ correspond au $c^{\text{ème}}$ coefficient du polynôme. L'équation (2.46) peut se réécrire comme un produit matriciel en considérant la matrice $\mathbf{Z}_{i,p} = [\mathbf{Z}_{i,p,1}, \dots, \mathbf{Z}_{i,p,\zeta}]$,

$$\begin{aligned} \mathbf{J}_{i,p}^{[j]} &= \mathbf{B}^{[j]} \mathbf{Z}_{i,p}, \\ \mathbf{B}^{[j]} &= [b_1(f_j), \dots, b_{\zeta}(f_j)] \otimes \mathbf{I}_2. \end{aligned} \quad (2.47)$$

Des stratégies de calibration multifréquentielle basées sur l'algorithme des directions alternées (ADMM) ont permis d'obtenir des algorithmes distribués à grande échelle, adaptés au grand volume de données imposé par les observations radio-interférométriques [14], [44]. Plus précisément dans [44], le problème de calibration, énoncé pour une modélisation gaussienne du bruit additif, est formulé comme un problème d'optimisation sous contrainte comme suit,

$$\hat{\mathbf{J}}^{[1]}, \dots, \hat{\mathbf{J}}^{[Q]}, \hat{\mathbf{Z}} = \underset{\mathbf{J}^{[1]}, \dots, \mathbf{J}^{[Q]}, \mathbf{Z}}{\operatorname{argmin}} \sum_{j=1}^Q g_j(\mathbf{J}^{[j]}) \quad \text{st } \forall j \in \llbracket 1, Q \rrbracket \mathbf{J}^{[j]} = \mathbf{B}^{[j]} \mathbf{Z}_{i,p}, \quad (2.48)$$

où $g_j(\mathbf{J}^{[j]}) = \sum_{pq} \|\mathbf{y}_{pq}^{[j]} - \mathbf{v}_{pq}^{[j]}(\mathbf{J}^{[j]})\|_2^2$. Le problème (2.48) est résolu par une méthode de Lagrangien Augmenté pour lequel un algorithme ADMM multi-agent est mis en œuvre [44], [45]. La méthodologie proposée permet une amélioration dans le coût computationnel de la calibration multifréquentielle par une distribution en fréquence. Par ailleurs, les résultats obtenus confirment la pertinence de la modélisation polynomiale effectuée sur la dépendance en fréquence des matrices de Jones.

De la même façon, [13] propose un algorithme de calibration distribué pour la calibration multifréquentielle en considérant une modélisation robuste du bruit additif. La méthodologie proposée considère un bruit additif gaussien composé. Les paramètres

de la distribution du bruit et les solutions de calibration sont estimés de manière alternée. Les paramètres du bruit sont estimés de manière analogue au modèle décrit en (2.39) et les solutions de calibration sont obtenues par un algorithme de type ADMM distribué [46].

2.3 Algorithmes d'imagerie en radio-interférométrie

Nous nous intéressons dans cette section aux méthodes d'imagerie radio-interférométrique en intensité, c'est-à-dire sans considérer les composantes polarisées des visibilités. Il s'agit alors d'obtenir une estimée de la distribution en intensité du champ de vue observé par le réseau d'antennes. Pour cela, nous considérons donc les visibilités en intensité, telles que définies en (2.20). Par ailleurs, nous nous plaçons sous les conditions du théorème de Van-Cittert Zernike, explicitées en (2.6). En conséquence, les visibilités mesurées correspondent à des coefficients de la DFT de l'image 2D du ciel observé et s'expriment comme suit,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (2.49)$$

où $\mathbf{H} \in \mathbb{C}^{K \times N}$ est l'opérateur de DFT, $\mathbf{x} \in \mathbb{R}^{N \times 1}$ représentent l'image du ciel que l'on cherche à reconstruire et \mathbf{n} est un bruit additif.

2.3.1 L'algorithme CLEAN

L'algorithme CLEAN est l'une des méthodes de reconstruction d'images les plus utilisées en radio-interférométrie. Il s'agit d'un algorithme de déconvolution itératif agissant sur la *dirty image*. Le principe général de l'algorithme, proposé originellement par Högbom en 1974 [25], réside dans l'hypothèse d'un ciel composé d'un nombre limité de sources ponctuelles. L'impact de chaque source, vue à travers la PSF de l'instrument, est soustrait jusqu'à obtenir un résiduel ne contenant plus que le bruit de mesure. En d'autres termes, l'algorithme vise à éliminer l'impact des lobes secondaires de la réponse du radio-interféromètre dans la *dirty image* de manière itérative. La popularité de l'algorithme CLEAN est notamment due à sa simplicité d'implémentation et son coût computationnel avantageux. En effet, le coût calculatoire à chaque itération ne dépend pas de la dimension du vecteur de visibilités, mais seulement de la taille de la *dirty image* et de la PSF choisie.

À chaque itération, la position du pic de l'image résiduelle (l_d, m_d) est déterminée, puis la puissance pour la direction considérée, $\hat{\sigma}_d^2$ est calculée en prenant en compte l'effet de la PSF. Une fraction, γ , de l'image contenant seulement une source à la direction déterminée, $\hat{\mathbf{x}}_q(l, m) = \gamma \hat{\sigma}_d^2 \delta_{l, l_d} \delta_{m, m_d}$ convoluée à la PSF, \mathcal{B} , est ensuite soustraite de l'image résiduelle. Le coefficient $\gamma \leq 1$ est appelé *gain de boucle* et définit la contribution de la source courante à l'image résiduelle. Il est en pratique choisi relativement petit afin d'atténuer l'impact des lobes secondaires sur l'image [25]. L'objectif est alors de minimiser l'énergie restante dans l'image résiduelle. La convergence est atteinte lorsque l'image résiduelle atteint un certain niveau de bruit. Le modèle de ciel est

l'image contenant l'ensemble des sources ponctuelles, $\hat{x} = \sum_d \hat{x}_d$. Enfin, l'image finale est calculée en ajoutant à l'image résiduelle le modèle de ciel convolué avec une PSF idéale, \mathcal{B}^* , généralement une gaussienne, permettant de lisser les artefacts causés par l'hypothèse de sources ponctuelles. Les étapes de l'algorithme CLEAN sont résumées dans l'algorithme 2.1.

Algorithm 2.1 L'algorithme CLEAN

input : $x_{\text{dirty}}, \mathcal{B}, \gamma, \epsilon$

output : x_{clean}, r

- 1: $r \leftarrow x_{\text{dirty}}$
 - 2: $\hat{x} \leftarrow \mathbf{0}$
 - 3: **while** $\|r\|_2^2 \leq \epsilon$ **do**
 - 4: $(l_d, m_d) \leftarrow \underset{k}{\operatorname{argmax}} r(l_k, m_k)$
 - 5: $\hat{\sigma}_d^2 \leftarrow r(l_d, m_d) / \mathcal{B}[0, 0]$
 - 6: $\hat{x}_d(l, m) \leftarrow \gamma \hat{\sigma}_d^2 \delta_{l, l_d} \delta_{m, m_d}$
 - 7: $\hat{x} \leftarrow \hat{x} + \hat{x}_d$
 - 8: $r \leftarrow r - \hat{x}_d * \mathcal{B}$
 - 9: **end while**
 - 10: $x_{\text{clean}} \leftarrow r + \hat{x} * \mathcal{B}^*$
-

En effet, l'hypothèse de sources ponctuelles effectuée ne correspond pas à la réalité des sources célestes observées. Des variantes de l'algorithme CLEAN ont alors été proposées pour pallier ces difficultés tout en gardant une efficacité computationnelle avantageuse. En particulier, des versions de CLEAN multi-échelles [27] permettent d'obtenir des images reconstruites compatibles avec des émissions de sources diffuses et ont conduit à une implémentation couramment utilisée décrite dans [26]. Pour cela, chaque source est supposée centrée et d'étendue spatiale définie par un facteur d'échelle, conduisant au modèle de ciel suivant,

$$x(l, m) = \sum_d \sigma_d^2 K(l - l_d, m - m_d, \alpha_d), \quad (2.50)$$

où $K(\cdot)$ est un noyau symétrique de facteur d'échelle α_d pour la $d^{\text{ème}}$ source. L'algorithme CLEAN Multi-échelle propose alors de suivre le principe originel de CLEAN en considérant le modèle multi-échelle. Ainsi, à chaque itération, la recherche du maximum se fait parmi un set d'échelle $\alpha = [\alpha_1, \dots, \alpha_M]$. Les étapes de l'algorithme CLEAN multi-échelle sont résumées dans l'algorithme 2.2.

2.3.2 Imagerie radio-interférométrique par optimisation convexe

Néanmoins, malgré ses extensions récentes, l'algorithme CLEAN reste limité par les hypothèses fortes qu'il suppose, ce qui a conduit à l'émergence de méthodes basées sur de l'optimisation convexe pour l'imagerie radio-interférométrique. En effet, la formulation du problème d'imagerie comme un problème d'optimisation convexe a permis de

Algorithm 2.2 Algorithmme CLEAN Multi-échelle

input : $\mathbf{x}_{\text{dirty}}, \mathcal{B}, \gamma, \alpha, K(\cdot), \epsilon$ **output** : $\mathbf{x}_{\text{clean}}, \mathbf{r}$

```
1:  $\hat{\mathbf{x}} \leftarrow \mathbf{0}$ 
2:  $\mathbf{r} \leftarrow \mathbf{x}_{\text{dirty}}$ 
3: while  $\|\mathbf{r}\|_2^2 \leq \epsilon$  do
4:   for  $\alpha \in \alpha$  do
5:      $\mathbf{r}_\alpha \leftarrow K(\alpha) * \mathbf{r}$ 
6:      $p_\alpha \leftarrow \max(\mathbf{r}_\alpha)$ 
7:   end for
8:    $\tilde{\alpha} \leftarrow \operatorname{argmax}_\alpha \{p_\alpha\}$ 
9:    $(l_d, m_d) \leftarrow \operatorname{argmax}_k \mathbf{r}_{\tilde{\alpha}}[l_k, m_k]$ 
10:   $\hat{\mathbf{x}}_d(l, m) \leftarrow \gamma \hat{\sigma}_d^2 \delta_{l, l_d} \delta_{m, m_d} * K(\tilde{\alpha})$ 
11:   $\hat{\mathbf{x}} \leftarrow \hat{\mathbf{x}} + \hat{\mathbf{x}}_d$ 
12:   $\mathbf{r} \leftarrow \mathbf{r} - \hat{\mathbf{x}}_d * \mathcal{B}$ 
13: end while
14:  $\mathbf{x}_{\text{clean}} \leftarrow \mathbf{r} + \hat{\mathbf{x}} * \mathcal{B}^*$ 
```

considérer des a priori variés et plus proches de la réalité physique des signaux observés. En particulier, un a priori parcimonieux paraît adapté à la nature des images du ciel pour lesquelles l'énergie se retrouve concentrée sur des directions spécifiques [28], [47]. Le problème d'imagerie est alors formulé comme un problème d'optimisation convexe en incluant un terme d'attache aux données et un terme de régularisation,

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + r(\mathbf{x}), \quad (2.51)$$

où $r : \mathbb{R}^{N \times 1} \mapsto \mathbb{R}$ est une fonction de régularisation convexe définie par l'information a priori sur le signal observé. Le terme d'attache aux données considéré dans les approches proposées se base sur la norme euclidienne de l'erreur de modèle,

$$f(\mathbf{x}; \mathbf{y}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2. \quad (2.52)$$

Un signal est dit parcimonieux s'il peut être représenté dans un dictionnaire par peu de coefficients non nuls. Des travaux ont permis de montrer la pertinence de ce type de représentations dans le cadre de la radioastronomie [28], [31], [47]. La norme ℓ_1 permet d'obtenir une telle régularisation convexe en promouvant les solutions ayant peu de coefficients non nuls,

$$r(\mathbf{x}) = \alpha \|\Psi \mathbf{x}\|_1, \quad (2.53)$$

où $\Psi \in \mathbb{C}^{d \times N}$ est un dictionnaire sur lequel l'image \mathbf{x} est représentée et $\alpha > 0$ un paramètre de régularisation. La base de Dirac, $\Psi = \mathbf{I}_N$, représente un cas particulier où l'image est supposée contenir peu de coefficients non nuls et permet notamment de représenter les sources célestes d'intérêts en radioastronomie [28]. Par ailleurs, plusieurs

bases ont été proposées pour représenter la parcimonie des signaux célestes. Dans [47], une transformée en ondelette non décimée isotropique est considérée pour prendre en compte la nature des signaux célestes. Dans [31], une concaténation de plusieurs bases orthonormales est proposée comme alternative pour représenter la parcimonie de diverses structures célestes. En effet, les sources ponctuelles sont bien représentées en considérant une base de Dirac alors que des structures diffuses sont mieux représentées en considérant une parcimonie sur le gradient de l'image, notamment par l'utilisation de base d'ondelette de Haar. De plus, les approches basées sur de l'optimisation convexe pour l'imagerie radio-interférométrique permettent de contraindre les solutions à des images qui correspondent à une réalité physique. En particulier, une contrainte de positivité sur les solutions permet de garantir la réalité physique des images reconstruites dont les coefficients correspondent à des puissances. La régularisation associée se définit par la fonction indicatrice de l'ensemble des réels positifs,

$$\iota_{\mathbb{R}^+}(\mathbf{x}) = \begin{cases} 0 & \text{si } \mathbf{x} \geq 0 \\ +\infty & \text{sinon.} \end{cases} \quad (2.54)$$

Le problème de minimisation défini en (2.51) peut-être résolu par des algorithmes de gradient proximaux [48], [49]. Ce type de méthodes se place dans le cadre où la fonction objective à minimiser peut être décomposée en une somme de sous-problèmes. On s'intéresse notamment aux problèmes de la forme,

$$\underset{\mathbf{x}}{\text{minimiser}} f(\mathbf{x}) + g(\mathbf{x}), \quad (2.55)$$

où $f : \mathbb{R}^N \mapsto \mathbb{R}$, $g : \mathbb{R}^N \mapsto \mathbb{R} \cup \{+\infty\}$ sont des fonctions convexes et f est différentiable sur son espace de définition. Les méthodes proximales se basent sur l'utilisation de l'opérateur de proximité défini pour une fonction f convexe par,

$$\text{prox}_f(\mathbf{x}) = \underset{\mathbf{u}}{\text{argmin}} f(\mathbf{u}) + \frac{1}{2} \|\mathbf{x} - \mathbf{u}\|_2^2. \quad (2.56)$$

L'algorithme de gradient proximal est alors défini en actualisant la solution courante de (2.55) ainsi,

$$\mathbf{x}^{(m+1)} = \text{prox}_{\gamma^{(m)}g}(\mathbf{x}^{(m)} - \gamma^{(m)}\nabla f(\mathbf{x}^{(m)})). \quad (2.57)$$

Dans le cas où le dictionnaire considéré est une base orthonormale de sorte que $\Psi^H \Psi = \mathbf{I}_N$, le problème d'optimisation présenté en (2.51) peut être reformulé afin d'être résolu en considérant l'opérateur de proximité associé à la norme ℓ_1 ,

$$\hat{\mathbf{x}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \|\mathbf{y} - \mathbf{H}\Psi^H \boldsymbol{\theta}\|_2^2 + \alpha \|\boldsymbol{\theta}\|_1. \quad (2.58)$$

L'algorithme de gradient proximal correspondant s'apparente à un algorithme ISTA (voir Algorithme 2.3) [50],

$$\boldsymbol{\theta}^{(m)} = \mathcal{T}_\alpha \left(\boldsymbol{\theta}^{(m)} + \gamma \Psi \mathbf{H}^H (\mathbf{y} - \mathbf{H}\Psi^H \boldsymbol{\theta}^{(m)}) \right), \quad (2.59)$$

Algorithm 2.3 Algorithme ISTA pour l'imagerie radio-interférométrique [47]

input : $x_0, \mathbf{y}, \Psi, \alpha, \gamma$ **output :** \hat{x}

$$\hat{\theta} \leftarrow \Psi x_0$$

while critère d'arrêt non atteint **do**

$$\hat{\theta} \leftarrow \mathcal{T}_\alpha \left(\hat{\theta} + \gamma \Psi \mathbf{H}^H (\mathbf{y} - \mathbf{H} \Psi^H \hat{\theta}) \right)$$

end while

$$\hat{x} \leftarrow \Psi^H \hat{\theta}$$

où $\mathcal{T}_\alpha(\cdot)$ est l'opérateur de proximité associé à la norme ℓ_1 qui s'exprime par un seuillage doux,

$$\mathcal{T}_\alpha(\mathbf{x}) = \max(\mathbf{0}, |\mathbf{x}| - \alpha). \quad (2.60)$$

Alternativement, dans le cas où le dictionnaire ne constitue pas une base complète, le problème (2.51) n'est pas équivalent à (2.58). Des méthodes d'optimisation convexe permettent alors de résoudre le problème de minimisation [45], [48],

$$\underset{\mathbf{x}}{\text{minimiser}} f(\mathbf{x}) + g(\Psi \mathbf{x}). \quad (2.61)$$

Des méthodes basées sur le Lagrangien Augmenté permettent notamment de résoudre le problème de minimisation sous contrainte associée,

$$\underset{\mathbf{x}}{\text{minimiser}} f(\mathbf{x}) + g(\mathbf{z}) \quad \text{st } \mathbf{z} = \Psi \mathbf{x}. \quad (2.62)$$

Plus généralement, des méthodes considérant (2.58) comme un problème primal et définissant un problème dual associé, tel que défini dans [48], conduisent à des algorithmes permettant de résoudre (2.58) à grande échelle. En effet, des travaux ont permis d'implémenter ce type d'algorithme en tirant avantage de la forte séparabilité des fonctions objectives permettant une implémentation largement distribuée [29], [30], [51], [52]. Néanmoins, le principal inconvénient des méthodes d'imagerie présentées dans cette section réside dans leur coût calculatoire et le temps de convergence nécessaire pour obtenir une solution. En effet, les algorithmes itératifs, mis en œuvre dans la résolution des problèmes d'optimisation formulés, nécessitent d'évaluer à chaque itération le gradient du terme d'attache aux données exprimées en (2.51),

$$\nabla_{\mathbf{x}} f = \mathbf{H}^H (\mathbf{y} - \mathbf{H} \mathbf{x}). \quad (2.63)$$

Cette évaluation dépend de la dimension du vecteur de mesure, qui dans le contexte de la radio-interférométrie, est suffisamment élevé pour rendre le calcul du gradient prohibitif. Ainsi, des stratégies ont été développées pour rendre l'implémentation de ce type de méthodes possible. Dans [29], un préconditionneur est proposé pour réduire le nombre d'itérations nécessaires pour atteindre la convergence. Aussi, une approche distribuée basée sur un ADMM proposé dans [30], [52] permet de répartir les calculs

de manière efficace au sein de plusieurs agents. Enfin, une implémentation efficace du calcul du gradient est proposée dans [53].

Les méthodes présentées pour l'imagerie radio-interférométrique considèrent dans leur ensemble un terme d'attache aux données formulées par une norme ℓ_2 , en considérant un bruit additif gaussien [28]. Néanmoins, les radio-interféromètres sont soumis à un certain nombre d'effets perturbatifs ne pouvant être modélisés par un bruit additif gaussien. Nous discutons notamment dans la section suivante des interférences de radiofréquence pour lesquelles les méthodes d'atténuation ne sont pas suffisantes pour éliminer leur impact sur les données.

2.4 Méthodes d'atténuation des interférences de radiofréquence

Les interférences de radiofréquence sont l'ensemble des signaux interférents provenant des activités humaines. Les observations en radioastronomie ont lieu sur des bandes de fréquence dédiées pour lesquelles les émissions radio sont rares, voire restreintes. Néanmoins, un certain nombre de signaux interférents affectent les mesures de radioastronomie de manière significative. La plupart des RFI résultent de signaux de télécommunications provenant notamment de satellites de géolocalisation, de communications aériennes ou de réseaux de téléphonie mobile. Ces signaux interférents ont une structure particulière due à leurs utilisations et sont concentrés sur des bandes spectrales précises plus ou moins étroites. Aussi, un certain nombre de RFI proviennent d'émissions radio non intentionnelles, causées par des systèmes électriques voisins. Ce type d'émissions, notamment issues de lignes à haute tension, prennent la forme d'impulsions électromagnétiques et présentent un fort étalement spectral. Les RFI sont captées par les antennes du réseau à travers leurs lobes secondaires. Ainsi, du fait de leur forte directivité, leur impact n'est pas homogène sur toutes les antennes du réseau et n'est donc pas éliminé dans le calcul des corrélations.

L'impact de signaux interférents peut aller de la saturation des mesures dans le cas de RFI de forte puissance à des distorsions subtiles dans le cas de RFI de faible puissance. En particulier, les RFI de faible puissance à bande étroite peuvent être confondues avec des émissions sur raies spectrales provenant de sources célestes.

Une première solution mise en place pour atténuer l'impact des RFI sur les mesures de radioastronomie consiste à les détecter afin d'éliminer les données corrompues. Ainsi, l'outil présenté dans [54] permet d'inspecter manuellement des données radio-interférométriques à la recherche de RFI. En effet, les RFI de forte puissance ont tendance à saturer les antennes du radiotélescope et sont facilement détectables. Des méthodes algorithmiques basées sur des tests de vraisemblance ont été développées pour automatiser en grande partie la recherche et l'élimination de RFI dans les mesures [9]. Toutefois, cette approche conduit à une perte d'informations de part la suppression des données corrompues. Des méthodes permettant de réduire l'impact des RFI sur les bandes de fréquence affectées, sans recourir à une suppression de données, ont alors été développées. Nous présentons ici le modèle de données pour des RFI sur bande

étroite puis une méthode d'atténuation des RFI par filtrage spatial [8], [10].

2.4.1 Modèle de données des interférences

Une interférence de radiofréquence est modélisée par une source ponctuelle. Ainsi, pour un radio-interféromètre composé de P antennes, le signal en bande de base reçu aux bornes de la $p^{\text{ème}}$ antenne s'exprime,

$$e_p(t) = a_p s(t - \tau_p) e^{-j2\pi f_c \tau_p}, \quad (2.64)$$

où $s(t)$ représente l'enveloppe complexe du signal interférent, a_p est un gain modélisant l'atténuation du signal pour la $p^{\text{ème}}$ antenne et τ_p est le délai de propagation induit par la géométrie du réseau. Sans hypothèses sur la distance séparant les antennes de la source interférente, les différents gains d'atténuation n'ont pas de raison d'être identiques. Ce modèle peut-être simplifié si l'hypothèse de bande étroite est respectée, c'est-à-dire si le produit du délai de propagation, τ_p , avec la largeur de bande des canaux fréquentiels du système, Δ_f , respecte la condition suivante,

$$\Delta_f \tau_p \ll 1. \quad (2.65)$$

À titre d'exemple, deux antennes séparées de 3km conduiront à un délai de $10\mu s$ et donc à une condition sur la largeur de bande $\Delta_f \ll 100$ kHz qui apparaît réaliste au vue des radiotélescopes existants [6], [33]. Sous ces conditions, l'enveloppe complexe du signal interférent varie peu sur un intervalle de temps de l'ordre de τ_p et on peut donc écrire $s(t - \tau_p) \approx s(t)$. Ce qui conduit au modèle suivant sur le vecteur de mesure $\mathbf{e}(t) = [e_1(t), \dots, e_p(t)]^T$,

$$\mathbf{e}(t) = \mathbf{a}s(t) + \mathbf{n}, \quad (2.66)$$

où $\mathbf{a} = [a_1 e^{-j2\pi f_c \tau_1}, \dots, a_P e^{-j2\pi f_c \tau_P}]^T$ est la réponse du réseau interféromètre au signal interférent, appelée aussi *signature spatiale* et $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ un bruit additif gaussien. Comme formulé en (2.3), les visibilité mesurées correspondent aux entrées de l'estimé de la matrice de corrélation du vecteur de mesures, $\mathbf{e}(t)$, obtenu par moyennage,

$$\mathbf{R}_s = \mathbb{E}[\mathbf{e}\mathbf{e}^H] = \sigma_s^2 \mathbf{a}\mathbf{a}^H + \sigma^2 \mathbf{I}, \quad (2.67)$$

$$\hat{\mathbf{R}}_s = \frac{1}{T} \sum_{m=1}^T \mathbf{e}(t_m)\mathbf{e}(t_m)^H, \quad (2.68)$$

où σ_s^2 est la puissance de la source interférente. Le modèle sur la matrice de corrélation exprimé en (2.68) se généralise en présence de L sources interférentes indépendantes de signature spatiale, $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_L]$,

$$\mathbf{R}_S = \mathbf{A}\Sigma_S\mathbf{A}^H + \sigma^2 \mathbf{I}. \quad (2.69)$$

L'estimation des signatures spatiales des sources interférentes peut se faire par des antennes de références pointées dans les directions des sources à soustraire [8]. Néanmoins, dans la situation où les signatures spatiales ne sont pas disponibles, il est courant d'utiliser les vecteurs propres associés aux plus grandes valeurs propres de la matrice de corrélation. En effet, la composante de la matrice de corrélation liée aux sources

interférentes, $\mathbf{A}\Sigma_s\mathbf{A}^H$, est une matrice de rang L . La décomposition en valeurs propres de la matrice de corrélation s'écrit alors,

$$\mathbf{R}_S = \mathbf{U}\Delta\mathbf{U}^H = [\mathbf{U}_s \ \mathbf{U}_n] \begin{bmatrix} \Delta_s + \sigma^2\mathbf{I}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{P-L} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix}, \quad (2.70)$$

où $\Delta_s = \text{diag}(\lambda_1, \dots, \lambda_L)$ est la matrice diagonale contenant les valeurs propres des composantes liées aux sources interférentes et \mathbf{U} est la matrice orthonormale des vecteurs propres correspondants. La décomposition en valeurs propres de la matrice de corrélation permet alors d'identifier le sous-espace lié aux RFI, défini par \mathbf{U}_s [55], [56]. Ainsi, les valeurs propres liées aux sources interférentes peuvent être identifiées en effectuant un seuillage sur les valeurs propres de la matrice de corrélation.

2.4.2 Méthode de filtrage des RFI

La matrice de corrélation globale incluant les signaux célestes et une source interférente s'écrit,

$$\mathbf{R} = \mathbf{R}_c + \sigma_s^2 \mathbf{a}\mathbf{a}^H \sigma^2 \mathbf{I}, \quad (2.71)$$

où \mathbf{R}_c est la matrice de corrélation issue des sources célestes considérées indépendantes des sources interférentes et de faible puissance, comparée au signal interférent. Les méthodes de filtrage spatial visent à annuler l'énergie reliée à la signature spatiale des sources interférentes par la projection définie par les matrices,

$$\begin{aligned} \mathbf{P}_a &= \mathbf{a}(\mathbf{a}^H \mathbf{a})^{-1} \mathbf{a}^H, \\ \mathbf{P}_a^\perp &= \mathbf{I} - \mathbf{P}_a. \end{aligned} \quad (2.72)$$

Il apparaît naturellement que $\mathbf{P}_a^\perp \mathbf{a} = 0$ et donc,

$$\tilde{\mathbf{R}} = \mathbf{P}_a^\perp \mathbf{R} \mathbf{P}_a^\perp = \mathbf{P}_a^\perp \mathbf{R}_c \mathbf{P}_a^\perp + \sigma^2 \mathbf{P}_a^\perp. \quad (2.73)$$

Cette approche se généralise au cas où plus d'une source interférente est présente. Dans ce cas, la matrice de projection est définie par,

$$\mathbf{P}_A = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H. \quad (2.74)$$

Alternativement, il a aussi été proposé de réduire l'impact d'une source interférente en soustrayant sa contribution à la matrice de corrélation des mesures [8],

$$\tilde{\mathbf{R}} = \mathbf{R} - \sigma_s^2 \mathbf{a}\mathbf{a}^H = \mathbf{R}_c + \sigma^2 \mathbf{I}, \quad (2.75)$$

ce qui implique d'avoir un estimé précis de σ_s .

2.5 Conclusion et motivations des travaux de thèse

Les radio-interféromètres fonctionnent en mesurant des échantillons de la transformée de Fourier 2D de l'image du ciel observé. Ces grands instruments permettent

ainsi d'obtenir des informations sur des structures fines de sources cosmiques par une chaîne de traitement du signal incluant une calibration des mesures suivie d'une reconstruction d'images. Néanmoins, du fait de leur sensibilité, les antennes des réseaux interférométriques sont fortement affectées par la présence de RFI dans les observations. Des méthodes permettent de détecter les observations corrompues afin de les filtrer ou de les retirer. Ces méthodes sont limitées à la prise en compte de RFI de fortes puissances et dégradent de plus les observations, réduisant l'information utile. Aussi, la forte utilisation d'ondes radio dans de multiples applications technologiques conduit à une forte probabilité de présence de RFI dans les données interférométriques. Il est donc important de s'intéresser à des algorithmes évitant ces inconvénients et permettant de traiter le cas des RFI de faible puissance. L'objet de ce travail de thèse est alors de tenir compte de RFI de puissance quelconque, tant en calibration qu'en imagerie, par l'introduction de modélisations robustes des données interférométriques.

- Des méthodes de calibration par maximum de vraisemblance permettent de considérer des distributions plus robustes, afin de prendre en compte des bruits hétérogènes. Toutefois, ces modélisations ne permettent pas de représenter de manière précise la structure des RFI dans le contexte d'observations multifréquentielles. Nous proposons alors dans la section 3.2.2 une modélisation multifréquentielle des visibilitées permettant de prendre en compte la structure des RFI par un bruit additif de rang faible. Un algorithme SAGE pour l'estimation des paramètres de calibration est ensuite dérivé dans la section 3.3.1.
- Les méthodes de l'état de l'art en imagerie radio-interférométrique considèrent dans leur ensemble un bruit additif gaussien qui ne permet pas de prendre en compte la présence de RFI dans les données. Dans la section 4.1, nous modélisons la présence d'interférences par un bruit additif gaussien composé. Le problème d'imagerie est alors formulé comme un problème de maximum de vraisemblance régularisé pour lequel un algorithme EM est dérivé.
- Afin d'obtenir des images de haute résolution, les données provenant de réseaux interférométriques sont de grande dimension et nécessitent des algorithmes efficaces. Nous étendons alors l'algorithme EM proposé en considérant un jeu de données complètes spécifique, présenté dans la section 4.3, de sorte à obtenir une étape de maximisation pouvant être implémenter à faible coût computationnel. De plus, l'algorithme 4.1 est intégré dans une librairie Python dédiée à l'imagerie radio-interférométrique sur des données de grande dimension. Pour cela, l'étape de maximisation est mise en oeuvre par un algorithme forward-backward préconditionné, permettant une convergence accélérée. Des résultats sur données réelles démontrent alors de la pertinence de l'algorithme proposé pour des observations en présence de RFI.

3 - Calibration multifréquentielle en présence d'interférences

Dans ce chapitre, nous nous intéressons à la calibration multifréquentielle de radio-interféromètres en présence de RFI. Nous nous plaçons dans le cadre de calibration à partir de l'observation de sources de références connues, permettant d'estimer les effets systématiques affectant les mesures, afin de les prendre en compte à l'étape d'imagerie.

La particularité de la calibration de radio-interféromètres réside dans la modélisation des effets instrumentaux et environnementaux affectant les mesures. Ainsi, le formalisme de Jones permet de modéliser les visibilitées mesurées en tenant compte des différents phénomènes physiques ayant transformé l'onde émise le long de son chemin de propagation [36], [57]. Toutefois, la présence d'interférences dans les mesures peut grandement affecter les solutions de calibration [58]. En effet, l'une des principales limites de la sensibilité des radiotélescopes interféromètres est la présence de RFI, notamment due aux ondes radio d'origine humaine [8]. Ces dernières, provenant principalement de signaux de télécommunications, sont de plus ou moins fortes puissance. Dans le cas de RFI de forte puissance, des méthodes de détection et de filtrage permettent d'atténuer efficacement l'impact des RFI sur les observations de radioastronomie. Néanmoins, la présence d'interférences radio, même de faible puissance, impacte significativement les solutions de calibration, rendant les données plus difficiles à exploiter [58]. Ainsi une modélisation robuste des perturbations additives affectant les mesures est primordiale à l'étape de calibration. Cela a conduit à l'émergence d'algorithmes de calibration basés sur une estimation par maximum de vraisemblance, en considérant des modèles statistiques robustes à la présence de valeurs aberrantes dans les données [13], [14], [42], [59].

Toutefois, du fait de leur utilisation dans les activités humaines, les RFI admettent une structure particulière et peuvent s'étaler sur plusieurs canaux fréquentiels [10], [60]. Nous présentons alors dans ce chapitre, un modèle statistique des données interférométriques tenant compte de la structure des RFI. Nous proposons un modèle unifié à partir duquel nous dérivons un estimateur du maximum de vraisemblance des solutions de calibration, ainsi que du sous-espace des RFI, permettant une calibration robuste en présence d'interférences. En modélisant avec précision la contribution des RFI aux visibilitées, nous améliorons la robustesse du processus de calibration pour les RFI de faible et moyenne puissance. Ainsi, en considérant les RFI comme des ondes électromagnétiques d'origine humaine, nous prouvons que le sous-espace lié aux RFI produit une perturbation structurée de rang faible dans les visibilitées. Nous construisons alors un modèle pour des observations multifréquentielles en utilisant le formalisme des modèles à effets mixtes [61]-[63]. Plus précisément, nous considérons un modèle composé d'effets fixes représentant les sources célestes observées et d'effets aléatoires,

qui tiennent compte de la présence possible de RFI exclusivement sur quelques canaux fréquentiels. Nous obtenons un estimateur du maximum de vraisemblance pour le modèle introduit en utilisant l'algorithme d'Esperance-Maximisation sur Espaces Alternés (SAGE) [64]. Ainsi, nous dérivons deux espaces originaux de données latentes pour lesquels nous obtenons des solutions explicites des étapes de maximisation. Enfin, des simulations numériques permettent de confirmer la pertinence de la modélisation proposée, ainsi que les performances de l'algorithme SAGE résultant comparées aux méthodes de l'état de l'art.

3.1 Énonciation du problème de calibration

Nous rappelons dans cette section le modèle de données pour un radiotélescope interféromètre, ainsi que la formulation du problème de calibration dans un contexte multifréquentiel.

