UNIVERSITÉ DE MONTRÉAL

MÉTHODES DE RECONSTRUCTION EN TOMOGRAPHIE DE DIFFRACTION 3-D

CORENTIN FRIEDRICH
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(GÉNIE ÉLECTRIQUE)
SEPTEMBRE 2016

© Corentin Friedrich, 2016.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

MÉTHODES DE RECONSTRUCTION EN TOMOGRAPHIE DE DIFFRACTION 3-D

présentée par : FRIEDRICH Corentin
en vue de l'obtention du diplôme de : Philosophiæ Doctor

M. GOUSSARD Yves, Ph. D., membre et directeur de recherche
M. IDIER Jérôme, Ph. D., membre et directeur de recherche
M. BOURGUIGNON Sébastien, Ph. D., membre et codirecteur de recherche
M. LAURIN Jean-Jacques, Ph. D., membre
M. ANJOS Miguel, Ph. D., membre
M. CHAUMET Patrick, Ph. D., membre externe
M. THOMAS RODET, Ph. D., membre externe

RÉSUMÉ

Ce travail porte sur la tomographie micro-ondes, technique d'imagerie permettant de reconstruire une image tridimensionnelle des propriétés diélectriques d'un volume inconnu à partir d'une succession de mesures du champ électromagnétique diffracté par ce volume. Cette modalité d'imagerie est prometteuse dans un grand nombre d'applications (imagerie médicale, géophysique, contrôle non destructif de matériaux, ...) mais souffre d'un inconvénient majeur, un coût de calcul très élevé des algorithmes de reconstruction, qui freine son développement industriel.

Dans cette thèse, le problème d'imagerie micro-ondes est vu comme un problème inverse où l'on cherche à minimiser l'erreur entre les données et un modèle direct simulant la diffraction de l'onde en fonction des propriétés de l'objet inconnu. Le modèle physique est discrétisé par la méthode des moments. Ce problème est mal-posé, en particulier car le nombre d'inconnues est supérieur au nombre de mesures. La solution est donc définie par la minimisation d'un critère des moindres carrés pénalisé par une fonction de régularisation, dont l'optimisation requiert la mise en œuvre d'algorithmes itératifs d'optimisation locale. Le modèle direct est en outre non-linéaire, rendant la reconstruction plus difficile. Le calcul numérique de ce modèle direct, nécessaire au calcul du critère et de son gradient à chaque itération, concentre la majorité du coût de calcul de l'inversion. Il repose sur la résolution d'un grand nombre de systèmes linéaires.

Nous proposons dans cette thèse de réduire le coût de calcul de ces méthodes de reconstruction en se focalisant sur ces résolutions de systèmes. Pour cela, les travaux sont divisés en deux contributions. Tout d'abord, nous proposons une procédure permettant de réduire le nombre de systèmes à résoudre en s'appuyant sur la configuration du montage d'acquisition. La seconde contribution est d'accélérer les résolutions de systèmes à l'aide d'algorithmes par blocs. Ce type d'approche permet de résoudre conjointement plusieurs systèmes linéaires impliquant la même matrice opérateur.

Ces méthodes sont validées sur des simulations de problèmes 3D réalistes, puis appliquées à la reconstruction d'objets réels sur des données expérimentales de champs diffractés. Des résultats satisfaisants de reconstruction sont obtenus, où un gain d'un facteur deux sur le temps de calcul est obtenu, notamment sur les problèmes de reconstruction les plus difficiles.

ABSTRACT

This thesis is focused on microwave tomography. This imaging technique consists in estimating a three-dimensional mapping of the dielectric properties of an unknown volume from measurements of the electromagnetic field from a known incident wave and scattered by this volume. This is a promising technique that is used in various applications (medical imaging, geophysics, non- destructive testing, ...) but suffers from high computational costs. This is a reason why microwave imaging is not widely used in industry.

In this thesis, microwave imaging is considered as an inverse problem, where the error between the measurements and a forward model that describes the scattered field is minimized as a function of the properties of the volume. The physical model is discretized using the method of moments. This inverse problem is ill-posed because the number of unknowns is higher than the number of measurements. It is tackled through the minimization of a regularized least-squares cost function, which is addresed by local iterative optimization algorithms. Moreover, the forward model is non-linear. Thus, reconstruction is a difficult and costful procedure. The computation of the objective function and of its gradient requires the resolution of a high number of linear systems, which are performed at each iteration of the optimization algorithm and represent most of the computational cost.

In this thesis, we propose to reduce the computational costs of the reconstruction algorithms by focusing on the resolution of these linear systems. Two contributions are presented. The first one is a procedure in order to reduce the number of linear systems depending on the configuration of the measurement setup. The second contribution offers an efficient way to speed up the resolutions of the systems. We adapt block resolution algorithms, in order to jointly solve multiple linear systems involving a common operator matrix.

These methods are validated on simulated, realistic, 3D problems, and applied to the reconstruction of real objects from experimental measurements of scattered fields. satisfactory results are obtained, where the computation time can be reduced by a factor of two, in particular for the most difficult reconstruction problems.

TABLE DES MATIÈRES

RÉSUMÉ iv

ABSTRACT v

TABLE DES MATIÈRES vi

LISTE DES TABLEAUX xi

LISTE DES FIGURES xiii

LISTE DES SIGLES ET ABRÉVIATIONS xv

LISTE DES ANNEXES xvi

CHAPITRE 1 INTRODUCTION GÉNÉRALE 1
1.1 La tomographie micro-ondes (TMO) 1
1.2 Méthodes de reconstruction en TMO 2
1.3 Difficultés et objectifs de recherche 5
1.3.1 Optimiser le nombre de systèmes linéaires 6
1.3.2 Accélérer la résolution des systèmes linéaires 6
1.3.3 Étudier l'effet de la troncature et améliorer la convergence 6
1.4 Plan de la thèse 7
1.5 Cadre du travail 8

CHAPITRE 2 LE MODÈLE DIRECT 9
2.1 Modélisation physique de la diffraction d'ondes électromagnétiques 10
2.1.1 Équation différentielle régissant la propagation d'une onde électromagnétique 10
2.1.2 Solution intégrale de l'équation d'onde dans un milieu homogène . . . 11
2.1.3 Équations intégrales en TMO 12
2.2 Discrétisation du modèle physique et formulation matricielle 14
2.2.1 État de l'art des différentes approches de discrétisation des équations électromagnétiques 14
2.2.2 Discrétisation des équations intégrales par la méthode des moments . 17
2.3 Implémentation du problème direct 18

2.3.1 Résolution du système linéaire 19
2.3.2 Multiplication par $\mathbf{L}_x$: comment réduire l'espace mémoire 21
2.4 Validation du modèle direct 21
2.4.1 Comparaison à la solution analytique de Mie 22
2.4.2 Comparaison aux données réelles de l'Institut Fresnel 28
2.5 Conclusion 32

CHAPITRE 3 MÉTHODES D'INVERSION EN IMAGERIE MICRO-ONDES . . . 33
3.1 L'imagerie micro-ondes 33
3.2 L'inversion vue comme un problème d'optimisation 35
3.2.1 Formulation bilinéaire 35
3.2.2 Formulation non-linéaire 36
3.3 La régularisation 37
3.3.1 Régularisation "$\ell_2\ell_1$" 38
3.3.2 Régularisation additive ou multiplicative ? 39
3.4 Les algorithmes d'optimisation utilisés 40
3.5 Calcul du gradient pour le critère non-linéaire 44
3.6 Conclusion 46

CHAPITRE 4 ÉTUDE ALGÉBRIQUE DU GRADIENT ET RÉDUCTION DU NOMBRE DE SYSTÈMES LINÉAIRES 47
4.1 Mise en œuvre du calcul du jacobien 48
4.1.1 L'état adjoint pour le calcul du jacobien 49
4.1.2 Équivalence entre l'état adjoint et la formulation matricielle du jacobien 50
4.2 Mises en œuvre du calcul du gradient 52
4.2.1 Mise en œuvre directe 52
4.2.2 Mise en œuvre adjointe du jacobien 53
4.2.3 Choix de la formulation en fonction du montage d'imagerie 53
4.3 Théorème de réciprocité de Lorentz 54
4.4 Procédure finale d'optimisation du nombre de systèmes 56
4.5 Exemple de configurations présentes dans la littérature 57
4.5.1 Montage générique d'acquisition 57
4.5.2 Montage simulé de Zhang et Liu (2015) 58
4.5.3 Configuration d'acquisition des données Fresnel 58
4.6 Conclusion 59

CHAPITRE 5 APPROCHES ITÉRATIVES PAR BLOCS POUR LA RÉSOLUTION

DE SYSTÈMES MULTIPLES 60
5.1 Les approches de résolution de systèmes par blocs 61
5.2 Adaptation de l'approche Block-BiCGStab pour une application en TMO . . 63
5.2.1 L'algorithme Block-BiCGStab 63
5.2.2 Réglage des paramètres du Block-BiCGStab pour une application en TMO 65
5.2.3 Résultats justifiant le réglage de l'initialisation 67
5.3 Amélioration du Block-BiCGStab pour une architecture multi-cœurs 74
5.3.1 L'approche Partial-Block BiCGStab 74
5.3.2 Réglages des paramètres de l'algorithme Partial-Block BiCGStab . . 75
5.3.3 Effet des réglages du Partial-Block BiCGStab sur la convergence . . . 78
5.4 Exploitation des approches par blocs dans le processus d'inversion 80
5.5 Évaluation des coûts de calcul des différents algorithmes pour la résolution de problèmes directs 82
5.5.1 Influence du contraste 82
5.5.2 Influence de la fréquence 83
5.5.3 Influence du nombre de systèmes à résoudre 84
5.5.4 Effet de l'architecture multi-cœurs 85
5.6 Évaluation des approches pour la reconstruction 86
5.6.1 Procédure de reconstruction 86
5.6.2 Résultats de reconstruction 87
5.6.3 Comparaison des coûts de reconstruction selon les approches 88
5.7 Conclusion 92

CHAPITRE 6 VALIDATION DE NOTRE ALGORITHME D'INVERSION SUR DONNÉES RÉELLES 94
6.1 Présentation de la base de données de l'Institut Fresnel 95
6.1.1 La configuration du montage d'acquisition 95
6.1.2 Les objets de la base 96
6.2 Mise en œuvre de notre algorithme de reconstruction 98
6.2.1 Réciprocité de Lorentz et polarisations utilisées 99
6.2.2 Volume d'étude et discrétisation 100
6.2.3 Procédure par saut de fréquence 100
6.3 Résultats de reconstruction 101
6.4 Comparaison des coûts de calcul des différentes approches 117
6.5 Conclusion 118

CHAPITRE 7 VERS UNE GESTION EFFICACE DES EFFETS DE LA TRONCATURE SUR LA COHÉRENCE ENTRE LE CRITÈRE ET LE GRADIENT . . . 120
7.1 Mise en évidence de l'effet de la troncature sur la convergence de l'algorithme d'optimisation . . . 121
7.2 Expression du champ total dans un sous-espace . . . 124
7.2.1 Expression du champ . . . 124
7.2.2 Choix d'un sous-espace . . . 125
7.3 Définition du nouveau critère et calcul du gradient . . . 126
7.3.1 Critère approché . . . 126
7.3.2 Gradient correspondant au critère approché . . . 126
7.3.3 Vérification du gradient pour **B** inversible . . . 127
7.3.4 Étude de la variation du critère selon la direction du gradient . . . 128
7.4 Conclusion . . . 129

CHAPITRE 8 CONCLUSION ET PERSPECTIVES . . . 132
8.1 Synthèse des travaux . . . 132
8.1.1 Réduction du nombre de systèmes linéaires . . . 132
8.1.2 Accélération des résolutions de systèmes . . . 133
8.1.3 Validation de l'algorithme de reconstruction sur des données expérimentales . . . 133
8.1.4 Cohérence entre le critère et le gradient . . . 134
8.2 Limitations des solutions proposées . . . 134
8.3 Améliorations futures et pistes de recherche . . . 135
8.3.1 Étude de la formulation cohérente du gradient . . . 135
8.3.2 Études complémentaires de l'approche Partial-Block BiCGStab . . . 136
8.3.3 Approches de parallélisation de l'algorithme Block-BiCGStab . . . 136
8.3.4 Comparer notre algorithme de reconstruction à d'autres algorithmes fréquents en TMO . . . 137
8.3.5 Élargir le champ des applications . . . 137

RÉFÉRENCES . . . 139

ANNEXES . . . 152
A.1 Discrétisation de l'équation du domaine . . . 152
A.2 Discrétisation de l'équation d'observation . . . 155
B.1 Expression algébrique du gradient . . . 156
B.2 Expression du jacobien . . . 157

C.1 Rappel du critère approché 159
C.2 Calcul du gradient cohérent 159

LISTE DES TABLEAUX

Tableau 2.1 Erreurs quadratiques relatives entre les mesures du champ diffracté obtenues par le modèle numérique et celles obtenues par la solution analytique de Mie. La sphère est de diamètre $\lambda/2$ et la discrétisation du domaine en $20 \times 20 \times 20$ voxels. 24

Tableau 2.2 Erreurs quadratiques relatives entre les mesures du champ diffracté obtenues par le modèle numérique et celles obtenues par la solution analytique. La sphère est de diamètre $\lambda/2$ et la discrétisation du domaine en $30 \times 30 \times 30$ voxels. 25

Tableau 2.3 Erreurs quadratiques relatives entre les mesures du champ diffracté obtenues par le modèle numérique et celles obtenues par la solution analytique. La sphère est de diamètre λ et la discrétisation du domaine en $40 \times 40 \times 40$ voxels. 25

Tableau 2.4 Erreurs quadratiques relatives entre les données réelles et simulées pour l'objet TwoCubes. 29

Tableau 5.1 Nombre d'itérations du Block-BiCGStab en fonction du niveau de bruit ajouté à l'initialisation. L'initialisation est choisie égale aux champs incidents. 73

Tableau 5.2 Nombre d'itérations du Block-BiCGStab en fonction du niveau de bruit ajouté à l'initialisation. L'initialisation est ici égale aux champs totaux calculés pour le contraste 0.99χ. 73

Tableau 5.3 Temps de calcul et nombre d'itérations pour la résolution de 160 problèmes directs avec BiCGStab et Block-BiCGStab. Différents contrastes de l'objet simulé sont testés. 83

Tableau 5.4 Temps de calcul et nombre d'itérations pour la résolution de 160 problèmes directs avec BiCGStab et Block-BiCGStab. Différentes fréquences d'ondes incidentes sont testées. 84

Tableau 5.5 Temps de calcul et nombre d'itérations pour la résolution de 160 problèmes directs avec BiCGStab et Block-BiCGStab. Différentes valeurs de N_S sont testées. 85

Tableau 5.6 Temps de calcul et nombre d'itérations pour la résolution de 160 problèmes directs avec BiCGStab et Partial-Block BiCGStab. Différentes architectures matérielles sont testées. 86

Tableau 5.7 Temps de reconstruction pour l'objet simulé avec résolutions des systèmes par BiCGStab et Block-BiCGStab ($N_c = 1$) ou Partial-Block BiCGStab ($N_c = 4$). 89
Tableau 6.1 Temps de reconstruction pour les objets de la base Fresnel. 118

LISTE DES FIGURES

Figure 1.1 Schéma de montage d'acquisition en tomographie micro-ondes 3
Figure 2.1 Schéma d'un montage d'acquisition en tomographie micro-ondes (Pastorino, 2010). 9
Figure 2.2 Schéma du montage pour la comparaison avec la solution analytique de Mie. 23
Figure 2.3 Mesures du champ diffracté par une sphère de diamètre $\lambda/2$. 26
Figure 2.4 Suite de la figure 2.3. Mesures du champ diffracté par une sphère de diamètre $\lambda/2$. 27
Figure 2.5 Représentation de l'objet TwoCubes de la base de données Fresnel. Figure tirée de Geffrin et Sabouroux (2009). 28
Figure 2.6 Comparaison des données réelles et simulées pour l'objet TwoCubes à 5 GHz. 30
Figure 2.7 Comparaison des données réelles et simulées pour l'objet TwoCubes à 8 GHz. 31
Figure 5.1 Schéma du montage d'acquisition pour l'objet simulé. 69
Figure 5.2 Représentation 3-D des parties réelle et imaginaire du contraste de l'objet simulé. 70
Figure 5.3 Évolution des résidus relatifs des 160 résolutions de systèmes par BiCG-Stab (à gauche) et par Block-BiCGStab (à droite) en fonction des itérations. Les initialisations des systèmes sont choisies égales aux champs incidents $\boldsymbol{e}_i^{\text{inc}}$. 71
Figure 5.4 Évolution des résidus relatifs des 160 résolutions de systèmes par BiCG-Stab (à gauche) et par Block-BiCGStab (à droite) en fonction des itérations. Les initialisations des systèmes sont choisies égales aux champs incidents $\boldsymbol{e}_i^{\text{inc}}$ avec une composante aléatoire de 50 dB. 72
Figure 5.5 Schéma de représentation du montage simulé pour l'application de l'algorithme Partial-Block BiCGStab où les sources sont partitionnées en $P = 4$ groupes. 77
Figure 5.6 Matrice de corrélation des champs incidents provenant des 160 sources du montage simulé 79
Figure 5.7 Temps de calcul pour la résolution des 160 problèmes directs de l'objet simulé par Partial-Block BiCGStab en fonction du nombre de groupes P. Quatre architectures multi-cœurs sont présentées : $N_c = 1, 2, 4$ et 8. 80

Figure 5.8 30 tranches horizontales de la partie réelle de l'objet exact. 89
Figure 5.9 30 tranches horizontales de la partie réelle de l'objet reconstruit. . . . 90
Figure 5.10 30 tranches horizontales de la partie imaginaire de l'objet exact. . . . 91
Figure 5.11 30 tranches horizontales de la partie imaginaire de l'objet reconstruit. 92
Figure 6.1 L'objet TwoCubes de la base Fresnel. 96
Figure 6.2 L'objet TwoSpheres de la base Fresnel. 97
Figure 6.3 L'objet Cylinder de la base Fresnel. 97
Figure 6.4 L'objet CubeSpheres de la base Fresnel. 98
Figure 6.5 L'objet IsocaSphere de la base Fresnel. 99
Figure 6.6 Reconstruction de l'objet TwoCubes avec procédure par saut de fréquence. 104
Figure 6.7 Isosurface de la solution de la reconstruction TwoCubes. 105
Figure 6.8 Initialisation de la reconstruction de l'objet TwoCubes pour la dernière fréquence (soit 8 GHz). 106
Figure 6.9 Reconstruction de l'objet TwoSpheres avec procédure par saut de fréquence. 107
Figure 6.10 Isosurface de la solution de la reconstruction TwoSpheres. 108
Figure 6.11 Isosurface de la solution de la reconstruction TwoSpheres pour un volume d'étude plus grand. 109
Figure 6.12 Reconstruction de l'objet Cylinder avec procédure par saut de fréquence. 111
Figure 6.13 Isosurface de la solution de la reconstruction Cylinder. 112
Figure 6.14 Reconstruction de l'objet CubeSpheres avec procédure par saut de fréquence. 113
Figure 6.15 Isosurface de la solution de la reconstruction CubeSpheres. 114
Figure 6.16 Reconstruction de l'objet IsocaSphere avec procédure par saut de fréquence. 115
Figure 6.17 Isosurface de la solution de la reconstruction IsocaSphere. 116
Figure 7.1 Variation du critère approché dans la direction du gradient approché. 123
Figure 7.2 Variation du critère approché dans la direction du gradient exact du critère approché. 129
Figure 7.3 Zoom de la figure 7.2. 130

LISTE DES SIGLES ET ABRÉVIATIONS

BiCGStab	BiConjugate Gradient Stabilized
CSI	Contrast Source Inversion
DBIM	Distorted Born Iterative Method
FFT	Fast Fourier Transform
GN	Gauss-Newton
LM	Levenberg-Marquardt
MoM	Method of Moments
RSB	Rapport Signal sur Bruit
TMO	Tomographie micro-ondes

LISTE DES ANNEXES

Annexe A DISCRÉTISATION DE L'ÉQUATION INTÉGRALE ET FORMULATION MATRICIELLE 152
Annexe B CALCUL DU GRADIENT DU CRITÈRE NON-LINÉAIRE 156
Annexe C CALCUL DU GRADIENT COHÉRENT AU CRITÈRE APPROCHÉ 159

CHAPITRE 1 INTRODUCTION GÉNÉRALE

Ce chapitre d'introduction présente brièvement au lecteur la tomographie micro-ondes ainsi que les motivations de ce travail. Les principales difficultés seront exposées afin de mettre l'accent sur les objectifs et contributions de cette thèse.

1.1 La tomographie micro-ondes (TMO)

Les champs électromagnétiques sont utilisés depuis plusieurs décennies pour l'imagerie de milieux inconnus. La tomodensitométrie (ou *CT-scan*) utilise des rayons X afin de mesurer leur absorption par des matériaux. L'imagerie par résonnance magnétique (IRM) est basée sur l'utilisation de radiofréquences pour mesurer le temps de relaxation des moments magnétiques nucléaires. L'imagerie par courants de Foucault utilise des ondes électromagnétiques de l'ordre de quelques centaines de kilohertz afin de détecter par exemple des défauts dans des matériaux conducteurs. La tomographie micro-ondes (TMO) fait partie de ces techniques basées sur l'exploitation de champs électromagnétiques : il s'agit d'une technique d'imagerie permettant de réaliser une cartographie d'un volume inconnu à partir d'une succession de mesures de champs électromagnétiques provenant de sources micro-ondes (cf. figure 1.1). La plage de fréquences concernée par les micro-ondes s'étend de 300 MHz à 300 GHz. L'imagerie d'objets à ces fréquences exploite le phénomène physique de diffraction de l'onde ; contrairement aux rayons X où la fréquence des ondes est très grande devant la taille des objets, les objets à reconstruire en imagerie micro-ondes sont généralement de l'ordre de grandeur de quelques longueurs d'onde. L'onde incidente est alors diffractée par l'objet. Il s'agit donc d'utiliser l'information fournie par les mesures du champ diffracté pour déterminer les propriétés diélectriques — permittivité diélectrique et conductivité — en tout point du volume inconnu.

La communauté scientifique s'est appropriée cette technologie dans une grande gamme d'applications dont :

- l'imagerie médicale : par exemple la détection du cancer du sein (Barrière, 2008), le suivi médical à la suite d'un accident vasculaire cérébral (Scapaticci et al., 2012) ;
- la géophysique : par exemple la détection d'objets enfouis (Abubakar, 2000), la recherche d'hydrocarbures ou de gaz (Forte et al., 2012) ;
- le contrôle non destructif d'objets non pénétrables : par exemple le contrôle de l'humidité de silos à grain (Asefi et al., 2015), la qualité de structures en béton (Maaref et al., 2009).

Bien que concurrente d'autres modalités d'imagerie, la tomographie micro-ondes détient certains avantages, notamment le faible coût d'un système d'acquisition micro-ondes et une nocivité réduite pour l'homme (pour les applications biomédicales). En revanche, certaines difficultés intrinsèques au problème de reconstruction en TMO freinent le développement industriel de cette technologie. La principale difficulté est un coût de calcul excessif, notamment en 3-D, rendant difficile l'utilisation pratique d'une telle technique. Nous verrons dans la suite de ce chapitre quelles sont ces difficultés et nous proposerons dans ce manuscrit quelques outils permettant de réduire le coût de calcul.

1.2 Méthodes de reconstruction en TMO

La reconstruction d'images en tomographie micro-ondes est considérée ici comme la résolution d'un problème inverse où l'on cherche à minimiser l'erreur entre un modèle numérique représentant la diffraction d'une onde incidente par un objet de propriétés inconnues et les données mesurées au niveau des récepteurs. Ici, l'inconnue du problème est le contraste en tout point de l'objet, fonction qui dépend des propriétés diélectriques de cet objet. Formaliser le problème inverse comme la minimisation d'un critère d'adéquation aux données est un choix fréquent dans la littérature (Pastorino, 2010). Il est alors nécessaire de définir un modèle direct qui permet d'établir une relation entre les propriétés diélectriques du volume (ici, supposées connues), l'onde électromagnétique incidente et le champ diffracté au niveau des récepteurs. Étant donnée la complexité du problème, il n'existe pas de solution analytique au phénomène de diffraction d'onde par des objets quelconques. Nous avons alors fait le choix d'utiliser un modèle numérique basé sur la méthode des moments (Harrington et Harrington, 1996). Cette approche de discrétisation s'appuie sur une formulation intégrale de l'équation d'onde. Le modèle numérique obtenu conduit à un ensemble de deux équations algébriques. La première équation, appelée équation d'observation, exprime le champ diffracté au niveau des récepteurs en fonction du contraste et du champ total à l'intérieur du volume d'étude. Ici, le champ total, qui correspond à la somme du champ incident et du champ diffracté, est inconnu dans le volume d'étude. La seconde équation, nommée équation du domaine, permet alors d'exprimer le champ total dans le volume en fonction du contraste et de l'onde incidente. L'équation du domaine n'est pas linéaire et cette propriété représente la principale difficulté du problème inverse en TMO.