3.1.1 Modèle de données

Un radio-interféromètre mesure un estimé des corrélations entre les différentes antennes du réseau [1], [65]. En concaténant les signaux provenant de l'ensemble des capteurs, le vecteur de mesures s'écrit,

$$\mathbf{e}(t) = \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_2(t) \\ \vdots \\ \mathbf{e}_P(t) \end{bmatrix}, \quad (3.1)$$

où $\mathbf{e}_p(t) = \begin{bmatrix} \mathbf{e}_{a_p}(t) \\ \mathbf{e}_{b_p}(t) \end{bmatrix}$ est le vecteur de mesures correspondant à la $p^{\text{ème}}$ antenne. La matrice de corrélation du réseau sur un canal fréquentiel peut être définie comme une matrice par blocs de taille $2P \times 2P$,

$$\mathbf{V} = \mathbb{E} [\mathbf{e}(t)\mathbf{e}^H(t)], \quad (3.2)$$

où chaque bloc de taille 2×2 correspond à la corrélation entre la $p^{\text{ème}}$ et la $q^{\text{ème}}$ antenne,

$$\mathbf{V}_{pq} = \mathbb{E} [\mathbf{e}_p(t)\mathbf{e}_q^H(t)]. \quad (3.3)$$

On note alors \mathbf{V}_{pq} la matrice de visibilité pour le couple d'antenne (p, q) . L'équation de mesure du radio-interféromètre est ensuite obtenue en appliquant le formalisme de Jones, tel que défini en (2.21). Ainsi, étant données D sources de calibration, d'enveloppe complexe $s_1(t), \dots, s_D(t)$, le vecteur de mesures pour la $p^{\text{ème}}$ antenne du réseau interféromètre s'écrit,

$$\mathbf{e}_p(t) = \sum_{i=1}^D \mathbf{J}_{i,p} \mathbf{s}_i(t), \quad (3.4)$$

où $\mathbf{J}_{i,p}$ est la matrice de Jones associée à la $i^{\text{ème}}$ source et la $p^{\text{ème}}$ antenne. L'équation (3.3) peut alors se réécrire ainsi en effectuant les hypothèses suivantes [36] :

- les sources sont considérées indépendantes entre elles,
- les matrices de Jones sont supposées constantes sur le temps d'intégration,

$$\mathbf{V}_{pq} = \sum_{i=1}^D \mathbf{J}_{i,p} \mathbf{C}_i \mathbf{J}_{i,q}, \quad (3.5)$$

où $\mathbf{C}_i = \mathbb{E} [\mathbf{s}_i(t) \mathbf{s}_i(t)^H]$ est la matrice de cohérence de la $i^{\text{ème}}$ source. L'équation (3.5) constitue *l'équation de mesure du radio-interféromètre*. Le vecteur de visibilité pour la ligne de base formée par le couple d'antennes d'indices (p, q) est alors noté,

$$\mathbf{v}_{pq} = \text{vec}(\mathbf{V}_{pq}) = \sum_{i=1}^D (\mathbf{J}_{i,q}^* \otimes \mathbf{J}_{i,p}) \text{vec}(\mathbf{C}_i). \quad (3.6)$$

L'ensemble des vecteurs de visibilité sont ensuite concaténés dans un vecteur,

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_{12} \\ \vdots \\ \mathbf{v}_{(P-1)P} \end{bmatrix} \in \mathbb{C}^{4K \times 1}, \quad (3.7)$$

où $K = P(P-1)/2$ est le nombre de couple d'antennes pouvant être formé par le réseau interférométrique. En pratique, les visibilité sont issues de l'estimé de la matrice de covariance des échantillons sur un intervalle de temps choisi, de sorte que les sources puissent être considérées stationnaires et les matrices de Jones constantes,

$$\mathbf{R}_{pq} = \frac{1}{T} \sum_{t=1}^T \mathbf{e}_p(t) \mathbf{e}_q^H(t), \quad (3.8)$$

où T est le nombre d'échantillons. Pour T suffisamment grand, il est couramment accepté de modéliser le vecteur de mesures, $\mathbf{y}_{pq} = \text{vec}(\mathbf{R}_{pq})$, comme la somme d'un signal déterministe et d'un bruit blanc gaussien qui représente le résiduel [1][44][66] :

$$\mathbf{y}_{pq} = \text{vec}(\mathbf{R}_{pq}) = \mathbf{v}_{pq} + \mathbf{n}_{pq}, \quad \mathbf{n}_{pq} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_4). \quad (3.9)$$

L'équation (3.9) représente une modélisation du vecteur de mesures pour un canal fréquentiel. Cette modélisation s'étend sur des observations multifréquentielles sur Q canaux. Ainsi, une matrice de Jones est définie pour chaque antenne, canal fréquentiel et direction de calibration. On aboutit alors au modèle suivant sur les visibilité,

$$\mathbf{y}_{pq}^{[j]} = \mathbf{v}_{pq}^{[j]} + \mathbf{n}_{pq}^{[j]}, \quad \mathbf{n}_{pq}^{[j]} \sim \mathcal{CN}(0, \sigma^{[j]2} \mathbf{I}_4), \quad (3.10)$$

$$\mathbf{v}_{pq}^{[j]} = \sum_{i=1}^D (\mathbf{J}_{i,q}^{[j]*} \otimes \mathbf{J}_{i,p}^{[j]}) \text{vec}(\mathbf{C}_i^{[j]}), \quad (3.11)$$

où j désigne l'indice du canal fréquentiel de fréquence f_j . Le vecteur de mesures est alors la concaténation des vecteurs de visibilité associés à chaque canal fréquentiel,

$$\mathbf{y} = [\mathbf{y}^{[1]T}, \dots, \mathbf{y}^{[Q]T}]^T \in \mathbb{C}^{4QK \times 1}. \quad (3.12)$$

L'enjeu des algorithmes de calibration est alors d'estimer les matrices de Jones pour l'ensemble des D directions, P antennes et Q canaux de fréquences,

$$\mathbf{J} = \left\{ \mathbf{J}_{i,p}^{[j]} \mid (i,p) \in \llbracket 1, D \rrbracket \times \llbracket 1, P \rrbracket, j \in \llbracket 1, Q \rrbracket \right\}. \quad (3.13)$$

3.1.2 Formulation du problème de calibration

Le problème de calibration peut être formulé comme un problème de maximum de vraisemblance à partir du modèle spécifié en (3.9),

$$\hat{\mathbf{J}} = \underset{\mathbf{J}}{\operatorname{argmax}}(\mathcal{L}(\mathbf{J}; \mathbf{y})), \quad (3.14)$$

où $\mathcal{L}(\mathbf{J}; \mathbf{y})$ dénote la vraisemblance des matrices de Jones pour les données interférométriques mesurées. Dans le cas d'une modélisation du bruit additif par un bruit blanc gaussien, le problème inverse résultant est équivalent à un problème de moindres carrés [11],

$$\hat{\mathbf{J}} = \underset{\mathbf{J}}{\operatorname{argmax}}(\|\mathbf{y} - \mathbf{v}(\mathbf{J})\|_2^2). \quad (3.15)$$

Les matrices de Jones dépendent du canal fréquentiel considéré. Ainsi, le nombre de paramètres à estimer croît rapidement sur des observations multifréquentielles, rendant l'estimation difficile. Une approximation polynomiale des matrices de Jones en fonction de la fréquence permet alors de réduire le nombre de paramètres à estimer, tout en étant fidèle à la réalité physique des perturbations subites par le signal [14], [44]. La matrice de Jones correspondant à la $p^{\text{ème}}$ antenne, la $i^{\text{ème}}$ direction, sur le $j^{\text{ème}}$ canal de fréquence f_j s'écrit alors,

$$\begin{aligned} \mathbf{J}_{i,p}^{[j]} &= \sum_{c=1}^{\zeta} \mathbf{Z}_{i,p,c} b_c(f_j), \\ b_c(f_j) &= \left(\frac{f_j - f_0}{f_0} \right)^{c-1}, \end{aligned} \quad (3.16)$$

où f_0 est la fréquence centrale des observations, ζ est le degré du polynôme et $\mathbf{Z}_{i,p,c}$ correspond au $\zeta^{\text{ème}}$ coefficient du polynôme. L'équation (3.16) peut se réécrire comme un produit matriciel en considérant la matrice $\mathbf{Z}_{i,p} = [\mathbf{Z}_{p,i,1}, \dots, \mathbf{Z}_{p,i,\zeta}]^T$,

$$\begin{aligned} \mathbf{J}_{i,p}^{[j]} &= \mathbf{B}^{[j]} \mathbf{Z}_{i,p}, \\ \mathbf{B}^{[j]} &= [b_1(f_j), \dots, b_{\zeta}(f_j)] \otimes \mathbf{I}_2. \end{aligned} \quad (3.17)$$

Après avoir introduit le modèle multifréquentiel dans cette section, nous présentons dans la section suivante un modèle statistique des données interférométriques en présence de RFI.

3.2 Modélisation des interférences de radiofréquence

3.2.1 Cas monofréquentiel

Nous proposons de modéliser l'impact des RFI sur les visibilitées mesurées par un processus de rang faible. En effet, les RFI restent confinées dans sous-espace signal de faible rang [67]-[69]. Cela nous conduit à considérer le processus stochastique de rang faible suivant :

$$\mathbf{y}^{\text{RFI}} = \mathbf{W}\mathbf{o}, \quad (3.18)$$

où $\mathbf{y}^{\text{RFI}} \in \mathbb{C}^{4K \times 1}$ représente l'impact des RFI sur les visibilitées pour un canal fréquentiel, \mathbf{W} est une matrice de structure, modélisant un dictionnaire représentant la contribution de sources parasites, et $\mathbf{o} \sim \mathcal{CN}(\text{vec}(\mathbf{I}), \mathbf{I})$ est un vecteur aléatoire gaussien. Pour un vecteur de visibilitées corrompues par L sources d'interférences, le modèle de mesure pour un canal fréquentiel devient alors,

$$\mathbf{y} = \mathbf{v} + \mathbf{W}\mathbf{o} + \mathbf{n}, \quad \mathbf{o} \sim \mathcal{CN}(\text{vec}(\mathbf{I}), \mathbf{I}), \quad \mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}), \quad \rho \leq 4L^2, \quad (3.19)$$

où $\mathbf{W} \in \mathbb{C}^{4K \times 4L^2}$, $\rho = \text{rank}(\mathbf{W})$.

Dans ce qui suit, nous justifions le modèle proposé en (3.19) en considérant le modèle physique sous-jacent aux RFI. En effet, les RFI sont des signaux électromagnétiques d'origine humaine [60]. Ainsi, les RFI, les sources célestes et le bruit thermique sont supposés indépendants de sorte à ce que l'on puisse modéliser l'impact des RFI de manière isolée. Ainsi, la composante du signal reçu liée aux RFI à la $p^{\text{ème}}$ antenne peut s'exprimer à l'aide du formalisme de Jones,

$$\begin{aligned} \mathbf{e}_p^{\text{RFI}}(t) &= \sum_{i=1}^L \mathbf{J}_{i,p}^{\text{RFI}}(t) \mathbf{s}_i^{\text{RFI}}(t), \\ \mathbf{s}_i^{\text{RFI}}(t) &\sim \mathcal{CN}(0, \mathbf{C}^{\text{RFI}}), \end{aligned}$$

où $\mathbf{s}_i^{\text{RFI}}(t)$ représente une source interférente et $\mathbf{J}_{i,p}^{\text{RFI}}(t)$ l'ensemble des transformations subites par l'onde interférente depuis la $i^{\text{ème}}$ direction pour la $p^{\text{ème}}$ antennes. Son impact sur la matrice de corrélation peut alors s'exprimer en appliquant l'équation de mesure du radio-interféromètre définie en (3.5). Ainsi,

$$\mathbf{V}_{pq}^{\text{RFI}} = \mathbb{E} \left[\mathbf{e}_p^{\text{RFI}}(t) \mathbf{e}_q^{\text{RFI}H}(t) \right] = \sum_{i=1}^L \mathbf{J}_{i,p}^{\text{RFI}} \mathbf{C}_i^{\text{RFI}} \mathbf{J}_{i,q}^{\text{RFI}H}. \quad (3.20)$$

En écrivant cette expression sous forme de produit matriciel on obtient,

$$\mathbf{V}_{pq}^{\text{RFI}} = \mathbf{J}_p^{\text{RFI}} \mathbf{C}^{\text{RFI}} \mathbf{J}_q^{\text{RFI}H}, \quad (3.21)$$

où $\mathbf{J}_p^{\text{RFI}} \in \mathbb{C}^{2 \times 2L}$ et $\mathbf{C}^{\text{RFI}} \in \mathbb{C}^{2L \times 2L}$ sont des matrices par blocs contenant respectivement les matrices de Jones correspondant à toutes les directions pour la $p^{\text{ème}}$ antenne et l'ensemble des matrices de cohérence des sources interférentes,

$$\mathbf{J}_p^{\text{RFI}} = [\mathbf{J}_{1,p}^{\text{RFI}}, \dots, \mathbf{J}_{L,p}^{\text{RFI}}], \quad (3.22)$$

$$\mathbf{C}^{\text{RFI}} = \text{diag}(\mathbf{C}_1^{\text{RFI}}, \dots, \mathbf{C}_L^{\text{RFI}}). \quad (3.23)$$

Ainsi la matrice de corrélation du radio-interféromètre en présence de RFI s'exprime,

$$\mathbf{V}^{\text{RFI}} = \mathbf{J}^{\text{RFI}} \mathbf{C}^{\text{RFI}} \mathbf{J}^{\text{RFI}H}, \quad (3.24)$$

où $\mathbf{J}^{\text{RFI}} \in \mathbb{C}^{2P \times 2L}$ est la matrice par blocs contenant l'ensemble des matrices de Jones associées au modèle,

$$\mathbf{J}^{\text{RFI}} = \begin{bmatrix} \mathbf{J}_1^{\text{RFI}} \\ \vdots \\ \mathbf{J}_P^{\text{RFI}} \end{bmatrix}. \quad (3.25)$$

Notons \mathbf{R}^{RFI} la matrice de covariance empirique correspondant aux mesures effectuées par le radio-interféromètre,

$$\mathbf{R}^{\text{RFI}} = \frac{1}{T} \sum_{t=1}^T e^{\text{RFI}}(t) e^{\text{RFI}}(t)^H. \quad (3.26)$$

Par définition, la matrice \mathbf{R}^{RFI} est distribuée selon une loi de Wishart complexe de degré T et de moyenne \mathbf{V}^{RFI} . En tenant compte de la symétrie de \mathbf{V}^{RFI} , il est alors possible de réécrire la matrice de corrélation mesurée de la manière suivante,

$$\mathbf{R}^{\text{RFI}} = \mathbf{A} \mathbf{M} \mathbf{A}^H, \quad \mathbf{M} \sim \mathcal{W}(T, \mathbf{I}), \quad (3.27)$$

où $\mathbf{A} = \mathbf{J}^{\text{RFI}} \mathbf{C}^{\text{RFI}^{1/2}}$. La matrice de visibilité formée pour la ligne de base correspondant au couple d'antennes p et q peut alors s'exprimer,

$$\mathbf{R}_{pq}^{\text{RFI}} = \mathbf{A}_p \mathbf{M} \mathbf{A}_q^H, \quad \mathbf{A}_p = \mathbf{J}_p^{\text{RFI}} \mathbf{C}^{\text{RFI}^{1/2}}. \quad (3.28)$$

En considérant la matrice de corrélation vectorisée on obtient,

$$\mathbf{y}_{pq}^{\text{RFI}} = \text{vec}(\mathbf{R}_{pq}^{\text{RFI}}) = (\mathbf{A}_q^* \otimes \mathbf{A}_p) \text{vec}(\mathbf{M}). \quad (3.29)$$

Enfin, l'ensemble des visibilité sont concaténées dans un vecteur conduisant à une structure de rang faible,

$$\mathbf{W} = \begin{bmatrix} \mathbf{A}_2^* \otimes \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_P^* \otimes \mathbf{A}_{P-1} \end{bmatrix}, \quad (3.30)$$

$$\mathbf{y}^{\text{RFI}} = \mathbf{W} \text{vec}(\mathbf{M}),$$

où $\mathbf{M} \sim \mathcal{W}(T, \mathbf{I})$. Dans le contexte de la radio-interférométrie, le nombre d'échantillons utilisés pour évaluer l'estimé de la matrice de covariance est de l'ordre de $T \geq 10^4$. Dans ce cas, la distribution de Wishart se rapproche asymptotiquement d'une distribution gaussienne généralisée [70],

$$\sqrt{T} \text{vec}(\mathbf{M} - \mathbf{I}) \xrightarrow{d} \mathcal{GCN}(\mathbf{0}, \mathbf{I}, \mathbf{I}), \quad (3.31)$$

d'où la justification du modèle proposé en (3.19). Ainsi, en considérant suffisamment d'échantillons de manière à pouvoir appliquer le théorème central limite, l'impact des RFI sur les données peut être pris en compte par le modèle proposé en (3.19). La valeur maximale du rang de \mathbf{W} peut-être identifiée à partir de l'équation (3.30), $\rho_{\max} = 4L^2$.

3.2.2 Cas multifréquentiel

Dans la section précédente, nous n'avons considéré qu'un seul canal fréquentiel. On s'intéresse dans la suite à des observations radio-interférométriques sur Q canaux fréquents perturbés par des RFI. Nous modélisons la dépendance en fréquence des matrices de Jones par un polynôme de degré faible, tel que décrit en (3.16). Le vecteur de visibilité sans bruit pour $j^{\text{ème}}$ canal et la ligne de base formée par les $p^{\text{ème}}$ et $q^{\text{ème}}$ antennes devient,

$$\mathbf{v}_{pq}^{[j]}(\mathbf{Z}) = \sum_{i=1}^D ((\mathbf{B}^{[j]*} \mathbf{Z}_{i,q}^*) \otimes (\mathbf{B}^{[j]} \mathbf{Z}_{i,p})) \text{vec}(\mathbf{C}_i^{[j]}). \quad (3.32)$$

Le modèle exprimé à l'équation (3.19) peut être réécrit en utilisant le formalisme des modèles à effets mixtes, dans le but de prendre en compte la variabilité des observations sur différents canaux fréquents. Ainsi, la présence de RFI sur un canal de fréquence donné, ainsi que le bruit thermique associé, sont considérés comme des effets aléatoires et sont modélisés par $\mathbf{W}^{[j]} \mathbf{y}^{[j]} + \mathbf{n}^{[j]}$. D'autre part, la contribution des sources calibrées aux visibilité, dont le comportement en fréquence est supposé connu, est considérée comme un effet fixe. Le modèle résultant peut être écrit comme suit,

$$\mathbf{y}^{[j]} = \mathbf{v}^{[j]}(\mathbf{Z}) + \mathbf{W}^{[j]} \mathbf{o}^{[j]} + \mathbf{n}^{[j]}, \quad \mathbf{o}^{[j]} \sim \mathcal{CN}(\text{vec}(\mathbf{I}), \mathbf{I}) \text{ et } \mathbf{n}^{[j]} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}). \quad (3.33)$$

Idéalement, puisque une RFI peut être présente sur plusieurs canaux fréquents, une matrice de rang faible, $\mathbf{W}^{[j]}$, serait assignée à chaque canal fréquentiel affecté par des RFI. Pour réduire le coût computationnel et simplifier les calculs, nous réduisons la dimension du modèle en considérant une matrice commune de rang faible, \mathbf{W} , qui englobe la structure des RFI sur l'ensemble des canaux fréquents. Ainsi, le modèle est factorisé en un modèle à effet mixte plus simple, où $\mathbf{W}^{[j]} = \sigma^{[j]} \mathbf{W}$, \mathbf{W} étant partagé sur plusieurs canaux (absorbant la structure des RFI) et $\sigma^{[j]}$ représentent une pondération. En concaténant les visibilité mesurées sur tous les canaux fréquents, nous déduisons une modélisation des données radio-interférométriques corrompues par des RFI.

$$\mathbf{y} = \mathbf{v}(\mathbf{Z}) + \left((\sigma^{[1]}, \dots, \sigma^{[Q]}) \otimes \mathbf{W} \right) \mathbf{o} + \mathbf{n}, \quad (3.34)$$

$$\mathbf{o} = [\mathbf{o}^{[1]}, \dots, \mathbf{o}^{[Q]}], \quad (3.35)$$

avec $\mathbf{W} \in \mathbb{C}^{4K \times M}$, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$, $\mathbf{o}^{[j]} \sim \mathcal{CN}(\text{vec}(\mathbf{I}), \mathbf{I})$ et où les $\sigma^{[j]}, \forall j \in \llbracket 1, J \rrbracket$ représentent des facteurs de pondération. Le nombre de colonnes, M , de la matrice \mathbf{W} correspond à son rang maximal (et également au rang maximal de la matrice $\mathbf{W} \mathbf{W}^H$).

3.3 Calibration multifréquentielle par maximum de vraisemblance

3.3.1 Maximisation de la vraisemblance par un algorithme SAGE

Dans cette section, nous présentons un algorithme qui vise à estimer les paramètres inconnus du modèle spécifié par (3.35). L'ensemble des paramètres à estimer est

$$\Theta = \{\mathbf{Z}, \mathbf{W}, \sigma^2, \sigma^{[1]}, \dots, \sigma^{[Q]}\}. \quad (3.36)$$

Le paramètre d'intérêt dans le contexte de la calibration multifréquentielle de radio-interféromètres est l'ensemble des coefficients polynomiaux contenus dans \mathbf{Z} . Étant donné un vecteur de visibilité, \mathbf{y} , la vraisemblance logarithmique pour le modèle spécifié par (3.35) est,

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}; \mathbf{y}) &= -\log(|\sigma^2 \mathbf{I} + \boldsymbol{\Psi} \boldsymbol{\Psi}^H|) - (\mathbf{y} - \mathbf{v}(\mathbf{Z}) - \boldsymbol{\Psi}(\text{vec}(\mathbf{I}_{4K}) \otimes \mathbf{1}_Q))^H \\ &\quad (\sigma^2 \mathbf{I} + \boldsymbol{\Psi} \boldsymbol{\Psi}^H)^{-1} (\mathbf{y} - \mathbf{v}(\mathbf{Z}) - \boldsymbol{\Psi}(\text{vec}(\mathbf{I}_{4K}) \otimes \mathbf{1}_Q)), \\ \boldsymbol{\Psi} &= [\sigma^{[1]}, \dots, \sigma^{[Q]}] \otimes \mathbf{W}, \\ \boldsymbol{\theta} &= [\sigma^2, \text{vec}(\mathbf{Z}), \text{vec}(\mathbf{W}), \sigma^{[1]}, \dots, \sigma^{[Q]}], \end{aligned} \tag{3.37}$$

où $\mathbf{1}_Q = (1, \dots, 1)$.

La maximisation directe de la vraisemblance n'est pas tractable. Nous proposons donc d'utiliser un algorithme basé sur l'algorithme SAGE pour résoudre le problème de maximum de vraisemblance (MLE) associé au modèle proposé [64]. L'algorithme SAGE est une variante de l'algorithme EM. Dans sa forme la plus simple, l'algorithme EM est une approche itérative pour l'estimation par maximum de vraisemblance lorsque les observations peuvent être vues comme des données incomplètes. Plusieurs travaux sur les propriétés de convergence de l'algorithme EM ont permis de montrer qu'il augmente de manière itérative la vraisemblance, convergeant au moins vers un maximum local [71][72]. Cependant, l'algorithme EM est connu pour sa convergence lente et son étape de maximisation potentiellement difficile lorsque l'espace des paramètres est de grande dimension [73]. L'algorithme SAGE proposé dans [64] résout ces problèmes en alternant sur plusieurs espaces de données complètes. Dans [11], l'algorithme SAGE est notamment utilisé pour séparer les signaux selon différentes directions en supposant un bruit gaussien sur les visibilité. Dans ce chapitre, l'algorithme SAGE est utilisé pour séparer les signaux célestes et le sous-espace RFI par un modèle qui considère spécifiquement la présence de RFI. En pratique, l'algorithme SAGE consiste à définir plusieurs espaces de paramètres et leurs espaces de données complets correspondants, dans lesquels les procédures EM sont effectuées.

Ainsi, nous définissons deux ensembles de données complètes à partir desquels nous dérivons un algorithme SAGE. Pour chaque espace de données complètes, nous calculons l'espérance de la vraisemblance logarithmique des données complètes, étant donné les données incomplètes observées dans l'étape-E. Les paramètres courants sont obtenus à l'étape-M en maximisant l'espérance calculée à l'étape-E. L'ensemble des données complètes est défini de sorte que la vraisemblance de ses paramètres conduise à des maximisations explicites. Le but de l'algorithme proposé est d'estimer le sous-espace RFI, \mathbf{W} et $(\sigma^{[j]})_{1 \leq j \leq F}$, en utilisant un premier espace de données complètes et d'estimer les paramètres de calibration, \mathbf{Z} et σ^2 , avec un second espace de données complètes. Le premier ensemble de données complètes contient donc les visibilité mesurées ainsi que le sous-espace RFI,

$$\mathcal{X}_1 = \{\mathbf{y}, \mathbf{o}\}. \tag{3.38}$$

Le vecteur de paramètres associé définit le sous-espace RFI,

$$\boldsymbol{\theta}_1 = [\text{vec}(\mathbf{W}), \sigma^{[1]}, \dots, \sigma^{[Q]}]. \quad (3.39)$$

Le second espace de données complètes contient les contributions individuelles des sources aux visibilitées, auxquelles l'impact des RFI est soustrait,

$$\mathcal{X}_2 = \{(\mathbf{u}_i)_{1 \leq i \leq D} : \mathbf{u} = \mathbf{y} - \Psi \boldsymbol{o} = \sum_{i=1}^D \mathbf{u}_i\}, \quad (3.40)$$

où $\mathbf{u}_i = \mathbf{v}_i(\mathbf{Z}) + \mathbf{n}_i$, $\mathbf{v}_i(\mathbf{Z})$ étant le vecteur de visibilitées sans bruit de la $i^{\text{ème}}$ source et $\mathbf{n}_i \sim \mathcal{CN}(0, \beta_i \sigma^2 \mathbf{I})$ tel que $\sum_{i=1}^D \beta_i = 1$ et $\beta_i > 0$. Les paramètres de calibration sont mis à jour en utilisant ce second ensemble de données complètes,

$$\boldsymbol{\theta}_2 = [\text{vec}(\mathbf{Z}), \sigma^2]. \quad (3.41)$$

Les fonctions de substitution calculées à l'étape-E peuvent être exprimées comme suit,

$$\begin{aligned} Q_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^{(m)}) &= \mathbb{E}_{\mathcal{X}_1 | \mathbf{y}; \boldsymbol{\theta}_1^{(m)}} [\log(p(\mathbf{y}, \boldsymbol{o}; \boldsymbol{\theta}_1))], \\ Q_2(\boldsymbol{\theta}_2, \boldsymbol{\theta}_2^{(m)}) &= \sum_{i=1}^D \mathbb{E}_{\mathcal{X}_2^i | \mathbf{y}; \boldsymbol{\theta}_2^{(m)}} [\log(p(\mathbf{u}_i; \boldsymbol{\theta}_2))]. \end{aligned} \quad (3.42)$$

Proposition 1 *La fonction de substitution Q_1 se réduit à*

$$\begin{aligned} Q_1(\boldsymbol{\theta} | \boldsymbol{\theta}^{(m)}) &= A - \log|\sigma^2 \mathbf{I}_{4FK}| \\ &\quad - \sum_{f=1}^Q \left(\frac{1}{\sigma^2} \|\mathbf{y}^{[j]} - \mathbf{v}^{[j]} - \sigma^{[j]} \mathbf{W} \hat{\boldsymbol{o}}^{[j](m)}\|_2^2 + \frac{\sigma^{[j]^2}}{\sigma^2} \text{tr}(\mathbf{W}^H \mathbf{W} \hat{\boldsymbol{\Sigma}}_o^{[j](m)}) \right. \\ &\quad \left. + \text{tr}(\hat{\boldsymbol{\Sigma}}_o^{[j](m)}) + \|\hat{\boldsymbol{o}}^{[j](m)} - \text{vec}(\mathbf{I})\|_2^2 \right) \end{aligned} \quad (3.43)$$

où A est une constante indépendante des paramètres du modèle et,

$$\begin{aligned} \hat{\boldsymbol{o}}^{[j](m)} &= \text{vec}(\mathbf{I}) \\ &\quad + \sigma^{[j](m)} \mathbf{W}^{(m)H} (\sigma^{(m)^2} \mathbf{I} + \sigma^{[j](m)^2} \mathbf{W}^{(m)} \mathbf{W}^{(m)H})^{-1} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}^{(m)}) - \sigma^{[j](m)} \mathbf{W}^{(m)} \text{vec}(\mathbf{I})), \\ \hat{\boldsymbol{\Sigma}}_o^{[j](m)} &= \mathbf{I} - \sigma^{[j](m)} \mathbf{W}^{(m)H} (\sigma^{(m)^2} \mathbf{I} + \sigma^{[j](m)^2} \mathbf{W}^{(m)} \mathbf{W}^{(m)H})^{-1} \sigma^{[j](m)} \mathbf{W}^{(m)}. \end{aligned} \quad (3.44)$$

Démonstration voir l'annexe A.1.

La maximisation de Q_1 se fait par application des dérivées de Wirtinger [74] et conduit aux expressions explicites suivantes (voir A.2),

$$\tilde{\sigma}^{[j](m+1)} = \frac{\Re\left((\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H \mathbf{W}^{(m)} \hat{\mathbf{o}}^{[j](m)}\right)}{\text{tr}\left(\mathbf{W}^{(m)H} \mathbf{W}^{(m)} (\hat{\Sigma}_o^{[j](m)} + \hat{\mathbf{o}}^{[j](m)} \hat{\mathbf{o}}^{[j](m)H})\right)}, \quad (3.45)$$

$$\begin{aligned} \tilde{\mathbf{W}}^{(m+1)} &= \left(\sum_{f=1}^F \sigma^{[f]} \left(\hat{\mathbf{o}}^{[f](m)} (\mathbf{y}^{[f]} - \mathbf{v}^{[f]})^H \right)^H \right) \\ &\quad \left(\sum_{f=1}^F \sigma^{[f]2} \left(\hat{\Sigma}_o^{[f](m)} + \hat{\mathbf{o}}^{[f](m)} \hat{\mathbf{o}}^{[f](m)H} \right)^H \right)^{-1}. \end{aligned} \quad (3.46)$$

Il convient de mentionner qu'il existe une ambiguïté entre les mises à jour de \mathbf{W} et de $\sigma^{[1]}, \dots, \sigma^{[Q]}$. Afin de lever cette ambiguïté, nous choisissons d'imposer une contrainte de normalisation unitaire sur \mathbf{W} ,

$$\sigma^{[j](m+1)} = \|\tilde{\mathbf{W}}^{(m+1)}\|_2 \times \tilde{\sigma}^{[j](m+1)}, \quad (3.47)$$

$$\mathbf{W}^{(m+1)} = \frac{\tilde{\mathbf{W}}^{(m+1)}}{\|\tilde{\mathbf{W}}^{(m+1)}\|_2}. \quad (3.48)$$

Proposition 2 *La fonction de substitution Q_2 se réduit à*

$$\begin{aligned} Q_2(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) &= \sum_{i=1}^D C_{u_i} - \log |\beta_i \sigma^2 \mathbf{I}_{4FK}| - \frac{1}{\beta_i \sigma^2} \|\hat{\mathbf{u}}_i^{(m)} - \mathbf{v}_i(\mathbf{Z})\|_2^2 \\ &\quad - \frac{1}{\beta_i \sigma^2} \text{tr}(\hat{\Sigma}_{\mathbf{u}_i}^{(m)}), \end{aligned} \quad (3.49)$$

où C_{u_i} est une constante indépendante des paramètres du modèle et,

$$\begin{aligned} \hat{\mathbf{u}}_i^{(m)} &= \mathbf{v}_i(\mathbf{Z}^{(m)}) + \beta_i \sigma^2 (\sigma^2 \mathbf{I} + \boldsymbol{\Psi} \boldsymbol{\Psi}^H)^{-1} (\mathbf{y} - \mathbf{v}(\mathbf{Z}^{(m)}) - \boldsymbol{\Psi}(\text{vec}(\mathbf{I}_{4K}) \otimes \mathbf{1}_Q)), \\ \hat{\Sigma}_{\mathbf{u}_i}^{(m)} &= \beta_i \sigma^2 \mathbf{I}_{4FK} - \beta_i^2 \sigma^4 (\sigma^2 \mathbf{I} + \boldsymbol{\Psi} \boldsymbol{\Psi}^H)^{-1}. \end{aligned} \quad (3.50)$$

Démonstration voir l'annexe A.3.

La maximisation de (3.49) en utilisant les dérivées de Wirtinger conduit à l'expression suivante pour σ^2 (voir l'annexe A.4),

$$\sigma^{2(m+1)} = \frac{1}{4QDK} \sum_{i=1}^D \frac{1}{\beta_i} \left(\|\hat{\mathbf{u}}_i^{(m)} - \mathbf{v}_i(\mathbf{Z}^{(m)})\|_2^2 + \text{tr}(\hat{\Sigma}_{\mathbf{u}_i}^{(m)}) \right). \quad (3.51)$$

On maximise Q_2 par rapport à \mathbf{Z} à l'aide d'un algorithme de descente par coordonnées [75]. Pour cela, on note $\mathbf{F}_{i,q}^{[j]} = \mathbf{B}^{[j]*} \mathbf{Z}_{i,q}^{(m+1)*} \mathbf{C}_i^{[j]T}$, $\mathbf{G}_{i,q}^{[j]} = \mathbf{B}^{[j]*} \mathbf{Z}_{i,q}^{(m+1)*} \mathbf{C}_i^{[j]*}$ et $\hat{\mathbf{U}}_{i,pq}^{[j](m)} = \text{vec}^{-1}(\hat{\mathbf{u}}_{i,pq}^{[j](m)})$, ce qui conduit à la mise à jour suivante (voir l'annexe A.4),

$$\text{vec}(\mathbf{Z}_{i,p}^{(m+1)*}) = (\mathbf{M}_{i,p} + \tilde{\mathbf{M}}_{i,p})^{-1} \text{vec}(\mathbf{T}_{i,p}), \quad (3.52)$$

avec

$$\mathbf{M}_{i,p} = \sum_{j=1}^Q \sum_{q=p+1}^N \mathbf{F}_{i,q}^{[j]T} \mathbf{F}_{i,q}^{[j]*} \otimes \mathbf{B}^{[j]T} \mathbf{B}^{[j]*}, \quad (3.53)$$

$$\tilde{\mathbf{M}}_{i,p} = \sum_{f=1}^Q \sum_{q=1}^{p-1} \mathbf{G}_{i,q}^{[f]T} \mathbf{G}_{i,q}^{[f]*} \otimes \mathbf{B}^{[f]T} \mathbf{B}^{[f]*}, \quad (3.54)$$

$$\mathbf{T}_{i,p} = \sum_{f=1}^Q \sum_{q=p+1}^N \mathbf{B}^{[f]T} \hat{\mathbf{U}}_{i,pq}^{[f(m)]*} \mathbf{F}_{i,q}^{[f]} + \sum_{j=1}^Q \sum_{q=1}^{p-1} \mathbf{B}^{[j]T} \hat{\mathbf{U}}_{i,qp}^{[j(m)]T} \mathbf{G}_{i,q}^{[j]}. \quad (3.55)$$

L'algorithme SAGE consiste à appliquer les deux étapes de l'EM de manière itérative jusqu'à convergence. Nous considérons que la convergence est atteinte lorsque la variation de la log-vraisemblance entre deux itérations est inférieure à un seuil choisi $\epsilon > 0$.