En plus de la non-linéarité du problème inverse, celui-ci est dit mal posé au sens de Hadamard. Plus précisément, comme le nombre d'inconnues est généralement plus grand que le nombre de données disponibles, l'unicité de la solution n'est pas assurée. Il est alors fréquent de contraindre l'espace des solutions admissibles en considérant un terme de régularisation qui

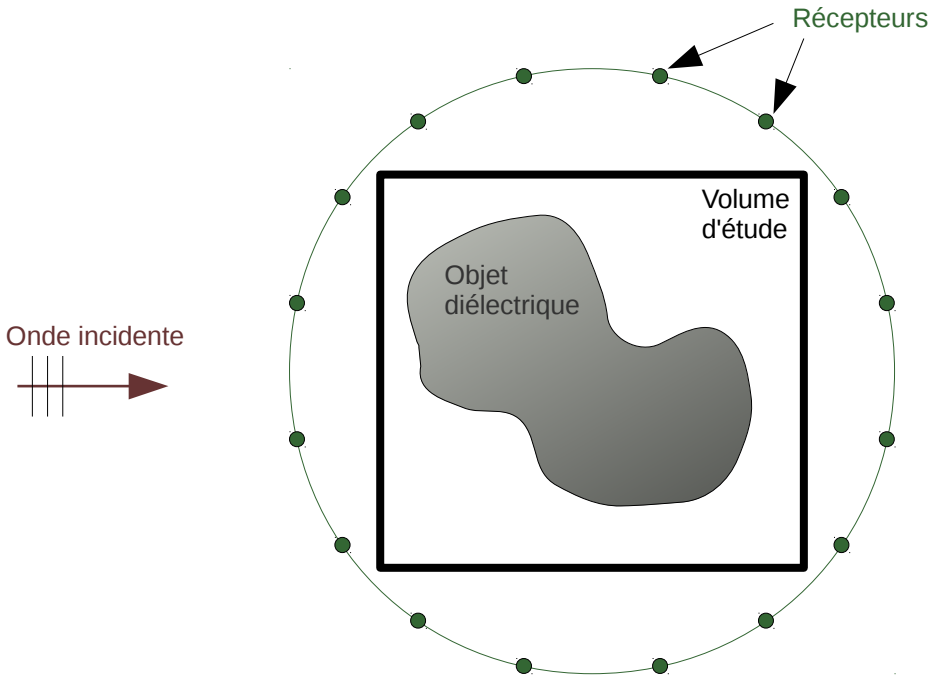


Figure 1.1 Schéma de montage d'acquisition en tomographie micro-ondes. Un objet diélectrique (inconnu) est contenu dans un volume d'étude que l'on cherche à reconstruire. Une succession d'ondes incidentes éclairent l'objet qui diffracte ces ondes. Le champ diffracté est mesuré au niveau des récepteurs.

permet de stabiliser la solution du problème et donc de converger vers un optimum acceptable.

Afin de contourner l'aspect non-linéaire du problème, plusieurs approximations linéaires du modèle ont été utilisées dans le but de réduire le coût de calcul et de simplifier la résolution du problème inverse. C'est le cas par exemple de l'approximation de Born qui considère que le champ total peut être approché au premier ordre par le champ incident, ce qui revient à linéariser l'équation du domaine. En revanche, le domaine de validité de cette approximation ne permet des reconstructions acceptables que pour des objets de petite taille électrique, c'est-à-dire de petite taille par rapport à la longueur d'onde, et faiblement contrastés. Des méthodes itératives basées sur des approximations linéaires successives ont également vu le jour, comme la méthode itérative de Born (Pastorino, 2010).

Avec le développement d'ordinateurs de plus en plus puissants, les chercheurs se sont concentrés sur la résolution du problème directemetn sous sa forme non-linéaire, c'est-à-dire avec la prise en compte des équations d'observation et du domaine. Le problème inverse peut alors être considéré comme la minimisation d'un terme d'adéquation aux données via l'équation

d'observation et une contrainte sur les champs totaux via l'équation du domaine. Une première approche fréquente consiste à relâcher la contrainte et à considérer la fonction coût comme l'assemblage d'un terme d'adéquation aux données et d'un terme de relaxation de la contrainte. Ce type d'approche prend la dénomination "modified gradient method" (MGM) ou la variante "Contrast Source Inversion" (CSI) (Kleinman et van den Berg, 1992; van den Berg et Kleinman, 1995, 1997; van den Berg et Abubakar, 2001). Avec cette formulation, les champs totaux deviennent des inconnues supplémentaires de la fonction coût. Cela conduit à un problème bilinéaire en fonction du contraste et des champs totaux soit un nombre d'inconnues très conséquent, notamment en 3-D. En revanche, la bilinéarité permet de contourner le calcul non-linéaire des champs totaux à travers l'équation du domaine.

Nous avons fait le choix de travailler avec une autre formulation très fréquente pour traiter le problème ; nous l'appellerons la formulation non-linéaire. L'idée est de combiner les deux équations du modèle afin d'obtenir une seule équation non-linéaire. De cette manière, la contrainte sur les champs totaux est respectée de façon exacte et le nombre d'inconnues est limité aux valeurs du contraste. En revanche, l'inconvénient majeur de ce type d'approche est que l'évaluation du critère requiert le calcul des champs totaux, opération coûteuse en grande dimension. Les champs totaux sont obtenus à partir de l'équation du domaine et nécessite d'inverser des systèmes linéaires de grande taille. Ces résolutions de systèmes sont au centre de l'attention dans cette thèse et la section 1.3 détaille les difficultés engendrées par ces résolutions.

La minimisation du critère régularisé peut être effectuée avec un grand nombre d'algorithmes d'optimisation. Comme le critère est non-linéaire, il pourrait sembler naturel d'utiliser des méthodes d'optimisation globale afin d'obtenir le minimum global de la fonction coût (Pastorino, 2007). En revanche, ces méthodes sont très coûteuses notamment à cause du grand nombre d'inconnues en 3-D. Les méthodes d'optimisation locale sont au contraire bien moins coûteuses mais elles ne garantissent la convergence que vers un minimum local. Malgré tout, elles sont très utilisées en pratique car elles permettent tout de même d'obtenir une solution convenable pour la majorité des problèmes traités en TMO. Parmi les algorithmes d'optimisation locale utilisés en imagerie micro-ondes, on peut citer l'algorithme du gradient conjugué non-linéaire (Chaumet et Belkebir, 2009), la méthode de Gauss-Newton (De Zaeytijd et al., 2007) ou encore l'algorithme de quasi-Newton BFGS (Hu et al., 2006).

La convergence des méthodes locales vers un minimum donné dépend fortement de l'initialisation de l'algorithme : l'algorithme converge vers un certain minimum si l'initialisation appartient au bassin d'attraction de ce minimum. L'initialisation est donc un paramètre très important dans la reconstruction. De plus, pour des objets de grande taille électrique et/ou

fortement contrastés, la non-linéarité du problème est d'autant plus importante. Pour initialiser correctement l'algorithme, plusieurs solutions existent comme le choix d'initialiser avec la solution de Born. Une solution efficace que nous avons utilisée consiste à réaliser l'acquisition de mesures à basses fréquences (c'est-à-dire avec une non-linéarité réduite) et d'appliquer une stratégie par saut de fréquence.

1.3 Difficultés et objectifs de recherche

Dans le contexte de l'inversion, les algorithmes d'optimisation itératifs nécessitent généralement l'évaluation du gradient et du critère à chaque itération afin d'estimer une direction et un pas de descente. Dans la formulation non-linéaire, le calcul du critère requiert la résolution de systèmes linéaires pour calculer les champs totaux. De même, le calcul du gradient nécessite la résolution de systèmes supplémentaires. À chaque itération de la reconstruction, un grand nombre de systèmes linéaires est donc à résoudre.

La taille de ces systèmes est conséquente. En effet, pour un volume 3-D discrétisé en N voxels, les systèmes linéaires à résoudre sont de taille $3N \times 3N$. Par exemple, un volume de $30 \times 30 \times 30$ voxels mène à un système de taille 81000×81000. Tout d'abord, une formulation implicite de la matrice opérateur est utilisée afin de réduire la mémoire requise au stockage des quantités matricielles. De plus, étant donnée la taille importante des systèmes, l'utilisation de méthodes itératives est nécessaire pour les résoudre. On peut citer une des méthodes les plus fréquentes pour ces résolutions de systèmes en TMO, la version stabilisée du gradient biconjugué (BiCGStab).

Pour réduire le coût de calcul de ces opérations, les résolutions par méthode itérative sont tronquées. Malgré cette troncature, ces opérations représentent tout de même la majorité du temps de calcul de l'inversion. Le coût de calcul extrêmement important des reconstructions 3-D en imagerie micro-ondes freine malheureusement son développement dans des applications industrielles ou médicales. De plus, la troncature des résolutions des systèmes introduit des erreurs d'approximation qui peuvent entraîner des difficultés de convergence de l'algorithme d'optimisation. Cela peut rendre l'inversion difficile et aboutir à un arrêt prématuré de l'algorithme de reconstruction.

L'objectif global de cette thèse est alors d'accélérer les méthodes de reconstruction 3-D en tomographie micro-ondes. Le travail s'est focalisé sur deux aspects : les résolutions de systèmes linéaires puisqu'il s'agit des opérations les plus chronophages dans l'inversion, et la troncature des systèmes.

Le travail présenté dans ce manuscrit est divisé en trois objectifs :

— le premier est d'optimiser le nombre de systèmes à résoudre dans le calcul du critère et du gradient ;
— le deuxième est d'accélérer la résolution de ces nombreux systèmes linéaires ;
— le troisième est d'étudier l'effet de la troncature des méthodes itératives de résolution des systèmes et d'améliorer la convergence de l'algorithme de reconstruction.

Ces trois objectifs sont détaillés ci-après.

1.3.1 Optimiser le nombre de systèmes linéaires

L'évaluation du critère requiert le calcul des champs totaux. Le nombre de ces champs dépend du nombre d'illuminations utilisé lors de l'acquisition des données. De même, le calcul du gradient nécessite la résolution de systèmes supplémentaires. Nous proposons une procédure afin d'optimiser le nombre de ces systèmes et donc de réduire le coût de calcul du critère et du gradient. Pour cela, nous nous appuyons sur deux techniques connues en imagerie micro-ondes mais aussi dans d'autres types d'application : le principe de réciprocité de Lorentz et la formulation de l'état adjoint.

1.3.2 Accélérer la résolution des systèmes linéaires

Les nombreux systèmes linéaires à résoudre à chaque évaluation du critère et du gradient ont la particularité de faire intervenir la même matrice opérateur. La communauté des mathématiques appliquées a développé depuis quelques décennies une série d'algorithmes par blocs permettant de résoudre plusieurs systèmes linéaires avec une matrice commune. Nous proposons alors d'adapter ce type d'approche, en particulier l'algorithme Block-BiCGStab (El Guennouni et al., 2003), à une utilisation en tomographie micro-ondes.

1.3.3 Étudier l'effet de la troncature et améliorer la convergence

L'idée de cette contribution est tout d'abord d'étudier l'effet de la troncature des résolutions de systèmes linéaires sur le comportement de l'algorithme de reconstruction. Nous proposons ensuite une nouvelle formulation du critère et du gradient permettant d'assurer la cohérence entre le critère et le gradient, malgré les approximations dues à la troncature des méthodes itératives d'inversion de systèmes. Comme le calcul du critère et du gradient nécessite des résolutions tronquées, on propose d'exprimer ces troncatures sous la forme de projections dans le sous-espace de Krylov engendré par les directions de descente de l'algorithme BiCGStab que nous exploitons pour assurer la cohérence entre le critère et le gradient. L'objectif est alors d'améliorer la convergence de l'algorithme d'optimisation, malgré une troncature des

résolutions de systèmes linéaires.

Ce dernier objectif n'a pas été autant développé que les deux premiers. Aussi, cette thèse se restreint à exposer une démarche méthodologique et présente quelques perspectives pour de futurs travaux.

1.4 Plan de la thèse

Le chapitre 2 présente le modèle numérique développé à des fins de reconstruction. Nous décrivons les formulations intégrales en continu et leur discrétisation après application de la méthode des moments. Nous mettons ensuite l'accent sur le calcul des champs totaux, opération au centre des contributions de cette thèse. Enfin, comme le modèle développé est nouveau pour les équipes engagées dans ce projet, nous validons le modèle numérique i) sur un objet simple pour lequel nous possédons une solution analytique, ii) en le comparant à des données expérimentales.

Le chapitre 3 présente le problème inverse vu comme la minimisation d'un critère régularisé. Après avoir comparé les approches bilinéaire et non-linéaire, nous nous concentrons sur la définition du critère non-linéaire régularisé et du calcul de son gradient.

Le chapitre 4 étudie la possibilité d'optimiser le nombre de systèmes à résoudre pour le calcul du critère et du gradient. Le théorème de réciprocité de Lorentz y est détaillé et une comparaison sur la mise en œuvre du calcul du gradient y est présentée. L'association de ces deux techniques aboutit à la proposition d'une procédure permettant de réduire le nombre de systèmes linéaires en fonction de la configuration d'acquisition.

Après avoir présenté l'approche Block-BiCGStab, algorithme itératif de résolution de systèmes linéaires multiples, le chapitre 5 s'attaque à l'utilisation de cette approche dans le contexte de TMO afin d'accélérer ces résolutions de systèmes. Nous proposons des adaptations efficaces de cette méthode pour la TMO et une généralisation de l'algorithme par blocs afin d'optimiser ses performances pour une architecture d'ordinateurs multi-cœurs. Enfin, des résultats comparatifs sont présentés aussi bien sur des calculs de champs totaux (modèle direct) que sur des reconstructions (problème inverse).

Afin de valider notre algorithme de reconstruction sur des données réelles, nous présentons dans le chapitre 6 un ensemble de données expérimentales acquises par l'Institut Fresnel et libre d'accès. Nous proposons de comparer nos résultats de reconstruction sur ces objets réels à ceux présentés dans la littérature. Enfin, nous mettons en avant nos contributions des chapitres précédents sur ces reconstructions pour évaluer le gain en coût de calcul.

Le chapitre 7 étudie l'effet de la troncature des méthodes itératives de résolution des systèmes

linéaires sur la convergence de l'algorithme de reconstruction. Une nouvelle formulation du critère et du gradient permettant d'assurer la cohérence est proposée afin d'améliorer la convergence de l'algorithme d'inversion.

Le chapitre 8 conclut ce manuscrit et propose des perspectives de poursuite des travaux de cette thèse.