Algorithm 3.1 Algorithme SAGE pour la calibration multifréquentielle en présence de RFI

Input : $\epsilon, D, P, Q, \mathbf{v}, \mathbf{y}, \sigma^{2(0)}, \mathbf{Z}^{(0)}, \mathbf{W}^{(0)}, \sigma^{[j(0)}, \dots, \sigma^{[j(0)}$.

Output : $\sigma^2, \mathbf{Z}, \mathbf{W}, \sigma_1, \dots, \sigma^{[j]}$

```

1: while  $(\mathcal{L}(\boldsymbol{\theta}^{(m+1)}; \mathbf{y}) - \mathcal{L}(\boldsymbol{\theta}^{(m)}; \mathbf{y})) \geq \epsilon$  do
2:    $\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\sigma}}$  obtenus à partir de (3.44)
3:   for  $j \in \llbracket 1, Q \rrbracket$  do
4:      $\tilde{\sigma}^{[j]}$  obtenu à partir de (3.45)
5:   end for
6:    $\tilde{\mathbf{W}}$  obtenu à partir de (3.46)
7:    $\sigma^{[1]}, \dots, \sigma^{[Q]}$  obtenu à partir de (3.47)
8:    $\mathbf{W}$  obtenu à partir de (3.48)
9:   for  $i \in \llbracket 1, D \rrbracket$  do
10:     $\hat{\mathbf{u}}_i, \hat{\boldsymbol{\Sigma}}_{\mathbf{u}_i}$  obtenu à partir de (3.50)
11:    for  $p$  in  $\llbracket 1, P \rrbracket$  do
12:       $\mathbf{Z}_{i,p}$  obtenu à partir de (3.52)
13:    end for
14:   end for
15:    $\sigma^2$  obtenu à partir de (3.51)
16: end while
17: Return :  $(\sigma^2, \mathbf{Z}, \mathbf{W}, \sigma^{[1]}, \dots, \sigma^{[Q]})$ 

```

3.3.2 Simulations numériques

Dans cette section, plusieurs simulations sont présentées afin d'étudier la robustesse du modèle, ainsi que les performances de l'algorithme de maximum de vraisemblance proposés dans ce chapitre. Pour ce faire, nous considérons un radio-interféromètre radio de $P = 8$ antennes, observant un ciel composé de $D = 2$ sources

de calibration sur $Q = 32$ canaux fréquentiels. Les bandes de fréquence et la fréquence centrale sont choisies de telle sorte que $\frac{f-f_0}{f_0} \in [-1, 1]$. Le ciel est composé de deux sources de calibrations non polarisées de puissance 100 Jy et 50 Jy. La matrice \mathbf{Z} est générée à l'aide d'une distribution uniforme, pour $i \in \llbracket 1, D \rrbracket$, $p \in \llbracket 1, P \rrbracket$, $\Re(\mathbf{Z}_{i,p}) \sim \mathcal{U}(0, 1)$ et $\Im(\mathbf{Z}_{i,p}) \sim \mathcal{U}(0, 1)$. Un nombre de 100 répétitions de Monte Carlo sont effectuées. Nous calculons pour chaque exécution l'erreur quadratique moyenne normalisée (NMSE) pour le paramètre d'intérêt, $\text{NMSE} = \frac{1}{100} \sum_{j=1}^{100} \frac{\|\mathbf{Z} - \hat{\mathbf{Z}}_j\|_2^2}{\|\mathbf{Z}\|_2^2}$. Nous montrons d'abord que notre méthode surpasse les méthodes de filtrage spatial [10]. Nous comparons ensuite notre algorithme à une méthode considérant un modèle de bruit robuste, à savoir la t -distribution [66]. Enfin, nous étudions la vitesse de convergence et l'influence du rang de \mathbf{W} pour l'algorithme proposé.

En présence de RFI de forte puissance sur certains canaux fréquentiels, il est possible d'estimer le sous-espace RFI en utilisant les plus grandes valeurs propres de la matrice de corrélation [10]. Cela permet alors de filtrer les visibilité corrompues. Une calibration peut ensuite être effectuée sur les visibilité filtrées. Nous comparons l'algorithme de calibration proposé qui ne nécessite pas d'étape de pré-traitement avec un algorithme de calibration classique effectué sur des visibilité filtrées [11]. Un ciel corrompu par 2 RFI, avec des paramètres de Stokes fixés à [100 10 50 30] et [50 0 0 0], est simulé.

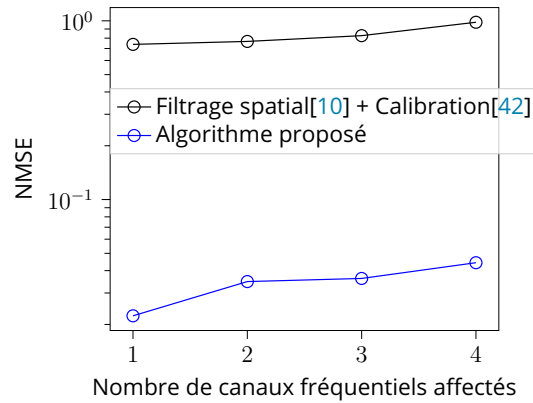


Figure 3.1 – Évolution du NMSE de \mathbf{Z} en fonction du nombre de canaux fréquentiels voisins affectés par les RFI

Dans la figure 3.1, nous représentons l'évolution du NMSE pour le paramètre d'intérêt en fonction du nombre de canaux fréquentiels affectés par des RFI. Il apparaît que l'algorithme de calibration proposé conduit à de meilleures performances. En effet, l'étape de filtrage des RFI peut conduire à une atténuation de certaines sources et être à l'origine de solutions de calibration biaisées. En revanche, l'algorithme proposé effectue la calibration sur les signaux originaux, tout en considérant le sous-espace RFI sans nécessité de pré-traitement. La t -distribution s'avère être une distribution de choix pour modéliser la possible présence de valeurs aberrantes dans les données [66][15][21]. Nous comparons donc le modèle proposé au modèle présenté dans [66],

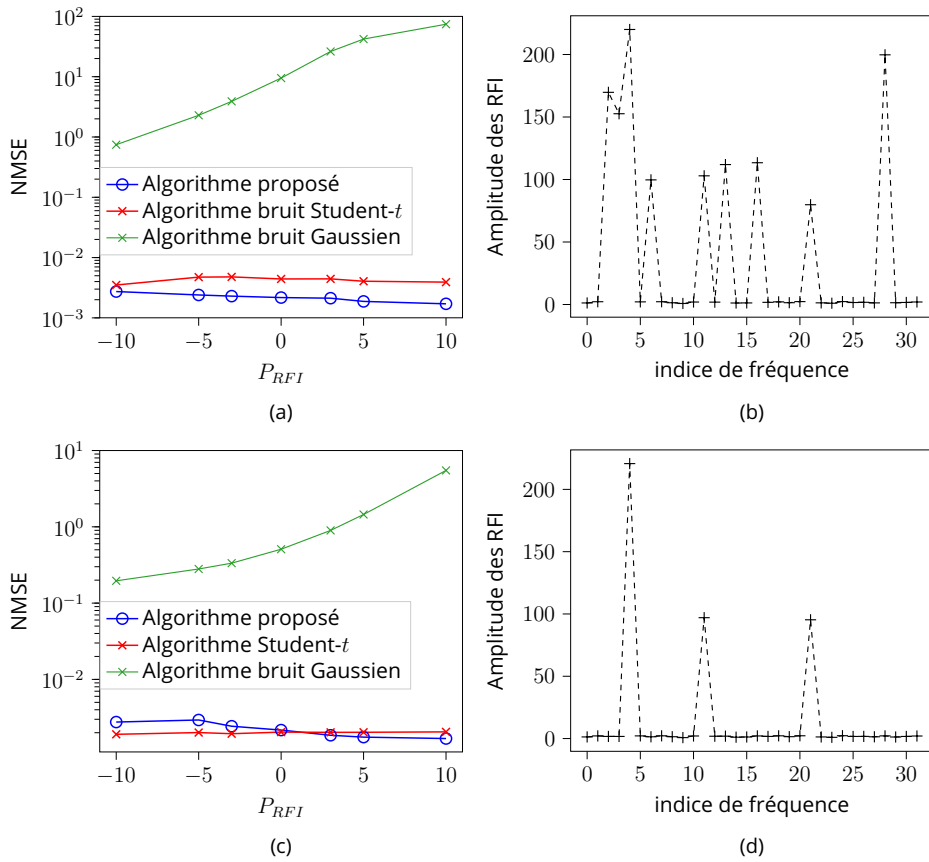


Figure 3.2 – Évolution du NMSE de Z en fonction de la puissance des RFI

en considérant une t -distribution comme bruit additif sur les visibilité. La figure 3.2 montre l'évolution du NMSE en fonction de la puissance des RFI, lorsque seulement quelques bandes de fréquence sont corrompues par des interférences. Des RFI d'une puissance allant de -10dB à 10dB sont ajoutées sur 10% et 30% des canaux fréquentiels. Les figures 3.2.b et 3.2.d présentent l'amplitude des visibilité sur toutes les bandes de fréquence pour une ligne de base lorsque respectivement 10% et 30% des canaux fréquentiels sont affectés par des RFI. La méthode proposée affiche, dans les figures 3.2.a et 3.2.c, des performances similaires aux méthodes de l'état de l'art, en considérant une t -distribution comme bruit additif, bien que de telles méthodes soient sujettes à un bon comportement dans ce contexte. En réalité, les RFI de faible puissance sont présentes sur plusieurs bandes de fréquence [7]. Par conséquent, nous avons simulé un ciel perturbé par des RFI faibles (-15 dB) sur chaque bande de fréquence. De plus, des RFI de forte puissance sont ajoutées à 10% des canaux fréquentiels avec une puissance variant de -10 dB à 10 dB, comme illustré dans la figure 3.3.b pour une ligne de base. Nous représentons dans la figure 3.3.a l'évolution du NMSE en fonction de la puissance des RFI fortes. On remarque une amélioration significative des performances pour l'algorithme

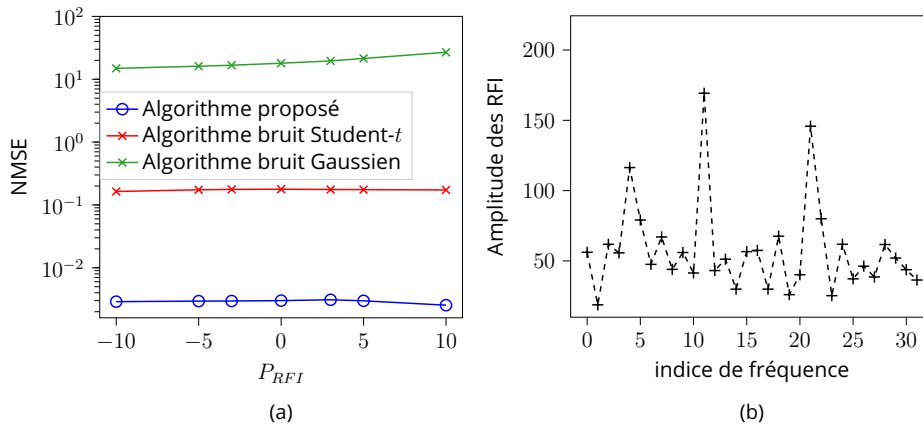


Figure 3.3 – Évolution du NMSE de \mathcal{Z} en fonction de la puissance des RFI de forte puissance en présence de RFI additionnelles de faible puissance sur chaque canal fréquentiel.

proposé par rapport à l'état de l'art. La contribution du modèle proposé réside dans la prise en compte de la structure RFI, ajoutant de la robustesse à la présence de RFI de formes variées.

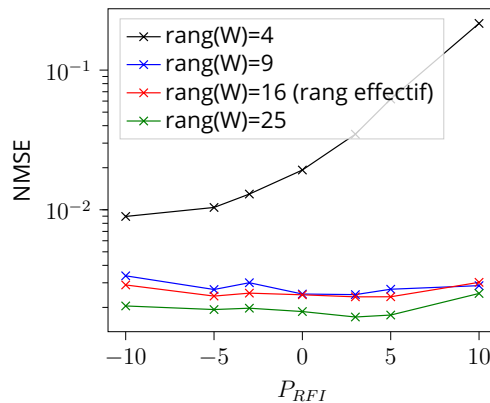


Figure 3.4 – Évolution du NMSE de \mathcal{Z} en fonction de la puissance des RFI pour plusieurs choix du rang de \mathbf{W}

Dans la figure 3.4, l'évolution du NMSE de \mathcal{Z} en fonction du rang choisi est affichée. Le rang effectif utilisé pour simuler les visibilité est $\rho_{\text{eff}} = 16$. Il peut être remarqué qu'une légère erreur d'estimation du rang n'affecte pas les performances de l'algorithme. En effet, l'estimation des paramètres de calibration dépend seulement de la matrice de covariance de la composante liée aux RFI, $\Phi = \mathbf{W}\mathbf{W}^H$. Ainsi, si le rang de \mathbf{W} est fortement sous-estimé, la matrice de covariance du bruit aura également un rang fortement sous-estimé, conduisant à une mauvaise reconstruction. Inversement, comme il n'y a pas de contrainte sur le rang de la matrice de structure des RFI estimé,

\mathbf{W} , la matrice de covariance du bruit peut encore être correctement estimée lorsque le nombre de colonnes de \mathbf{W} est surestimé. En effet, dans ce cas, \mathbf{W} contient des colonnes colinéaires qui n'affectent pas le rang de Φ . Ainsi, en pratique, nous observons que le paramètre d'intérêt, \mathbf{Z} , est correctement estimé, même lorsque le rang de \mathbf{W} est fortement sous-estimé.

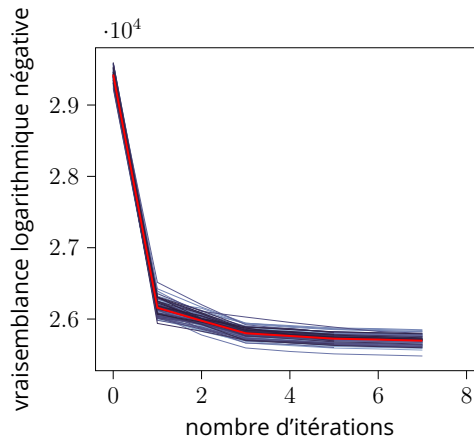


Figure 3.5 – Évolution de la log-vraisemblance négative en fonction du nombre d'itérations pour 100 réalisations de Monte Carlo.

La figure 3.5 présente l'évolution de la vraisemblance logarithmique négative en fonction du nombre d'itérations pour 100 réalisations. Il apparaît que seulement quelques itérations sont nécessaires pour que l'algorithme proposé converge, ce qui conduit à un coût calculatoire raisonnable.

3.4 Conclusion

Dans ce chapitre, nous avons présenté un modèle de données pour la calibration multifréquentielle de radio-interféromètres en présence de RFI. La composante des visibilités liée aux RFI est modélisée par un bruit gaussien de rang faible. Les solutions de calibration sont ensuite obtenues par un maximum de vraisemblance formulé à l'aide d'un algorithme d'Espérance Maximisation sur Espaces Alternés. Des mises à jour explicites de l'étape de maximisation ont pu être dérivées grâce à un choix judicieux de deux espaces de données latentes. Des simulations numériques ont permis de montrer que l'estimateur proposé apporte une robustesse à la présence d'interférences de radiofréquence et démontre un gain en performance comparé aux méthodes de l'état de l'art. De plus, des simulations ont montré qu'une surestimation du rang de la matrice de structure \mathbf{W} n'impacte pas les performances de l'algorithme.

4 - Imagerie interférométrique en présence d'interférences

Dans ce chapitre, nous présentons de nouvelles méthodes d'imagerie pour les radiotélescopes interféromètres en présence de RFI par maximum de vraisemblance. L'imagerie radio-interférométrique consiste à produire une estimée de la densité spectrale de puissance émise dans le spectre des ondes radio sur une portion de la voûte céleste. Nous nous plaçons dans le cadre du théorème de Van-Cittert Zernike par lequel les visibilités mesurées sont modélisées par une DFT de l'image du ciel. Ainsi, la reconstruction d'image peut être effectuée par synthèse de Fourier, en appliquant une DFT inverse aux visibilités mesurées. Cela conduit à une image reconstruite contenant des artefacts dus à la couverture incomplète du plan de Fourier, appelée *dirty image*.

Plusieurs générations d'algorithmes d'imagerie ont permis d'améliorer la qualité des images reconstruites. L'algorithme CLEAN propose de reconstruire l'image du ciel en résolvant un problème de déconvolution, en supposant le ciel composé de sources ponctuelles indépendantes [25]-[27]. Des approches basées sur de l'optimisation convexe permettent de prendre en compte de manière plus précise la nature des images du ciel en considérant des a priori parcimonieux [28], [30], [31].

Les méthodes d'imagerie actuelles considèrent un bruit additif gaussien. Pourtant, l'hypothèse de gaussianité du bruit ne permet pas de représenter fidèlement les observations de radioastronomie, notamment du fait de la présence d'interférences. Dans cette perspective, nous proposons un modèle statistique robuste à la présence de RFI dans les données, à partir duquel nous dérivons un algorithme d'imagerie basé sur une estimation du maximum de vraisemblance régularisée. Plus précisément, nous modélisons la présence d'interférences par des distributions issues de la classe des gaussiennes composées. Dans ce cas, la maximisation de la vraisemblance n'est pas tractable nous conduisant à dériver un algorithme EM pour l'imagerie radio-interférométrique.

Les données interférométriques ont un volume important du fait du nombre d'antennes des réseaux interférométriques et des longues durées d'observation. L'algorithme CLEAN permet notamment une imagerie à faible coût computationnel en agissant directement sur la *dirty image* et donc dont la complexité ne dépend pas de la dimension du vecteur de visibilités. Les approches basées sur de l'optimisation convexe nécessitent d'évaluer le gradient du modèle direct de manière itérative et nécessitent une attention particulière dans le cas de mesure en grande dimension. Une factorisation de l'opérateur permet notamment de réduire le coût calculatoire lié à l'évaluation du modèle, mais reste prohibitive pour un volume de données important. Des implémentations distribuées de ces approches ont notamment été proposées pour palier cette difficulté [29], [30], [52], [76].

Nous proposons alors d'adapter l'algorithme EM proposé en considérant une fac-

torisation de l'opérateur direct du modèle de données permettant d'introduire un opérateur de FFT. L'étape-M de l'EM est alors conduite sur une variable latente dont la maximisation peut se faire à faible coût, ne dépendant pas de la dimension du vecteur de visibilité. Des simulations numériques illustrent les performances des algorithmes proposés en présence d'interférences, comparativement aux algorithmes de l'état de l'art.

4.1 Modèle de données en présence de RFI

Les radio-interféromètres sont sujets à diverses perturbations additives qui ne peuvent pas être modélisées efficacement par une distribution gaussienne. Ainsi, les RFI représentent les ondes radio artificielles pouvant affecter de manière significative les données observées [8] et nécessitent une modélisation spécifique [77].

Cependant, à notre connaissance, une telle modélisation n'a pas été utilisée dans le contexte de l'imagerie radio-interférométrique. En réponse, nous considérons la famille de distributions gaussiennes composées, définie comme le produit de deux variables aléatoires indépendantes. Plus précisément, un composant gaussien appelé le *speckle*, $\mathbf{t} = [t[1], \dots, t[K]]^T$, où K est le nombre de visibilité mesurées, et une *texture* aléatoire positive, τ [78],

$$\mathbf{n}[k] = \sqrt{\tau} \mathbf{t}[k], \quad (4.1)$$

où $\mathbf{t} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ et $\tau \sim p(\tau)$. Dans cette section, nous présentons un algorithme générique d'imagerie radio-interférométrique en tenant compte d'un bruit gaussien composé pour modéliser la présence de RFI. Le modèle direct s'exprime comme suit,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{\Omega}^{-1/2} \mathbf{t}, \quad (4.2)$$

où \mathbf{y} est le vecteur de visibilité, $\mathbf{\Omega} = \text{diag}(\tau_1^{-1}, \dots, \tau_K^{-1})$, $\mathbf{t} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ et $\forall i \in [1, K] \tau_i \sim p(\tau_i)$. L'opérateur direct \mathbf{H} , est décrit dans la section 2.1.

Remarque *Dans le modèle suivant, nous considérons une texture pour chaque visibilité mesurée, ce qui représente la variabilité de la variance du bruit à travers les mesures. Notons que la texture peut être considérée comme partagée entre un ensemble de visibilité pour modéliser une variabilité spécifique du bruit additif. Par exemple, une observation composée de T intégrations peut être modélisée par une texture partagée entre T visibilité, pour représenter la variabilité temporelle du bruit additif. Inversement, la variabilité spatiale du bruit additif peut être modélisée par une texture partagée entre un ensemble de visibilité mesurées en utilisant la même ligne de base à différents moments. Ces deux exemples correspondraient alors des textures corrélées [79].*

Nous pouvons exprimer la vraisemblance des données observées comme suit,

$$\mathcal{L}(\mathbf{x}; \mathbf{y}, \sigma^2) = \sum_{k=1}^K \log \left(\int_0^\infty (\pi \sigma^2 \tau)^{-1} \exp((\sigma^2 \tau)^{-1} \delta[k]^2) p(\tau) d\tau \right), \quad (4.3)$$

avec $\delta[k] = |\mathbf{y}[k] - \mathbf{h}_k^T \mathbf{x}|$, où \mathbf{h}_k est la $k^{\text{ème}}$ ligne de \mathbf{H} .

Le problème de maximum de vraisemblance régularisé résultant pour l'imagerie radio-interférométrique est formulé en tenant compte d'une régularisation sur le vecteur de l'image du ciel,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}} (\mathcal{L}(\mathbf{x}; \mathbf{y}) + r(\mathbf{x})), \quad (4.4)$$

où $r(\cdot)$ est une fonction de régularisation convexe permettant de favoriser une parcimonie des images du ciel, en accord avec la nature des sources observées. En particulier, la norme ℓ_1 appliquée sur une certaine base est couramment utilisée pour modéliser la parcimonie inhérente des images du ciel [28].

4.2 Algorithme EM pour l'imagerie en présence de RFI

Le problème d'estimation du maximum de vraisemblance régularisé, tel que présenté ci-dessous, est souvent intractable pour de telles distributions. Dans de tels cas, l'algorithme EM offre une alternative en augmentant itérativement la vraisemblance par la maximisation de l'espérance d'une vraisemblance complète [72].

Pour ce faire, nous considérons comme espace de données complet $\boldsymbol{\xi} = (\mathbf{y}, \tau_1, \dots, \tau_K)$, à partir duquel nous dérivons un algorithme EM. Les paramètres d'intérêt comprennent l'image du ciel et la variance globale du bruit, $\boldsymbol{\theta} = (\mathbf{x}, \sigma^2)$. L'espérance de la log-vraisemblance des données complètes, sachant les données observées incomplètes et l'estimation courante de l'image du ciel, est calculée dans une étape d'espérance (étape-E) comme suit :

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) = \mathbb{E}_{(\mathbf{y}, \tau_1, \dots, \tau_K) | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\log(p(\mathbf{y}, \tau_1, \dots, \tau_K | \boldsymbol{\theta}))]. \quad (4.5)$$

Cette dernière est ensuite maximisée dans une étape de maximisation (étape-M) pour produire la nouvelle estimation de l'image du ciel $\mathbf{x}^{(m+1)}$. Ainsi, l'algorithme EM met en œuvre de manière itérative une étape-E suivie d'une étape-M jusqu'à ce que la convergence soit atteinte.

Proposition 3 *L'étape-E pour les données complètes considérées s'exprime comme suit :*

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) = \sum_{k=1}^K \log(\sigma^2) - \log(\hat{\omega}_k^{(m)}) + \frac{\hat{\omega}_k^{(m)}}{\sigma^2} |\mathbf{y} - \mathbf{h}_k^T \mathbf{x}|^2, \quad (4.6)$$

$$\hat{\omega}_k^{(m)} = \mathbb{E}_{\tau | \mathbf{y}[k], \boldsymbol{\theta}^{(m)}} [\tau^{-1}] = \int_{\mathbb{R}^+} \frac{1}{\tau} p(\tau | \mathbf{y}[k], \boldsymbol{\theta}^{(m)}) d\tau \quad (4.7)$$

Démonstration Voir l'annexe B.1.

L'étape-M consiste à résoudre le problème d'optimisation suivant :

$$\hat{\boldsymbol{\theta}}^{(m+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \left(Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) + r(\boldsymbol{\theta}) \right), \quad (4.8)$$

où $\hat{\Omega}^{(m)} = \text{diag}(\hat{\omega}_1^{(m)}, \dots, \hat{\omega}_K^{(m)})$. Le problème d'optimisation obtenu peut être résolu de manière alternée pour produire la nouvelle estimation de l'image du ciel $\mathbf{x}^{(m+1)}$ et la nouvelle estimation de la variance du bruit $\sigma^{2(m+1)}$:

$$\mathbf{x}^{(m+1)} = \underset{\mathbf{x}}{\text{argmin}} \left((\mathbf{y} - \mathbf{H}\mathbf{x})^H \hat{\Omega}^{(m)} (\mathbf{y} - \mathbf{H}\mathbf{x}) + r(\mathbf{x}) \right), \quad (4.9)$$

$$\sigma^{2(m+1)} = \frac{1}{K} \|\mathbf{y} - \mathbf{H}\mathbf{x}^{(m+1)}\|_{\hat{\Omega}^{(m)}}^2. \quad (4.10)$$

Remarque La résolution de (4.9) pour obtenir une image du ciel correspond à un problème de moindres carrés pondérés régularisé par $r(\mathbf{x})$. Pour que l'algorithme EM converge vers un maximum local de la vraisemblance, il suffit que l'étape-M augmente la fonctionnelle $Q(\mathbf{x}|\mathbf{x}^{(m)})$ [72]. En conséquence, l'étape-M peut être résolue à l'aide d'algorithmes d'imagerie fournissant au moins une solution approximative de (4.9) [73]. Cela permet d'améliorer la robustesse à la présence d'interférences dans les données de la plupart des algorithmes d'imagerie radio-interférométrique existants, avec une mise en œuvre légère et des propriétés de convergence connues. Plus précisément, l'étape-M de l'algorithme d'imagerie proposé peut être mise en œuvre à l'aide de méthodes basées sur CLEAN, ainsi que d'algorithmes proximaux, tout en fournissant une estimation du maximum de vraisemblance de l'image du ciel robuste à la présence d'interférences dans les données.

Algorithm 4.1 EM pour l'imagerie radio-interférométrique en présence de RFI

Entrée : \mathbf{y}, \mathbf{H}

Sortie : $\hat{\mathbf{x}}, \hat{\sigma}^2$

Initialisation : $\hat{\mathbf{x}} \leftarrow \mathbf{x}_0, \hat{\sigma}^2 \leftarrow \sigma_0^2$

- 1: **while** critère d'arrêt non atteint **do**
 - 2: Étape-E :
 - 3: $\hat{\omega}_k^{(m)}$ obtenu à partir de (4.7) pour $n \in \llbracket 1, K \rrbracket$
 - 4: Étape-M :
 - 5: $\hat{\mathbf{x}}^{(m+1)}$ obtenu à partir de (4.9)
 - 6: $\hat{\sigma}^{2(m+1)}$ obtenu à partir de (4.10)
 - 7: **end while**
-

4.3 Factorisation du modèle direct pour un algorithme EM accéléré

Les mesures radio-interférométriques sont des échantillons de la transformée de Fourier discrète (DFT) de l'image du ciel aux coordonnées uv continues. Ces derniers sont échantillonnés sur une grille non uniforme, qui dépend des lignes de base du réseau interférométrique. Par conséquent, le modèle direct pour l'imagerie radio-interférométrique est une transformée de Fourier non uniforme qui ne peut pas être calculée efficacement à l'aide de la FFT [1]. De plus, les mesures réelles en radio-interférométrie sont composées d'un grand nombre de visibilités, ce qui entraîne des

coûts calculatoires élevés pour l'évaluation du modèle direct. En réponse à cela, la plupart des algorithmes d'imagerie radio-interférométrique considèrent une factorisation du modèle direct qui introduit un opérateur de FFT [34],

$$\mathbf{H} \approx \mathbf{S}\mathbf{F}, \quad (4.11)$$

où $\mathbf{S} \in \mathbb{R}^{K \times N}$ représente un opérateur de convolution et $\mathbf{F} \in \mathbb{C}^{N \times N}$ est une matrice de Fourier sur une grille uniforme [1], [34]. Dans cette section, nous exploitons cette factorisation de l'opérateur radio-interférométrique pour concevoir un algorithme EM avec une étape-M plus rapide, qui est effectuée sur les visibilités qui se trouvent sur une grille uniforme. Nous considérons comme variable latente les visibilités virtuelles qui auraient été produites par un radio-interféromètre dont les coordonnées uv sont situées sur une grille uniforme, à savoir,

$$\mathbf{z} = \mathbf{F}\mathbf{x} + \mathbf{e}, \quad (4.12)$$

où $\mathbf{e} \sim \mathcal{CN}(0, \sigma_e^2 \mathbf{I})$ est un bruit gaussien complexe. Dans ce cas, le modèle direct est une transformée de Fourier uniforme qui peut être calculée efficacement à l'aide de la FFT. Étant donné la composante de texture du bruit gaussien composé, $\mathbf{\Omega}$, nous pouvons exprimer les visibilités mesurées en fonction de la variable latente \mathbf{z} en tirant parti de la factorisation de l'opérateur radio-interférométrique,

$$\begin{aligned} \mathbf{y}_{|\Omega} &= \mathbf{S}\mathbf{F}\mathbf{x} + \mathbf{\Omega}^{-1/2}\mathbf{t} = \mathbf{S}\mathbf{z} + \mathbf{n}', \\ \mathbf{n}'_{|\Omega} &= \mathbf{\Omega}^{-1/2}\mathbf{t} - \mathbf{S}\mathbf{e}, \end{aligned} \quad (4.13)$$

où $\mathbf{n}'_{|\Omega} \sim \mathcal{CN}(0, \mathbf{\Sigma}')$ est un bruit gaussien complexe avec une matrice de covariance $\mathbf{\Sigma}' = \sigma^2 \mathbf{\Omega}^{-1} - \sigma_e^2 \mathbf{S}\mathbf{S}^T$.

Remarque Notons que pour que la décomposition du bruit obtenue soit valide, la matrice de covariance de $\mathbf{n}'_{|\Omega}$ doit être semi-définie positive, c'est-à-dire, $\sigma_e^2 \leq \frac{\sigma^2}{\lambda_{\max} \tau_{\max}}$, où λ_{\max} est la plus grande valeur propre de $\mathbf{S}^T \mathbf{S}$ et $\tau_{\max} = \max(\tau_1, \dots, \tau_N)$.

Dans ce cas, l'étape-M de l'algorithme EM est réalisée sur la variable latente \mathbf{z} qui se trouve sur une grille uniforme. Cela donne un modèle direct qui peut être calculé efficacement à l'aide de la FFT et qui ne dépend pas de la dimension du vecteur de visibilité \mathbf{y} . La variance du bruit σ^2 ne bénéficie pas de cette factorisation et peut être estimée en utilisant l'espace des données complètes défini dans la section précédente. Par conséquent, nous proposons d'estimer les paramètres inconnus, $\boldsymbol{\theta} = (\mathbf{x}, \sigma^2)$, en utilisant un SAGE [64]. Plus précisément, cela consiste à alterner entre l'estimation de \mathbf{x} et de σ^2 en utilisant des espaces de données complètes distincts. Ainsi, nous définissons les espaces de données complètes suivants pour estimer l'image du ciel :

$$\boldsymbol{\xi}^{[1]} = (\mathbf{y}, \mathbf{z}, \tau_1, \dots, \tau_N). \quad (4.14)$$

La variance du bruit, σ^2 , est estimée en utilisant l'espace de données complètes défini dans la section précédente :

$$\boldsymbol{\xi}^{[2]} = (\mathbf{y}, \tau_1, \dots, \tau_N). \quad (4.15)$$

L'algorithme EM associé s'exprime alors de la manière suivante :

$$\mathbf{x}^{(m+1)} = \underset{\mathbf{x}}{\operatorname{argmin}} \mathbb{E}_{\boldsymbol{\xi}^{[1]}|\mathbf{y}, \mathbf{x}^{(m)}} \left[-\log p(\boldsymbol{\xi}^{[1]}; \mathbf{x}) \right], \quad (4.16)$$

$$\sigma^{2(m+1)} = \underset{\sigma^2}{\operatorname{argmin}} \mathbb{E}_{\boldsymbol{\xi}^{[2]}|\mathbf{y}, \sigma^{2(m)}} \left[-\log p(\boldsymbol{\xi}^{[2]}; \sigma^2) \right]. \quad (4.17)$$

Proposition 4 *L'étape-E pour les données complètes considérées est donnée par :*

$$Q_1(\mathbf{x}|\mathbf{x}^{(m)}) \propto \mathbb{E}_{\boldsymbol{\xi}^{[1]}|\mathbf{y}, \mathbf{x}^{(m)}} [\log p(\mathbf{z} | (\tau_n)_{n \leq N}; \mathbf{x})] \propto (\hat{\mathbf{z}}^{(m)} - \mathbf{F}\mathbf{x})^H (\hat{\mathbf{z}}^{(m)} - \mathbf{F}\mathbf{x}) \quad (4.18)$$

$$Q_2(\sigma^2|\sigma^{2(m)}) \propto K \log(\sigma^2) + \frac{1}{\sigma^2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_{\hat{\boldsymbol{\Omega}}^{(m)}}^2 \quad (4.19)$$

$$\hat{\mathbf{z}}^{(m)} = \mathbb{E}_{\boldsymbol{\xi}^{[1]}|\mathbf{y}, \mathbf{x}^{(m)}} [\mathbf{z}] = \mathbf{F}\mathbf{x}^{(m)} + \frac{\sigma_e^2}{\sigma^2} \mathbf{S}^T \hat{\boldsymbol{\Omega}}^{(m)} (\mathbf{y} - \mathbf{H}\mathbf{x}^{(m)}) \quad (4.20)$$

$$\hat{\omega}_k^{(m)} = \mathbb{E}_{\tau|\mathbf{y}[k], \mathbf{x}^{(m)}} [\tau^{-1}] = \int_{\mathbb{R}^+} \frac{1}{\tau} p(\tau|\mathbf{y}[k], \mathbf{x}^{(m)}) d\tau, \quad (4.21)$$

où $\hat{\boldsymbol{\Omega}}^{(m)} = \operatorname{diag}(\hat{\omega}_1^{(m)}, \dots, \hat{\omega}_N^{(m)})$.