1.5 Cadre du travail

Ce travail a été réalisé dans le cadre d'une thèse en cotutelle entre l'École Polytechnique de Montréal (Canada) et l'École Centrale de Nantes (France). Cette collaboration n'est pas nouvelle : depuis 2005, trois thèses sur la tomographie de diffraction ont été réalisées entre le département de Génie Électrique (Montréal) et l'équipe "Analyse et Décision en Traitement du Signal et de l'Image" (ADTSI, Nantes) (Barrière, 2008; Trillon, 2010; Vautrin, 2011). Parmi ces trois projets, un seul a concerné l'imagerie micro-ondes : Barrière (2008) a travaillé sur l'accélération d'algorithmes 2-D pour la détection du cancer du sein par TMO. Cette thèse se place dans la pousuite de ce travail à travers des algorithmes de reconstruction efficaces permettant de passer à l'échelle 3-D.

Le problème de TMO mêle aussi bien des notions d'électromagnétisme que de traitement de signal et de problèmes inverses. L'intérêt de cette cotutelle est de profiter des compétences des équipes dans les deux domaines afin d'avoir une vision globale du problème.

CHAPITRE 2 LE MODÈLE DIRECT

La tomographie micro-ondes (TMO) s'appuie sur la diffraction des ondes électromagnétiques à travers un objet inconnu afin d'estimer les propriétés diélectriques de celui-ci. Pour cela il est nécessaire de définir un modèle liant les propriétés de l'objet à l'onde diffractée. Ce modèle est appelé *modèle direct*. La figure 2.1 présente schématiquement un montage d'acquisition en tomographie micro-ondes. Le modèle direct consiste à exprimer le champ électromagnétique diffracté aux points de mesure (récepteurs) en fonction de l'onde incidente et de l'objet diélectrique. Dans la littérature, l'expression *problème direct* est fréquemment utilisée pour représenter le calcul du champ total. Nous verrons ce calcul plus en détail dans la section 2.3.

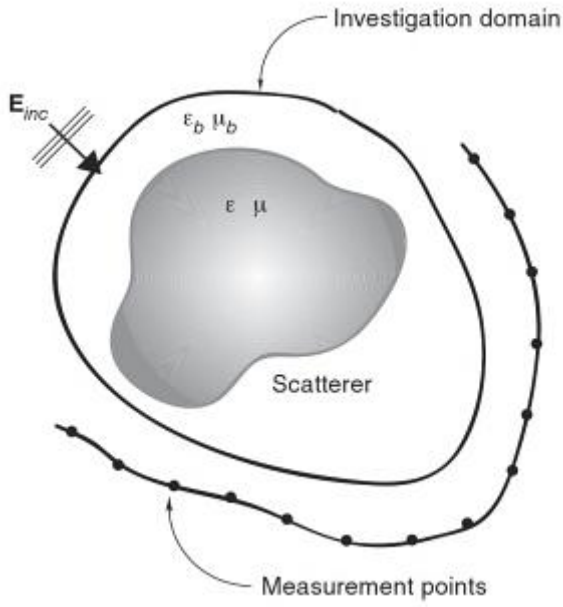


Figure 2.1 Schéma d'un montage d'acquisition en tomographie micro-ondes (Pastorino, 2010). Une onde incidente éclaire l'objet diffractant (en gris) plongé dans un volume d'étude. Le champ diffracté par l'objet est mesuré en différents points de mesure

Ce chapitre se décompose en quatre sections.

— La section 2.1 présente le modèle direct sous la forme continue et nous justifions les hypothèses prises en compte dans ce modèle.

— La section 2.2 décrit brièvement les différentes méthodes numériques utilisées en tomographie micro-ondes pour discrétiser le problème. Nous nous concentrerons sur la méthode retenue pour notre problème, soit la méthode des moments.

— Nous verrons dans la section 2.3 comment la résolution du problème direct est mise

en œuvre en pratique ; ce point est au cœur des contributions de cette thèse.
— Enfin, nous validerons notre modèle numérique dans la section 2.4 en le comparant : i) à la solution analytique de Mie qui est une solution exacte de la diffraction d'une onde plane par une sphère, puis ii) à un ensemble de mesures expérimentales acquises au sein de l'Institut Fresnel de Marseille où l'objet éclairé est connu.

Dans ce document, les notations suivantes seront adoptées :
— notation classique pour les scalaires (exemples : x, ϵ)
— notation en gras et italique pour les vecteurs (exemples : $\boldsymbol{r}$, $\boldsymbol{E}$)
— notation en gras et écriture romaine pour les matrices (exemples : $\mathbf{G}$, $\mathbf{E}$)

2.1 Modélisation physique de la diffraction d'ondes électromagnétiques

2.1.1 Équation différentielle régissant la propagation d'une onde électromagnétique

Les équations de Maxwell régissent le comportement des ondes électromagnétiques et l'équation d'onde est à la base des modèles analytiques généralement utilisés en tomographie de diffraction. L'équation d'onde, obtenue en combinant les équations de Maxwell, permet de lier le champ électromagnétique aux propriétés diélectriques du milieu qu'il traverse. Dans la suite de ce manuscrit, l'équation d'onde dans le domaine fréquentiel sera utilisée. Ce choix s'explique par le fait qu'en TMO, il est avantageux de travailler à une fréquence fixe. En effet, cela permet de simplifier l'expression du modèle direct. De plus, comme les propriétés diélectriques des objets dépendent généralement de la fréquence de l'onde d'excitation, se placer dans le domaine harmonique permet de ne considérer les propriétés des objets qu'à la fréquence d'étude. Dans le cadre de l'inversion, cela permet donc de réduire le nombre d'inconnues. Bien que certaines méthodes de reconstruction utilisent une approche temporelle (Li et al., 2005; Takenaka et al., 2003; Ali et Moghaddam, 2010), les méthodes s'appuyant sur une formulation fréquentielle sont bien plus répandues. On supposera une fréquence angulaire ω et une dépendance temporelle en $\exp(-j\omega t)$. Les champs vectoriels tridimensionnels seront exprimés sous la forme $\overrightarrow{\boldsymbol{U}}(\boldsymbol{r},t) = \mathrm{Re}\left\{\overrightarrow{\boldsymbol{U}}(\boldsymbol{r})\exp(-j\omega t)\right\}$ où $\boldsymbol{r}$ représente un point de l'espace. L'équation d'onde pour le champ électrique en régime harmonique, dite équation d'Helmholtz, s'écrit comme suit :

$$\nabla \times \nabla \times \overrightarrow{\boldsymbol{E}}(\boldsymbol{r}) - \omega^2 \mu\epsilon \overrightarrow{\boldsymbol{E}}(\boldsymbol{r}) = j\omega\mu \overrightarrow{\boldsymbol{J}_S}(\boldsymbol{r}) \tag{2.1}$$

où $\overrightarrow{\boldsymbol{E}}(\boldsymbol{r})$ est le champ électrique au point $\boldsymbol{r}$ généré par une densité de courant $\overrightarrow{\boldsymbol{J}_S}$. Les constantes physiques ϵ et μ sont, respectivement, la permittivité diélectrique et la perméabilité magnétique du milieu homogène. Le terme $\nabla\times$ représente l'opérateur rotationnel.

La plupart des milieux et objets étant amagnétiques, on supposera dans la suite que la perméabilité magnétique est constante et égale à la perméabilité du vide : $\mu = \mu_0 = 4\pi 10^{-7}$ kg.m.A^{-2}.s^{-2}. La permittivité diélectrique ϵ, qui représente ici la permittivité complexe, peut être décomposée en un terme réel représentant la permittivité relative ϵ_r et un terme imaginaire proportionnel à la conductivité du milieu σ. Elle est exprimée sous la forme :

$$\epsilon = \epsilon_0 \left(\epsilon_r + j\frac{\sigma}{\omega\epsilon_0}\right) \tag{2.2}$$

où ϵ_0 est la permittivité du vide.

L'équation (2.1) ne fait intervenir que le champ électrique, bien qu'il serait possible d'établir l'équation d'onde avec le champ magnétique. La formulation à partir du champ magnétique n'entraîne pas de différences majeures dans les formulations analytiques que nous verrons plus tard. Le choix du champ électrique est le plus courant et la plupart des montages d'acquisition en TMO utilise des mesures électriques. Certains auteurs ont tout de même travaillé à partir de mesures magnétiques (Abubakar et van den Berg, 2004).

2.1.2 Solution intégrale de l'équation d'onde dans un milieu homogène

Il existe de nombreux modèles numériques en tomographie micro-ondes. Nous verrons dans la section suivante que la discrétisation d'un modèle continu s'appuie sur une formulation du problème physique choisie en conséquence. En TMO, deux grandes familles de méthodes cohabitent : les méthodes dites différentielles et les méthodes dites intégrales. Les méthodes différentielles sont basées sur la discrétisation de l'équation différentielle d'onde (2.1) et les méthodes intégrales s'appuient sur la solution intégrale de l'équation d'onde, autrement appelée *Eletric Field Integral Equation* (EFIE). Ces deux classes de méthodes ont chacune leurs avantages et inconvénients que nous détaillerons dans la section suivante. Dans cette thèse, nous avons fait le choix d'une approche intégrale ; nous justifierons ce choix plus tard. C'est la raison pour laquelle nous présentons ici la formulation intégrale utilisée en TMO.

Dans un premier temps, nous allons exprimer la solution intégrale à l'équation d'onde (2.1). En supposant un milieu homogène infini, le champ électrique $\overrightarrow{\boldsymbol{E}}$ s'écrit sous la forme intégrale (Pastorino, 2010)

$$\overrightarrow{\boldsymbol{E}}(\boldsymbol{r}) = j\omega\mu \int_{V_S} \overline{\overline{\mathbf{G}}}(\boldsymbol{r}, \boldsymbol{r}')\overrightarrow{\boldsymbol{J}_S}(\boldsymbol{r}')d\boldsymbol{r}'. \tag{2.3}$$

Le volume V_S est le volume contenant la source définie par la densité de courant $\overrightarrow{\boldsymbol{J}_S}$. Le tenseur $\overline{\overline{\mathbf{G}}}$ d'ordre 2 représente le tenseur dyadique de Green homogène

$$\overline{\overline{\mathbf{G}}} = (\mathbf{I} + \frac{1}{k^2}\nabla\nabla)\, g(\boldsymbol{r}, \boldsymbol{r}'). \tag{2.4}$$

On définit la fonction de Green scalaire $g(\boldsymbol{r}, \boldsymbol{r}')$

$$g(\boldsymbol{r}, \boldsymbol{r}') = \frac{\exp(jk\|\boldsymbol{r} - \boldsymbol{r}'\|)}{4\pi\|\boldsymbol{r} - \boldsymbol{r}'\|}, \tag{2.5}$$

k étant le nombre d'onde : $k = \omega^2\mu\epsilon$. Notons que le tenseur de Green est un noyau de convolution symétrique. L'équation (2.3) représente alors le produit de convolution de la densité de courant $\overrightarrow{\boldsymbol{J}_S}$ par le noyau de convolution $\overline{\overline{\mathbf{G}}}$.

2.1.3 Équations intégrales en TMO

Puisque l'imagerie micro-ondes a pour objet la reconstruction de milieux électromagnétiquement non homogènes, il est nécessaire d'exprimer l'équation intégrale (2.3) avec une présence d'inhomogénéités. Supposons un milieu de référence homogène défini par sa permittivité ϵ_b et sa perméabilité μ_0. Une source émet un champ incident $\overrightarrow{\boldsymbol{E}}^{\text{inc}}$ de pulsation ω. On définit le nombre d'onde dans le milieu homogène $k_b = \sqrt{\omega^2\mu_0\epsilon_b}$. Un objet diélectrique de permittivité $\epsilon(\boldsymbol{r})$ est inclus dans un volume V. Le champ électrique total $\overrightarrow{\boldsymbol{E}}^{\text{tot}}$ est décomposé en la somme du champ incident $\overrightarrow{\boldsymbol{E}}^{\text{inc}}$ correspondant au champ généré par la source dans le milieu de référence (sans objet) et d'un champ diffracté $\overrightarrow{\boldsymbol{E}}^{\text{scat}}$ provenant de la diffraction de l'onde incidente par cet objet diélectrique :

$$\overrightarrow{\boldsymbol{E}}^{\text{tot}} = \overrightarrow{\boldsymbol{E}}^{\text{inc}} + \overrightarrow{\boldsymbol{E}}^{\text{scat}}. \tag{2.6}$$

Principe des volumes équivalents

D'après le principe des volumes équivalents (Pastorino, 2010), le champ diffracté $\overrightarrow{\boldsymbol{E}}^{\text{scat}}$ est considéré comme étant équivalent à un champ généré par une densité de courant dans le volume V de l'objet et ce dans le milieu de référence. Cela revient à remplacer l'objet diélectrique par un ensemble de densités de courant suivant la relation

$$\overrightarrow{\boldsymbol{J}_{\text{eq}}}(\boldsymbol{r}) = -j\omega(\epsilon(\boldsymbol{r}) - \epsilon_b)\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}). \tag{2.7}$$

Dans le cas particulier où l'objet a la même permittivité que le fond ($\epsilon(\boldsymbol{r}) = \epsilon_b$), on obtient $\overrightarrow{\boldsymbol{J}_{\text{eq}}} = \overrightarrow{\mathbf{0}}$ et donc le champ diffracté est bien nul. Dans le cas d'une inhomogénéité ($\epsilon(\boldsymbol{r}) \neq \epsilon_b$),

en reprenant la solution intégrale (2.3), on en déduit que

$$\begin{aligned}\overrightarrow{\boldsymbol{E}}^{\text{scat}}(\boldsymbol{r}) &= j\omega\mu_0 \int_V \overline{\overline{\mathbf{G}}}(\boldsymbol{r},\boldsymbol{r}')\overrightarrow{\boldsymbol{J}_{\text{eq}}}(\boldsymbol{r}')d\boldsymbol{r}' \\ &= k_b^2 \int_V \overline{\overline{\mathbf{G}}}(\boldsymbol{r},\boldsymbol{r}')\,\chi(\boldsymbol{r}')\,\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}')d\boldsymbol{r}'. \end{aligned} \tag{2.8}$$

On définit ici le contraste en tout point de l'objet χ

$$\chi(\boldsymbol{r}) = \frac{\epsilon(\boldsymbol{r}) - \epsilon_b}{\epsilon_b}. \tag{2.9}$$

On peut donc exprimer le champ total $\overrightarrow{\boldsymbol{E}}^{\text{tot}}$ en fonction du champ incident $\overrightarrow{\boldsymbol{E}}^{\text{inc}}$ et du contraste de l'objet $\chi(\boldsymbol{r})$:

$$\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}) = \overrightarrow{\boldsymbol{E}}^{\text{inc}}(\boldsymbol{r}) + k_b^2 \int_V \overline{\overline{\mathbf{G}}}(\boldsymbol{r},\boldsymbol{r}')\,\chi(\boldsymbol{r}')\,\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}')d\boldsymbol{r}'. \tag{2.10}$$

Équations du domaine et d'observation en TMO

Pour définir le modèle direct, on cherche à modéliser le champ diffracté au niveau des antennes réceptrices placées en dehors du volume V. Supposons un récepteur placé en $\boldsymbol{r}^R$. À partir de l'équation (2.8), on définit le champ mesuré au récepteur par

$$\overrightarrow{\boldsymbol{E}}^{\text{scat}}(\boldsymbol{r}^R) = k_b^2 \int_V \overline{\overline{\mathbf{G}}}(\boldsymbol{r}^R,\boldsymbol{r}')\,\chi(\boldsymbol{r}')\,\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}')d\boldsymbol{r}'\ ,\quad \forall \boldsymbol{r}^R \notin V. \tag{2.11}$$

Cette équation est appelée *équation d'observation* puisqu'elle fait le lien entre les observations du champ diffracté au niveau des récepteurs et le contraste. Le champ total $\overrightarrow{\boldsymbol{E}}^{\text{tot}}$ dans le domaine V est en revanche inconnu. On définit alors l'*équation du domaine*

$$\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}) = \overrightarrow{\boldsymbol{E}}^{\text{inc}}(\boldsymbol{r}) + k_b^2 \int_V \overline{\overline{\mathbf{G}}}(\boldsymbol{r},\boldsymbol{r}')\,\chi(\boldsymbol{r}')\,\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}')d\boldsymbol{r}'\ ,\quad \forall \boldsymbol{r} \in V. \tag{2.12}$$

Notons que cette équation est implicite puisque le champ total intervient aussi bien dans le terme de gauche que dans l'intégrale. Les équations d'observation (2.11) et du domaine (2.12) forment le modèle direct continu permettant de modéliser le champ diffracté $\overrightarrow{\boldsymbol{E}}^{\text{scat}}(\boldsymbol{r}^R)$ aux récepteurs en fonction du contraste de l'objet χ et du champ incident $\overrightarrow{\boldsymbol{E}}^{\text{inc}}$. Ces deux équations intégrales sont à l'origine des méthodes de discrétisation intégrales.

2.2 Discrétisation du modèle physique et formulation matricielle

Les phénomènes électromagnétiques sont, sauf cas particuliers, trop complexes pour calculer une solution analytique à partir des équations continues. On entend par *cas particulier* des configurations simples comme par exemple la solution analytique de Mie qui permet de calculer la diffraction d'une onde plane par une sphère homogène, ou comme la diffraction d'une onde plane par un cylindre homogène infini. Hormis ces exemples spécifiques, la complexité des problèmes oblige à discrétiser les quantités impliquées afin de résoudre de façon numérique ces équations.

Dans cette section, nous verrons les différentes méthodes numériques utilisées en tomographie micro-ondes. Nous détaillerons ensuite la méthode choisie, la méthode des moments, qui nous mènera à la formulation matricielle du modèle direct.

2.2.1 État de l'art des différentes approches de discrétisation des équations électromagnétiques

Le choix de la méthode numérique pour modéliser les phénomènes électromagnétiques est très important et dépend du problème traité. Un mauvais choix peut conduire à des résultats erronés ou à des coûts de calcul trop importants. De nombreuses méthodes numériques ont été développées pour modéliser une grande variété de problèmes électromagnétiques. Cette science a pris le nom de *Computational Electromagnetics*. On peut citer par exemple l'optique ondulatoire (*Physical Optics*, Akhmanov et Nikitin (1997)) ou l'approche UTD (*Uniform Theory of Diffraction*, Kouyoumjian et Pathak (1974)) qui sont des modèles valides pour de très faibles longueurs d'ondes par rapport aux objets. En tomographie micro-ondes, les objets étudiés sont généralement de taille de l'ordre de grandeur de quelques longueurs d'onde. Comme vu précédemment, deux familles de méthodes numériques sont très fréquentes en TMO : les méthodes différentielles et les méthodes intégrales.

Ces deux types d'approche s'opposent sur un certain nombre de points que nous allons détailler :

— Le milieu de référence. Comme nous l'avons vu précédemment, on fait l'hypothèse dans la formulation intégrale que le milieu de référence est homogène (de nombre d'onde k_b). La solution intégrale fait alors intervenir le tenseur de Green homogène. Avoir une configuration où le milieu de référence n'est pas homogène rend le problème bien plus complexe car cela requiert le calcul d'un tenseur de Green inhomogène qui perd les propriétés utiles (symétrie, convolution) du tenseur homogène. En revanche,

ce problème ne se pose pas pour les approches différentielles qui peuvent prendre en compte un milieu de référence hétérogène.

— Les conditions limites. La formulation intégrale fait l'hypothèse d'un milieu infini ce qui suppose que le milieu est ouvert : l'énergie des sources est supposée se disperser à l'infini (condition de rayonnement de Sommerfeld). Les méthodes intégrales sont donc bien adaptées pour les problèmes ouverts mais ne conviennent pas à des problèmes fermés comme des cuves métalliques par exemple. Au contraire, les méthodes différentielles prennent facilement en compte des contraintes de bords pour des milieux fermés. En revanche, considérer des milieux ouverts nécessitent l'introduction de conditions limites d'absorption qui compliquent le problème (Berenger, 1994). En conclusion, les méthodes intégrales sont mieux adaptées aux milieux infinis, tandis que les méthodes différentielles sont mieux adaptées aux milieux fermés.

— Le volume de discrétisation. Les approches intégrales considèrent uniquement le domaine de l'objet pour la discrétisation : toute zone où le contraste est nul ne nécessite pas d'être prise en compte. Cela se remarque dans les équations (2.11) et (2.12) où l'intégrale n'agit que sur le volume V où le contraste est non nul. Ce volume à discrétiser est donc restreint aux inhomogénéités dans l'espace. Au contraire, les approches différentielles nécessitent de prendre en compte aussi bien le volume des objets que les sources et les récepteurs. Cela revient en général à discrétiser un volume bien plus conséquent contenant tout le dispositif d'acquisition. Pour des problèmes 3-D de grande taille, le nombre de voxels peut rapidement exploser avec les méthodes différentielles.

— Les quantités matricielles. L'équation intégrale du domaine (2.12) montre que le calcul du champ total en un point nécessite de connaître celui-ci sur tout le volume V. Lors de la discrétisation dans les méthodes intégrales, cela implique que les quantités matricielles intervenant dans le modèle numérique sont pleines (c'est-à-dire avec peu de valeurs nulles). L'espace mémoire nécessaire pour stocker ces quantités peut être conséquent. On verra dans la suite qu'il est tout de même possible de réduire l'espace mémoire en profitant des propriétés de convolution du tenseur de Green. Au contraire, les méthodes différentielles sont basées sur une approximation locale des opérateurs différentiels. Cela implique que le champ total en un point ne dépend que de son voisinage. Algébriquement, les quantités matricielles impliquées dans un modèle numérique différentiel sont alors creuses (c'est-à-dire contenant beaucoup de valeurs

nulles). L'espace mémoire nécessaire s'en retrouve réduit et des algorithmes adaptés à des problèmes parcimonieux peuvent être appliqués.

Parmi ces deux grandes familles, nous pouvons citer quelques méthodes numériques utilisées en TMO. Davidson (2010) détaille certaines de ces méthodes. Pour les approches intégrales :

- la méthode des moments (MoM) (Harrington et Harrington, 1996; Maniatis et al., 2000; Zhang et al., 2003a; Hu et al., 2005; Franceschini et al., 2005; Eyraud et al., 2009; De Zaeytijd et Franchois, 2009; Pastorino, 2010; Barrière et al., 2011; Rubaek et al., 2011; Oliveri et al., 2012; Abubakar et al., 2012a; Estatico et al., 2013)
- la méthode *Discrete Dipole Approximation* (DDA) (Goodman et al., 1991; Draine et Flatau, 1994; Chaumet et al., 2004; Yurkin et Hoekstra, 2007; Meaney et al., 2012; Grzegorczyk et al., 2012; Flatau et Draine, 2014)
- *Fast Multipole Method* (FMM) (Geng et al., 1999; Darve, 2000; Wu et al., 2005)

Pour les approches différentielles :

- la méthode des différences finies (FDM) (Zainud-Deen et al., 2008; Abubakar et al., 2008a; Fhager et al., 2009; Gilmore et al., 2009a; Fang et al., 2010; Golnabi et al., 2011; Al Sharkawy et El-Ocla, 2013)
- la méthode des éléments finis (FEM) (Rekanos et Tsiboukis, 1999; Fang et al., 2004; Soleimani et al., 2006; Meaney et al., 2007; Zakaria et al., 2010; Zakaria et LoVetri, 2011)

Comme nous venons de le voir, le choix du type d'approche dépend grandement du problème à traiter. Bien que les deux types soient applicables théoriquement pour la modélisation en TMO en 3-D, il est plus fréquent de trouver des modèles numériques basées sur des approches intégrales. Nous avons donc fait le choix d'utiliser un schéma numérique basé sur la méthode des moments. Plusieurs arguments justifient ce choix. Tout d'abord, nous disposons déjà d'une expertise concernant les méthodes intégrales dans le département de Génie Électrique de l'École Polytechnique Montréal et au laboratoire IRCCyN de l'École Centrale de Nantes à travers la thèse Barrière (2008) qui utilise un modèle numérique basée sur la méthode des moments en 2-D. Le développement d'un modèle 3-D semble donc naturel. De plus, nous nous sommes intéressés à la base de données Fresnel dont la difficulté pour reconstruire certains objets est un réel défi. Le montage d'acquisition et le problème à traiter ne permettent pas d'utiliser facilement des approches différentielles. Le choix d'une méthode intégrale était donc indispensable. Le choix de la méthode des moments semblait alors le plus judicieux.

Nous allons maintenant voir l'application de la méthode des moments pour la discrétisation des équations intégrales en TMO.

2.2.2 Discrétisation des équations intégrales par la méthode des moments

Le volume d'étude V est discrétisé en volumes élémentaires et on fait l'hypothèse que les quantités en jeu dans les équations intégrales sont constantes sur les voxels du volume discrétisé. On suppose que le volume V est un parallélépipède discrétisé en N voxels cubiques. Le contraste $\chi(\boldsymbol{r})$ et le champ total $\overrightarrow{\boldsymbol{E}}^{\text{tot}}$ sont alors décomposés sur des fonctions de base $f_n(\boldsymbol{r})$:

$$\chi(\boldsymbol{r}) = \sum_{n=1}^{N} \chi_n f_n(\boldsymbol{r}), \tag{2.13}$$

$$\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}) = \sum_{n=1}^{N} \overrightarrow{\boldsymbol{E}}_n^{\text{tot}} f_n(\boldsymbol{r}), \tag{2.14}$$

où les fonctions de base $f_n(\boldsymbol{r})$ valent 1 sur le voxel n et 0 ailleurs. L'équation du domaine (2.12) peut alors être réécrite comme

$$\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}_m) = \overrightarrow{\boldsymbol{E}}^{\text{inc}}(\boldsymbol{r}_m) + k_b^2 \Delta V \sum_{n=1}^{N} \overline{\overline{\mathbf{G}}}(\boldsymbol{r}_m, \boldsymbol{r}_n)\, \chi_n\, \overrightarrow{\boldsymbol{E}}_n^{\text{tot}} \ , \quad \forall \boldsymbol{r}_m \in V. \tag{2.15}$$

Le scalaire ΔV représente ici le volume d'un voxel. On souhaite maintenant regrouper les valeurs des trois composantes du champ total pour tous les voxels du domaine V. On choisit les points de l'espace $\boldsymbol{r}_m$ et $\boldsymbol{r}_n$ de façon à ce qu'ils coïncident avec les centres des N voxels. L'équation (2.15) est réduit à un ensemble de $3N$ équations linéaires que l'on peut écrire sous forme matricielle :

$$\boldsymbol{e}^{\text{tot}} = \boldsymbol{e}^{\text{inc}} + \mathbf{G}_D \mathbf{X} \boldsymbol{e}^{\text{tot}}. \tag{2.16}$$

Les vecteurs $\boldsymbol{e}^{\text{tot}}$ et $\boldsymbol{e}^{\text{inc}}$ contiennent respectivement les trois composantes spatiales des champs total et incident en tous les voxels du domaine V. Ces vecteurs sont donc de longueur $3N$. La matrice $\mathbf{X}$ de taille $3N \times 3N$ contient les valeurs du contraste sur les N voxels :

$$\mathbf{X} = \begin{bmatrix} \text{diag}(\boldsymbol{x}) & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \text{diag}(\boldsymbol{x}) & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & \text{diag}(\boldsymbol{x}) \end{bmatrix}. \tag{2.17}$$

où $\mathbf{0}_n$ est la matrice nulle de taille $N \times N$ et $\boldsymbol{x}$ le vecteur de longueur N contenant le contraste sur tous les voxels : $\boldsymbol{x} = [\chi_1, \ldots, \chi_N]^{\text{t}}$. La matrice $\mathbf{G}_D$ de taille $3N \times 3N$ représente la discrétisation du tenseur de Green pour tous les couples $(\boldsymbol{r}_m, \boldsymbol{r}_n)$. L'annexe A détaille la démarche permettant d'arriver à l'équation (2.16) et donne l'expression des quantités intervenant dans la formulation matricielle.