Démonstration *La démonstration est donnée dans l'annexe B.2.*

Les étapes de maximisation sont alors exprimées comme suit :

$$\hat{\mathbf{x}}^{(m+1)} = \underset{\mathbf{x}}{\operatorname{argmax}} (\|\hat{\mathbf{z}} - \mathbf{F}\mathbf{x}\|_2^2 + r(\mathbf{x})), \quad (4.22)$$

$$\sigma^{2(m+1)} = \frac{1}{K} \|\mathbf{y} - \mathbf{H}\mathbf{x}^{(m+1)}\|_{\hat{\boldsymbol{\Omega}}^{(m)}}^2. \quad (4.23)$$

Remarque *Par rapport à l'algorithme EM proposé dans la section 4.2, l'étape-M conduisant à l'estimation de l'image du ciel est exprimée dans une base de Fourier complète, ce qui permet d'utiliser la FFT pour réduire considérablement son coût computationnel. En effet, de cette manière, la complexité de calcul d'une itération repose sur une multiplication par une matrice de Fourier et est mise en œuvre par des algorithmes de FFT en $N \log(N)$, où N est la dimension de l'image. Le coût computationnel résultant est donc indépendant du nombre de visibilités mesurées. En comparaison, la complexité de calcul pour l'évaluation de l'opérateur direct nécessaire dans les algorithmes d'imagerie de l'état de l'art, \mathbf{H} , est de $O(KN \log(N))$ avec, K étant le nombre de visibilités.*

Algorithm 4.2 SAGE pour l'imagerie radio-interférométrique accéléré en présence d'interférences radioélectriques

Entrée : $\mathbf{y}, \mathbf{S}, \mathbf{F}, \sigma_e^2$

Sortie : $\hat{\mathbf{x}}$

Initialisation : $\hat{\mathbf{x}} \leftarrow \mathbf{x}_0$

- 1: **while** critère d'arrêt non atteint **do**
 - 2: $\hat{\omega}_n^{(m)}$ obtenu à partir de (4.21) pour $k \in \llbracket 1, K \rrbracket$ # SAGE étape 1 - Étape-E
 - 3: $\hat{z}^{(m)}$ obtenu à partir de (4.20) # SAGE étape 1 - Étape-E
 - 4: $\hat{\mathbf{x}}^{(m+1)}$ obtenu à partir de (4.22) # SAGE étape 1 - Étape-M
 - 5: $\hat{\omega}_n^{(m+\frac{1}{2})}$ obtenu à partir de (4.21) pour $k \in \llbracket 1, K \rrbracket$ # SAGE étape 2 - Étape-E
 - 6: $\hat{\sigma}^{2(m+1)}$ obtenu à partir de (4.23) # SAGE étape 2 - Étape-M
 - 7: **end while**
-

4.4 Résultats numériques

4.4.1 Simulation des visibilitées

Dans cette section, nous comparons l'algorithme EM proposé à des algorithmes d'imagerie radio-interférométrique de l'état de l'art. Nous simulons des images du ciel de taille 256×256 pixels. Plus précisément, nous considérons des images du ciel composées d'ellipsoïdes gaussiens conformes à la forme attendue des galaxies. Le nombre de sources dans les images du ciel est échantillonné à partir d'une distribution de Poisson avec une moyenne de 15 sources par image. La puissance des sources est discrétisée à partir d'une distribution uniforme entre 0 et 10 en unité Jy. L'échelle des ellipsoïdes gaussiens est discrétisée à partir d'une distribution uniforme entre 5 et 20 pixels. L'angle de position est échantillonné à partir d'une distribution uniforme entre 0 et π .

Les visibilitées sont simulées à partir des images du ciel en utilisant le package Python *ducc0* [35]. Nous utilisons les positions des antennes du réseau radio-interférométrique Meerkat avec 64 antennes et 1 canal de fréquence à 1,4 GHz pour une observation de 8 heures. L'annexe D présente le processus de simulation des visibilitées en détail. Un bruit gaussien est ajouté aux visibilitées avec un rapport signal sur bruit de 10dB. Nous simulons 1 000 RFI par temps d'intégration en tant que sources ponctuelles supplémentaires dans la scène. La puissance de chaque RFI est définie par rapport à la puissance de la source, $P_{\text{RFI, dB}} = 10 \log(\frac{P_{\text{RFI}}}{P_0})$, où P_0 est la puissance de la source et P_{RFI} est la puissance de la source interférente.

4.4.2 Le cas de la t -distribution

Nous considérons le cas de la distribution gamma inverse pour la texture de la distribution gaussienne composée, ce qui conduit à une t -distribution pour le bruit additif. Cela conduit à une étape-E pour les algorithmes proposés qui peut être calculée analytiquement,

$$\hat{\tau}_k^{(m)} = \frac{\nu + 1}{\nu + \sigma^{-2} |\mathbf{y}[k] - \mathbf{h}_k^T \mathbf{x}^{(m)}|^2}, \quad (4.24)$$

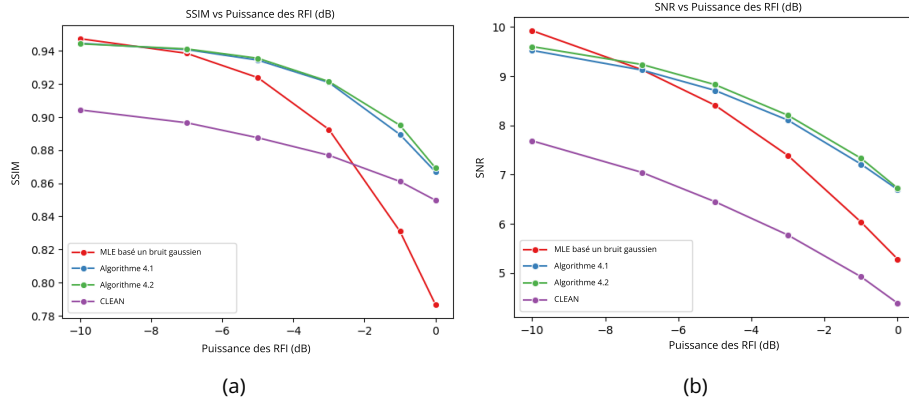


Figure 4.1 – Évolution de (a) la SSIM et (b) du SNR des images en fonction la puissance des RFI.

où ν est le nombre de degrés de liberté de la t -distribution et h_k est la $k^{\text{ème}}$ ligne de \mathbf{H} .

4.4.3 Régularisation dans le domaine de l'image

Nous choisissons comme régularisation la norme ℓ_1 , qui favorise la parcimonie dans le domaine de l'image, conformément à la nature des images du ciel [28]. De plus, nous ajoutons une contrainte de positivité sur les pixels de l'image pour assurer la validité physique de la solution, ce qui conduit à la fonction de régularisation suivante,

$$r(\mathbf{x}) = \alpha \|\mathbf{x}\|_1 + \iota_{\mathbb{R}^+}(\mathbf{x}), \quad (4.25)$$

où α est le paramètre de régularisation et $\iota_{\mathbb{R}^+}$ est la fonction indicatrice,

$$\iota_{\mathbb{R}^+}(\mathbf{x}) = \begin{cases} 0 & \text{si } \mathbf{x} \in \mathbb{R}^+ \\ +\infty & \text{sinon.} \end{cases} \quad (4.26)$$

L'étape-M est ensuite résolue par un algorithme de gradient proximal [49]. Pour cela, l'opérateur proximal de la fonction de régularisation, $r(\cdot)$, s'écrit,

$$\text{prox}_r(\mathbf{x}) = \max(0, \mathbf{x} - \alpha). \quad (4.27)$$

4.4.4 Résultats de simulation

Nous comparons les algorithmes EM proposés aux algorithmes d'imagerie standard qui considèrent une norme ℓ_2 comme terme d'attache aux données, ainsi qu'à l'algorithme multi-échelle proposé dans [27]. Nous évaluons les performances des différents algorithmes en termes de SNR (rapport signal sur bruit) défini par : $\text{SNR} = 10 \log_{10} \left(\frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2} \right)$, où \mathbf{x} est l'image originale et $\hat{\mathbf{x}}$ est l'image reconstruite. Nous utilisons également l'indice de similarité structurale (SSIM) pour évaluer les performances des algorithmes d'imagerie en termes de qualité d'image [80]. La figure 4.1 montre le SNR et le SSIM des différents algorithmes en fonction de la puissance simulée des interférences, et démontre les performances améliorées en précisions des algorithmes EM

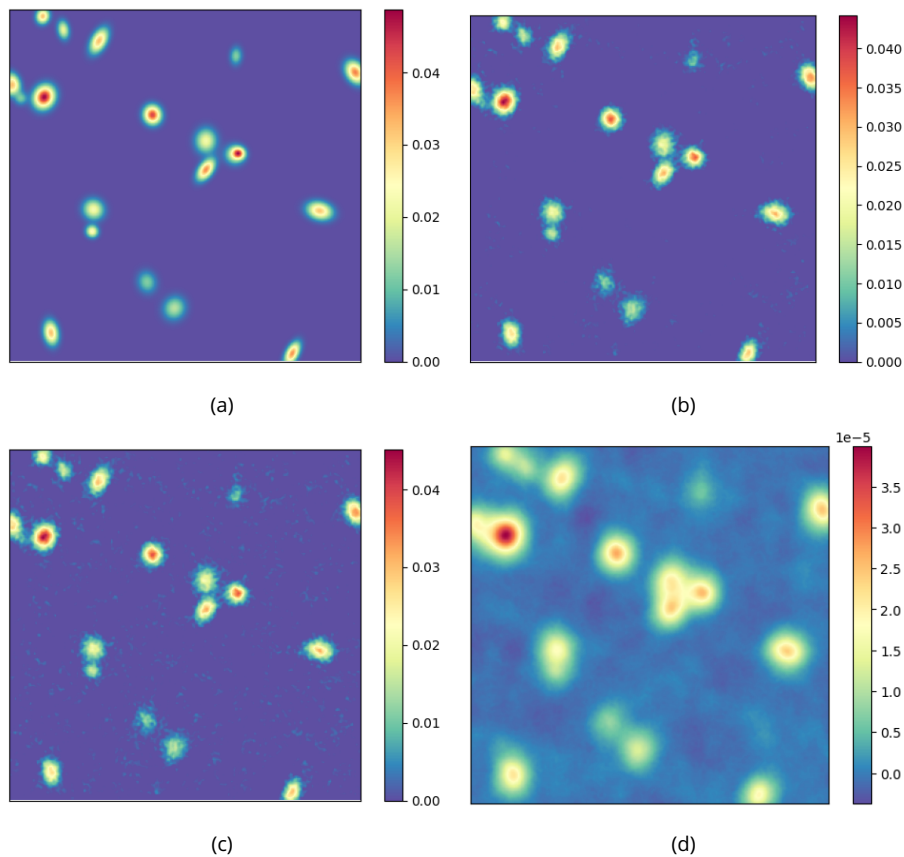


Figure 4.2 – Exemples d’images reconstruites pour un ciel simulé de 25 sources en présence de RFI de puissance -5dB. (a) Image originale, (b) Algorithme EM proposé, (c) Algorithme d’imagerie basé sur un bruit gaussien, (d) *dirty image*.

proposés, par rapport aux méthodes de l’état de l’art. Dans les figures 4.2 et 4.3, nous présentons les images reconstruites par les différents algorithmes pour une image simulée du ciel comprenant 25 sources et une puissance d’interférence de -5 dB.

4.4.5 Efficacité computationnelle

La figure 4.4 met en évidence l’efficacité computationnelle de l’algorithme EM accéléré, proposé dans l’algorithme 4.2 par rapport à l’algorithme 4.1. Elle montre le temps CPU requis pour l’étape-M de l’algorithme EM en fonction du nombre de visibilité mesurées et met en évidence l’efficacité computationnelle améliorée de l’algorithme EM accéléré, par rapport à l’algorithme 4.1. Dans le cas de l’algorithme 4.1, le temps de calcul requis pour l’étape-M de l’algorithme EM est linéaire par rapport au nombre de visibilité mesurées. En revanche, dans le cas de l’algorithme EM accéléré proposé, le temps de calcul requis pour l’étape-M est indépendant du nombre de visibilité mesurées et dépend uniquement de la taille de l’image à reconstruire. Par ailleurs, nous observons que le temps de calcul requis pour l’étape-M de l’algorithme EM accéléré est plus long que le temps de calcul requis pour CLEAN, mais permet d’apporter une

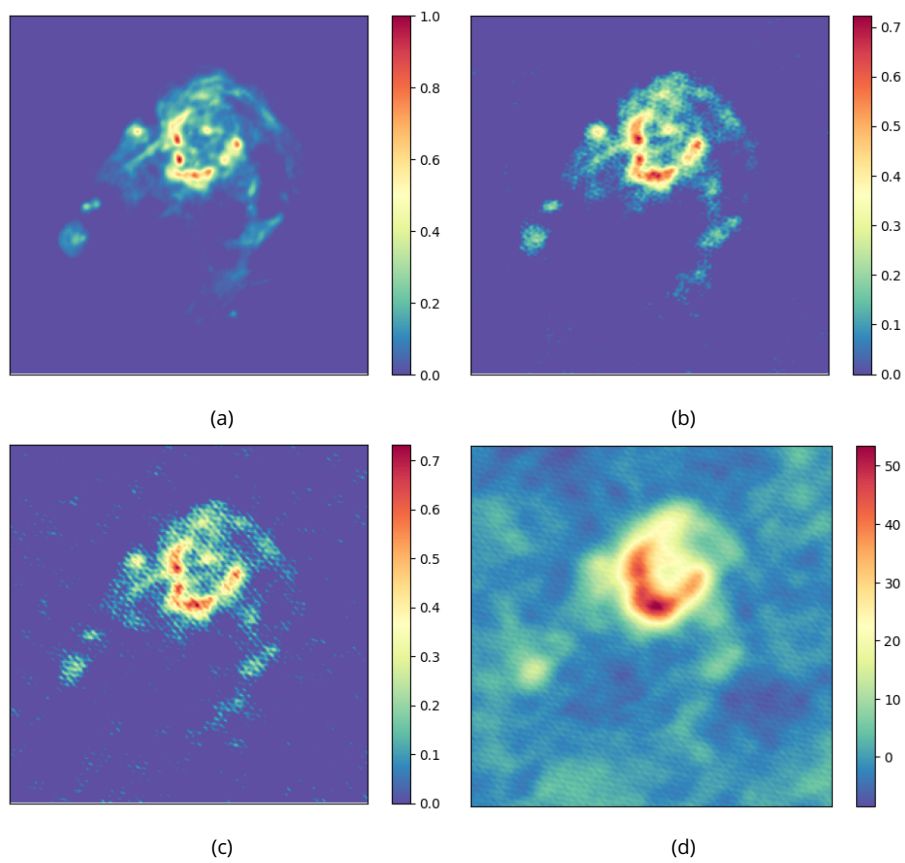


Figure 4.3 – Images de la radiogalaxie *M31* reconstruite par différents algorithmes en présence de RFI de puissance -4dB. (a) Image originale, (b) Algorithme EM proposé, (c) Algorithme d'imagerie basé sur un bruit gaussien, (d) *dirty image*.

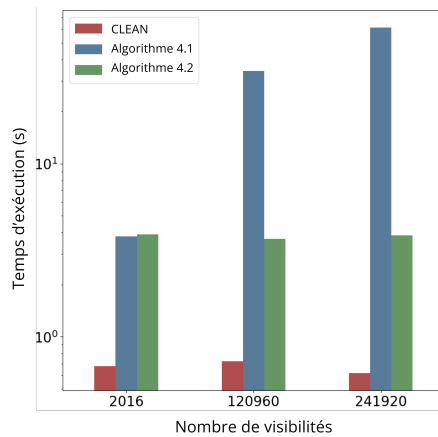


Figure 4.4 – Temps d’exécution des différents algorithmes d’imagerie en fonction du nombre de visibilités.

amélioration de la qualité de l’image reconstruite.

4.5 EM déroulé : Travaux en cours et résultats préliminaires

4.5.1 Contexte

Dans les sections précédentes, nous avons présenté deux algorithmes EM pour des modèles de bruit gaussien composé. Nous avons explicité le calcul de l’étape-E dans le cas particulier de la t -distribution. Toutefois, cette étape-E est difficile à calculer dans le cas général d’une distribution gaussienne composée.

Par ailleurs, la prise en compte de la distribution du bruit peut être mise en oeuvre au travers de méthodes basées sur un apprentissage statistique [81], [82]. Toutefois, dans le contexte de la radio-interférométrie, la plupart des architectures neuronales introduites sont complexes et produisent des poids qui ne sont pas directement interprétables, contrairement aux méthodes d’optimisation basées sur une modélisation physique du signal.

En ce sens, nous proposons une méthodologie d’apprentissage informée pour l’imagerie radio-interférométrique. Pour cela, nous considérons le cadre introduit dans [83] qui a conduit à l’émergence d’architectures neuronales interprétables, dérivées à partir d’algorithmes itératifs. En particulier, nous proposons d’appliquer cette méthodologie d’apprentissage informée à l’algorithme 4.1. Ainsi, nous considérons une architecture neuronale dépliée de l’algorithme EM, où chaque couche est définie par une itération de l’algorithme EM. L’étape-E de l’algorithme EM est remplacée par une architecture neuronale spécifique, dont les poids sont appris à partir d’une base d’apprentissage.

L’avantage de ces approches est de produire une architecture explicable à partir d’un algorithme basé sur une modélisation des données. De plus, l’inférence est plus rapide, comparée à un algorithme itératif, en raison d’un apprentissage supervisé [84]. De

plus cela nous permet d'identifier les poids liés à la statistique du bruit, ce qui conduit à un apprentissage ciblé du réseau et à faible coût. La robustesse de la méthodologie proposée à la présence d'interférences est illustrée sur des données simulées. Les poids de l'architecture résultante, pour lesquels un apprentissage supervisé est conduit, pourront être directement interprétables à partir du modèle dont est issu l'algorithme itératif. Ces derniers correspondent à la minimisation d'un critère obtenu à partir d'une modélisation physique. Il a été montré que de telles architectures conduisent à un apprentissage à plus bas coût et à une accélération de l'inférence [84].

4.5.2 Architecture déroulée

Nous proposons de déplier l'algorithme 4.1 et d'en apprendre l'étape-E de manière supervisée. Nous considérons une régularisation par la norme ℓ_1 de l'image du ciel, $r(\mathbf{x}) = \alpha \|\mathbf{x}\|_1$, où $\alpha > 0$ est le paramètre de régularisation. Pour une itération de l'algorithme 4.1, la fonction de mise à jour peut se réécrire comme une fonction paramétrique,

$$\begin{aligned}\hat{\omega}^{(m+1)} &= f_{\Theta}(\mathbf{y}, \hat{\mathbf{x}}^{(m)}, \hat{\omega}^{(m)}), \\ \hat{\mathbf{x}}^{(m+1)} &= \text{prox}_r \left(\hat{\mathbf{x}}^{(m)} + \gamma \mathbf{H}^H \hat{\Omega}^{(m+1)} (\mathbf{y} - \mathbf{H} \hat{\mathbf{x}}^{(m)}) \right),\end{aligned}\tag{4.28}$$

où $\text{prox}_r(\mathbf{x}) = \text{sign}(\mathbf{x}) \max(0, |\mathbf{x}| - \alpha)$ est la fonction de seuillage doux et $f_{\Theta}(\mathbf{y}, \hat{\mathbf{x}}^{(m)}, \hat{\omega}^{(m)})$ est une architecture neuronale dont Θ contient l'ensemble des paramètres à apprendre. La fonction paramétrée, $f_{\Theta}(\cdot)$, remplace le calcul d'espérance en (4.7). L'architecture obtenue correspond à une architecture neuronale récurrente. Néanmoins, les réseaux de neurones récurrents sont sujets à une dispersion du gradient lors de l'apprentissage [85]. En conséquence, nous proposons l'architecture suivante pour $f_{\Theta}(\cdot)$, à partir de la méthodologie introduite dans [86], illustrée en figure 4.5,

$$\mathbf{h}^{(m)} = \text{Sigmoid}(\mathbf{W}_0 \|\mathbf{y} - \mathbf{H} \mathbf{x}^{(m)}\|_2^2),\tag{4.29}$$

$$f_{\Theta}(\mathbf{y}, \hat{\mathbf{x}}^{(m)}, \hat{\omega}^{(m)}) = \text{ReLU}(\mathbf{W}_1 \hat{\omega}^{(m)} + \mathbf{W}_2 \mathbf{h}^{(m)}),\tag{4.30}$$

où $\mathbf{W}_0 \in \mathbb{R}^{K \times K}$, $\mathbf{W}_1 \in \mathbb{R}^{K \times K}$ et $\mathbf{W}_2 \in \mathbb{R}^{K \times K}$ représentent des couches linéaires, K étant la dimension du vecteur de mesure. L'architecture proposée, inspirée des architectures récurrentes à activation [86], est la combinaison linéaire de la pondération calculée à l'itération précédente, $\hat{\omega}^{(m)}$, et de la sortie de la couche cachée $\mathbf{h} \in [0, 1]$. Une fonction d'activation ReLU permet de stabiliser le gradient lors de l'entraînement [86]. Par conséquent, aucune hypothèse sur la distribution du bruit n'est effectuée. L'apprentissage sur des données perturbées par des bruits non gaussiens permet d'obtenir une architecture robuste à de telles perturbations. L'image obtenue peut être interprétée comme la solution d'un problème des moindres carrés pondéré pour un modèle linéaire gaussien où les pondérations sont obtenues par 4.30.

4.5.3 Résultats numériques

Dans cette section, nous présentons des résultats numériques et comparons les performances du réseau déplié sur 10 itérations avec l'algorithme EM proposé en terme de

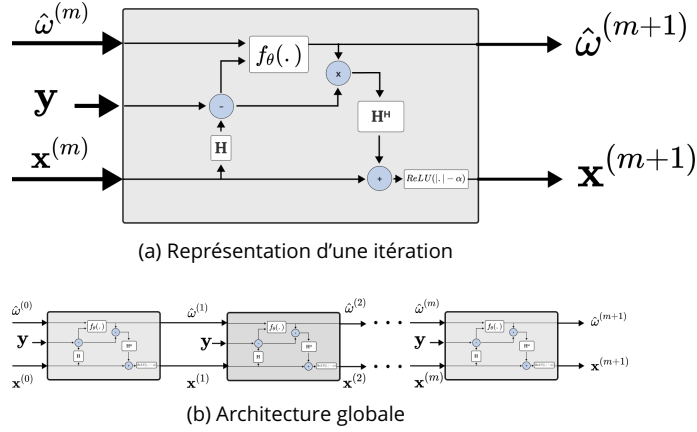


Figure 4.5 – Diagramme de l'architecture neuronale dépliée

SNR et de SSIM.

L'ensemble de données d'entraînement est construit à partir d'un modèle de sources célestes. En effet, les émissions d'ondes radio provenant de galaxies correspondent à des ellipsoïdes gaussiennes et ont été efficacement utilisées comme données d'entraînement dans des approches supervisées [81]. Nous dérivons alors un ensemble de données d'entraînement à partir d'un ensemble de modèles de ciel composés de sources gaussiennes. Plus précisément, nous considérons un ensemble de données avec 1 000 images de taille 64×64 pixels, composées d'un certain nombre de sources gaussiennes aléatoires, issu d'une distribution uniforme, $N_{\text{sources}} \sim \mathcal{U}(2, 10)$. De même, pour chaque source, la puissance et l'échelle sont issues d'une distribution uniforme, $P \sim \mathcal{U}(0.5, 10)$ et $\sigma \sim \mathcal{U}(2, 3)$. Pour chaque image simulée, les visibilité sans bruit correspondantes sont générées à partir de l'opérateur direct, tel que défini dans (2.11). Nous considérons pour cela un réseau interférométrique dont les positions des antennes ont été tirées aléatoirement. La figure 4.6 montre le plan uv correspondant. Un bruit issu d'une mixture de gaussiennes composées variées est ensuite ajouté aux données d'entraînement et permet de simuler la présence d'interférences dans les mesures. Plus précisément, on considère une mixture de gaussiennes composées dont la texture est issue d'une distribution gamma, d'une inverse gamma ou d'une gaussienne inversée [87]. Dans ce cas, l'étape-E de l'algorithme EM proposé ne peut être calculée de manière explicite. Nous utilisons l'erreur quadratique moyenne comme fonction coût pour l'entraînement, $l(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2$, effectué sur 1 000 *epochs*. Nous utilisons l'optimiseur Adam [88] avec un pas d'apprentissage de 0.0001. La différenciation automatique est effectuée avec Pytorch (1.13.1).

La figure 4.7 permet de comparer les métriques sur un jeu de données de test perturbées par une mixture de gaussiennes composées pour l'architecture dépliée et l'algorithme EM considérant un bruit t -distribué.

Pour l'ensemble des métriques considérées, nous affichons les performances de l'algorithme EM pour 100 itérations, ainsi que celles de l'architecture dépliée sur 10 ité-

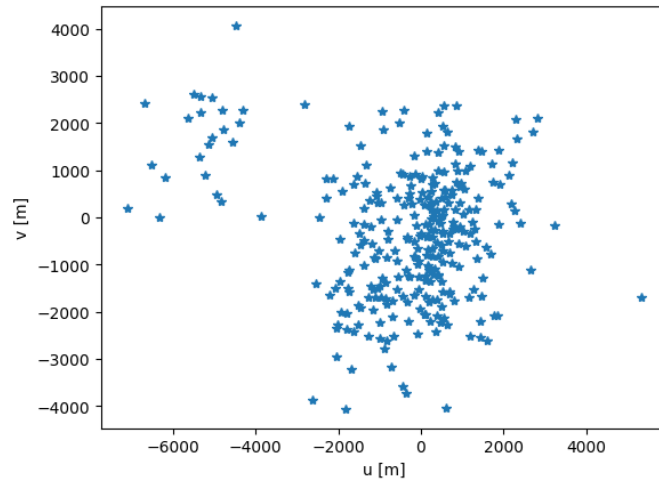


Figure 4.6 – Plan uv utilisé pour simuler les jeux de données d’entraînement et de test.

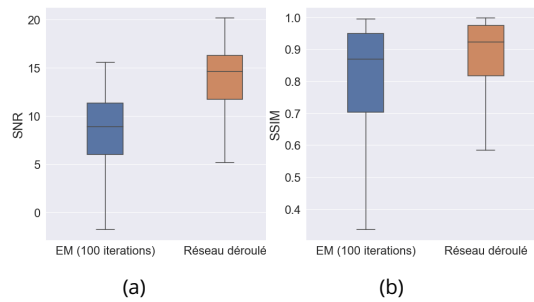


Figure 4.7 – Comparaison des performances de l’algorithme EM pour 100 itérations avec l’architecture neuronale proposée dépliée sur 10 itérations sur (a) le SNR et (b) la SSIM.

rations. Les performances sont évaluées sur un jeu de test composé de 1000 images. L’algorithme déplié apporte de meilleures performances pour l’ensemble des métriques à un coût computationnel moindre. Nous affichons sur la figure 4.8 l’image du ciel estimée par l’algorithme EM proposé pour 10 et 100 itérations, ainsi que les résultats obtenus à l’aide du réseau déplié.

En complément, nous avons simulé des observations en utilisant les positions des antennes du réseau radio-interférométrique Meerkat. Après un apprentissage sur les données simulées, nous montrons en figure 4.9 les images reconstruite par l’algorithme déplié sur 10 itérations pour des observations simulées en utilisant les positions des antennes du réseau radio-interférométrique Meerkat. Nous observons que l’algorithme déplié permet d’obtenir une image très fidèle à l’image originale, malgré la présence d’interférences et le faible nombre d’itérations. La *dirty image* affichée en 4.8.b a servi d’initialisation.

4.6 Conclusion

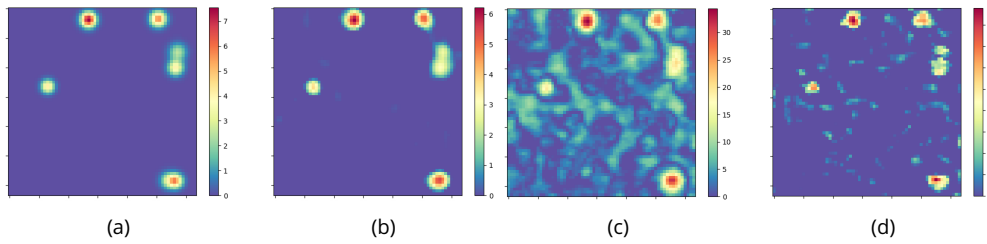


Figure 4.8 – (a) Image à reconstruire, (b) Image obtenue en utilisant l’architecture dépliée sur 10 itérations, (c) Image estimée par 10 itérations de l’algorithme EM considérant une t -distribution, (d) Image estimée par 100 itérations de l’algorithme EM considérant une t -distribution

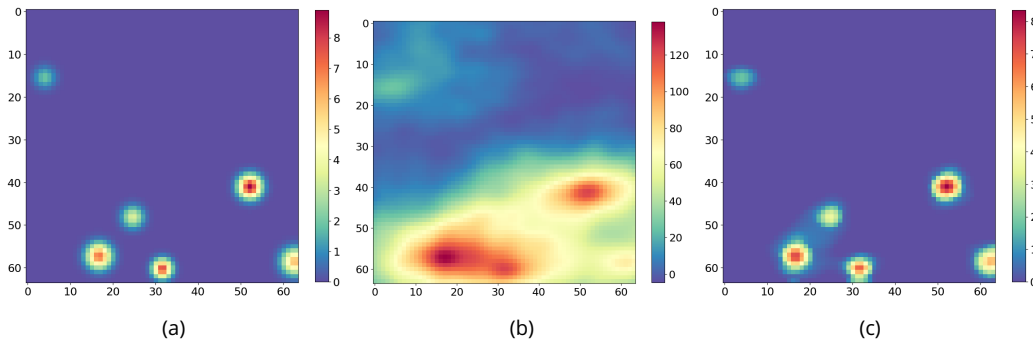


Figure 4.9 – Image reconstruite par l’algorithme déplié sur 10 itérations pour des observations simulées en utilisant les positions des antennes du réseau radio-interférométrique Meerkat. La figure (a) correspond à l’image originale, (b) à la *dirty image* et (c) à l’image reconstruite par l’algorithme déplié sur 10 itérations.

Dans ce chapitre, nous avons développé deux algorithmes d’espérance maximisation pour la reconstruction d’images radio-interférométriques en présence d’interférences de radiofréquence. Pour cela, le problème d’imagerie est formulé comme un problème de maximum de vraisemblance régularisé en considérant un bruit additif issu d’une distribution gaussienne composée. Cette modélisation a permis de prendre en compte la présence d’interférences dans les données, mais conduit à un problème de maximum de vraisemblance intractable dans le cas général.

Nous avons alors proposé un premier algorithme d’espérance maximisation qui permet une implémentation simple à partir de méthodes d’imagerie existantes, en ajoutant une robustesse à la présence d’interférence dans les données. Nous avons ensuite proposé une extension de cet algorithme EM qui tire partie de la factorisation de l’opérateur direct du modèle pour obtenir un algorithme à plus faible coût calculatoire. Des simulations numériques ont montré les performances supérieures des algorithmes proposés, apportant une meilleure précision par rapport aux méthodes de l’état de l’art en présence d’interférences.

Enfin, nous avons présenté une architecture neuronale dépliée, développée à partir de l’algorithme 4.1. En particulier, nous avons proposé d’apprendre l’étape-E de l’algorithme EM de manière supervisée. Des résultats préliminaires ont montré l’apport de

l'apprentissage informée pour l'imagerie radio-interférométrique en présence de RFI. En effet, nous illustrons les performances de l'architecture apprise sur des images de ciel simulées. Comparativement aux méthodes de maximum de vraisemblance, l'algorithme appris montre de meilleurs résultats sur plusieurs métriques.

5 - Traitement de données réelles

Ce chapitre présente les travaux menés en collaboration avec l'Observatoire de Radioastronomie d'Afrique du Sud. Ces travaux ont conduit à l'intégration de l'algorithme 4.1, proposé dans la section 4.2, dans la librairie Python dénommée *pfbclean*¹ au RATT dédiée à l'imagerie interférométrique sur des données de grande dimension.

Nous introduisons d'abord l'écosystème logiciel en radioastronomie pour le traitement de données réelles. Nous présentons ensuite la librairie *pfbclean*. Il s'agit d'une librairie Python permettant une implémentation efficace d'algorithmes pour l'imagerie radio-interférométrique. En particulier, *pfbclean* implémente un algorithme *forward-backward* pour l'imagerie radio-interférométrique multifréquentielle [29]. Ce dernier correspond à un maximum de vraisemblance régularisé par une norme $\ell_{2,1}$ permettant d'imposer une parcimonie jointe sur les pixels des images dans une base d'ondelettes, tout en garantissant une douceur sur la dimension spectrale. L'algorithme 4.1, proposé en section 4.2, est ainsi intégré en considérant à l'étape de maximisation l'algorithme *forward-backward* proposé dans [29]. Nous illustrons ensuite les performances de l'algorithme EM obtenues sur des données réelles, issues d'une observation du *Very Large Array telescope* (VLA) corrompues par des RFI. Nous montrons alors l'apport de la modélisation proposée et de l'algorithme obtenu sur la prise en compte des RFI à l'étape d'imagerie.