Tout comme l'équation du domaine, l'équation intégrale d'observation (2.11) peut être for-

mulée de façon discrète. Supposons un ensemble de N_R récepteurs en dehors du volume V où le champ diffracté est mesuré. En suivant la même démarche que pour l'équation du domaine, l'équation d'observation devient

$$\boldsymbol{e}^{\text{scat}} = \mathbf{G}_o \mathbf{X} \boldsymbol{e}^{\text{tot}}, \tag{2.18}$$

où le vecteur $\boldsymbol{e}^{\text{scat}}$ de taille N_R correspond aux valeurs du champ diffracté au niveau des N_R récepteurs et la matrice $\mathbf{G}_o$ de taille $N_R \times 3N$ correspond à la discrétisation du tenseur de Green entre les N voxels du domaine et les N_R points de mesure. Les quantités algébriques intervenant dans l'équation d'observation (2.18) sont également détaillées dans l'annexe A.

La modélisation sous forme discrète se résume alors aux deux équations matricielles suivantes :

$$\begin{cases} \boldsymbol{e}^{\text{scat}} &= \mathbf{G}_o \mathbf{X} \boldsymbol{e}^{\text{tot}} \\ \boldsymbol{e}^{\text{tot}} &= (\mathbf{I} - \mathbf{G}_D \mathbf{X})^{-1} \boldsymbol{e}^{\text{inc}}, \end{cases} \tag{2.19}$$

où $\mathbf{I}$ représente la matrice identité de taille $3N$. Ces deux équations matricielles peuvent être combinées afin d'exprimer, en une seule équation, les mesures du champ diffracté en fonction de la grandeur d'intérêt, le contraste :

$$\boldsymbol{e}^{\text{scat}} = \mathbf{G}_o \mathbf{X} \left(\mathbf{I} - \mathbf{G}_D \mathbf{X}\right)^{-1} \boldsymbol{e}^{\text{inc}}. \tag{2.20}$$

Le calcul du champ total $\boldsymbol{e}^{\text{tot}}$ nécessite de résoudre un système linéaire souvent appelé *problème direct* dans la littérature. Dans la suite de ce manuscrit, nous poserons $\mathbf{L}_x = \mathbf{I} - \mathbf{G}_D \mathbf{X}$ afin de simplifier les notations. Le modèle direct peut être simplifié :

$$\boldsymbol{e}^{\text{scat}} = \mathbf{G}_o \mathbf{X} \, \mathbf{L}_x^{-1} \, \boldsymbol{e}^{\text{inc}}, \quad \text{avec } \mathbf{L}_x = \mathbf{I} - \mathbf{G}_D \mathbf{X}. \tag{2.21}$$

2.3 Implémentation du problème direct

De nombreux algorithmes de reconstruction en tomographie micro-ondes se fondent sur une formulation du modèle direct similaire à (2.21) et nécessitent alors la résolution du système linéaire intervenant dans cette formulation. La résolution du système $\mathbf{L}_x \boldsymbol{e}^{\text{tot}} = \boldsymbol{e}^{\text{inc}}$ est une opération très coûteuse, notamment pour des problèmes 3-D de grande taille. Deux questions se posent. Comment résoudre efficacement le système linéaire ? Peut-on stocker cette matrice ? Nous répondrons dans cette section à ces questions afin de comprendre comment le problème direct est résolu actuellement dans la littérature.

2.3.1 Résolution du système linéaire

Le coût de la résolution du problème direct est très important et nous verrons dans le chapitre suivant qu'il s'agit de l'opération la plus coûteuse dans le processus d'inversion. Réduire le temps de calcul de cette opération est donc un défi que les auteurs en TMO ont cherché à relever. Le chapitre 5 propose justement une solution à ce challenge. Comme la grande taille de la matrice $\mathbf{L}_x$ empêche une inversion directe, les auteurs se sont focalisés sur l'utilisation d'algorithmes itératifs qu'ils tronquent afin de réduire le temps de calcul.

De nombreux algorithmes itératifs ont été utilisés pour résoudre le modèle direct en TMO :

- le gradient conjugué (Sarkar et al., 1986; Zwamborn et van den Berg, 1992)
- le gradient biconjugué (Gan et Chew, 1995)
- *Quasi-Minimal Residual method* (QMR) et la variante *Transpose-Free* (TFQMR) (Ellis, 2002; Wang et Jin, 1998)
- la version stabilisée du gradient biconjugué (BiCGStab) (Xu et al., 2002; Zhang et al., 2003a; De Zaeytijd et al., 2007; Abubakar et al., 2012a)
- une variante QMR du BiCGStab (QMRCGStab) (Hursan et Zhdanov, 2002; Chaumet et Rahmani, 2009)
- *Generalized Product-type methods based on the Biconjugate Gradient* (GPBiCG) (Zhang, 1997; Fujino, 2002; Chaumet et Rahmani, 2009; Chaumet et al., 2013; Flatau et Draine, 2014)
- la généralisation de la méthode de minimisation du résidu (GMRES) (Saad et Schultz, 1986; Drogoudis et al., 2009; Hu et al., 2011).

L'algorithme BiCGStab, développé par Van der Vorst (1992) et Gutknecht (1993) est devenu l'algorithme de référence depuis une quinzaine d'années pour la résolution des systèmes linéaires en TMO. Xu et al. (2002) ont montré que BiCGStab était très efficace comparé à d'autres algorithmes itératifs (BiCG et QMR). L'utilisation de cet algorithme dans le domaine de la tomographie micro-ondes a été initiée par Xu et al. (2002); Zhang et al. (2003a) pour la résolution du modèle direct. Puis cette méthode a été reprise pour le problème inverse par de nombreux auteurs comme De Zaeytijd et al. (2007); Abubakar et al. (2012a). L'algorithme BiCGStab est également très utilisé dans d'autres applications comme le calcul de surfaces équivalentes radar (*Radar-Cross Section, RCS*) par exemple (Zhang et al., 2003b; Ke-Ming et Yu-Fa, 2010) ou la magnétotellurique (Liu et al., 2009; Li et al., 2016). Dans la suite de ce manuscrit, nous nous intéresserons au BiCGStab et à sa variante par blocs. C'est pourquoi nous allons décrire plus précisément cet algorithme.

L'algorithme BiCGStab est présenté dans l'Algorithme 1. Il requiert deux multiplications par

$\mathbf{L}_x$ à chaque itération (lignes 5 et 8 de l'algorithme). Ces opérations sont les plus coûteuses et concentrent la grande majorité du temps de calcul de la résolution. L'algorithme BiCGStab nécessite le réglage de trois paramètres : l'initialisation du champ total à calculer, un vecteur $\tilde{\boldsymbol{r}}_0$ et un seuil de tolérance pour le critère d'arrêt. Concernant l'initialisation, quand aucune information n'est connue sur la solution du champ total, le choix d'initialisation se porte généralement sur le champ incident. En revanche, si une information *a priori* est donnée sur la solution, il est possible d'initialiser avec cette information. Pour le choix du vecteur $\tilde{\boldsymbol{r}}_0$, les auteurs proposent en général de choisir ce vecteur de façon arbitraire et beaucoup choisissent $\tilde{\boldsymbol{r}}_0 = \boldsymbol{r}^{(0)}$ (Xu et al., 2002). Enfin, l'algorithme est arrêté quand le résidu relatif est inférieur à un certain seuil fixé par l'opérateur (ligne 15 de l'algorithme). La valeur de la tolérance dépend de la précision souhaitée de la résolution. Plus la tolérance est stricte, plus la solution sera proche de la solution exacte $\mathbf{L}_x^{-1}\boldsymbol{e}^{\text{inc}}$ mais plus le nombre d'itérations avant convergence sera élevé. Il s'agit donc d'un compromis entre précision de la solution et temps de calcul. Dans le cadre de l'inversion, tronquer les systèmes linéaires trop tôt peut entraîner des instabilités dues aux erreurs d'approximations. Malheureusement, il est rare que les auteurs précisent le niveau de troncature utilisé dans leurs méthodes de reconstruction. L'effet de la troncature sera au cœur du chapitre 7 où nous y verrons un moyen d'assurer une cohérence dans l'inversion malgré la résolution tronquée de ces systèmes linéaires.

Algorithme 1 : Algorithme BiCGStab

1: Pour une initialisation $\boldsymbol{e}^{\text{tot}(0)}$, on définit $\boldsymbol{r}^{(0)} = \boldsymbol{e}^{\text{inc}} - \mathbf{L}_x\boldsymbol{e}^{\text{tot}(0)}$ et $\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)}$
2: Choisir un vecteur arbitraire $\tilde{\boldsymbol{r}}_0$
3: $k = 0$
4: **repeat**
5: $\quad \boldsymbol{v} = \mathbf{L}_x\boldsymbol{p}^{(k)}$
6: $\quad \alpha = (\tilde{\boldsymbol{r}}_0^{\dagger}\boldsymbol{r}^{(k)})/(\tilde{\boldsymbol{r}}_0^{\dagger}\boldsymbol{v})$
7: $\quad \boldsymbol{s} = \boldsymbol{r}^{(k)} - \alpha\boldsymbol{v}$
8: $\quad \boldsymbol{t} = \mathbf{L}_x\boldsymbol{s}$
9: $\quad \omega = \boldsymbol{t}^{\dagger}\boldsymbol{s}/\boldsymbol{t}^{\dagger}\boldsymbol{t}$
10: $\quad \boldsymbol{e}^{\text{tot}(k+1)} = \boldsymbol{e}^{\text{tot}(k)} + \alpha\boldsymbol{p}^{(k)} + \omega\boldsymbol{s}$
11: $\quad \boldsymbol{r}^{(k+1)} = \boldsymbol{s} - \omega\boldsymbol{t}$
12: $\quad \beta = (\alpha\tilde{\boldsymbol{r}}_0^{\dagger}\boldsymbol{r}^{(k+1)})/(\omega\tilde{\boldsymbol{r}}_0^{\dagger}\boldsymbol{r}^{(k)})$
13: $\quad \boldsymbol{p}^{(k+1)} = \boldsymbol{r}^{(k+1)} + \beta(\boldsymbol{p}^{(k)} - \omega\boldsymbol{v})$
14: $\quad k \leftarrow k + 1$
15: **until** $\|\boldsymbol{r}^{(k)}\|/\|\boldsymbol{e}^{\text{inc}}\| <$ tolérance

2.3.2 Multiplication par $\mathbf{L}_x$: comment réduire l'espace mémoire

La résolution du système linéaire fait intervenir un grand nombre de multiplications par la matrice $\mathbf{L}_x$ (2 multiplications par itération de BiCGStab). Prenons par exemple un domaine d'étude discrétisé en $30\times30\times30$ voxels. La matrice $\mathbf{L}_x$ est de taille $3N\times3N$, ce qui représente une matrice 81000×81000. Si l'on souhaite stocker directement cette matrice dans la mémoire de l'ordinateur en nombre à virgule flottante à double précision (soit 8 octets par nombre), il faudrait environ 100 Go de mémoire uniquement pour stocker la matrice $\mathbf{L}_x$. Cela est bien évidemment impossible sur des machines personnelles actuelles où la mémoire vive dépasse rarement une vingtaine de gigaoctets. En revanche, $\mathbf{L}_x$ ayant une structure particulière, il est possible de réduire fortement l'espace mémoire nécessaire.

Rappelons que la matrice $\mathbf{L}_x$ vaut $\mathbf{I}-\mathbf{G}_D\mathbf{X}$ où $\mathbf{G}_D$ est une matrice pleine de taille $3N\times3N$. Cette matrice provient de la discrétisation du tenseur de Green qui est un noyau de convolution. La multiplication de $\mathbf{L}_x$ avec un vecteur quelconque $\boldsymbol{u}$ revient à effectuer l'opération :

$$\mathbf{L}_x\boldsymbol{u} = (\mathbf{I}-\mathbf{G}_D\mathbf{X})\boldsymbol{u} \tag{2.22}$$

$$= \boldsymbol{u}-\mathbf{G}_D(\mathbf{X}\boldsymbol{u}). \tag{2.23}$$

Le produit par $\mathbf{L}_x$ peut être effectué comme le produit de convolution entre le noyau représenté par $\mathbf{G}_D$ et le vecteur $(\mathbf{X}\boldsymbol{u})$. Plutôt que de stocker la matrice $\mathbf{G}_D$ (qui est de la même taille que $\mathbf{L}_x$), il suffit de stocker le noyau de convolution discret 3-D correspondant de taille $(2N+1)\times(2N+1)\times(2N+1)$. Ce produit de convolution discret, que l'on retrouve dans l'équation (2.15) à travers la somme sur les voxels du domaine, peut être évalué à partir de transformées de Fourier. Le produit de $\mathbf{G}_D$ avec un vecteur est alors réalisé avec des routines `FFT` (*Fast Fourier Transform*). Cette astuce est nécessaire pour résoudre le problème direct 3-D en grande taille et elle est utilisée dans de nombreux algorithmes de reconstruction (Sarkar et al., 1986; Zwamborn et van den Berg, 1992; Gan et Chew, 1995; Wang et Jin, 1998; De Zaeytijd et al., 2008; Bjelogrlic et al., 2015). Abubakar et van den Berg (2004) et Zhang et al. (2003a) décrivent plus en détail les opérations permettant de développer les produits avec $\mathbf{L}_x$.

2.4 Validation du modèle direct

Bien que nous ayons une certaine expertise au sujet de la méthode des moments parmi les équipes de recherche encadrant cette thèse, seul le modèle 2-D avait été développé par Barrière (2008). Le modèle 3-D présenté précédemment est complètement nouveau. Il est important que le modèle numérique soit le plus représentatif des phénomènes physiques étudiés. Il est donc nécessaire de s'assurer de la validité de ce modèle dans les gammes d'utilisation :

fréquences utilisées, taille et contraste des objets à reconstruire, etc. Dans cette section, le modèle numérique est comparé à :

— la solution analytique de Mie (Balanis, 2012). Il s'agit d'une solution analytique permettant de connaître le champ diffracté d'une onde plane incidente par une sphère diélectrique homogène. Cette solution permet de valider le modèle et d'observer les limites de celui-ci en fonction de la discrétisation et du contraste de la sphère ;

— la base de données de l'Institut Fresnel (Geffrin et Sabouroux, 2009). Cette base fournit un ensemble de mesures expérimentales de champ électrique diffracté par des objets purement diélectriques.

2.4.1 Comparaison à la solution analytique de Mie

La solution de Mie exprime analytiquement le champ diffracté par une sphère diélectrique homogène pour une onde plane incidente. La forme analytique fait intervenir une somme infinie de termes impliquant des fonctions sphériques d'Hankel. Plusieurs ouvrages présentent cette solution dont Balanis (2012). On peut faire varier plusieurs paramètres dont la fréquence de l'onde incidente, la taille de la sphère et le contraste de celle-ci.

Pour comparer le modèle numérique à la solution de Mie, nous avons utilisé une onde incidente de fréquence égale à 1 GHz ($\lambda = 30$ cm) et différentes configurations de sphère dont nous avons fait varier le diamètre et le contraste. Dans les exemples suivants, une onde incidente plane se déplaçant dans la direction z et polarisée linéairement selon x se propage dans l'air. La sphère est placée au centre du repère. Nous avons choisi de positionner des récepteurs sur deux cercles centrés à l'origine, de rayon 2λ et dans les plans (xOz) et (yOz). Nous visualisons le champ diffracté $\boldsymbol{e}_x^{\text{scat}}$ sur le cercle du plan (yOz) et les champs diffractés $\boldsymbol{e}_x^{\text{scat}}$ et $\boldsymbol{e}_z^{\text{scat}}$ sur le cercle du plan (xOz). La figure 2.2 représente l'emplacement de la sphère et des récepteurs.

Le modèle numérique nécessite le calcul du champ total dans le volume et celui-ci est réalisé par l'algorithme BiCGStab. La tolérance a été fixée à 10^{-6}, ce qui donne une précision suffisante sur la solution. Concernant la discrétisation, il est nécessaire d'avoir un maillage suffisamment fin pour représenter correctement les variations des champs électromagnétiques dans le volume. Il est préconisé de prendre au moins 10 voxels par longueur d'onde afin d'éviter des erreurs de discrétisation trop importantes. Dans la suite, nous avons fait en sorte que la discrétisation respecte cette condition.

Enfin, pour quantifier l'erreur entre le modèle numérique et la solution analytique, nous avons

fait le choix de mesurer l'erreur quadratique suivante :

$$\text{erreur} = \frac{\sum_i |y_i - \hat{y}_i|^2}{\sum_i |\hat{y}_i|^2}, \tag{2.24}$$

où y_i représente les mesures du champ diffracté obtenues par le modèle numérique et $\hat{y}_i$ les mêmes mesures calculées par la solution analytique.

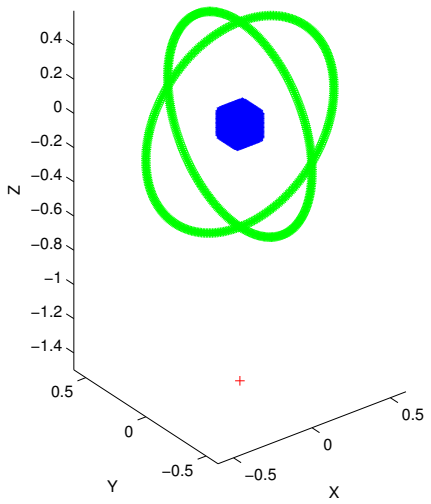


Figure 2.2 Schéma du montage pour la comparaison avec la solution analytique de Mie. Le volume cubique discrétisé est représenté en bleu. Les deux cercles verts dans les plans (xOz) et (yOz) représentent les points de mesure.

Sphère de diamètre $\lambda/2$

Le premier objet est une sphère de diamètre $\lambda/2 = 15$ cm placée dans l'air. On a choisi de discrétiser un volume d'étude cubique de côté $\lambda/2$ contenant la sphère en $20 \times 20 \times 20$ voxels : la résolution dans l'air est donc de $\lambda/40 = 7.5$ mm.

Plusieurs contrastes sont testés : une sphère purement diélectrique à contraste faible, une sphère purement diélectrique à contraste élevé, une sphère conductrice à contraste faible et une sphère conductrice à contraste élevé. Ces quatre sphères sont résumées comme suit :

— $\epsilon_r = 3$, $\sigma = 0$, soit un contraste $\chi = 2$. La longueur d'onde dans la sphère est d'environ $\lambda_S \approx 17$ cm, soit une résolution du maillage dans la sphère d'environ $\lambda_S/23$.

— $\epsilon_r = 3$, $\sigma = 0.1$, soit un contraste $\chi = 2 + 1.8j$. La longueur d'onde dans la sphère est

d'environ $\lambda_S \approx 17$ cm, soit une résolution du maillage dans la sphère d'environ $\lambda_S/22$.

— $\epsilon_r = 10, \sigma = 0$, soit un contraste $\chi = 9$. La longueur d'onde dans la sphère est d'environ $\lambda_S \approx 9.5$ cm, soit une résolution du maillage dans la sphère d'environ $\lambda_S/13$.

— $\epsilon_r = 10$, $\sigma = 1$, soit un contraste $\chi = 9 + 18j$. La longueur d'onde dans la sphère est d'environ $\lambda_S \approx 7.7$ cm, soit une résolution du maillage dans la sphère d'environ $\lambda_S/10$.

L'amplitude des champs diffractés est donnée dans les figures 2.3 et 2.4 et le tableau 2.1 récapitule les erreurs quadratiques entre le modèle numérique et la solution analytique. Tout d'abord, on note que le modèle numérique est globalement très proche de la solution analytique. Pour de faibles contrastes, les erreurs sont très faibles. Pour les contrastes plus élevés, le modèle ne colle pas parfaitement à la solution analytique. On remarque aussi que l'erreur est bien plus élevée pour un objet purement diélectrique ($\chi = 9$) que pour un objet ayant un contraste complexe ($\chi = 9 + 18j$). Nous n'avons pas d'explication pour ce phénomène. Dans le cadre de la tomographie micro-ondes, il est très difficile de reconstruire des objets aussi contrastés que les sphères $\chi = 9$ et $\chi = 9 + 18j$. Cela conforte le fait que le modèle numérique utilisé convient pour un objectif de reconstruction.

Contraste χ	$\boldsymbol{e}_x^{\text{scat}}$ sur (yOz)	$\boldsymbol{e}_x^{\text{scat}}$ sur (xOz)	$\boldsymbol{e}_z^{\text{scat}}$ sur (xOz)
$\chi = 2$	0.05%	0.08%	0.20%
$\chi = 2 + 1.8j$	0.07%	0.06%	0.08%
$\chi = 9$	13.8%	15.5%	3.0%
$\chi = 9 + 18j$	1.2%	1.1%	1.0%

Tableau 2.1 Erreurs quadratiques relatives entre les mesures du champ diffracté obtenues par le modèle numérique et celles obtenues par la solution analytique de Mie. La sphère est de diamètre $\lambda/2$ et la discrétisation du domaine en $20 \times 20 \times 20$ voxels.

Nous avons voulu étudier l'effet de la résolution du maillage sur le modèle numérique. Le même volume est discrétisé avec un maillage plus fin que précédemment : $30 \times 30 \times 30$ voxels plutôt que $20 \times 20 \times 20$ voxels, soit une résolution de $\lambda/60 = 5$ mm. Le tableau 2.2 présente les erreurs quadratiques avec ce nouveau maillage. Comparé au tableau 2.1, on remarque que les erreurs sont nettement plus faibles avec cette résolution. On peut expliquer cette différence par : i) le fait que la résolution plus fine permet de réduire les erreurs d'approximation dans le modèle physique, ii) le fait que la sphère est mieux approchée par le maillage cartésien pour la résolution plus faible.

En conclusion, le contraste joue un rôle important dans la précision du modèle numérique : plus le contraste est élevé, moins le modèle est précis. En revanche, la résolution du maillage

influe sur la qualité du modèle : raffiner le maillage permet de réduire les erreurs.

Contraste χ	$\boldsymbol{e}_x^{\text{scat}}$ sur (yOz)	$\boldsymbol{e}_x^{\text{scat}}$ sur (xOz)	$\boldsymbol{e}_z^{\text{scat}}$ sur (xOz)
$\chi = 2$	0.02%	0.02%	0.03%
$\chi = 2 + 1.8j$	0.05%	0.04%	0.04%
$\chi = 9$	3.2%	3.9%	0.76%
$\chi = 9 + 18j$	0.64%	0.61%	0.54%

Tableau 2.2 Erreurs quadratiques relatives entre les mesures du champ diffracté obtenues par le modèle numérique et celles obtenues par la solution analytique. La sphère est de diamètre $\lambda/2$ et la discrétisation du domaine en $30 \times 30 \times 30$ voxels.

Sphère de diamètre λ

Nous prenons maintenant une sphère plus grande de diamètre $\lambda = 30$ cm avec les mêmes contrastes que précédemment. Le volume cubique de côté λ est discrétisé en $40 \times 40 \times 40$ voxels ce qui correspond à la même résolution que dans l'exemple précédent, soit 7.5 mm. Le tableau 2.3 présente les erreurs entre les deux modèles. On remarque que la taille de l'objet a une influence sur l'erreur : pour un même contraste et une même résolution, l'erreur est globalement plus élevée pour la sphère de grande taille.

Contraste χ	$\boldsymbol{e}_x^{\text{scat}}$ sur (yOz)	$\boldsymbol{e}_x^{\text{scat}}$ sur (xOz)	$\boldsymbol{e}_z^{\text{scat}}$ sur (xOz)
$\chi = 2$	0.12%	0.11%	0.09%
$\chi = 2 + 1.8j$	0.02%	0.02%	0.11%
$\chi = 9$	8.5%	11.6%	28.9%
$\chi = 9 + 18j$	0.54%	0.74%	2.5%

Tableau 2.3 Erreurs quadratiques relatives entre les mesures du champ diffracté obtenues par le modèle numérique et celles obtenues par la solution analytique. La sphère est de diamètre λ et la discrétisation du domaine en $40 \times 40 \times 40$ voxels.