5.1 L'écosystème logiciel en radioastronomie

La chaîne de traitement de données en radio-interférométrie se décompose en trois étapes : i) *L'inspection et le marquage* des données est un premier prétraitement permettant d'identifier les données corrompues pour lesquelles un filtrage ou une ablation des corrélations est effectuée, ii) La *calibration* consiste ensuite à estimer et corriger les différents effets perturbatifs affectant les données à partir d'observations sur des sources de référence. iii) Enfin, *l'imagerie* est la reconstruction de l'image du ciel à partir de mesures calibrées et conduit à l'estimation d'un modèle de ciel pour la portion de la voûte céleste observée. Toutefois, ces différentes étapes peuvent être appliquées plusieurs fois le long de la chaîne de traitement. Les boucles de *self-calibration* nécessitent notamment d'effectuer de manière itérative une calibration suivie d'une imagerie, afin d'affiner la calibration pour la direction d'observation. Cela nécessite alors d'effectuer un grand nombre de lectures et d'écritures sur les mesures acquises, créant le besoin d'une structure de données efficace en grande dimension. En effet, dans le cas des grands radiotélescopes interféromètres de nouvelle génération, le taux de mémorisation des données est de l'ordre de plusieurs Tera octets par seconde [89], conduisant à des jeux de données de plus en plus volumineux et nécessitant des systèmes dédiés

1. <https://github.com/ratt-ru/pfb-clean>

Nom de la table	Description
DATA	Visibilités mesurées
UVW	Coordonnées des vecteurs de ligne de base
ANTENNA1	Indice de la première antenne des lignes de base
ANTENNA2	Indice de la seconde antenne des lignes de base
FLAG	Masque à appliquer aux visibilités
DATA_DESC_ID	Descripteur indiquant la bande spectrale pour chaque visibilité
TIME	Temps de mesure des corrélations
FIELD_ID	Descripteur indiquant la direction d'observation pointée
WEIGHT	Pondération à appliquer aux visibilités
INTERVAL	Temps d'intégration

Table 5.1 – Table principale (MAIN) d'un *measurement set*

pour leur traitement.

Ainsi, la construction de radio-interféromètres a conduit au développement d'outils logiciels spécifiques à la radioastronomie. La suite logicielle CASA [90] pour *Common Astronomy Software Application* est la principale suite logicielle pour le traitement de données provenant des radiotélescopes interféromètres ALMA (*Atacama Large Millimeter/submillimeter Array*) et VLA et contient de nombreux outils dédiés aux traitements de données radio-interférométriques. Le développement de CASA a conjointement abouti à un format de données unifié pour les observations de radioastronomie : le *measurement set* [91], [92]. Il s'agit d'un modèle de base de données permettant une manipulation efficace des données radio-interférométriques volumineuses, qui constitue aujourd'hui le principal format de données pour l'acquisition et le traitement de données radio-interférométriques. Un *measurement set* se matérialise par une collection de tables contenant les corrélations mesurées ainsi que des métadonnées utiles à leur traitement. Le tableau 5.1 décrit les principaux champs compris dans la table principale d'un jeu d'observations.

Il inclut notamment un champ *FLAG* qui permet de spécifier les données corrompues ne devant pas être utilisées. De plus, le format *measurement set* offre la possibilité de créer des sous-tables spécifiques permettant d'intégrer des métadonnées supplémentaires. La figure 5.1 est un schéma de la hiérarchie des *measurement set* utilisés pour les observations provenant du LoFAR.

Par ailleurs, la construction de nouveaux radiotélescopes interféromètres à très grande échelle conduit à l'émergence d'une nouvelle génération d'outils logiciels. Le développement du LoFAR a notamment permis l'élaboration d'outils spécifiques [54]. Ainsi, dans [9], un logiciel d'inspection et de détection des RFI dans les données est décrit. Une implémentation efficace d'un algorithme CLEAN pour l'imagerie radio-interférométrique communément utilisée est développée dans [26].

Aussi, le projet de construction du *Square Kilometer Array* (SKA) génère de nouvelles

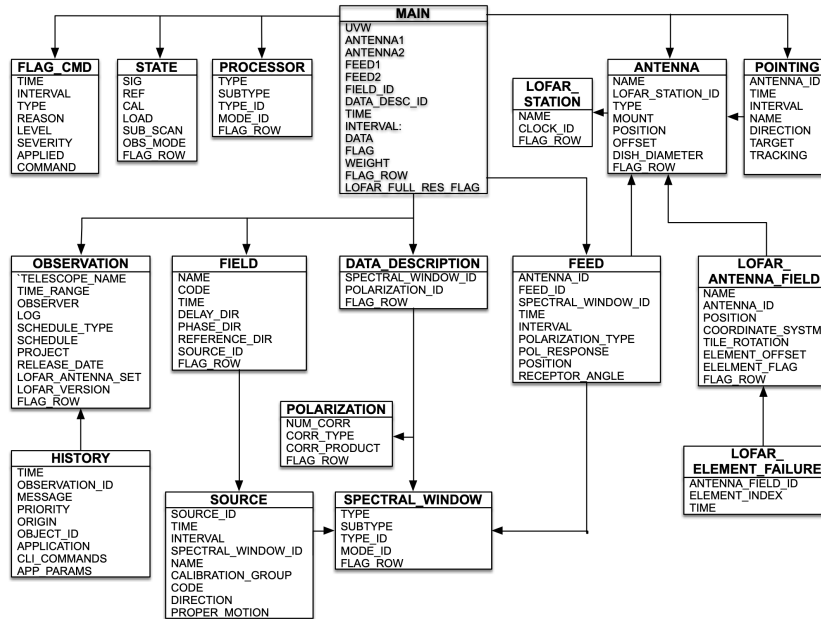


Figure 5.1 – Hiérarchie du format *measurement set* utilisé pour les données provenant du LoFAR.

contraintes pour le traitement des données. En effet, le projet vise à augmenter drastiquement le nombre d’antennes pour construire un réseau interférométrique de grande échelle. Plusieurs bibliothèques logicielles sont alors en cours de développement pour permettre une instrumentation et une reconstruction d’image efficace dans le contexte de données de grande dimension. Ainsi, le *Rhodes Centre for Radio Astronomy Techniques and Technologies (RATT)* a été établi en 2012 autour du projet SKA. Il a pour mission de conduire une recherche approfondie sur les méthodes algorithmiques et le développement d’outils logiciels nécessaires pour maximiser l’impact scientifique des radiotélescopes interféromètres de nouvelle génération. Le RATT propose alors plusieurs bibliothèques logicielles *open source* pour le traitement de données radio-interférométriques en grande dimension². En particulier, nous présentons dans la section 5.2, la bibliothèque *pfbclean* qui propose des implémentations parallèles d’algorithmes d’imagerie pouvant être distribuées sur un ordinateur haute performance. En particulier, nous présentons dans les sections suivantes l’intégration de l’algorithme 4.1 dans *pfbclean* pour lequel un algorithme *forward-backward* pour l’imagerie radio-interférométrique multifréquentiel [29] permet d’effectuer l’étape de maximisation.

5.2 Intégration dans la bibliothèque Python *pfbclean*

2. <https://github.com/ratt-ru>

Nom du module	Description
<code>init</code>	Initialisation des sous-produits et de l'architecture de données nécessaire à la imagerie
<code>grid</code>	Calcul de la <i>dirty image</i> et du <i>dirty beam</i>
<code>spotless</code>	Imagerie interférométrique par un algorithme <i>forward-backward</i> [29]

Table 5.2 – Exemples de modules implémentés dans *pfb-clean*

Cette section présente dans un premier temps la librairie Python *pfb-clean*³ qui implémente notamment l'algorithme d'imagerie multifréquentielle proposé dans [29]. Nous introduisons donc le modèle de données multifréquentielles et le problème d'optimisation associé, puis nous formulons l'algorithme *forward-backward* préconditionné proposé pour l'imagerie dans [29]. Celui-ci permet un calcul efficace du gradient de la fonction coût et une convergence accélérée, conduisant à une implémentation à plus faible coût computationnel. Nous présentons ensuite l'algorithme EM multifréquentiel intégrant l'algorithme présenté dans son étape de maximisation, permettant une robustesse à la présence de RFI.

5.2.1 La librairie *pfb-clean*

La librairie Python *pfb-clean* s'insère dans un nouveau paradigme lié au développement des grands radiotélescopes interféromètres que sont les réseaux SKA et MeerKat [33], [89]. Ces réseaux interférométriques voient leur nombre d'antennes croître significativement. La quantité de visibilités mesurées augmente alors drastiquement (de manière quadratique [93]), créant le besoin d'outils computationnels modernes permettant de traiter les données en un temps raisonnable. Le RATT propose alors une suite d'outils Python permettant le développement d'algorithmes de radioastronomie haute performance et pouvant être distribués efficacement [93]. En particulier, *pfb-clean* a la vocation d'agréger plusieurs outils algorithmiques spécifiques à l'imagerie radio-interférométrique. La librairie se divise ainsi en plusieurs modules indépendants. Le tableau 5.2 montre des exemples de modules implémentés dans *pfb-clean*. En lien avec le formalisme développé au RATT, les modules peuvent ensuite être combinés dans des chaînes de traitement globales pouvant être distribuées sur plusieurs ressources matérielles.

Nous présentons dans les sections suivantes l'algorithme implémenté pour l'imagerie multifréquentielle dans le module nommé `spotless`, puis l'intégration de l'algorithme EM proposé dans la section 4.2 à travers un module spécifique.

5.2.2 Modèle de données multifréquentielles

3. <https://github.com/ratt-ru/pfb-clean>

Le modèle de données introduit dans la section 2.3 est étendu à des données multifréquentielles. Le vecteur de mesure s'exprime alors pour le canal de fréquence f_j comme suit,

$$\mathbf{y}^{[j]} = \mathbf{H}^{[j]}\mathbf{x}^{[j]} + \mathbf{n}^{[j]}, \quad (5.1)$$

où $\mathbf{H}^{[j]}$ est l'opérateur de DFT défini dans la section 2.3.2 pour la longueur d'onde $\lambda_j = \frac{c}{f_j}$. Le modèle décrit en (5.1) peut-être formulé comme un modèle linéaire en considérant la concaténation de chaque composante spectrale,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (5.2)$$

avec $\mathbf{y} = [\mathbf{y}^{[1]}; \dots; \mathbf{y}^{[Q]}]$, $\mathbf{x} = [\mathbf{x}^{[1]}; \dots; \mathbf{x}^{[Q]}]$, $\mathbf{n} = [\mathbf{n}^{[1]}; \dots; \mathbf{n}^{[Q]}]$, pour Q canaux fréquentiels. L'opérateur direct, \mathbf{H} , s'exprime par une matrice diagonale par bloc, où chaque bloc correspond à une matrice de DFT, $\text{diag}(\mathbf{H}^{[1]}, \dots, \mathbf{H}^{[Q]})$. Le problème d'optimisation convexe permettant de reconstruire le cube d'images, \mathbf{x} , est formulé en considérant une régularisation imposant une contrainte de douceur sur la dimension spectrale [94]-[98],

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\text{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + r(\mathbf{x}), \quad (5.3)$$

où $r(\cdot)$ est une fonction de régularisation convexe. Ce problème d'optimisation correspond à un maximum de vraisemblance régularisée en présence de bruit gaussien. En particulier, la norme $\ell_{2,1}$ permet d'imposer une parcimonie jointe sur les pixels des images tout en garantissant une douceur sur la dimension spectrale [94], [95]. La régularisation choisie est alors la somme entre la norme $\ell_{2,1}$ sur le cube d'images et la fonction indicatrice de l'espace des réels positifs. La norme $\ell_{2,1}$ est définie comme la norme ℓ_1 du vecteur contenant les normes de Frobenius sur la dimension spectrale pour chaque pixel,

$$\|\mathbf{X}\|_{\ell_{2,1}} = \sum_{n=1}^N \|\mathbf{x}_n\|_2, \quad (5.4)$$

où \mathbf{x}_n est la $n^{\text{ème}}$ ligne de \mathbf{X} et $\mathbf{X} = [\mathbf{x}^{[1]}, \dots, \mathbf{x}^{[Q]}] \in \mathbb{R}^{N \times Q}$ est la matrice contenant sur chaque colonne l'image pour un canal fréquentiel. Le problème d'optimisation associé s'écrit alors,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\text{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \iota_{\mathbb{R}^+}(\mathbf{X}) + \alpha \|\Psi \mathbf{X}\|_{\ell_{2,1}}, \quad (5.5)$$

où $\alpha > 0$ est un hyperparamètre. La matrice Ψ est la concaténation de plusieurs bases d'ondelettes permettant d'imposer une contrainte de parcimonie sur chaque image du cube en adéquation avec la nature des sources célestes observées. En pratique, le dictionnaire, Ψ , contient la base de Dirac et les six premières ondelettes de Daubechies tel que proposé dans [31].

5.2.3 Algorithme *forward-backward* préconditionné pour l'imagerie

Le problème d'optimisation défini en (5.3) est résolu par un algorithme de type *forward-backward* préconditionné [99],

$$\mathbf{x}^{(k+\frac{1}{2})} = \mathbf{x}^{(k)} - \gamma \mathbf{U}^{-1} \nabla_{\mathbf{x}} f(\mathbf{x}^{(k)}), \quad (5.6)$$

$$\mathbf{x}^{(k+1)} = \text{prox}_{\gamma r}^{\mathbf{U}}(\mathbf{x}^{(k+\frac{1}{2})}) = \underset{\mathbf{x}}{\text{argmin}} r(\mathbf{x}) + \frac{1}{2\gamma} (\mathbf{x} - \mathbf{x}^{(k+\frac{1}{2})})^H \mathbf{U} (\mathbf{x} - \mathbf{x}^{(k+\frac{1}{2})}), \quad (5.7)$$

où $f(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2$ et $r(\mathbf{x}) = \alpha \|\Psi \mathbf{X}\|_{\ell_{2,1}} + \iota_{\mathbb{R}^+}(\mathbf{X})$ et \mathbf{U} est une matrice semi-définie positive. L'utilisation du préconditionneur, \mathbf{U} , permet de réduire le nombre d'itérations nécessaires pour arriver à la convergence et donc de réduire le coût calculatoire de l'algorithme [99], [100]. Toutefois, dans ce cas l'opérateur proximal en (5.7) ne peut pas s'exprimer de manière explicite et doit être calculé par un algorithme spécifique [101]. Le choix du préconditionneur est soumis à un certain nombre de contraintes [99], [100], [102]. La hessienne du terme d'attache aux données, $f(\cdot)$, est choisie ici car f est deux fois différentiable [99], ce qui conduit à l'expression suivante pour \mathbf{U} ,

$$\mathbf{U} = \mathbf{H}^H \mathbf{H}. \quad (5.8)$$

Toutefois, l'utilisation directe de la hessienne n'est pas pratique car elle n'est pas nécessairement inversible et nécessite un trop gros nombre de ressources calculatoires à chaque itération. Il est alors proposé d'ajouter un faible multiple, $\eta > 0$, de l'identité conduisant à une approximation de la hessienne (*diagonal loading*),

$$\mathbf{U} = \mathbf{H}^H \mathbf{H} + \eta \mathbf{I}. \quad (5.9)$$

Par ailleurs, le calcul du gradient et de la hessienne est effectué par le théorème de convolution de Fourier, en tirant partie du fait que l'opérateur direct peut s'exprimer par une DFT sur chaque image du cube,

$$\nabla_{\mathbf{x}} f \approx \mathcal{B} * \mathbf{x} - \mathbf{x}_{\text{dirty}}, \quad (5.10)$$

$$\mathbf{U}\mathbf{x} = \mathcal{B} * \mathbf{x} + \eta \mathbf{I}, \quad (5.11)$$

où \mathcal{B} est le *dirty beam* tel que défini en (2.14) et $\mathbf{x}_{\text{dirty}} = \mathbf{H}^H \mathbf{y}$. L'équation (5.10) correspond effectivement au gradient du terme d'attache aux données lorsque les conditions du théorème de Van-Cittert Zernike sont respectées (voir section 2.1). Dans le cas de large champs de vue, le théorème de Van-Cittert Zernike n'est plus valide et (5.10) est une approximation du gradient.

5.2.4 Intégration de l'algorithme EM

L'algorithme EM présenté dans la section 4.2 pour l'imagerie radio-interférométrique en présence de RFI est intégré dans *pfbclean*. Pour cela, la modélisation proposée dans la section 4.1 est étendue à des observations multifréquentielles. Ainsi, nous considérons le modèle suivant,

$$\mathbf{y}^{[j]} = \mathbf{H}^{[j]} \mathbf{x}^{[j]} + \Omega^{[j]1/2} \mathbf{t}^{[j]}, \quad (5.12)$$

Algorithm 5.1 Imagerie multifréquentielle par maximum de vraisemblance régularisée en présence de bruit gaussien [29]

input : $\mathbf{x}^{(0)}, \mathbf{x}_{\text{dirty}}, \mathcal{B}, \gamma$

output : $\hat{\mathbf{x}}$

- 1: $\hat{\mathbf{x}} \leftarrow \mathbf{x}^{(0)}$
 - 2: **while** critère d'arrêt **do**
 - 3: $\nabla_{\mathbf{x}} f \leftarrow \mathcal{B} * \hat{\mathbf{x}} - \mathbf{x}_{\text{dirty}}$
 - 4: $\hat{\mathbf{x}} \leftarrow \underset{\mathbf{x}}{\operatorname{argmin}} r(\mathbf{x}) + \frac{1}{2\gamma} (\mathbf{x} - \hat{\mathbf{x}})^H \mathbf{U} (\mathbf{x} - \hat{\mathbf{x}})$
 - 5: **end while**
-

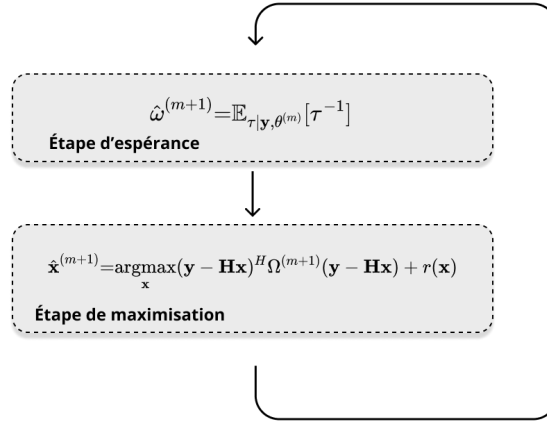


Figure 5.2 – Schéma de l’algorithme EM intégré dans *pfb-clean*. L’étape de maximisation est mise en oeuvre par l’algorithme 5.1.

où $\Omega^{[j]} = \operatorname{diag}(\tau_1^{[j]-1}, \dots, \tau_K^{[j]-1})$, $\mathbf{t}^{[j]} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ et $\forall i \in \llbracket 1, K \rrbracket, j \in \llbracket 1, Q \rrbracket, \tau_i^{[j]} \sim p(\tau)$. Additionnellement au modèle proposé en (4.2), (5.12) permet de prendre en compte la variabilité de la puissance des RFI sur la dimension spectrale. L’algorithme EM permettant d’estimer l’image du ciel, \mathbf{x} et la variance du bruit σ^2 , s’exprime alors comme suit,

$$\hat{\omega}_k^{[j](m+1)} = \mathbb{E}_{\tau|\mathbf{y}_k^{[j]}, \hat{\mathbf{x}}^{(m)}} [\tau^{-1}], \quad (5.13)$$

$$\mathbf{x}^{(m+1)} = \underset{\mathbf{x}}{\operatorname{argmin}} \left((\mathbf{y} - \mathbf{H}\mathbf{x})^H \hat{\Omega}^{(m+1)} (\mathbf{y} - \mathbf{H}\mathbf{x}) + r(\mathbf{x}) \right), \quad (5.14)$$

$$\sigma^{2(m+1)} = \frac{1}{KQ} \|\mathbf{y} - \mathbf{H}\mathbf{x}^{(m+1)}\|_{\hat{\Omega}^{(m+1)}}^2, \quad (5.15)$$

avec $\hat{\Omega}^{(m+1)} = \operatorname{diag}(\hat{\omega}_1^{[1](m+1)}, \dots, \hat{\omega}_K^{[1](m+1)}, \dots, \hat{\omega}_1^{[Q](m+1)}, \dots, \hat{\omega}_K^{[Q](m+1)})$. Les étapes de l’algorithme sont illustrées sous forme de schéma dans la figure 5.2.

L'algorithme présenté dans la section 4.2 considère un bruit additif gaussien composé, sans préciser de distribution explicite pour le paramètre de texture. Nous choisissons une distribution Inverse Gamma pour la texture, ce qui conduit à modéliser le bruit additif par une t -distribution. En effet, il a été montré que cette modélisation du bruit additif est pertinente dans le contexte de la radioastronomie pour la prise en compte de valeurs aberrantes [13], [21]. Dans ce cas, l'étape d'espérance se calcule de manière explicite comme montré dans la section 4.4.2,

$$\hat{\omega}_k^{[j](m)} = \frac{\nu + 1}{\nu + \frac{1}{\sigma^{(m)2}} |\mathbf{y}[k][j] - \mathbf{h}_k^{[j]} \mathbf{x}^{[j]}|^2}, \quad (5.16)$$

où $\nu > 2$ est le degré de liberté de la t -distribution. En pratique nous avons fixé $\nu = 5$ de manière empirique. Conformément à la méthodologie explicitée dans la remarque 2, l'algorithme EM est intégré en considérant le module `spotless` pour l'étape de maximisation. En effet, ce dernier permet de résoudre le moindre carré pondéré exprimé en (5.14). L'étape d'espérance est alors intégrée dans le module `grid`, puis l'étape de maximisation est effectuée par le module `spotless`. En effet, le module `grid` est utilisé pour calculer la *dirty image*, $\mathbf{x}_{\text{dirty}}$, et le *dirty beam*, \mathcal{B} , qui permettent d'approximer l'opérateur direct, \mathbf{H} , dans le module `spotless`. Les pondérations calculées en (5.16), sont alors mises à jour et appliquées à l'opérateur direct dans le module `grid`. L'algorithme EM est alors mis en oeuvre en combinant les deux modules de manière itérative.

Algorithm 5.2 EM multifréquentiel pour l'imagerie en présence de RFI

input : $\mathbf{x}^{(0)}$, $\mathbf{x}_{\text{dirty}}$, \mathcal{B} , γ , ν

output : $\hat{\mathbf{x}}$

- 1: **while** critère d'arrêt **do**
 - 2: **Étape E :**
 - 3: $\hat{\omega}_k^{[j](m)}$ calculé en (5.16) $\forall k \in \llbracket 1, K \rrbracket, j \in \llbracket 1, Q \rrbracket$.
 - 4: **Étape M :**
 - 5: $\hat{\mathbf{x}}^{(m)}$ obtenu par l'algorithme 5.1.
 - 6: **end while**
-

5.3 Imagerie sur données réelles en présence de RFI

5.3.1 Observation par le VLA du rémanent de supernova G55.7+3.4

Nous utilisons le *measurement set* provenant des observations décrites dans [103]. Il s'agit de l'observation d'une source radio, nommée *SNR G55.7+3.4*, découverte en 1977 et classifiée comme étant un rémanent de supernova galactique [103], [104]. Les observations ont été menées en utilisant 20 antennes du VLA dans la configuration D sur la bande L (1-2 GHz). La direction d'observation pour la source observée est 19 :21 :40 RA 21.45.00 DEC J2000⁴. Les mesures ont initialement été effectuées sur 8 bandes spectrales dont 4 ont été retirées à cause de la présence de RFI de forte puissance. Chaque

Bande spectrale	Fréquence centrale	Largeur de bande
0	1.256 GHz	2 MHz
1	1.384 GHz	2 MHz
2	1.648 GHz	2 MHz
3	1.776 GHz	2 MHz

Table 5.3 – Bandes spectrales considérées pour l’observation de la source radio *SNR G55.7+3.4* [103]

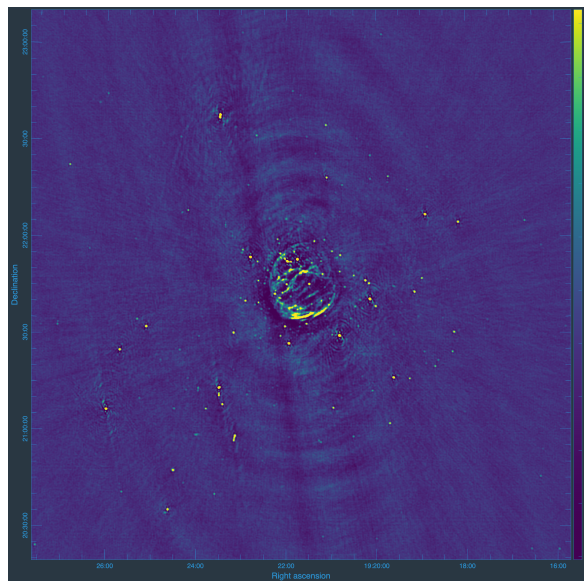


Figure 5.3 – *Dirty image* de la source radio *SNR G55.7+3.4* obtenue à partir des données du VLA [103].

bande spectrale est divisée en 64 canaux fréquentiels de largeur de bande de 2 MHz. L’observation a été menée sur 7 heures et les corrélations ont été obtenues avec un temps d’intégration d’une seconde. Une haute résolution temporelle et fréquentielle a été choisie afin de pouvoir détecter la présence de RFI dans les observations. Les visibilités corrompues par des RFI ont alors été marquées. Le masque résultant est sauvegardé dans le champ *FLAG* du *measurement set*. Le jeu de données final a ensuite été réduit en effectuant un moyennage résultant en une résolution temporelle de 10 secondes et un nombre total de visibilités disponible de 1 280 640. Les caractéristiques des bandes spectrales considérées sont résumées dans le tableau 5.3. La figure 5.3 montre la *dirty image* obtenue pour cette observation.

5.3.2 Impact des RFI sur l’imagerie

4. Les coordonnées sont données dans le repère équatorial (voir Annexe C) à la date de références J2000. L’ascension droite (RA) est mesurée en heure :minute :seconde et la déclinaison (DEC) en degré.arcminute.arcseconde

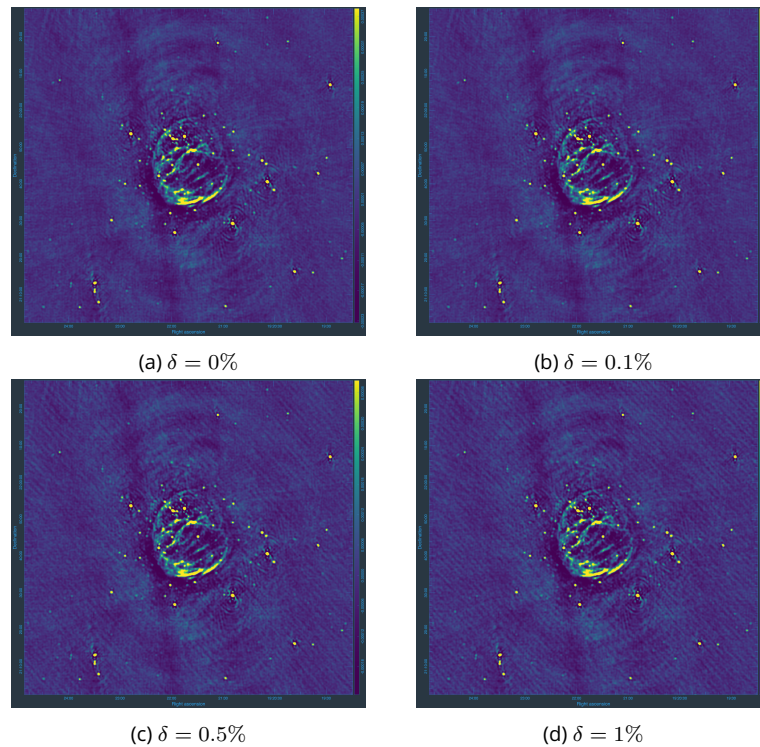


Figure 5.4 - Impact de la présence de RFI sur la *dirty image*. Sur chaque image une proportion, $\delta \in \{1\%, 0.5\%, 0.1\%\}$, de visibilités masquées contenant des RFI est ajoutée.

Les RFI se matérialisent comme des sources additionnelles et impactent les observations à travers les lobes secondaires des antennes du réseau interférométrique [1], [8], [21]. Par ailleurs, la durée de ce type de signaux interférents est en général courte du fait de leur nature, liée notamment à leur utilisation dans les technologies de télécommunication [7]. En conséquence, la présence de RFI varie selon la position des antennes et le temps d'observation. La présence de RFI de forte puissance apparaît nettement dans la *dirty image*. En effet, les visibilités sont des coefficients de la DFT de l'image du ciel. Ainsi, une visibilité corrompue par des RFI de forte puissance perturbe, en l'augmentant, la composante liée à la fréquence spatiale correspondante. Cela conduit alors à l'apparition de bandes périodiques dans la *dirty image*. Afin d'illustrer l'impact des RFI sur les observations de radioastronomie, nous ajoutons aux visibilités utiles une proportion, $\delta \in [0, 100\%]$, des visibilités marquées comme corrompues, du fait de la présence de RFI et spécifié par le champ *FLAG* du *measurement set*. Ainsi, nous construisons trois *measurement set* pour lesquels respectivement $\delta = 0.1\%$, $\delta = 0.5\%$ et $\delta = 1\%$.

La figure 5.4 montre les *dirty images* obtenues pour les différentes valeurs de δ . Des artefacts sur la *dirty image* dus aux RFI ajoutées apparaissent dès $\delta = 0.1\%$. Des bandes périodiques, causées par la présence de RFI de forte puissance, apparaissent dès $\delta = 0.5\%$ et avec une plus grande intensité pour $\delta = 1\%$.

La figure 5.5 montre l'image obtenue en appliquant, au *measurement set* original, l'algorithme *forward-backward* présenté dans la section précédente (Algorithme 5.1) à un maximum de vraisemblance régularisée en présence de bruit gaussien. Nous comparons ensuite les images obtenues par l'Algorithme 5.1 présenté dans la section précédente et l'algorithme EM proposé (Algorithme 5.2) sur des *measurement set* pour lesquels des visibilités corrompues ont été ajoutées. L'algorithme EM implémenté, dont l'étape de maximisation est mise en œuvre par l'algorithme 5.1, est alors initialisé à partir des images obtenues par ce dernier. Nous fixons comme critère d'arrêt un seuil, $\epsilon = 0.0001$, sur la variation relative de l'image reconstruite,

$$\frac{\|\mathbf{x}^{(m+1)} - \mathbf{x}^{(m)}\|}{\|\mathbf{x}^{(m)}\|} \leq \epsilon. \quad (5.17)$$

En pratique nous observons que 10 itérations de l'EM suffisent. La figure 5.6 montre les images obtenues par les deux algorithmes pour $\delta = 0.1\%$, $\delta = 0.5\%$ et $\delta = 1\%$. Nous observons une nette amélioration des images par l'application de l'algorithme EM proposé.

Pour $\delta = 0.5\%$ et $\delta = 1\%$, des lignes caractéristiques de la présence de RFI apparaissent clairement sur les images obtenues par l'algorithme 5.1. En effet, la direction pointée par les antennes du réseau interférométrique varie au cours de l'observation afin de suivre la source observée sur la sphère céleste en compensant la rotation de la Terre. Ainsi, une source interférente présente dans l'environnement proche du radiotélescope apparaît comme une source mobile sur l'image reconstruite. Ces lignes, liées à la présence de RFI, sont éliminées sur les images obtenues par l'algorithme EM proposé et la source imagée apparaît plus clairement, se rapprochant des résultats obtenus sur le *measurement set* original. La figure 5.7 souligne cette observation pour $\delta = 0.5\%$.

L'effet des RFI sur les images obtenues par l'algorithme 5.1 pour $\delta = 0.1\%$ est léger. Néanmoins, nous montrons sur la figure 5.8 que l'image reconstruite par l'algorithme EM permet la reconstruction de sources, ayant été atténuées par l'algorithme 5.1 en présence des RFI.

5.4 Conclusion

Ce chapitre présente les résultats obtenus dans le cadre d'une collaboration avec l'Observatoire de Radioastronomie d'Afrique du Sud. Nous avons présenté un algorithme maximum de vraisemblance régularisée en présence de bruit gaussien pour l'imagerie radio-interférométrique sur des observations multifréquentielles. La régularisation prend en compte une douceur spectrale et une parcimonie dans la base d'ondelette. L'optimisation est assurée par un algorithme *forward-backward* préconditionné. Nous introduisons la librairie Python *pfb-clean* dans laquelle l'algorithme d'imagerie présenté est implémenté à faible coût computationnel dans un contexte de calculs distribués en grande dimension.

Nous expliquons ensuite l'intégration de l'algorithme EM présenté dans la section 4.2 dans la librairie *pfb-clean*. Pour cela, l'algorithme *forward-backward* est intégré

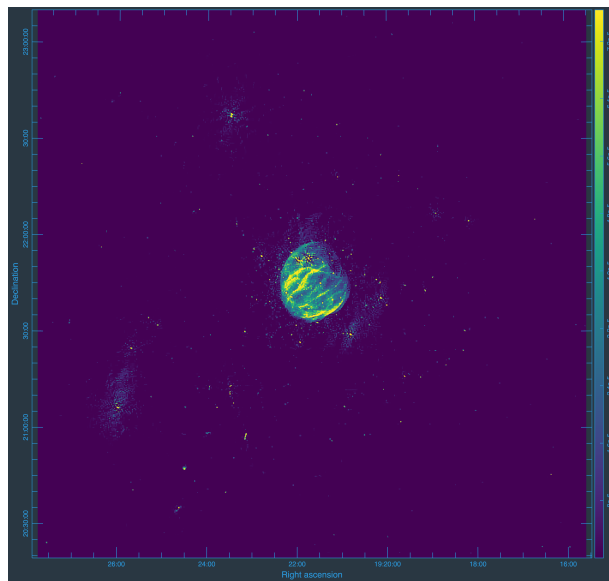


Figure 5.5 – Image obtenue en appliquant l’algorithme 5.1 au *measurement set* original ($\delta = 0\%$).

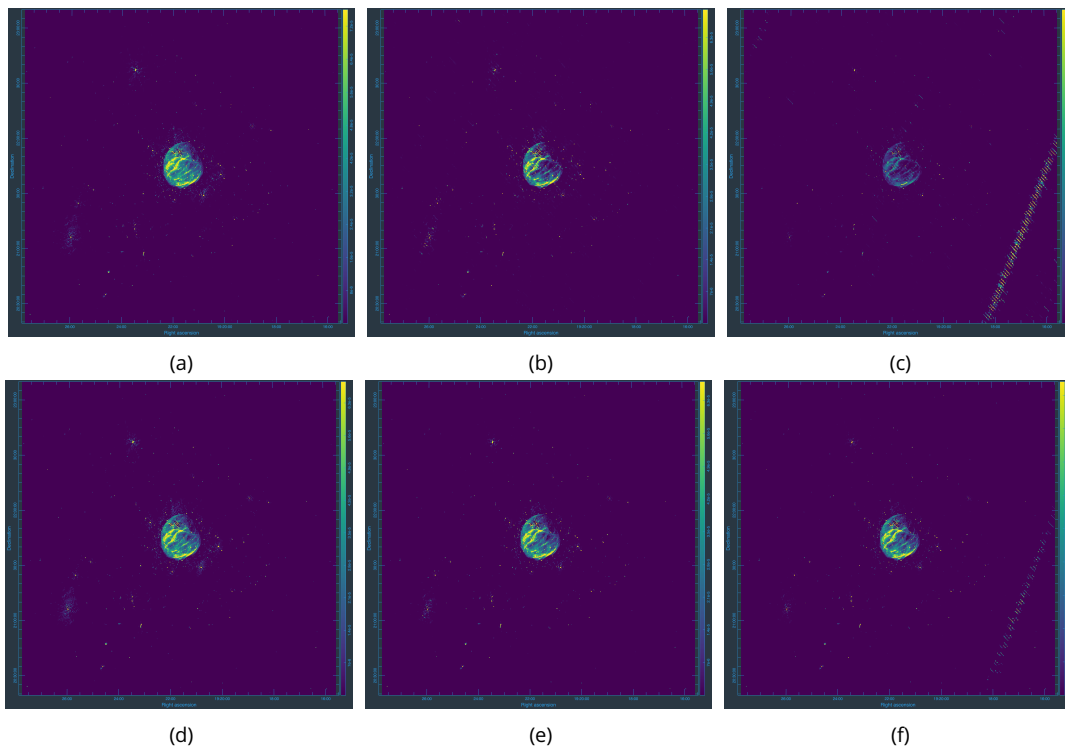


Figure 5.6 – Images obtenues par l’algorithme 5.1 ((a) $\delta = 0.1\%$, (b) $\delta = 0.5\%$, (c) $\delta = 1\%$) et par l’algorithme EM proposé ((d) $\delta = 0.1\%$, (e) $\delta = 0.5\%$, (f) $\delta = 1\%$)

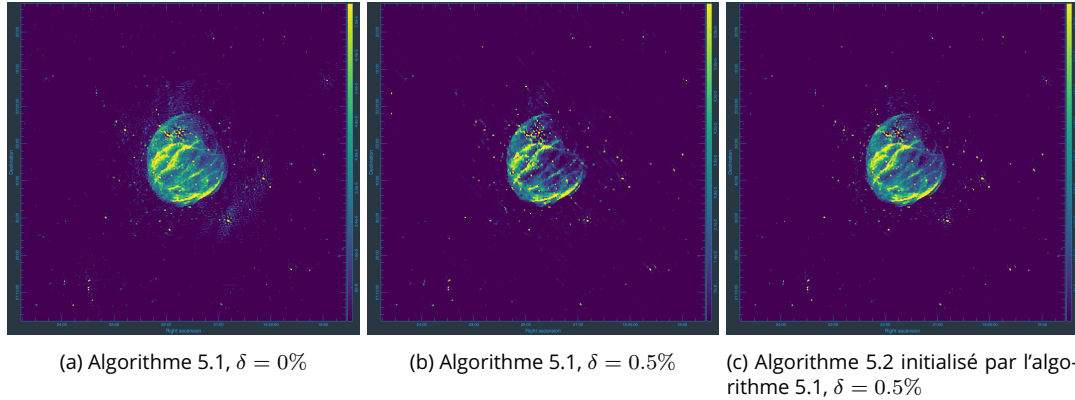


Figure 5.7 – Images obtenues par l'algorithme 5.1 pour le *measurement set* original ($\delta = 0\%$) (a) et pour $\delta = 0.5\%$ (b). L'image obtenue par l'algorithme EM proposé pour $\delta = 0.5\%$ est montrée en (c). Les lignes caractéristiques de la présence de RFI apparaissent en (b). L'image obtenue par l'algorithme EM proposé permet de les éliminer en (c) et obtenir une image plus proche de celle obtenue sur le *measurement set* original (a).

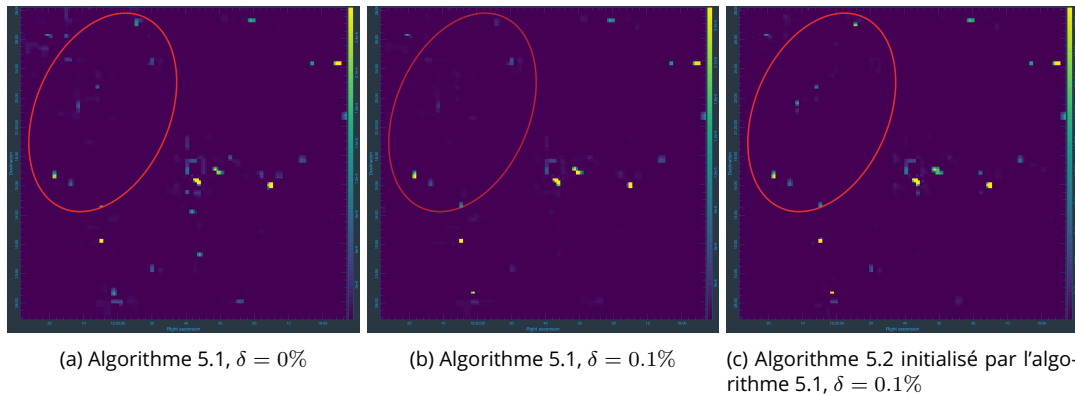


Figure 5.8 – Images obtenues par l'algorithme 5.1 pour le *measurement set* original ($\delta = 0\%$) (a) et pour $\delta = 0.1\%$ (b). L'image obtenue par l'algorithme EM proposé pour $\delta = 0.1\%$ est montrée en (c). Les sources faibles (dans la zone entourée en rouge) présentes dans l'image obtenue sur le *measurement set* original (a), sont atténuées par l'algorithme 5.1 en présence de RFI (b). L'algorithme EM proposé permet de les reconstruire en (c).

comme étape de maximisation de l'algorithme EM et permet d'ajouter une robustesse à la présence de RFI.

Enfin, nous exposons les expériences menées sur un jeu de données réelles provenant du VLA pour l'étude de l'impact des RFI sur les images reconstruites. Les observations étant perturbées par des RFI, nous avons pu montrer l'apport de l'algorithme EM proposé sur la prise en compte des RFI dans l'étape d'imagerie.

6 - Conclusion et perspectives

6.1 Conclusion

Les travaux présentés dans cette thèse traitent de méthodes de calibration et d'imagerie dans le contexte de la radio-interférométrie en présence de RFI. Ainsi, nous avons proposé plusieurs outils méthodologiques permettant d'atténuer l'impact des RFI sur les différentes étapes de la chaîne de traitement des données radio-interférométriques. Pour cela, nous avons formulé des modélisations statistiques permettant de décrire la présence de RFI dans les données radio-interférométriques. Nous avons ensuite proposé des algorithmes de maximum de vraisemblance pour l'estimation des paramètres. En particulier, nous avons dérivé plusieurs algorithmes de types EM en tirant parti de la structure des modèles proposés pour réduire le coût computationnel des algorithmes proposés.

Nous avons d'abord proposé une modélisation multifréquentielle des données radio-interférométriques dans le contexte de la calibration dans la section 3.2. Plus spécifiquement, nous avons modélisé la composante liée aux RFI par un bruit gaussien de rang faible, en adéquation avec le modèle de propagation qui leur correspond. Nous avons ensuite conçu un algorithme SAGE pour l'estimation des paramètres de calibration par maximum de vraisemblance qui est détaillé dans la section 3.3. Pour cela, nous avons originalement dérivé deux jeux de données complètes permettant d'obtenir des mises à jours explicites des paramètres aux étapes de maximisation.

Afin de prendre en compte la présence de RFI lors de l'étape d'imagerie, nous avons proposé de modéliser le bruit additif par des distributions gaussiennes composées. Cela nous a permis de développer, dans la section 4.2, un algorithme EM dont l'étape de maximisation correspond à un problème de moindres carrées pondérées régularisé. Ainsi, cela nous permet d'améliorer la robustesse des algorithmes d'imagerie à la présence de RFI.

Nous avons ensuite étendu l'algorithme EM proposé en tirant partie d'une factorisation de l'opérateur direct intégrant un opérateur de FFT. Cela nous a permis de considérer une décomposition du bruit additif conduisant à l'introduction de visibilité latentes pour lesquelles l'imagerie peut être réalisée à faible coût calculatoire. Nous développons alors un jeu de données complètes, spécifié dans la section 4.3, dont l'étape de maximisation de l'algorithme EM correspondant peut être résolue à moindre coût computationnel. En effet, dans ce cas, la dimension du vecteur de données latentes ne dépend pas de la dimension du vecteur de mesures et l'opérateur direct est implémentable efficacement par FFT.

En outre, nous avons conçu une architecture neuronale dépliée développée à partir de l'algorithme 4.1. Nous avons présenté, dans la section 4.5, des résultats préliminaires permettant de souligner l'apport d'un apprentissage informé pour l'imagerie radio-interférométrique en présence de RFI.

Enfin, en collaboration avec l'Observatoire de Radioastronomie d'Afrique du Sud, nous avons intégré l'algorithme EM, proposé dans la section 4.2, dans la librairie Python *pfbclean*. Ces travaux ont abouti à un algorithme d'imagerie multifréquentielle robuste à la présence de RFI décrit dans la section 5.2. Des résultats sur données réelles permettent alors de démontrer l'intérêt de la modélisation proposée pour la prise en compte des RFI à l'étape d'imagerie.

Ainsi, dans cette thèse, nous avons apporté de nouvelles modélisations statistiques pour décrire la présence de RFI dans les données interférométriques. Cela nous a permis de proposer des algorithmes de maximum de vraisemblance originaux et robustes tant pour l'imagerie que pour la calibration. Les différents algorithmes EM développés, grâce aux choix pertinents de données latentes, ou encore aux techniques d'optimisation utilisées, ont donné lieu à des implémentations efficaces qui sont pertinentes dans le contexte de la radioastronomie. En effet, cette potentialité a été à l'origine d'une collaboration avec l'équipe de radioastronomie du RATT. Les résultats obtenus dans le cadre de cette collaboration sur signaux réels sont prometteurs pour l'imagerie en présence d'interférences.

6.2 Perspectives

Ces travaux ouvrent la voie à plusieurs perspectives de recherche et d'implémentations sur des plateformes de radioastronomie. En effet, les nouveaux grands radiotélescopes interféromètres, tels que le *Square Kilometer Array* (SKA), ont vocation à produire des quantités de données considérables. Ces derniers seront d'une sensibilité accrue, les rendant ainsi plus vulnérables aux RFI. Il est alors nécessaire de développer des méthodes robustes et à faible coût pour la calibration et l'imagerie radio-interférométrique. Pour cela, les travaux présentés dans cette thèse peuvent être étendus afin de permettre d'améliorer la robustesse et l'efficacité des algorithmes de traitement du signal, dans le contexte des nouveaux grands radiotélescopes.

Ainsi, la méthodologie introduite dans cette thèse peut être intégrée sur des pipelines de traitement de données. Le développement de méthodes de calibration et d'imagerie jointe peut notamment être envisagé. En effet, la calibration et l'imagerie sont généralement effectués par des algorithmes indépendants. Pourtant, ces deux étapes sont intimement liées. En particulier, les erreurs de calibration peuvent conduire à des artefacts dans les images reconstruites, conduisant à une charge de calcul supplémentaire pour les atténuer [12], [21]. Ainsi, la calibration et l'imagerie peuvent être formulées comme un problème d'optimisation conjoint [105]. Les modélisations proposées dans cette thèse peuvent alors être intégrées dans un problème d'optimisation conjoint pour la calibration et l'imagerie. En particulier, si l'on dispose d'un a priori sur l'image, le problème de calibration et d'imagerie jointes pourra être formulé comme un problème de maximum a posteriori (MAP) dont la résolution pourra être réalisée par des algorithmes de type EM.

L'application de l'apprentissage profond pour l'imagerie en radioastronomie est un

domaine de recherche en pleine expansion [81], [82], [106], [107]. En effet, la quantité de données disponibles permet d'envisager des apprentissages à grande échelle. Des méthodes dites de *plug-and-play* sont proposées pour l'imagerie radio-interférométrique [82], [106]-[108]. Ces méthodes utilisent un réseau de neurones débruiteurs entraîné. Ce dernier est ensuite intégré dans un algorithme de type primal-dual permettant de résoudre le problème d'optimisation décrit dans la section 2.3.2, associé à l'imagerie. Ces méthodologies permettent de définir a priori sur la distribution des images du ciel de manière implicite par un entraînement supervisé, et tirent notamment parti des évolutions récentes des architectures neuronales débruiteuses [109]. Ainsi, dans ce cas, il n'est plus nécessaire de définir une fonction de régularisation spécifique. Néanmoins, ces méthodes se basent sur un modèle considérant un bruit additif gaussien.

Pour dépasser le cadre gaussien, des méthodes de *plug-and-play* intégrés à des algorithmes de maximum de vraisemblance, associée à un modèle statistique des données radio-interférométriques, peuvent alors être envisagées. Cette méthodologie pourrait être dans un premier temps appliquée à l'étape d'imagerie avant de considérer une calibration et une imagerie conjointes par un MAP comme évoquées précédemment. Dans ce cas, la distribution a priori sur les images du ciel pourra être définie par un réseau de neurones débruiteurs entraîné sur des images du ciel simulées à partir de modèles de ciel plus complexes.

Parmi les travaux en cours, nous pouvons citer :

- Un article de revue décrivant les méthodologies et l'implémentation de la librairie python *pfbclean* est en cours de rédaction et intègre la prise en compte des RFI par l'algorithme décrit dans la section 5.2.
- En collaboration avec l'Observatoire de Radioastronomie d'Afrique du Sud, des travaux complémentaires sont en cours pour étudier la pertinence du modèle proposé dans la section 4.1 pour atténuer l'impact d'erreurs de calibration dans les images reconstruites. En effet, dans le cadre de la calibration d'effets dépendants de la direction, les algorithmes de calibration classiques conduisent à des artefacts dans les images reconstruites dû à l'absence de sources de références [12], [21]. Les boucles de self-calibration, venant palier cette absence d'information, alternent alors entre l'estimation des paramètres de calibration et l'imagerie de la scène observée afin de converger vers une image fidèle à la scène observée. Actuellement, cette stratégie itérative demande un grand nombre d'itérations et se trouve de ce fait coûteuse en temps de calcul. La modélisation proposée peut alors permettre d'accélérer les boucles de self-calibration, conduisant à une convergence plus rapide vers une image fidèle à la scène observée. En effet, les erreurs introduites par la calibration peuvent être vues comme des distorsions des visibilités. Un article de revue est en cours de rédaction sur ces travaux [110].
- Les résultats préliminaires présentés dans la section 4.5 montrent l'apport d'un apprentissage informé pour l'imagerie radio-interférométrique en présence de

RFI. Ainsi, nous envisageons de comparer les performances de l'algorithme déroulé proposé dans la section 4.5 avec des méthodes d'apprentissage profond pour l'imagerie radio-interférométrique [81], [82], [106], [107]. De plus, une étude sur l'extension de cette méthode à des données simulées à partir de modèles de ciel plus complexes est en cours. Un article de revue est en cours de rédaction pour présenter l'algorithme et les résultats obtenus [111].

Annexe A - Calibration multifréquentielle en présence d'interférences

A.1 Calcul de l'étape d'espérance pour le premier jeu de données complètes

Nous cherchons à calculer,

$$Q_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^{(m)}) = \mathbb{E}_{\mathcal{X}_1 | \mathbf{y}; \boldsymbol{\theta}^{(m)}} [\log(p(\mathbf{y}, \boldsymbol{o}; \boldsymbol{\theta}_1))], \quad (\text{A.1})$$

où $\boldsymbol{\theta}_1 = [\text{vec}(\mathbf{W}), \sigma^{[1]}, \dots, \sigma^{[Q]}]$. Les vecteurs \mathbf{y} et \boldsymbol{o} peuvent être décomposés en des vecteurs fréquentiels mutuellement indépendants de la sorte,

$$\mathbf{y} = [\mathbf{y}^{[1]T}, \dots, \mathbf{y}^{[Q]T}], \quad (\text{A.2})$$

$$\boldsymbol{o} = [\boldsymbol{o}^{[1]T}, \dots, \boldsymbol{o}^{[Q]T}]. \quad (\text{A.3})$$

Le vecteur fréquentiel d'indice j , s'exprime alors ainsi,

$$\mathbf{y}^{[j]} = \mathbf{v}^{[j]} + \sigma^{[j]} \mathbf{W} \boldsymbol{o}^{[j]} + \mathbf{n}^{[j]}. \quad (\text{A.4})$$

La fonction Q_1 peut alors être réécrite comme suit,

$$Q_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^{(m)}) = \sum_{j=1}^Q \mathbb{E}_{\mathcal{X}_1 | \mathbf{y}; \boldsymbol{\theta}_1^{(m)}} [\log(p(\mathbf{y}^{[j]}, \boldsymbol{o}^{[j]}; \boldsymbol{\theta}_1))]. \quad (\text{A.5})$$

Le modèle défini en (3.35) permet de déduire les distributions conditionnelles suivantes,

$$\mathbf{y}^{[j]} | \boldsymbol{o}^{[j]} \sim \mathcal{CN}(\mathbf{v}^{[j]}(\mathbf{Z}) + \sigma^{[j]} \mathbf{W} \boldsymbol{o}^{[j]}, \sigma^2 \mathbf{I}_{4K}) \quad (\text{A.6})$$

$$\boldsymbol{o}^{[j]} \sim \mathcal{CN}(\text{vec}(\mathbf{I}), \mathbf{I}). \quad (\text{A.7})$$

Les vraisemblances logarithmiques associées à $\mathbf{y}^{[j]} | \boldsymbol{o}^{[j]}$ et $\mathbf{y}^{[j]}$ s'écrivent respectivement,

$$\begin{aligned} \log(p(\mathbf{y}^{[j]} | \boldsymbol{o}^{[j]}; \boldsymbol{\theta}_1)) &= -4QK \log(\pi) - \log |\sigma^2 \mathbf{I}| \\ &\quad - \frac{1}{\sigma^2} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}) - \sigma^{[j]} \mathbf{W} \boldsymbol{o}^{[j]})^H (\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}) - \sigma^{[j]} \mathbf{W} \boldsymbol{o}^{[j]}), \end{aligned} \quad (\text{A.8})$$

$$\log(p(\boldsymbol{o}^{[j]}; \boldsymbol{\theta}_1)) = -\rho \log(\pi) - \log |\mathbf{I}_\rho| - \sum_{j=1}^Q (\boldsymbol{o}^{[j]} - \text{vec}(\mathbf{I}))^H (\boldsymbol{o}^{[j]} - \text{vec}(\mathbf{I})).$$

Cela nous permet d'exprimer la fonctionnelle $Q_1(\boldsymbol{\theta}_1 \mid \boldsymbol{\theta}_1^{(m)})$ ainsi,

$$\begin{aligned}
Q_1(\boldsymbol{\theta}_1 \mid \boldsymbol{\theta}_1^{(m)}) &= \sum_{j=1}^Q \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} \left[\log(p(\mathbf{y}^{[j]} \mid \mathbf{o}^{[j]}; \boldsymbol{\theta}_1)) + \log(p(\mathbf{o}^{[j]}; \boldsymbol{\theta}_1)) \right] \\
&\propto \sum_{j=1}^Q -\log |\sigma^2 \mathbf{I}_{4K}| \\
&\quad - \frac{1}{\sigma^2} \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} \left[(\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}) - \sigma^{[j]} \mathbf{W} \mathbf{o}^{[j]})^H (\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}) - \sigma^{[j]} \mathbf{W} \mathbf{o}^{[j]}) \right] \\
&\quad - \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} \left[(\mathbf{o}^{[j]} - \text{vec}(\mathbf{I}))^H (\mathbf{o}^{[j]} - \text{vec}(\mathbf{I})) \right] \\
&\propto -Q \log |\sigma^2 \mathbf{I}_{4K}| + \sum_{j=1}^Q \frac{2\sigma^{[j]}}{\sigma^2} \Re((\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}))^H \mathbf{W} \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} [\mathbf{o}^{[j]}]) \\
&\quad - \frac{1}{\sigma^2} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}))^H (\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z})) - \frac{\sigma^{[j]^2}}{\sigma^2} \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} [\mathbf{o}^{[j]H} \mathbf{W}^H \mathbf{W} \mathbf{o}^{[j]}] \\
&\quad - \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} [\mathbf{o}^{[j]H} \mathbf{o}^{[j]}] + 2 \text{vec}(\mathbf{I})^H \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} [\mathbf{o}^{[j]}] - \text{vec}(\mathbf{I})^H \text{vec}(\mathbf{I})
\end{aligned} \tag{A.9}$$

où \propto indique l'égalité à une constante additive près. Nous intéressons ensuite à la moyenne et à la covariance du vecteur $\mathbf{o}^{[j]}$,

$$\hat{\mathbf{o}}^{[j]} = \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} [\mathbf{o}^{[j]}], \tag{A.10}$$

$$\hat{\Sigma}_{\mathbf{o}}^{[j]} = \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} [(\mathbf{o}^{[j]} - \hat{\mathbf{o}}^{[j]})(\mathbf{o}^{[j]} - \hat{\mathbf{o}}^{[j]})^H]. \tag{A.11}$$

Les vecteur $\mathbf{o}^{[j]}$ et $\mathbf{y}^{[j]}$ étant conjointement gaussien de moyenne respective,

$$\boldsymbol{\mu}_{\mathbf{o}^{[j]}} = \text{vec}(\mathbf{I}), \quad \boldsymbol{\mu}_{\mathbf{y}^{[j]}} = \mathbf{v}^{[j]}(\mathbf{Z}) + \sigma^{[j]} \mathbf{W} \text{vec}(\mathbf{I}), \tag{A.12}$$

les matrices de covariances $\Sigma_{\mathbf{o}\mathbf{o}}$, $\Sigma_{\mathbf{y}\mathbf{y}}$ and $\Sigma_{\mathbf{y}\mathbf{o}}$ s'écrivent,

$$\begin{aligned}
\Sigma_{\mathbf{o}\mathbf{o}} &= \mathbb{E} [(\mathbf{o}^{[j]} - \boldsymbol{\mu}_{\mathbf{o}^{[j]}})(\mathbf{o}^{[j]} - \boldsymbol{\mu}_{\mathbf{o}^{[j]}})^H] = \mathbf{I} \\
\Sigma_{\mathbf{y}\mathbf{y}} &= \mathbb{E}_{\mathbf{y}} [(\mathbf{y}^{[j]} - \boldsymbol{\mu}_{\mathbf{y}^{[j]}})(\mathbf{y}^{[j]} - \boldsymbol{\mu}_{\mathbf{y}^{[j]}})^H] = \sigma^{(m)^2} \mathbf{I}_{4K} + \sigma^{[j]2^{(m)}} \mathbf{W}^{(m)} \mathbf{W}^{(m)H} \\
\Sigma_{\mathbf{o}\mathbf{y}} &= \Sigma_{\mathbf{y}\mathbf{o}}^H = \mathbb{E} [(\mathbf{o}^{[j]} - \boldsymbol{\mu}_{\mathbf{o}^{[j]}})(\mathbf{y}^{[j]} - \boldsymbol{\mu}_{\mathbf{y}^{[j]}})^H] \\
&= \mathbb{E} [(\mathbf{o}^{[j]} - \text{vec}(\mathbf{I}))(\mathbf{y}^{[j]} - \mathbf{v}^{[j]} - \sigma^{[j]} \mathbf{W} \text{vec}(\mathbf{I}))^H] \\
&= \mathbb{E} [(\mathbf{o}^{[j]} - \text{vec}(\mathbf{I}))(\sigma^{[j]^{(m)}} \mathbf{W}^{(m)} \mathbf{o}^{[j]} + \mathbf{n}^{[j]})^H] \\
&= \mathbb{E} [\mathbf{o}^{[j]} (\sigma^{[j]^{(m)}} \mathbf{W}^{(m)} \mathbf{o}^{[j]})^H] + \mathbb{E} [\mathbf{o}^{[j]} \mathbf{n}^{[j]H}] - \text{vec}(\mathbf{I}) (\sigma^{[j]^{(m)}} \mathbf{W}^{(m)} \mathbb{E} [\mathbf{o}^{[j]})^H \\
&\quad - \text{vec}(\mathbf{I}) \mathbb{E} [\mathbf{n}^{[j]H}].
\end{aligned} \tag{A.13}$$

Sachant que $\mathbf{o}^{[j]}$ et $\mathbf{n}^{[j]}$ sont indépendants et,

$$\mathbb{E} \left[\mathbf{n}^{[j]} \right] = \mathbf{0}, \quad (\text{A.14})$$

$$\mathbb{E} \left[\mathbf{y}^{[j]} \mathbf{n}^{[j]H} \right] = \mathbf{0}. \quad (\text{A.15})$$

On obtient alors,

$$\begin{aligned} \Sigma_{\mathbf{o}\mathbf{y}} &= \Sigma_{\mathbf{y}\mathbf{o}}^H = \mathbb{E} \left[\mathbf{o}^{[j]} (\sigma^{[j](m)} \mathbf{W}^{(m)} \mathbf{o}^{[j]})^H \right] - \text{vec}(\mathbf{I}) (\sigma^{[j](m)} \mathbf{W}^{(m)} \mathbb{E} \left[\mathbf{o}^{[j]} \right])^H \\ &= \sigma^{[j](m)} \mathbf{W}^{(m)H} \end{aligned} \quad (\text{A.16})$$

en utilisant [112, p36] les expressions analytiques de $\hat{\Sigma}_o^{[j]}$ and $\hat{\delta}^{[j]}$ sont obtenues,

$$\begin{aligned} \hat{\delta}^{[j]} &= \boldsymbol{\mu}_{\mathbf{o}^{[j]}} + \Sigma_{\mathbf{o}\mathbf{y}} \Sigma_{\mathbf{y}\mathbf{y}}^{-1} (\mathbf{y}^{[j]} - \boldsymbol{\mu}_{\mathbf{y}^{[j]}}) \\ &= \text{vec}(\mathbf{I}) + \sigma^{[j](m)} \mathbf{W}^{(m)H} (\sigma^{(m)2} \mathbf{I}_{4K} + \sigma^{[j]2(m)} \mathbf{W}^{(m)} \mathbf{W}^{(m)H})^{-1} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}^{(m)}) - \sigma^{[j](m)} \mathbf{W}^{(m)} \text{vec}(\mathbf{I})) \\ \hat{\Sigma}_o^{[j]} &= \Sigma_{\mathbf{o}\mathbf{o}} - \Sigma_{\mathbf{o}\mathbf{y}} \Sigma_{\mathbf{y}\mathbf{y}}^{-1} \Sigma_{\mathbf{y}\mathbf{o}} \\ &= \mathbf{I} - \sigma^{[j](m)} \mathbf{W}^{(m)H} (\sigma^{(m)2} \mathbf{I}_{4K} + \sigma^{[j]2(m)} \mathbf{W}^{(m)} \mathbf{W}^{(m)H})^{-1} \sigma^{[j]} \mathbf{W}^{(m)}. \end{aligned} \quad (\text{A.17})$$

Les équations (A.10) et (A.11) sont alors réécrites comme suit,

$$\begin{aligned} \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} \left[\mathbf{o}^{[j]H} \mathbf{o}^{[j]} \right] &= \text{tr}(\hat{\Sigma}_o^{[j]}) + \hat{\delta}^{[j]H} \hat{\delta}^{[j]} \\ \mathbb{E}_{\mathbf{o}|\mathbf{y};\boldsymbol{\theta}_1^{(m)}} \left[\mathbf{o}^{[j]H} (\sigma^{[j]2} \mathbf{W}^H \mathbf{W}) \mathbf{o} \right] &= \text{tr}((\sigma^{[j]2} \mathbf{W}^H \mathbf{W}) \hat{\Sigma}_o^{[j]}) + \hat{\delta}_f^H (\sigma^{[j]2} \mathbf{W}^H \mathbf{W}) \hat{\delta}^{[j]} \end{aligned} \quad (\text{A.18})$$

Enfin, Q_1 s'écrit,

$$\begin{aligned} Q_1(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_1^{(m)}) &\propto -Q \log |\sigma^2 \mathbf{I}_{4K}| + \sum_{j=1}^Q \frac{2}{\sigma^2} \Re((\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H \sigma^{[j]} \mathbf{W} \hat{\delta}^{[j]}) \\ &\quad - \frac{1}{\sigma^2} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H (\mathbf{y}^{[j]} - \mathbf{v}^{[j]}) - \frac{\sigma^{[j]2}}{\sigma^2} (\text{tr}(\mathbf{W}^H \mathbf{W} \hat{\Sigma}_o^{[j]}) + \hat{\delta}^{[j]H} (\sigma^{[j]2} \mathbf{W}^H \mathbf{W}) \hat{\delta}^{[j]}) \\ &\quad - \text{tr}(\hat{\Sigma}_o^{[j]}) - \hat{\delta}^{[j]H} \hat{\delta}^{[j]} + 2 \text{vec}(\mathbf{I})^H \hat{\delta}^{[j]} - \text{vec}(\mathbf{I})^H \text{vec}(\mathbf{I}) \\ &\propto -F \log |\sigma^2 \mathbf{I}_{4K}| - \sum_j \left(\frac{1}{\sigma^2} \|\mathbf{y}^{[j]} - \mathbf{v}^{[j]} - \sigma^{[j]} \mathbf{W} \hat{\delta}^{[j]}\|_2^2 \right. \\ &\quad \left. + \|\hat{\delta}^{[j]} - \text{vec}(\mathbf{I})\|_2^2 + \frac{\sigma^{[j]2}}{\sigma^2} \text{tr}(\mathbf{W}^H \mathbf{W} \hat{\Sigma}_o^{[j]}) + \text{tr}(\hat{\Sigma}_o^{[j]}) \right) \end{aligned} \quad (\text{A.19})$$

A.2 Étape de maximisation pour le premier jeu de données complètes

Nous calculons les dérivées partielles de $Q_1(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)})$ par rapport à $\sigma^{[j]}$, pour $j \in [1; Q]$

$$\frac{\partial Q_1(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)})}{\partial \sigma^{[j]}} = -\frac{2\sigma^{[j]}}{\sigma^2} \text{tr}(\mathbf{W}^H \mathbf{W}(\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]H} \hat{\boldsymbol{\delta}}^{[j]})) + \frac{2}{\sigma^2} \Re\left((\mathbf{y}^{[j]} - \mathbf{v}^{[j]}(\mathbf{Z}))^H \mathbf{W} \hat{\boldsymbol{\delta}}^{[j]}\right) \quad (\text{A.20})$$

Nous résolvons $\frac{\partial Q_1(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)})}{\partial \sigma^{[j]}} = 0$. Cela permet d'obtenir,

$$\hat{\sigma}^{[j]} = \frac{\Re\left((\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H \mathbf{W} \hat{\boldsymbol{\delta}}^{[j]}\right)}{\text{tr}(\mathbf{W}^H \mathbf{W}(\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]} \hat{\boldsymbol{\delta}}^{[j]H}))} \quad (\text{A.21})$$

$Q_1(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)})$ s'écrit en fonction de \mathbf{W} ainsi,

$$\begin{aligned} Q_1(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)}) &\approx \frac{1}{\sigma^2} \sum_{j=1}^Q \left(2\Re\left((\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H \sigma^{[j]} \mathbf{W} \hat{\boldsymbol{\delta}}^{[j]}\right) - \text{tr}(\sigma^{[j]2} \mathbf{W}^H \mathbf{W} \hat{\boldsymbol{\Sigma}}_o^{[j]}) - \hat{\boldsymbol{\delta}}^{[j]H} \sigma^{[j]2} \mathbf{W}^H \mathbf{W} \hat{\boldsymbol{\delta}}^{[j]} \right) \\ &\approx + \frac{1}{\sigma^2} \sum_{j=1}^Q \left(\text{tr}(\mathbf{W} \sigma^{[j]} \hat{\boldsymbol{\delta}}^{[j]} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H) + \text{tr}(\mathbf{W}^H \sigma^{[j]} \hat{\boldsymbol{\delta}}^{[j]H} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})) \right. \\ &\quad \left. - \text{tr}(\sigma^{[j]2} \mathbf{W}^H \mathbf{W}(\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]} \hat{\boldsymbol{\delta}}^{[j]H})) \right) \end{aligned} \quad (\text{A.22})$$

Le gradient de Q_1 par rapport à \mathbf{W} est calculé par dérivation de Wirtinger [74]

$$\begin{aligned} \frac{\partial}{\partial \mathbf{W}} \left(\text{tr}(\mathbf{W} \sigma^{[j]} \hat{\boldsymbol{\delta}}^{[j]} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H) \right) &= (\sigma^{[j]} \hat{\boldsymbol{\delta}}^{[j]} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H)^T \\ \frac{\partial}{\partial \mathbf{W}} \left(\text{tr}(\sigma^{[j]} \hat{\boldsymbol{\delta}}^{[j]} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H \mathbf{W}^H) \right) &= \mathbf{0} \\ \frac{\partial}{\partial \mathbf{W}} \left(\text{tr}(\sigma^{[j]2} \mathbf{W}^H \mathbf{W}(\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]} \hat{\boldsymbol{\delta}}^{[j]H})) \right) &= \sigma^{[j]2} \mathbf{W}^* (\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]} \hat{\boldsymbol{\delta}}^{[j]H})^T. \end{aligned} \quad (\text{A.23})$$

En définitive, le gradient de Q_1 par rapport à \mathbf{W} s'écrit,

$$\begin{aligned} \frac{\partial}{\partial \mathbf{W}} (Q_1(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)})) &= \sum_{j=1}^Q \left((\sigma^{[j]} \hat{\boldsymbol{\delta}}^{[j]} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H)^T - \sigma^{[j]2} \mathbf{W}^* (\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]} \hat{\boldsymbol{\delta}}^{[j]H})^T \right) \\ &= \sum_{j=1}^Q \sigma^{[j]} (\hat{\boldsymbol{\delta}}^{[j]} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H)^T - \mathbf{W}^* \sum_{j=1}^Q \sigma^{[j]2} (\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]} \hat{\boldsymbol{\delta}}^{[j]H})^T \end{aligned} \quad (\text{A.24})$$

Ainsi, résoudre $\frac{\partial Q_1(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)})}{\partial \mathbf{W}} = 0$ conduit à,

$$\begin{aligned} \mathbf{W}^* &= \left(\sum_{j=1}^Q \sigma^{[j]} \left(\hat{\boldsymbol{\delta}}^{[j]} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H \right)^T \right) \left(\sum_{j=1}^Q \sigma^{[j]^2} \left(\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]} \hat{\boldsymbol{\delta}}^{[j]H} \right)^T \right)^{-1} \\ \mathbf{W} &= \left(\sum_{j=1}^Q \sigma^{[j]} \left(\hat{\boldsymbol{\delta}}^{[j]} (\mathbf{y}^{[j]} - \mathbf{v}^{[j]})^H \right)^H \right) \left(\sum_{j=1}^Q \sigma^{[j]^2} \left(\hat{\boldsymbol{\Sigma}}_o^{[j]} + \hat{\boldsymbol{\delta}}^{[j]} \hat{\boldsymbol{\delta}}^{[j]H} \right)^H \right)^{-1}. \end{aligned} \quad (\text{A.25})$$

A.3 Calcul de l'étape d'espérance pour le second jeu de données complètes

A partir de l'équation (3.40), nous pouvons écrire,

$$\mathbf{u}_i \sim \mathcal{CN}(\mathbf{v}_i(\mathbf{Z}), \beta_i \sigma^2 \mathbf{I}_{4QK}) \text{ avec } \sum_{i=1}^D \beta_i = 1$$

La log-vraisemblance du second jeu de données complètes s'écrit alors,

$$\log(p(\mathbf{u}_i; \boldsymbol{\theta})) = C_{\mathbf{u}_i} - \log |\beta_i \sigma^2 \mathbf{I}_{4QK}| - \frac{1}{\beta_i \sigma^2} (\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z}))^H (\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z})). \quad (\text{A.26})$$

Ainsi, Q_2 peut être exprimé comme suit,

$$\begin{aligned} Q_2(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)}) &= \sum_{i=1}^D \mathbb{E}_{\mathbf{u}_i \mid \mathbf{y}; \boldsymbol{\theta}^{(m)}} [\log(p(\mathbf{u}_i; \boldsymbol{\theta}))] \\ &= \sum_{i=1}^D \mathbb{E}_{\mathbf{u}_i \mid \mathbf{y}} \left[C_{\mathbf{u}_i} - \log |\beta_i \sigma^2 \mathbf{I}_{4QK}| - \frac{1}{\beta_i \sigma^2} (\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z}))^H (\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z})) \right] \\ &= \sum_{i=1}^D C_{\mathbf{u}_i} - \log |\beta_i \sigma^2 \mathbf{I}_{4QK}| - \frac{1}{\beta_i \sigma^2} \mathbb{E}_{\mathbf{u}_i \mid \mathbf{y}} [(\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z}))^H (\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z}))] \\ &= \sum_{i=1}^D C_{\mathbf{u}_i} - \log |\beta_i \sigma^2 \mathbf{I}_{4QK}| - \frac{1}{\beta_i \sigma^2} \mathbb{E}_{\mathbf{u}_i \mid \mathbf{y}} [\mathbf{u}_i^H \mathbf{u}_i] + \frac{2}{\beta_i \sigma^2} \Re \left(\mathbf{v}_i(\mathbf{Z})^H \mathbb{E}_{\mathbf{u}_i \mid \mathbf{y}} [\mathbf{u}_i] \right) \\ &\quad - \frac{1}{\beta_i \sigma^2} \mathbf{v}_i^H \mathbf{v}_i. \end{aligned} \quad (\text{A.27})$$

Notons $\hat{\mathbf{u}}_i = \mathbb{E}_{\mathbf{u}_i \mid \mathbf{y}; \boldsymbol{\theta}^{(m)}} [\mathbf{u}_i]$ et $\hat{\boldsymbol{\Sigma}}_{\mathbf{u}_i} = \mathbb{E}_{\mathbf{u}_i \mid \mathbf{y}; \boldsymbol{\theta}^{(m)}} [(\mathbf{u}_i - \hat{\mathbf{u}}_i)(\mathbf{u}_i - \hat{\mathbf{u}}_i)^H]$ les moyennes et covariances du jeu de données complètes sachant les données observées et l'estimé courant des paramètres de calibration. Nous définissons le vecteur gaussien multivarié $\mathbf{z}_i = [\mathbf{u}_i, \mathbf{y}]$. Les moyennes de \mathbf{u}_i et \mathbf{y} sont respectivement

$$\boldsymbol{\mu}_{\mathbf{u}_i} = \mathbf{v}_i(\mathbf{Z}), \boldsymbol{\mu}_{\mathbf{y}} = \mathbf{v}(\mathbf{Z}) + \boldsymbol{\Psi}(\text{vec}(\mathbf{I}_{4K}) \otimes \mathbf{1}_Q).$$

$$\begin{aligned}
\Sigma_{\mathbf{u}_i \mathbf{u}_i} &= \mathbb{E}_{\mathbf{u}_i} [(\mathbf{u}_i - \boldsymbol{\mu}_{\mathbf{u}_i})(\mathbf{u}_i - \boldsymbol{\mu}_{\mathbf{u}_i})^H] = \mathbb{E}_{\mathbf{u}_i} [(\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z}))(\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z}))^H] = \beta_i \sigma^2 \mathbf{I}_{4QK} \\
\Sigma_{\mathbf{y} \mathbf{y}} &= \mathbb{E}_{\mathbf{y}} [(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^H] = (\sigma^2 \mathbf{I} + \boldsymbol{\Psi} \boldsymbol{\Psi}^H) \\
\Sigma_{\mathbf{y} \mathbf{u}_i} &= \Sigma_{\mathbf{u}_i \mathbf{y}}^H = \mathbb{E} [(\mathbf{u}_i - \boldsymbol{\mu}_{\mathbf{u}_i})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^H] \\
&= \mathbb{E} \left[(\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z})) \left(\sum_{k=1}^D (\mathbf{u}_k - \mathbf{v}_k(\mathbf{Z})) + \boldsymbol{\Psi} \mathbf{o} \right)^H \right] \\
&= \sum_{k=1}^D \mathbb{E} [(\mathbf{u}_i - \mathbf{v}_i(\mathbf{Z}))(\mathbf{u}_k - \mathbf{v}_k)^H]
\end{aligned} \tag{A.28}$$

Sachant que les différents vecteur \mathbf{u}_i sont indépendant entre eux,

$$\begin{aligned}
\Sigma_{\mathbf{y} \mathbf{u}_i} &= \Sigma_{\mathbf{u}_i \mathbf{y}}^H = \Sigma_{\mathbf{u}_i \mathbf{u}_i} + \sum_{k=1, k \neq i}^D \mathbb{E} [(\mathbf{u}_i - \mathbf{v}_i)] \mathbb{E} [(\mathbf{u}_k - \mathbf{v}_k)] \\
&= \Sigma_{\mathbf{u}_i \mathbf{u}_i}
\end{aligned} \tag{A.29}$$

En utilisant [112, p36], Nous calculons les expressions analytiques de $\hat{\Sigma}_{\mathbf{u}_i}$ et $\hat{\mathbf{u}}_i$,

$$\begin{aligned}
\hat{\mathbf{u}}_i &= \boldsymbol{\mu}_{\mathbf{u}_i} + \Sigma_{\mathbf{u}_i \mathbf{y}} \Sigma_{\mathbf{y} \mathbf{y}}^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}}) = \mathbf{v}_i(\mathbf{Z}) + \beta_i \sigma^2 (\sigma^2 \mathbf{I} + \boldsymbol{\Psi} \boldsymbol{\Psi}^H)^{-1} (\mathbf{y} - \mathbf{v}(\mathbf{Z}) - \boldsymbol{\Psi} (\text{vec}(\mathbf{I}_{4K}) \otimes \mathbf{1}_Q)) \\
\hat{\Sigma}_{\mathbf{u}_i} &= \Sigma_{\mathbf{u}_i \mathbf{u}_i} + \Sigma_{\mathbf{u}_i \mathbf{y}} \Sigma_{\mathbf{y} \mathbf{y}}^{-1} \Sigma_{\mathbf{y} \mathbf{u}_i} = \beta_i \sigma^2 \mathbf{I}_{4QK} - \beta_i^2 \sigma^4 (\sigma^2 \mathbf{I} + \boldsymbol{\Psi} \boldsymbol{\Psi}^H)^{-1}
\end{aligned} \tag{A.30}$$

Ainsi, nous pouvons écrire,

$$\begin{aligned}
\mathbb{E}_{\mathbf{u}_i | \mathbf{y}} [\mathbf{u}_i] &= \hat{\mathbf{u}}_i \\
\mathbb{E}_{\mathbf{u}_i | \mathbf{y}} [\mathbf{u}_i^H \mathbf{u}_i] &= \text{tr}(\hat{\Sigma}_{\mathbf{u}_i}) + \hat{\mathbf{u}}_i^H \hat{\mathbf{u}}_i.
\end{aligned} \tag{A.31}$$

Enfin,

$$Q_2(\boldsymbol{\theta} | \boldsymbol{\theta}^{(m)}) = \sum_{i=1}^D C_{u_i} - \log |\beta_i \sigma^2 \mathbf{I}_{4QK}| - \frac{1}{\beta_i \sigma^2} \|\hat{\mathbf{u}}_i - \mathbf{v}_i(\mathbf{Z})\|_F^2 - \frac{1}{\beta_i \sigma^2} \text{tr}(\hat{\Sigma}_{\mathbf{u}_i}). \tag{A.32}$$

A.4 Étape de maximisation pour le second jeu de données complètes

Nous calculons d'abord le gradient de Q_2 par rapport à σ^2 :

$$\begin{aligned}
\frac{\partial Q_2}{\partial \sigma^2} &= \sum_{i=1}^D -\frac{4QK}{\sigma^2} + \frac{1}{\beta_i \sigma^4} \left(\|\hat{\mathbf{u}}_i - \mathbf{v}_i(\mathbf{Z})\|_F^2 + \text{tr}(\hat{\Sigma}_{\mathbf{u}_i}) \right) \\
&= \frac{1}{\sigma^4} \sum_{i=1}^D -4QK \sigma^2 + \frac{1}{\beta_i} \left(\|\hat{\mathbf{u}}_i - \mathbf{v}_i(\mathbf{Z})\|_F^2 + \text{tr}(\hat{\Sigma}_{\mathbf{u}_i}) \right) \\
&= \frac{1}{\sigma^4} \left(-4QK D \sigma^2 + \sum_{i=1}^D \frac{1}{\beta_i} \left(\|\hat{\mathbf{u}}_i - \mathbf{v}_i(\mathbf{Z})\|_F^2 + \text{tr}(\hat{\Sigma}_{\mathbf{u}_i}) \right) \right)
\end{aligned} \tag{A.33}$$

La mise à jour de σ^2 est alors obtenue en annulant le gradient calculé,

$$\sigma^2 = \frac{1}{4QKD} \sum_{i=1}^D \frac{1}{\beta_i} \left(\|\hat{\mathbf{u}}_i - \mathbf{v}_i(\mathbf{Z})\|_F^2 + \text{tr}(\hat{\Sigma}_{\mathbf{u}_i}) \right) \quad (\text{A.34})$$

Afin d'obtenir les mise à jour des paramètres de calibration, \mathbf{Z} , Q_2 est exprimé pour souligner la dépendance en \mathbf{Z} ,

$$Q_2(\mathbf{Z}) \propto - \sum_{i=1}^D \frac{1}{\beta_i \sigma^2} \|\hat{\mathbf{u}}_i - \mathbf{v}_i(\mathbf{Z})\|_F^2, \quad (\text{A.35})$$

Les vecteurs $\hat{\mathbf{u}}_i$ et $\mathbf{v}_i(\mathbf{Z}_i)$ peuvent être divisés en Q sous-vecteurs,

$$\hat{\mathbf{u}}_i = [\hat{\mathbf{u}}_i^{[j]T}, \dots, \hat{\mathbf{u}}_i^{[j]T}]^T, \quad (\text{A.36})$$

$$\mathbf{v}_i(\mathbf{Z}) = [\mathbf{v}_i^{[j]}(\mathbf{Z})^T, \dots, \mathbf{v}_i^{[j]}(\mathbf{Z})^T]^T. \quad (\text{A.37})$$

Ainsi Q_2 peut être exprimé comme suit,

$$Q_2(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(m)}) \propto - \sum_{i=1}^D \sum_{j=1}^Q \frac{1}{\beta_i \sigma^2} \|\hat{\mathbf{u}}_i^{[j]} - \mathbf{v}_i^{[j]}(\mathbf{Z})\|_F^2. \quad (\text{A.38})$$

Nous considérons,

$$\phi(\mathbf{Z}_i) = \sum_{j=1}^Q \frac{1}{\beta_i \sigma^2} \|\hat{\mathbf{u}}_i^{[j]} - \mathbf{v}_i^{[j]}(\mathbf{Z})\|_F^2$$

où $\mathbf{v}_i^{[j]}$ est la composante du bruit provenant de la $i^{\text{ème}}$ source. Aussi, $\mathbf{v}_i^{[j]}$ est partitionné en $K = \frac{P(P-1)}{2}$ vecteurs de dimension 4,

$$\mathbf{v}_i^{[j]} = [\mathbf{v}_{i,12}^{[j]T}, \mathbf{v}_{i,13}^{[j]T}, \dots, \mathbf{v}_{i,P(P-1)}^{[j]T}]^T \mathbf{v}_{i,pq}^{[j]} = \left(\mathbf{B}^{[j]*} \otimes \mathbf{B}^{[j]} \right) (\mathbf{Z}_{i,q}^{[j]*} \otimes \mathbf{Z}_{i,p}) \text{vec}(\mathbf{C}_i^{[j]}). \quad (\text{A.39})$$

De la même façon, nous notons $\hat{\mathbf{u}}_{i,pq}^{[j]}$ le vecteur de dimension 4 tel que,

$$\hat{\mathbf{u}}_i^{[j]} = [\hat{\mathbf{u}}_{i,12}^{[j]}, \hat{\mathbf{u}}_{i,13}^{[j]}, \dots, \hat{\mathbf{u}}_{i,(P-1)P}^{[j]}]. \quad (\text{A.40})$$

Ainsi, $\phi(\mathbf{Z}_i)$ s'écrit comme suit,

$$\phi(\mathbf{Z}_i) = \sum_{j=1}^Q \sum_{(p,q) \in [1;P]^2 \mid p < q} \|\hat{\mathbf{u}}_{i,pq}^{[j]} - \mathbf{v}_{i,pq}^{[j]}(\mathbf{Z}_i)\|_F^2 \quad (\text{A.41})$$

où \mathbf{Z}_i est une matrice de taille $2MP \times 2$ exprimé comme une matrice par bloc de taille $2M \times 2$,

$$\mathbf{Z}_i = \begin{bmatrix} \mathbf{Z}_{i,1} \\ \mathbf{Z}_{i,2} \\ \vdots \\ \mathbf{Z}_{i,P} \end{bmatrix} \quad \text{où } \mathbf{Z}_{i,p} \in \mathbb{C}^{2\zeta \times 2} \text{ et } \mathbf{Z}_{i,p} = \begin{bmatrix} \mathbf{Z}_{i,p,1} \\ \mathbf{Z}_{i,p,2} \\ \vdots \\ \mathbf{Z}_{i,p,\zeta} \end{bmatrix} \quad \text{où } \mathbf{Z}_{i,p,c} \in \mathbb{C}_{2 \times 2} \quad (\text{A.42})$$

Nous rappelons que,

$$\begin{aligned} \mathbf{v}_{i,pq}^{[j]} &= ((\mathbf{B}^{[j]*} \mathbf{Z}_{i,q}^* \otimes (\mathbf{B}^{[j]} \mathbf{Z}_{i,p})) \text{vec}(\mathbf{C}_i^{[j]})) \\ &= \text{vec}(\mathbf{B}^{[j]} \mathbf{Z}_{i,p} \mathbf{C}_i^{[j]} \mathbf{Z}_{i,q}^H \mathbf{B}^{[j]H}) \end{aligned} \quad (\text{A.43})$$

En définissant $\hat{\mathbf{U}}_{i,pq}^{[j]} = \text{vec}^{-1}(\hat{\mathbf{u}}_{i,pq}^{[j]})$, nous pouvons exprimer $\phi(\mathbf{Z}_i)$ ainsi,

$$\phi(\mathbf{Z}_i) = \sum_{j=1}^Q \sum_{(p,q) \in [1:P]^2 | p < q} \|\hat{\mathbf{U}}_{i,pq}^{[j]} - \mathbf{B}^{[j]} \mathbf{Z}_{i,p} \mathbf{C}_i^{[j]} \mathbf{Z}_{i,q}^H \mathbf{B}^{[j]H}\|_F^2. \quad (\text{A.44})$$

$\phi(\mathbf{Z}_i)$ est minimisé par rapport à \mathbf{Z}_i par une descente de coordonnées [75] en minimisant de manière alternée $\phi(\mathbf{Z}_i)$ par rapport à $\mathbf{Z}_{i,p}$ en fixant $\mathbf{Z}_{i,q}$ pour $p \neq q$. Soit $p \in [1 : P]$, nous notons,

$$f_{i,pq}^{[j]}(\mathbf{Z}_{i,p}) = \|\hat{\mathbf{U}}_{i,pq}^{[j]} - \mathbf{B}^{[j]} \mathbf{Z}_{i,p} \mathbf{C}_i^{[j]} \mathbf{Z}_{i,q}^H \mathbf{B}^{[j]H}\|_F^2 \quad (\text{A.45})$$

$$g_{i,pq}^{[j]}(\mathbf{Z}_{i,p}) = \|\hat{\mathbf{U}}_{i,qp}^{[j]} - \mathbf{B}^{[j]} \mathbf{Z}_{i,q} \mathbf{C}_i^{[j]} \mathbf{Z}_{i,p}^H \mathbf{B}^{[j]H}\|_F^2. \quad (\text{A.46})$$

La fonction objective $\phi(\mathbf{Z}_{i,p})$ devient alors,

$$\phi(\mathbf{Z}_{i,p}) \propto \sum_{j=1}^Q \sum_{q=p+1}^P f_{i,pq}^{[j]}(\mathbf{Z}_{i,p}) + \sum_{j=1}^Q \sum_{q=1}^{p-1} g_{i,pq}^{[j]}(\mathbf{Z}_{i,p}). \quad (\text{A.47})$$

Les gradients de f et g par rapport à $\mathbf{Z}_{i,p}$ sont calculés en utilisant [74],

$$\frac{\partial f_{i,pq}^{[j]}}{\partial \mathbf{Z}_{i,p}} = \mathbf{B}^{[j]T} (\hat{\mathbf{U}}_{i,pq}^{[j]*} - \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{F}_{i,q}^{[j]H}) \mathbf{F}_{i,q}^{[j]} \quad (\text{A.48})$$

$$\frac{\partial g_{i,pq}^{[j]}}{\partial \mathbf{Z}_{i,p}} = \mathbf{B}^{[j]T} (\hat{\mathbf{U}}_{i,qp}^{[j]T} - \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{G}_{i,q}^{[j]H}) \mathbf{G}_{i,q}^{[j]}. \quad (\text{A.49})$$

Ainsi, le gradient de $\phi(\mathbf{Z}_{i,p})$ par rapport à $\mathbf{Z}_{i,p}$ s'exprime ainsi,

$$\begin{aligned} \frac{\partial \phi}{\partial \mathbf{Z}_{i,p}} &= \sum_{j=1}^Q \sum_{q=p+1}^P \mathbf{B}^{[j]T} (\hat{\mathbf{U}}_{i,pq}^{[j]*} - \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{F}_{i,q}^{[j]H}) \mathbf{F}_{i,q}^{[j]} \\ &\quad + \sum_{j=1}^Q \sum_{q=1}^{p-1} \mathbf{B}^{[j]T} (\hat{\mathbf{U}}_{i,qp}^{[j]T} - \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{G}_{i,q}^{[j]H}) \mathbf{G}_{i,q}^{[j]} \\ &= \sum_{j=1}^Q \sum_{q=p+1}^P \mathbf{B}^{[j]T} \hat{\mathbf{U}}_{i,pq}^{[j]*} \mathbf{F}_{i,q}^{[j]} + \sum_{j=1}^Q \sum_{q=1}^{p-1} \mathbf{B}^{[j]T} \hat{\mathbf{U}}_{i,qp}^{[j]T} \mathbf{G}_{i,q}^{[j]} \\ &\quad - \sum_{j=1}^Q \sum_{q=p+1}^P \mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{F}_{i,q}^{[j]H} \mathbf{F}_{i,q}^{[j]} \\ &\quad - \sum_{j=1}^Q \sum_{q=1}^{p-1} \mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{G}_{i,q}^{[j]H} \mathbf{G}_{i,q}^{[j]}. \end{aligned} \quad (\text{A.50})$$

Notons,

$$\mathbf{T}_{i,p} = \sum_{j=1}^Q \sum_{q=p+1}^P \mathbf{B}^{[j]T} \hat{\mathbf{U}}_{i,pq}^{[j]*} \mathbf{F}_{i,q}^{[j]} + \sum_{j=1}^Q \sum_{q=1}^{p-1} \mathbf{B}^{[j]T} \hat{\mathbf{U}}_{i,qp}^{[j]T} \mathbf{G}_{i,q}^{[j]}.$$

Le gradient de $\phi(\mathbf{Z}_{i,p})$ par rapport à $\mathbf{Z}_{i,p}$ devient alors,

$$\frac{\partial \phi}{\mathbf{Z}_{i,p}} = \mathbf{T}_{i,p} - \sum_{j=1}^Q \sum_{q=p+1}^P \mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{F}_{i,q}^{[j]H} \mathbf{F}_{i,q}^{[j]} - \sum_{j=1}^Q \sum_{q=1}^{p-1} \mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{G}_{i,q}^{[j]H} \mathbf{G}_{i,q}^{[j]}.$$
(A.51)

Annuler le gradient de $\phi(\mathbf{Z}_{i,p})$ par rapport à $\mathbf{Z}_{i,p}$ est équivalent à annuler le gradient de $\text{vec}(\phi(\mathbf{Z}_{i,p}))$ par rapport à $\text{vec}(\mathbf{Z}_{i,p})$,

$$\begin{aligned} \text{vec} \left(\frac{\partial \phi}{\mathbf{Z}_{i,p}} \right) &= \text{vec}(\mathbf{T}_{i,p}) - \sum_{j=1}^Q \sum_{q=p+1}^P \text{vec}(\mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{F}_{i,q}^{[j]H} \mathbf{F}_{i,q}^{[j]}) \\ &\quad - \sum_{j=1}^Q \sum_{q=1}^{p-1} \text{vec}(\mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \mathbf{Z}_{i,p}^* \mathbf{G}_{i,q}^{[j]H} \mathbf{G}_{i,q}^{[j]}) \\ &= \text{vec}(\mathbf{T}_i) - \sum_{j=1}^Q \sum_{q=p+1}^P \left(\mathbf{F}_{i,q}^{[j]T} \mathbf{F}_{i,q}^{[j]*} \otimes \mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \right) \text{vec}(\mathbf{Z}_{i,p}^*) \\ &\quad - \sum_{j=1}^Q \sum_{q=1}^{p-1} \left(\mathbf{G}_{i,q}^{[j]T} \mathbf{G}_{i,q}^{[j]*} \otimes \mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \right) \text{vec}(\mathbf{Z}_{i,p}^*). \end{aligned}$$
(A.52)

Enfin, nous obtenons une solution pour $\mathbf{Z}_{i,p}$ tel que $\text{vec} \left(\frac{\partial \phi}{\mathbf{Z}_{i,p}} \right) = 0$,

$$\text{vec}(\mathbf{Z}_{i,p}^*) = (\mathbf{M}_{i,p} + \tilde{\mathbf{M}}_{i,p})^{-1} \text{vec}(\mathbf{T}_{i,p}),$$

$$\text{avec } \mathbf{M}_{i,p} = \sum_{j=1}^Q \sum_{q=p+1}^P \left(\mathbf{F}_{i,q}^{[j]T} \mathbf{F}_{i,q}^{[j]*} \otimes \mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \right)$$

$$\tilde{\mathbf{M}}_{i,p} = \sum_{j=1}^Q \sum_{q=1}^{p-1} \left(\mathbf{G}_{i,q}^{[j]T} \mathbf{G}_{i,q}^{[j]*} \otimes \mathbf{B}^{[j]T} \mathbf{B}^{[j]*} \right)$$

$$\mathbf{T}_{i,p} = \sum_{j=1}^Q \sum_{q=p+1}^P \mathbf{B}^{[j]T} \hat{\mathbf{U}}_{i,pq}^{[j]*} \mathbf{F}_{i,q}^{[j]} + \sum_{j=1}^Q \sum_{q=1}^{p-1} \mathbf{B}^{[j]T} \hat{\mathbf{U}}_{i,qp}^{[j]T} \mathbf{G}_{i,q}^{[j]}$$

$$\mathbf{F}_{i,q}^{[j]} = \mathbf{B}^{[j]*} \mathbf{Z}_{i,q}^* \mathbf{C}_i^{[j]T}$$

$$\mathbf{G}_{i,q}^{[j]} = \mathbf{B}^{[j]*} \mathbf{Z}_{i,q}^* \mathbf{C}_i^{[j]*}.$$

Annexe B - Imagerie interférométrique en présence d'interférences

B.1 Calcul de l'étape E de l'algorithme 4.1

Étant donnée le modèle défini en (4.2), la vraisemblance complète est donnée par,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) = \mathbb{E}_{(\mathbf{y}, \tau_1, \dots, \tau_K) | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\log(p(\mathbf{y}, \tau_1, \dots, \tau_K | \boldsymbol{\theta}))] \quad (\text{B.1})$$

Comme nous cherchons à maximiser $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)})$ par rapport à $\boldsymbol{\theta}$, nous pouvons ignorer les termes qui n'en dépendent pas, conduisant à l'expression suivante,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) \propto \sum_{k=1}^K \mathbb{E}_{\tau_k | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\log(p(\mathbf{y}[k] | \tau_k, \boldsymbol{\theta}))] \quad (\text{B.2})$$

Le modèle défini en (4.2) est un modèle gaussien composé, ainsi, la distribution conditionnelle de $\mathbf{y}[k]$ sachant τ_k et $\boldsymbol{\theta}$ est une distribution gaussienne de moyenne $\mathbf{h}_k \mathbf{x}$ et de variance $\tau_k \sigma^2$, où \mathbf{h}_k est la $k^{\text{ème}}$ ligne de l'opérateur direct \mathbf{H} . L'expression de la log-vraisemblance conditionnelle est alors donnée par,

$$\log(p(\mathbf{y}[k] | \tau_k, \boldsymbol{\theta})) = -\log(\pi \tau_k \sigma^2) - \frac{1}{\tau_k \sigma^2} |\mathbf{y}[k] - \mathbf{h}_k^T \mathbf{x}|^2. \quad (\text{B.3})$$

Nous pouvons calculer l'espérance de la log-vraisemblance conditionnelle s'écrit comme suit,

$$\mathbb{E}_{\tau_k | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\log(p(\mathbf{y}[k] | \tau_k, \boldsymbol{\theta}))] \propto -\log(\sigma^2) - \frac{\hat{\omega}_k^{(m)}}{\sigma^2} |\mathbf{y}[k] - \mathbf{h}_k^T \mathbf{x}|^2 \quad (\text{B.4})$$

où $\hat{\omega}_k^{(m)} = \mathbb{E}_{\tau_k | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\tau_k^{-1}]$. La fonction objective $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)})$ s'écrit alors,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) \propto \sum_{k=1}^K -\log(\sigma^2) - \sum_{k=1}^K \frac{\hat{\omega}_k^{(m)}}{\sigma^2} |\mathbf{y}[k] - \mathbf{h}_k^T \mathbf{x}|^2. \quad (\text{B.5})$$

B.2 Calcul de l'étape E de l'algorithme 4.2

Nous considérons le modèle de données défini en (4.2) et la décomposition du bruit définie en (4.13). Nous supposons la condition explicitée dans la section 4.3 est satisfaite. Nous cherchons à calculer l'espérance de la log-vraisemblance des données complètes sachant le vecteur de mesures et l'estimé courant des paramètres $\boldsymbol{\theta}^{(m)}$,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) = \mathbb{E}_{\xi | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\log(p(\mathbf{y}, \mathbf{z}, \boldsymbol{\Omega} | \boldsymbol{\theta}))] \quad (\text{B.6})$$

Nous pouvons ignorer les termes qui ne dépendent pas de x , conduisant à l'expression suivante,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) \propto \mathbb{E}_{z, \Omega | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\log(p(z|\Omega, \boldsymbol{\theta}))] \quad (\text{B.7})$$

Étant donnée la modélisation de $z|\Omega$, nous pouvons calculer la log-vraisemblance conditionnelle de z sachant Ω et $\boldsymbol{\theta}$,

$$\log(p(z|\Omega, \boldsymbol{\theta})) \propto \frac{1}{\sigma_e^2} \|z - \mathbf{F}\mathbf{x}\|^2 \quad (\text{B.8})$$

Ainsi, la log-vraisemblance conditionnelle s'écrit comme suit,

$$\mathbb{E}_{z, \Omega | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\log(p(z|\Omega, \boldsymbol{\theta}))] = \frac{1}{\sigma_e^2} \|\mathbb{E}_{z, \Omega | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [z] - \mathbf{F}\mathbf{x}\|^2 + C_{\boldsymbol{\theta}} \quad (\text{B.9})$$

L'espérance conditionnelle de z est calculée en utilisant le théorème de l'espérance totale

$$\hat{z}^{(m)} = \mathbb{E}_{\Omega | \mathbf{y}, \boldsymbol{\theta}^{(m)}} \left[\mathbb{E}_{z | \mathbf{y}, \Omega, \boldsymbol{\theta}^{(m)}} [z] \right]. \quad (\text{B.10})$$

De plus, z_{Ω} et \mathbf{y}_{Ω} sont conjointement gaussien, ainsi nous pouvons calculer l'espérance conditionnelle de z sachant \mathbf{y} et Ω ainsi [112],

$$\mathbb{E}_{z | \mathbf{y}, \Omega, \boldsymbol{\theta}^{(m)}} [z] = \mathbf{F}\mathbf{x}^{(m)} + \frac{\sigma_e^2}{\sigma^2} \mathbf{S}^H \Omega (\mathbf{y} - \mathbf{S}\mathbf{F}\mathbf{x}^{(m)}) \quad (\text{B.11})$$

Cela conduit à l'expression suivante pour l'espérance conditionnelle de z ,

$$\hat{z}^{(m)} = \mathbf{F}\mathbf{x}^{(m)} + \frac{\sigma_e^2}{\sigma^2} \mathbf{S}^H \hat{\Omega}^{(m)} (\mathbf{y} - \mathbf{S}\mathbf{F}\mathbf{x}^{(m)}) \quad (\text{B.12})$$

avec,

$$\hat{\Omega}^{(m)} = \mathbb{E}_{\Omega | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\Omega]. \quad (\text{B.13})$$

Les textures τ_1, \dots, τ_K sont supposé indépendantes, ainsi la matrice de covariance Ω est diagonale et son espérance est données par,

$$\begin{aligned} \hat{\Omega}^{(m)} &= \text{diag}(\hat{\omega}_1^{(m)}, \dots, \hat{\omega}_K^{(m)}) \\ \hat{\omega}_k^{(m)} &= \mathbb{E}_{\tau_k | \mathbf{y}, \boldsymbol{\theta}^{(m)}} [\tau_k^{-1}] \end{aligned} \quad (\text{B.14})$$

La fonction objective $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)})$ s'écrit alors,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) \propto \|\hat{z}^{(m)} - \mathbf{F}\mathbf{x}^{(m)}\|^2. \quad (\text{B.15})$$

Annexe C - Systèmes de coordonnées utilisés en radioastronomie

C.1 Systèmes de coordonnées équatoriales

Un **système de coordonnées équatoriales** est un système de coordonnées en 3 dimensions dont les composantes ne dépendent pas de la position de l'observateur et de l'instant de l'observation. Il sert à repérer les objets célestes les uns par rapport aux autres. Les directions sont exprimées en coordonnées sphériques (α, δ) , où α est l'ascension droite et δ la déclinaison, aussi noté RA et DEC pour *Right Ascension* et *Declination*. L'axe Z est dirigé vers le pôle nord céleste. L'équateur céleste est alors défini comme le plan orthogonal à Z passant par le centre de la sphère céleste. L'écliptique est le plan correspondant à l'orbite de la Terre dans son mouvement relatif au Soleil. Ce dernier intersecte l'équateur céleste en deux points, appelés points vernaux et correspondant aux équinoxes d'automne et de printemps. L'axe X est alors choisi comme étant le point vernal d'équinoxe, noté γ , et l'axe Y est l'axe orthogonal orienté vers l'est. Ainsi, les points de références du système de coordonnées équatoriales sont le **pôle nord céleste**, le **point vernal** et l'**équateur céleste**.

Les **coordonnées horaires** sont définies par rapport à un lieu d'observation. Elles permettent de décrire la position d'un objet sur la sphère céleste en fonction de l'instant et de la latitude du lieu d'observation. Les directions sont exprimées en coordonnées sphériques (H, δ) , où H est l'angle horaire et δ la déclinaison. Le repère horaire est une rotation du repère équatorial autour de l'axe Z d'un angle dépendant de la latitude du lieu d'observation et de la position du point vernal. En effet, la position du point vernal varie en fonction du temps. Ainsi, pour un lieu d'observation donné, Le temps sidéral local, LST, est l'angle horaire mesuré par rapport au point vernal. Le temps sidéral permet alors de convertir les coordonnées équatoriales en coordonnées horaires,

$$H = \text{LST} - \alpha. \quad (\text{C.1})$$

La figure C.1 illustre les systèmes de coordonnées équatoriales et horaires.

Le repère ICRF (*International Celestial Reference Frame*) décrit les différentes convention adoptées par l'IAU (*International Astronomical Union*)¹ pour normaliser le système de coordonnées équatoriales. Il inclut la position du point vernal et du pôle nord céleste. L'IERS (*International Earth Rotation and Reference Systems Service*)², en collaboration avec l'IAU, est l'organisme en charge de la réalisation de ce repère.

C.2 Systèmes de coordonnées géocentriques

-
1. <https://www.iau.org/>
 2. <https://www.iers.org/>

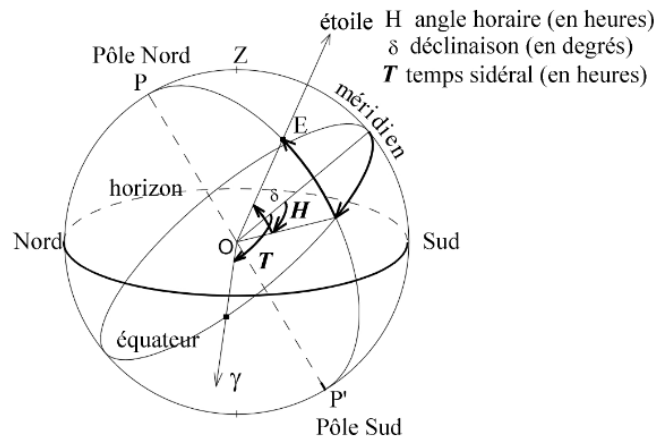


Figure C.1 – Systèmes de coordonnées équatoriales et horaires.

Les système de coordonnées géocentrique permettent de positionner des objets à la surface de la Terre. On parle de repère ECEF pour *Earth-centered Earth-fixed coordinate system*.

Le repère ITRF *International Terrestrial Reference Frame*³ est un repère ECEF dont l'origine est au centre de masse de la Terre. L'axe Oz relie le pôle sud au pôle nord, l'axe Ox est dirigé vers le méridien du Greenwich et l'axe Oy est orthogonal à l'axe Ox dans la direction est. Ce dernier est maintenu par l'IERS dont la dernière réalisation date de l'année 2020 (ITRF2020)⁴.

C.3 Systèmes de coordonnées horizontales

Un système de coordonnées horizontales est un système de coordonnées locales pouvant être utilisé par un observateur au sol. Ce système de coordonnées permet de définir la direction pointée par un télescope par rapport plan tangent à la surface de la Terre au lieu d'observation.

Les directions sont exprimées en coordonnées sphériques (a, h) , où a est l'azimuth et h est la hauteur. L'azimuth est l'angle entre le nord et la direction pointée par le télescope. La hauteur est l'angle entre le plan tangent à la surface de la Terre et la direction pointée par le télescope. La figure C.2 illustre le système de coordonnées horizontales.

3. <https://itrf.ign.fr/en/homepage>

4. <https://itrf.ign.fr/en/solutions/itrf2020>

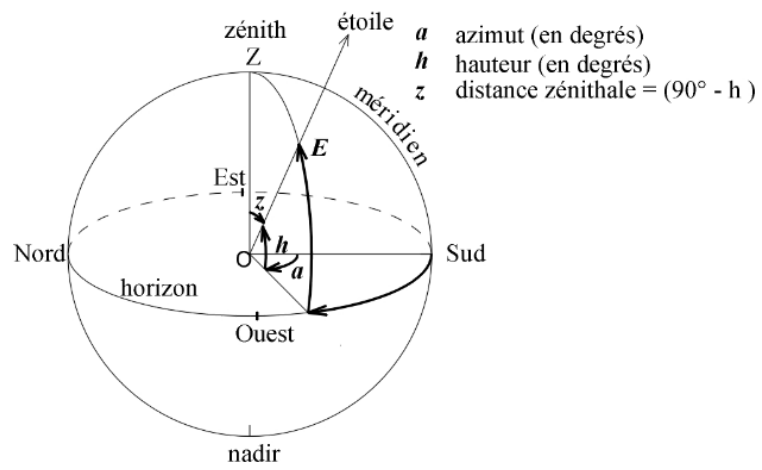


Figure C.2 – Systèmes de coordonnées horizontales.

Annexe D - Simulation des visibilités

D.1 Choix de la résolution de l'image du ciel

Les visibilités sont simulées à partir d'une image du ciel discrétisée et des coordonnées du plan uv . L'ensemble des coordonnées du plan uv est défini par les coordonnées des lignes de base du réseau interférométrique considéré. On note U et V les ensembles des coordonnées u et v disponible, respectivement,

$$U = \{u_1, u_2, \dots, u_{N_u}\}, V = \{v_1, v_2, \dots, v_{N_v}\} \quad (\text{D.1})$$

Chaque pixel de l'image du ciel discrétisée correspond à des cosinus directeurs l et m échantillonnés avec un pas Δ_l et Δ_m ,

$$l_k = \left(-\frac{N}{2} + k\right) \Delta_l, m_k = \left(-\frac{N}{2} + k\right) \Delta_m \quad (\text{D.2})$$

où $k \in \{0, 1, \dots, N-1\}$ et N est le nombre de pixels de l'image du ciel. On se place sous les conditions du théorème de Van-Cittert Zernike, ainsi, la visibilité, $v[u, v]$, pour les coordonnées (u, v) du plan uv peut être simulée en effectuant une DFT de l'image du ciel discrétisée,

$$v[u, v] = \sum_{k=0}^{N-1} I[l_k, m_k] \exp(-2\pi j (ul_k + vm_k)), \quad (\text{D.3})$$

où $I[l_k, m_k]$ est l'intensité du $k^{\text{ème}}$ pixel de l'image du ciel discrétisée. Les pas d'échantillonnage Δ_l et Δ_m sont choisis de sorte à respecter le théorème de Shannon,

$$\Delta_l \leq \frac{1}{2u_{\max}}, \Delta_m \leq \frac{1}{2v_{\max}} \quad (\text{D.4})$$

où u_{\max} et v_{\max} sont les coordonnées maximales des lignes de base du réseau interférométrique considéré,

$$u_{\max} = \max(U), v_{\max} = \max(V). \quad (\text{D.5})$$

D.2 Projection des lignes de base sur le plan uv

Les coordonnées du plan uv sont une projection des vecteurs ligne de base sur le plan orthogonal à la direction d'observation. Les vecteurs ligne de base sont définis dans le repère ITRF et sont exprimés en mètres. On note B l'ensemble des vecteurs ligne de base,

$$B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_K\}, \quad (\text{D.6})$$

où K est le nombre de lignes de base du réseau interférométrique considéré.

La direction d'observation est définie par son angle horaire, H , et sa déclinaison, δ . Les coordonnées des vecteurs ligne de base sont projetées sur le plan orthogonal à la

direction d'observation en effectuant deux rotations : une rotation selon l'axe Oz pour aligner les vecteurs ligne de base avec l'angle horaire de la source observée et une rotation selon l'axe Ox pour aligner les vecteurs ligne de base avec la direction d'observation,

$$\mathbf{b}'_k = \mathbf{R}_x\left(\frac{\pi}{2} - \delta\right)\mathbf{R}_z\left(\frac{\pi}{2} - H\right)\mathbf{b}_k, \quad (\text{D.7})$$

où,

$$\mathbf{R}_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad \mathbf{R}_z(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (\text{D.8})$$

et donc,

$$\mathbf{R}_x\left(\frac{\pi}{2} - \delta\right)\mathbf{R}_z\left(\frac{\pi}{2} - H\right) = \begin{pmatrix} \sin(H) & \cos(H) & 0 \\ -\sin(\delta)\cos(H) & \sin(H)\sin(\delta) & \cos(\delta) \\ \cos(H)\cos(\delta) & -\cos(\delta)\sin(H) & \sin(\delta) \end{pmatrix}. \quad (\text{D.9})$$

D.3 Earth-rotation synthesis

La rotation de la Terre induit un mouvement de la direction pointée par le réseau de radiotélescopes. Ainsi, la direction d'observation varie au cours du temps et les coordonnées des vecteurs ligne de base projetés sur le plan orthogonal à la direction d'observation varient également.

Afin de simuler l'effet de la rotation de la Terre, on considère que la source observée se trouve au zénith sur le lieu d'observation et que son angle horaire varie de $-\frac{T}{2}$ à $\frac{T}{2}$, où T est la durée de l'observation en heure. La source apparaît donc à l'horizon Est, se déplace jusqu'au zénith et disparaît à l'horizon Ouest.

Les lignes de bases sont donc projetées sur le plan orthogonal à la direction d'observation pour chaque angle horaire de l'observation, $H_t = t\Delta_t$, où $t \in \{0, 1, \dots, T-1\}$ et Δ_t est le pas de temps de l'observation en heure.

Bibliographie

- [1] A. R. Thompson, J. M. Moran et G. W. Swenson, *Interferometry and synthesis in radio astronomy*. Springer Nature, 2017.
- [2] M. P. van Haarlem, M. W. Wise, A. Gunst et al., « LOFAR : The low-frequency array, » *Astronomy & astrophysics*, t. 556, A2, 2013.
- [3] T. Shimwell, C. Tasse, M. Hardcastle et al., « The LOFAR two-metre sky survey-II. First data release, » *Astronomy & Astrophysics*, t. 622, A1, 2019.
- [4] K. Akiyama, A. Alberdi, W. Alef et al., « First M87 event horizon telescope results. II. Array and instrumentation, » *The Astrophysical Journal Letters*, t. 875, n° 1, p. L2, 2019.
- [5] P. E. Dewdney, P. J. Hall, R. T. Schilizzi et T. J. L. Lazio, « The square kilometre array, » *Proceedings of the IEEE*, t. 97, n° 8, p. 1482-1496, 2009.
- [6] P. Zarka, M. Tagger, L. Denis et al., « NenUFAR : Instrument description and science case, » in *2015 International Conference on Antenna Theory and Techniques (ICATT)*, IEEE, 2015, p. 1-6.
- [7] P. Fridman et W. Baan, « RFI mitigation methods in radio astronomy, » *Astronomy & Astrophysics*, t. 378, n° 1, p. 327-344, 2001.
- [8] W. A. Baan, « Implementing RFI mitigation in radio science, » *Journal of Astronomical Instrumentation*, t. 8, n° 01, p. 1 940 010, 2019.
- [9] A. Offringa, A. De Bruyn, S. Zaroubi et M. Biehl, « A LOFAR RFI detection pipeline and its first results, » *arXiv preprint arXiv :1007.2089*, 2010.
- [10] A. Leshem et A.-J. van der Veen, « Introduction to interference mitigation techniques in radio astronomy, » in *Perspectives on Radio Astronomy : Technologies for Large Antenna Arrays*, 2000, p. 201.
- [11] S. Kazemi, S. Yatawatta, S. Zaroubi et al., « Radio interferometric calibration using the SAGE algorithm, » *Monthly Notices of the Royal Astronomical Society*, t. 414, n° 2, p. 1656-1666, 2011.
- [12] T. L. Grobler, C. Nunhokee, O. M. Smirnov, A. Van Zyl et A. De Bruyn, « Calibration artefacts in radio interferometry-I. Ghost sources in Westerbork Synthesis Radio Telescope data, » *Monthly Notices of the Royal Astronomical Society*, t. 439, n° 4, p. 4030-4047, 2014.
- [13] V. Ollier, M. N. El Korso, A. Ferrari, R. Boyer et P. Larzabal, « Robust distributed calibration of radio interferometers with direction dependent distortions, » *Signal Processing*, t. 153, p. 348-354, 2018.

- [14] M. Brossard, M. N. El Korso, M. Pesavento et al., « Parallel multi-wavelength calibration algorithm for radio astronomical arrays, » *Signal Processing*, t. 145, p. 258-271, 2018.
- [15] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal et M. Pesavento, « Robust calibration of radio interferometers in non-Gaussian environment, » *IEEE Transactions on Signal Processing*, t. 65, n° 21, p. 5649-5660, 2017.
- [16] M. Brossard, V. Ollier, M. El Korso, R. Boyer et P. Larzabal, *Fast decentralized multi-frequency calibration for doa estimation with distributed sensors*, 2020.
- [17] B. Mériaux, X. Zhang, M. N. El Korso et M. Pesavento, « Iterative marginal maximum likelihood DOD and DOA estimation for MIMO radar in the presence of SIRP clutter, » *Signal Processing*, t. 155, p. 384-390, 2019, issn : 0165-1684. adresse : <https://www.sciencedirect.com/science/article/pii/S0165168418303189>.
- [18] X. Zhang, M. N. El Korso et M. Pesavento, « MIMO radar target localization and performance evaluation under SIRP clutter, » *Signal Processing*, t. 130, p. 217-232, 2017.
- [19] X. Zhang, M. N. El Korso et M. Pesavento, « Maximum likelihood and maximum a posteriori direction-of-arrival estimation in the presence of SIRP noise, » in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2016, p. 3081-3085.
- [20] X. Zhang, M. N. El Korso et M. Pesavento, « MIMO radar performance analysis under K-distributed clutter, » in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2014, p. 5287-5291.
- [21] U. Sob, H. L. Bester, O. Smirnov, J. Kenyon et T. Grobler, « Radio interferometric calibration using a complex Student's t-distribution and Wirtinger derivatives, » *Monthly Notices of the Royal Astronomical Society*, t. 491, n° 1, p. 1026-1042, 2020.
- [22] B. Mériaux, C. Ren, A. Breloy, M. N. El Korso et P. Forster, « Matched and mismatched estimation of kronecker product of linearly structured scatter matrices under elliptical distributions, » *IEEE Transactions on Signal Processing*, t. 69, p. 603-616, 2020.
- [23] B. Meriaux, C. Ren, M. N. El Korso, A. Breloy et P. Forster, « Robust estimation of structured scatter matrices in (mis) matched models, » *Signal Processing*, t. 165, p. 163-174, 2019.
- [24] B. Mériaux, C. Ren, M. N. El Korso, A. Breloy et P. Forster, « Asymptotic Performance of Complex M -Estimators for Multivariate Location and Scatter Estimation, » *IEEE Signal Processing Letters*, t. 26, n° 2, p. 367-371, 2019.

- [25] J. Hogbom, « Aperture synthesis with a non-regular distribution of interferometer baselines, » *Astronomy and Astrophysics Supplement Series*, t. 15, p. 417, 1974.
- [26] A. Offringa et O. Smirnov, « An optimized algorithm for multiscale wide-band deconvolution of radio astronomical images, » *Monthly Notices of the Royal Astronomical Society*, t. 471, n° 1, p. 301-316, 2017.
- [27] T. J. Cornwell, « Multiscale CLEAN deconvolution of radio synthesis images, » *IEEE Journal of selected topics in signal processing*, t. 2, n° 5, p. 793-801, 2008.
- [28] Y. Wiaux, L. Jacques, G. Puy, A. M. Scaife et P. Vandergheynst, « Compressed sensing imaging techniques for radio interferometry, » *Monthly Notices of the Royal Astronomical Society*, t. 395, n° 3, p. 1733-1742, 2009.
- [29] L. Bester, A. Repetti, S. Perkins, O. M. Smirnov et J. S. Kenyon, « A Practical Preconditioner for Wide-field Continuum Imaging of Radio Interferometric Data, » in *Astronomical Society of the Pacific Conference Series*, sér. Astronomical Society of the Pacific Conference Series, t. 532, juill. 2022, p. 199.
- [30] A. Onose, R. E. Carrillo, A. Repetti et al., « Scalable splitting algorithms for big-data interferometric imaging in the SKA era, » *Monthly Notices of the Royal Astronomical Society*, t. 462, n° 4, p. 4314-4335, 2016.
- [31] R. E. Carrillo, J. D. McEwen et Y. Wiaux, « Sparsity averaging reweighted analysis (SARA) : a novel algorithm for radio-interferometric imaging, » *Monthly Notices of the Royal Astronomical Society*, t. 426, n° 2, p. 1223-1234, 2012.
- [32] A.-J. van der Veen et S. J. Wijnholds, « Signal processing tools for radio astronomy, » *Handbook of Signal Processing Systems*, p. 421-463, 2013.
- [33] J. Jonas et MeerKAT Team, « The MeerKAT Radio Telescope, » in *MeerKAT Science : On the Pathway to the SKA*, jan. 2016, 1, p. 1.
- [34] H. Ye, S. F. Gull, S. M. Tan et B. Nikolic, « Optimal gridding and degrading in radio interferometry imaging, » *Monthly Notices of the Royal Astronomical Society*, t. 491, n° 1, p. 1146-1159, 2020.
- [35] Philipp Arras and Martin Reinecke and Rüdiger Westermann and Torsten A. Enßlin, « Efficient wide-field radio interferometry response, » *Astronomy and Astrophysics*, t. 646, A58, 2021.
- [36] O. M. Smirnov, « Revisiting the radio interferometer measurement equation-I. A full-sky Jones formalism, » *Astronomy & Astrophysics*, t. 527, A106, 2011.
- [37] J. Wang, L. Bacharach, M. N. El Korso et P. Larzabal, « Barankin Bound vs Cramér-Rao Bound for Interferometric-Like Array Design at Low SNR, » in *2023 31st European Signal Processing Conference (EUSIPCO)*, IEEE, 2023, p. 1554-1558.

- [38] J. Wang, L. Bacharach, P. Larzabal et M. N. El Korso, « A Comparison of Antenna Placement Criteria Based on the Cramér-Rao and Barankin Bounds for Interferometer Arrays, » *Available at SSRN 4549498*,
- [39] F. De Gasperin, M. Mevius, D. Rafferty, H. Intema et R. Fallows, « The effect of the ionosphere on ultra-low-frequency radio-interferometric observations, » *Astronomy & Astrophysics*, t. 615, A179, 2018.
- [40] S. van der Tol et A.-J. van der Veen, « Ionospheric calibration for the LOFAR radio telescope, » in *2007 International Symposium on Signals, Circuits and Systems*, IEEE, t. 2, 2007, p. 1-4.
- [41] D. W. Marquardt, « An algorithm for least-squares estimation of nonlinear parameters, » *Journal of the society for Industrial and Applied Mathematics*, t. 11, n° 2, p. 431-441, 1963.
- [42] S. Kazemi et S. Yatawatta, « Robust radio interferometric calibration using the t-distribution, » *Monthly Notices of the Royal Astronomical Society*, t. 435, n° 1, p. 597-605, 2013.
- [43] D. Kushary, *The EM algorithm and extensions*, 1998.
- [44] S. Yatawatta, « Distributed radio interferometric calibration, » *Monthly Notices of the Royal Astronomical Society*, t. 449, n° 4, p. 4506-4514, 2015.
- [45] S. P. Boyd et L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [46] V. Ollier, M. N. El Korso, A. Ferrari, R. Boyer et P. Larzabal, « Robust calibration of radio interferometers in multi-frequency scenario, » in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2018, p. 3494-3498.
- [47] F. Li, T. J. Cornwell et F. de Hoog, « The application of compressive sampling to radio astronomy-i. deconvolution, » *Astronomy & Astrophysics*, t. 528, A31, 2011.
- [48] H. H. Bauschke, P. L. Combettes et al., *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2011, t. 408.
- [49] N. Parikh, S. Boyd et al., « Proximal algorithms, » *Foundations and trends® in Optimization*, t. 1, n° 3, p. 127-239, 2014.
- [50] D. L. Donoho, « De-noising by soft-thresholding, » *IEEE transactions on information theory*, t. 41, n° 3, p. 613-627, 1995.
- [51] R. E. Carrillo, V. Kartik, J.-P. Thiran et Y. Wiaux, « A scalable algorithm for radio-interferometric imaging, » in *Signal Processing with Adaptive Sparse Structured Representations 2015*, 2015.

- [52] L. Pratley, J. D. McEwen, M. d’Avezac et al., « Robust sparse image reconstruction of radio interferometric observations with PURIFY, » *Monthly Notices of the Royal Astronomical Society*, t. 473, n° 1, p. 1038-1058, 2018.
- [53] N. Monnier, F. Orieux, N. Gac et al., « Fast Sky to Sky Interpolation for Radio Interferometric Imaging, » in *2022 IEEE International Conference on Image Processing (ICIP)*, 2022, p. 1571-1575.
- [54] F. de Gasperin, T. J. Dijkema, A. Drabent et al., « Systematic effects in LOFAR data : A unified calibration strategy, » *Astronomy and Astrophysics*, t. 622, A5, A5, fév. 2019. arXiv : [1811.07954](https://arxiv.org/abs/1811.07954) [[astro-ph](https://arxiv.org/archive/astro).IM].
- [55] M. Haardt, M. Pesavento, F. Roemer et M. N. El Korso, « Subspace methods and exploitation of special array structures, » in *Academic Press Library in Signal Processing*, t. 3, Elsevier, 2014, p. 651-717.
- [56] R. B. Abdallah, A. Breloy, M. N. El Korso et D. Lautru, « Bayesian signal subspace estimation with compound Gaussian sources, » *Signal Processing*, t. 167, p. 107310, 2020.
- [57] J. Hamaker, J. Bregman et R. Sault, « Understanding radio polarimetry. I. Mathematical foundations, » *Astronomy and Astrophysics Supplement Series*, t. 117, n° 1, p. 137-147, 1996.
- [58] A.-J. Boonstra, S. J. Wijnholds, S. van der Tol et B. Jeffs, « Calibration, sensitivity and RFI mitigation requirements for LOFAR, » in *Proceedings.(ICASSP’05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, IEEE, t. 5, 2005, p. v-869.
- [59] S. Yatawatta, « Spatially constrained direction-dependent calibration, » *Monthly Notices of the Royal Astronomical Society*, t. 510, n° 2, p. 2718-2723, 2022.
- [60] J.J. Lemmon, « Wideband model of man-made HF noise and interference, » *Radio Science*, t. 32, n° 2, p. 525-539, 1997.
- [61] M. Lavielle, *Mixed effects models for the population approach : models, tasks, methods and tools*. CRC press, 2014.
- [62] A. H. Ferrer, M. N. El Korso, A. Breloy et G. Ginolhac, « Robust mean and covariance matrix estimation under heterogeneous mixed-effects model with missing values, » *Signal Processing*, t. 188, p. 108195, 2021.
- [63] A. Hippert-Ferrer, M. N. El Korso, A. Breloy et G. Ginolhac, « Robust low-rank covariance matrix estimation with a general pattern of missing values, » *Signal Processing*, t. 195, p. 108460, 2022.
- [64] J. A. Fessler et A. O. Hero, « Space-alternating generalized expectation-maximization algorithm, » *IEEE Transactions on signal processing*, t. 42, n° 10, p. 2664-2677, 1994.

- [65] S. J. Wijnholds et A.-J. van der Veen, « Fundamental imaging limits of radio telescope arrays, » *IEEE Journal of Selected Topics in Signal Processing*, t. 2, n° 5, p. 613-623, 2008.
- [66] S. Yatawatta, « Stochastic calibration of radio interferometers, » *Monthly Notices of the Royal Astronomical Society*, t. 493, n° 4, p. 6071-6078, 2020.
- [67] Y. Sun, A. Breloy, P. Babu et al., « Low-Complexity Algorithms for Low Rank Clutter Parameters Estimation in Radar Systems, » *IEEE Transactions on Signal Processing*, t. 64, n° 8, p. 1986-1998, 2016.
- [68] M. Rangaswamy, F. Lin et K. Gerlach, « Robust adaptive signal processing methods for heterogeneous radar clutter scenarios, » in *Proceedings of the 2003 IEEE Radar Conference (Cat. No. 03CH37474)*, 2003, p. 265-272. doi : [10.1109/NRC.2003.1203412](https://doi.org/10.1109/NRC.2003.1203412).
- [69] H. Yang, C. Chen, S. Chen, F. Xi et Z. Liu, « SAR RFI Suppression for Extended Scene Using Interferometric Data via Joint Low-Rank and Sparse Optimization, » *IEEE Geoscience and Remote Sensing Letters*, t. 18, n° 11, p. 1976-1980, 2021.
- [70] M. Mahot, F. Pascal, P. Forster et J.-P. Ovarlez, « Asymptotic properties of robust complex covariance matrix estimates, » *IEEE Transactions on Signal Processing*, t. 61, n° 13, p. 3348-3356, 2013.
- [71] A. P. Dempster, N. M. Laird et D. B. Rubin, « Maximum likelihood from incomplete data via the EM algorithm, » *Journal of the Royal Statistical Society : Series B (Methodological)*, t. 39, n° 1, p. 1-22, 1977.
- [72] C. J. Wu, « On the convergence properties of the EM algorithm, » *The Annals of statistics*, p. 95-103, 1983.
- [73] Z. Wang, Q. Gu, Y. Ning et H. Liu, « High dimensional expectation-maximization algorithm : Statistical optimization and asymptotic normality, » *arXiv preprint arXiv :1412.8729*, 2014.
- [74] A. Hjørungnes et D. Gesbert, « Complex-valued matrix differentiation : Techniques and key results, » *IEEE Transactions on Signal Processing*, t. 55, n° 6, p. 2740-2746, 2007.
- [75] J. Friedman, T. Hastie, H. Höfling et R. Tibshirani, « Pathwise Coordinate Optimization, » *The Annals of Applied Statistics*, t. 1, n° 2, p. 302-332, 2007, issn : 19326157. (visité le 22/07/2023).
- [76] L. Pratley, « Radio Astronomy Image Reconstruction in the Big Data Era, » thèse de doct., UCL (University College London), 2019.
- [77] Y. Mhiri, M. N. El Korso, A. Breloy et P. Larzabal, « Multifrequency Array Calibration in Presence of Radio Frequency Interferences, » *Signal Processing*, p. 108 613, 2022.

- [78] E. Ollila, D. E. Tyler, V. Koivunen et H. V. Poor, « Compound-Gaussian clutter modeling with an inverse Gaussian texture distribution, » *IEEE Signal Processing Letters*, t. 19, n° 12, p. 876-879, 2012.
- [79] Ç. Candan et F. Pascal, « Covariance Matrix Estimation of Texture Correlated Compound-Gaussian Vectors for Adaptive Radar Detection, » *IEEE Transactions on Aerospace and Electronic Systems*, t. 59, n° 3, p. 3009-3020, 2023.
- [80] Z. Wang, A. C. Bovik, H. R. Sheikh et E. P. Simoncelli, « Image quality assessment : from error visibility to structural similarity, » *IEEE transactions on image processing*, t. 13, n° 4, p. 600-612, 2004.
- [81] L. Connor, K. L. Bouman, V. Ravi et G. Hallinan, « Deep radio-interferometric imaging with POLISH : DSA-2000 and weak lensing, » *Monthly Notices of the Royal Astronomical Society*, t. 514, n° 2, p. 2614-2626, 2022.
- [82] M. Terris, A. Dabbech, C. Tang et Y. Wiaux, « Image reconstruction algorithms in radio interferometry : From handcrafted to learned regularization denoisers, » *Monthly Notices of the Royal Astronomical Society*, t. 518, n° 1, p. 604-622, 2023.
- [83] K. Gregor et Y. LeCun, « Learning fast approximations of sparse coding, » in *Proceedings of the 27th international conference on international conference on machine learning*, 2010, p. 399-406.
- [84] V. Monga, Y. Li et Y. C. Eldar, « Algorithm unrolling : Interpretable, efficient deep learning for signal and image processing, » *IEEE Signal Processing Magazine*, t. 38, n° 2, p. 18-44, 2021.
- [85] R. Pascanu, T. Mikolov et Y. Bengio, « On the difficulty of training recurrent neural networks, » in *International conference on machine learning*, Pmlr, 2013, p. 1310-1318.
- [86] Y. Bengio, N. Boulanger-Lewandowski et R. Pascanu, « Advances in optimizing recurrent networks, » in *2013 IEEE international conference on acoustics, speech and signal processing*, IEEE, 2013, p. 8624-8628.
- [87] V. Ollier, M. N. El Korso, A. Ferrari, R. Boyer et P. Larzabal, « Bayesian calibration using different prior distributions : an iterative maximum a posteriori approach for radio interferometers, » in *2018 26th European Signal Processing Conference (EUSIPCO)*, IEEE, 2018, p. 2673-2677.
- [88] D. P. Kingma et J. Ba, « Adam : A method for stochastic optimization, » *arXiv preprint arXiv :1412.6980*, 2014.
- [89] F. Acero, J. -. Acquaviva, R. Adam et al., *French SKA White Book - The French Community towards the Square Kilometre Array*, 2018. arXiv : [1712 . 06950](https://arxiv.org/abs/1712.06950) [[astro-ph.IM](https://arxiv.org/abs/1712.06950)].

- [90] B. Bean, S. Bhatnagar, S. Castro et al., « CASA, Common Astronomy Software Applications for Radio Astronomy, » *Publications of the Astronomical Society of the Pacific*, t. 134, n° 1041, p. 114-501, 2022.
- [91] A. Kemball et M. Wieringa, « MeasurementSet definition version 2.0, » URL : <http://casa.nrao.edu/Memos/229.html>, 2000.
- [92] G. N. van Diepen, « Casacore Table Data System and its use in the MeasurementSet, » *Astronomy and Computing*, t. 12, p. 174-180, 2015.
- [93] S. J. Perkins, H. L. Bester, B. Hugo, J. S. Kenyon et O. M. Smirnov, « A dask distributed radio astronomy reduction framework, » in *2021 XXXIVth General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, IEEE, 2021, p. 1-3.
- [94] A. Abdulaziz, A. Dabbech, A. Onose et Y. Wiaux, « A low-rank and joint-sparsity model for hyper-spectral radio-interferometric imaging, » in *2016 24th European Signal Processing Conference (EUSIPCO)*, IEEE, 2016, p. 388-392.
- [95] A. Abdulaziz, A. Dabbech et Y. Wiaux, « Wideband super-resolution imaging in radio interferometry via low rankness and joint average sparsity models (HyperSARA), » *Monthly Notices of the Royal Astronomical Society*, t. 489, n° 1, p. 1230-1248, 2019.
- [96] J. Deguignet, A. Ferrari, D. Mary et C. Ferrari, « Distributed multi-frequency image reconstruction for radio-interferometry, » in *2016 24th European Signal Processing Conference (EUSIPCO)*, IEEE, 2016, p. 1483-1487.
- [97] A. Ferrari, J. Deguignet, C. Ferrari et al., « Multi-frequency image reconstruction for radio interferometry. A regularized inverse problem approach, » *arXiv preprint arXiv:1504.06847*, 2015.
- [98] S. Wenger et M. Magnor, « A sparse reconstruction algorithm for multi-frequency radio images, » *Computer Graphics Lab, TU Braunschweig, Tech. Rep*, 2014.
- [99] E. Chouzenoux, J.-C. Pesquet et A. Repetti, « Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function, » *Journal of Optimization Theory and Applications*, t. 162, n° 1, p. 107-132, 2014.
- [100] P. L. Combettes et B. C. Vũ, « Variable metric forward-backward splitting with applications to monotone inclusions in duality, » *Optimization*, t. 63, n° 9, p. 1289-1318, 2014.
- [101] L. Condat, « A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, » *Journal of optimization theory and applications*, t. 158, n° 2, p. 460-479, 2013.

- [102] A. Repetti et Y. Wiaux, « Variable metric forward-backward algorithm for composite minimization problems, » *SIAM Journal on Optimization*, t. 31, n° 2, p. 1215-1241, 2021.
- [103] S. Bhatnagar, U. Rau, D. Green et M. Rupen, « Expanded Very Large Array Observations of Galactic Supernova Remnants : Wide-field Continuum and Spectral-index Imaging, » *The Astrophysical Journal Letters*, t. 739, n° 1, L20, p. L20, sept. 2011. arXiv : [1106.2796](https://arxiv.org/abs/1106.2796) [[astro-ph.IM](#)].
- [104] W. Goss, U. Schwartz, S. Siddesh et K. Weiler, « A high-resolution investigation of the shell source G55. 7+ 3.4 at 610 and 1415 MHz, » *Astronomy and Astrophysics*, vol. 61, no. 1, Oct. 1977, p. 93-98. *Research supported by the Nederlandse Organisatie voor Zuiver-Wetenschappelijk Onderzoek.*, t. 61, p. 93-98, 1977.
- [105] A. Repetti, J. Birdi, A. Dabbech et Y. Wiaux, « Non-convex optimization for self-calibration of direction-dependent effects in radio interferometric imaging, » *Monthly Notices of the Royal Astronomical Society*, t. 470, n° 4, p. 3981-4006, 2017.
- [106] A. Repetti, M. Terris, Y. Wiaux et J.-C. Pesquet, « Dual Forward-Backward Unfolded Network for Flexible Plug-and-Play, » in *2022 30th European Signal Processing Conference (EUSIPCO)*, IEEE, 2022, p. 957-961.
- [107] A. Dabbech, M. Terris, A. Jackson et al., « First AI for deep super-resolution wide-field imaging in radio astronomy : unveiling structure in ESO 137-006, » *The Astrophysical Journal Letters*, t. 939, n° 1, p. L4, 2022.
- [108] J.-C. Pesquet, A. Repetti, M. Terris et Y. Wiaux, « Learning maximally monotone operators for image recovery, » *SIAM Journal on Imaging Sciences*, t. 14, n° 3, p. 1206-1237, 2021.
- [109] K. Zhang, W. Zuo, Y. Chen, D. Meng et L. Zhang, « Beyond a gaussian denoiser : Residual learning of deep cnn for image denoising, » *IEEE transactions on image processing*, t. 26, n° 7, p. 3142-3155, 2017.
- [110] Y. Mhiri, L. Bester et O. Smirnov, « *t*-distribution based imaging in radiointerferometry : theory and validation on MeerKat data, » *en cours de rédaction*,
- [111] Y. Mhiri, M. N. El Korso, A. Breloy et P. Larzabal, « Unrolled EM algorithm for radio-interferometric imaging in the presence of interference, » *en cours de rédaction*,
- [112] T. W. Anderson, *An introduction to multivariate statistical analysis*. Wiley New York, 1958, t. 2.