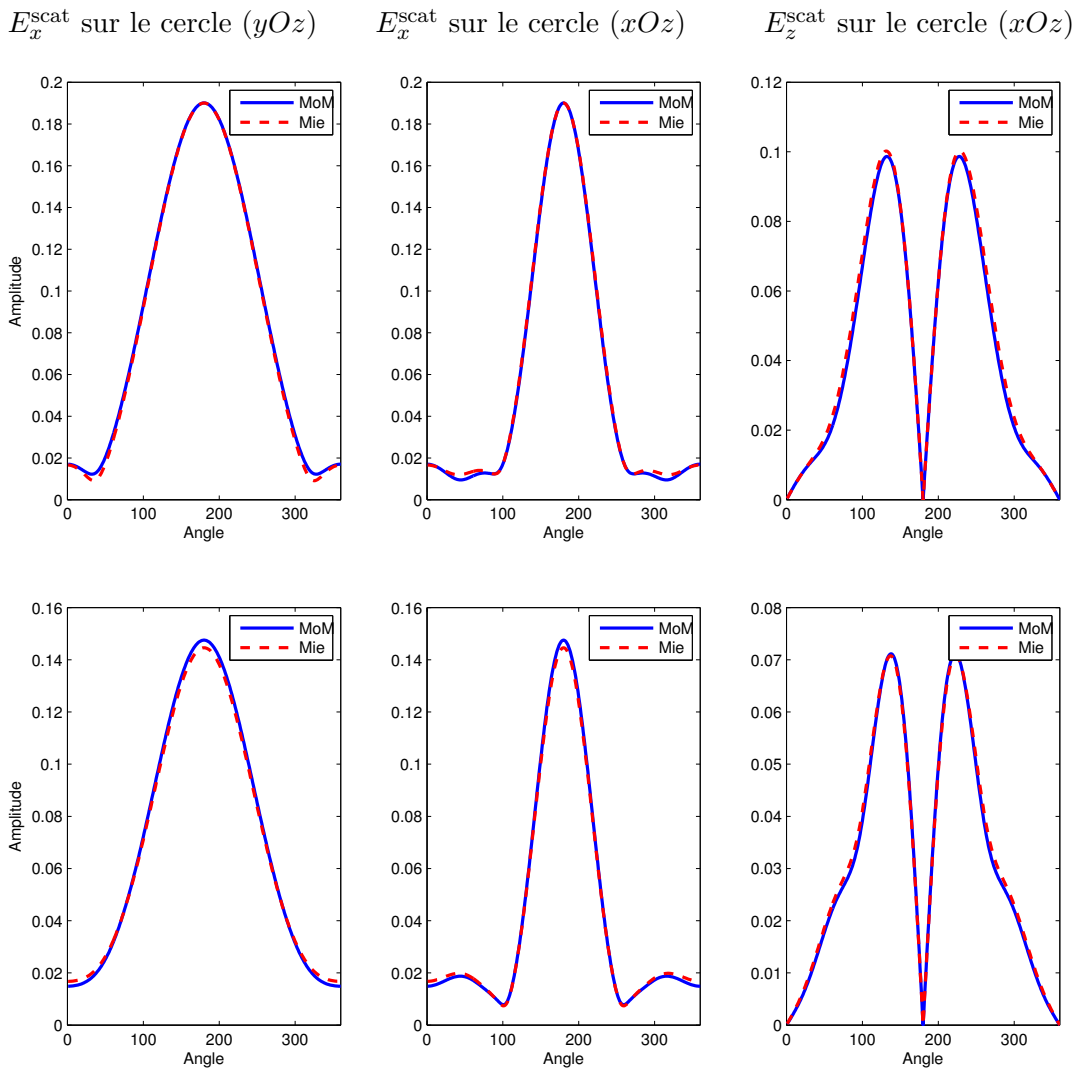


Figure 2.3 Mesures du champ diffracté par une sphère de diamètre $\lambda/2$. La permittivité est $\chi = 2$ en haut et $\chi = 2 + 1.8j$ en bas. La sphère est discrétisée en $20 \times 20 \times 20$ voxels. En trait plein bleu, les mesures obtenues par le modèle numérique (MoM) ; en pointillés rouges, la solutions analytique de Mie.

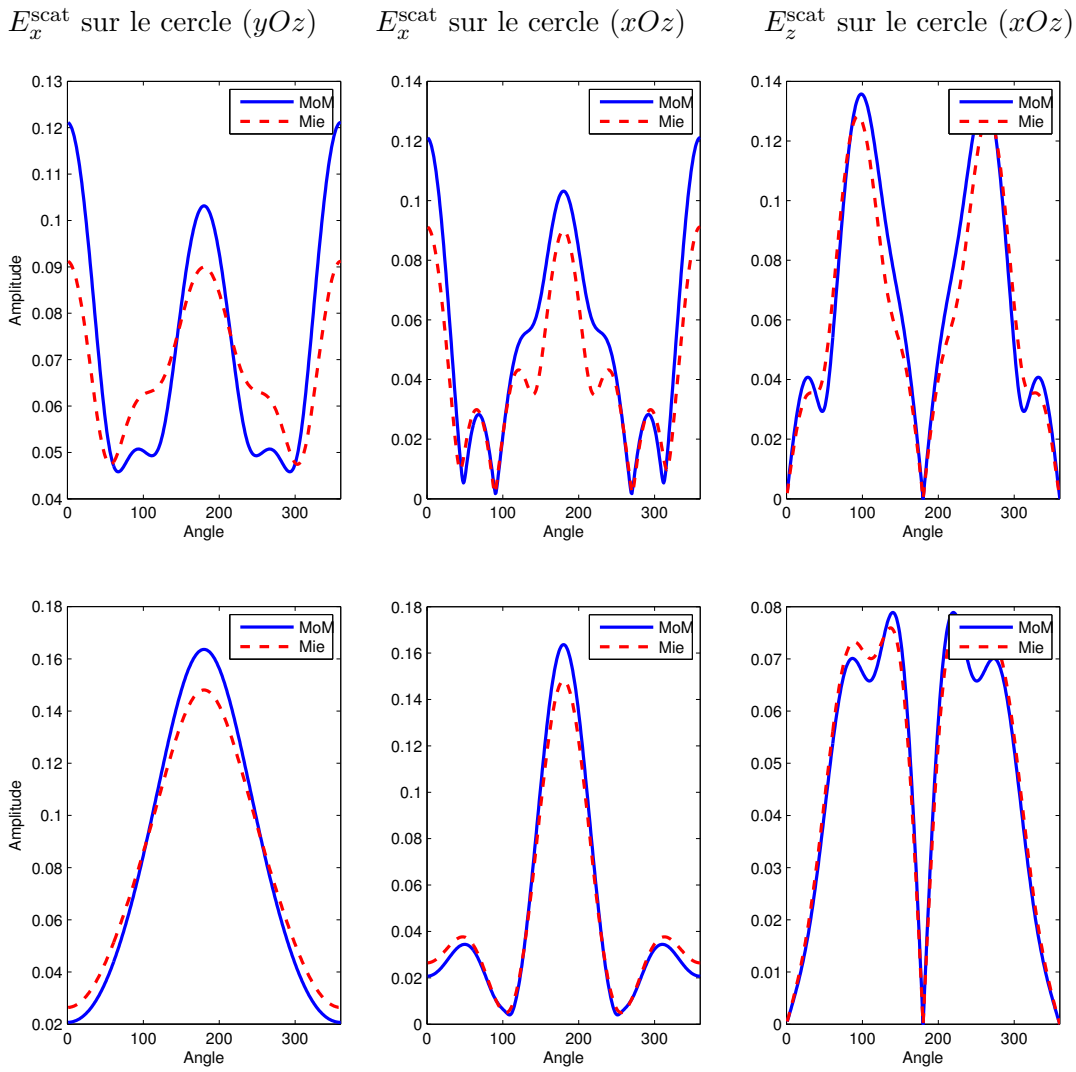


Figure 2.4 Suite de la figure 2.3. Mesures du champ diffracté par une sphère de diamètre $\lambda/2$. La permittivité est $\chi = 9$ en haut et $\chi = 9 + 18j$ en bas. La sphère est discrétisée en $20 \times 20 \times 20$ voxels. En trait plein bleu, les mesures obtenues par le modèle numérique (MoM) ; en pointillés rouges, la solutions analytique de Mie.

2.4.2 Comparaison aux données réelles de l'Institut Fresnel

Nous avons également souhaité valider le modèle en s'appuyant sur des données réelles. L'Institut Fresnel (Marseille, France) propose librement un ensemble de fichiers de mesures issues d'expérimentations réalisées en chambre anéchoïque. Ces données ont été présentées au monde scientifique dans l'objectif de valider les méthodes d'inversion et de reconstruire des objets de différentes formes, tailles et propriétés diélectriques. Toutes les informations nécessaires à la reproductibilité des expérimentations sont données dans Geffrin et Sabouroux (2009). Le chapitre 6 se focalise sur la reconstruction de ces objets et les détails du montage d'acquisition y sont donnés.

Pour valider notre modèle, nous avons reproduit les expériences de l'Institut Fresnel et comparé les mesures expérimentales aux valeurs obtenues par le modèle numérique. Ici, nous donnons uniquement les résultats pour l'objet TwoCubes. La figure 2.5 est tirée de Geffrin et Sabouroux (2009) et présente l'objet. L'objet est composé de deux cubes purement diélectriques de 25 mm de côté et d'une permittivité relative de 2.35. Le volume d'étude englobant les deux cubes est discrétisé avec une résolution de 2.5 mm. Les figures 2.6 et 2.7 présentent les parties réelle et imaginaire des données aux 243 récepteurs pour les 36 sources aux fréquences 5 GHz et 8 GHz. On observe que les valeurs du champ diffracté obtenues par le modèle numérique sont très proches des données expérimentales.

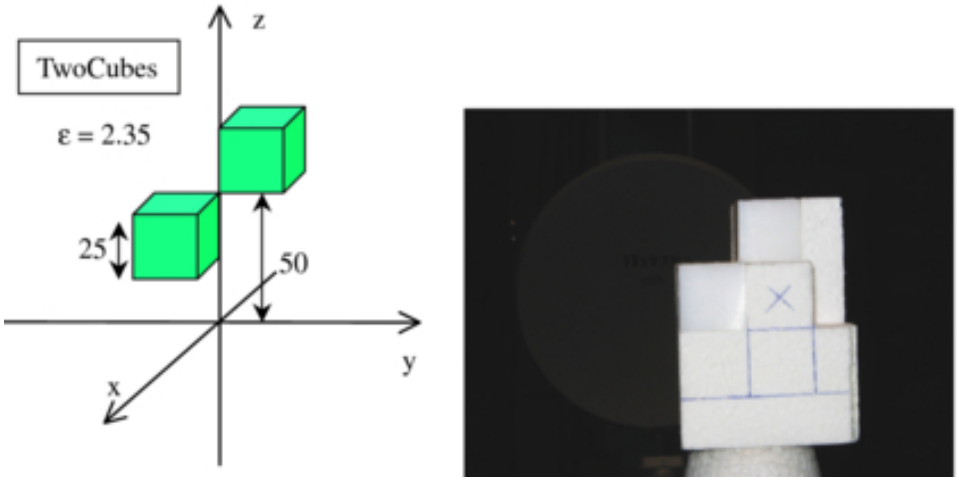


Figure 2.5 Représentation de l'objet TwoCubes de la base de données Fresnel. Figure tirée de Geffrin et Sabouroux (2009). L'objet TwoCubes est composé de deux cubes en contact par un angle. Les deux cubes sont identiques : ils mesurent 25 mm de côté et ont une permittitivité relative de 2.35.

Pour mesurer quantitativement ces erreurs, nous avons calculé l'erreur quadratique. Contrairement aux comparaisons avec la solution analytique de Mie, nous comparons ici à des don-

nées expérimentales. Les données étant bruitées, il est important de comprendre que nous ne cherchons pas à obtenir une erreur nulle mais plutôt à obtenir une variance faible entre le modèle et les mesures expérimentales. D'après les chercheurs de l'Institut Fresnel, le modèle numérique développé au sein de leur équipe affichait une erreur entre 5 et 10% par rapport aux mesures expérimentales pour les différentes fréquences utilisées. Le tableau 2.4 donne les erreurs pour les différentes mesures. Malgré quelques mesures où les erreurs atteignent 17-18%, la plupart des erreurs sont comprises entre 5 et 10%. Les erreurs les plus importantes apparaissent pour les plus faibles fréquences et pour les plus hautes fréquences. Une explication possible de ces erreurs pour les faibles fréquences est que les valeurs du champ diffracté sont faibles alors que le bruit de mesure est constant pour toutes les fréquences. Il semble alors normal que ces forts pourcentages correspondent à la variance du bruit de mesure. Pour les hautes fréquences, une explication éventuelle est qu'à cette fréquence, l'objet est de grande taille électrique et la résolution est grande par rapport à la longueur d'onde. Tout comme pour la solution analytique, le modèle numérique pourrait montrer ses limites, ce qui justifierait ces erreurs plus élevées que pour les données aux autres fréquences.

Fréquence	Erreur	Fréquence	Erreur
3 GHz	16.6 %	5.75 GHz	6.4 %
3.25 GHz	12.5 %	6 GHz	8.3%
3.5 GHz	17.0 %	6.25 GHz	8.6 %
3.75 GHz	9.3 %	6.5 GHz	8.0 %
4 GHz	8.9 %	6.75 GHz	11.5 %
4.25 GHz	7.8 %	7 GHz	9.0 %
4.5 GHz	7.4 %	7.25 GHz	13.0 %
4.75 GHz	7.4 %	7.5 GHz	11.9 %
5 GHz	5.9 %	7.75 GHz	18.4 %
5.25 GHz	6.1 %	8 GHz	16.1 %
5.5 GHz	5.7 %		

Tableau 2.4 Erreurs quadratiques relatives entre les données réelles et simulées pour l'objet TwoCubes.

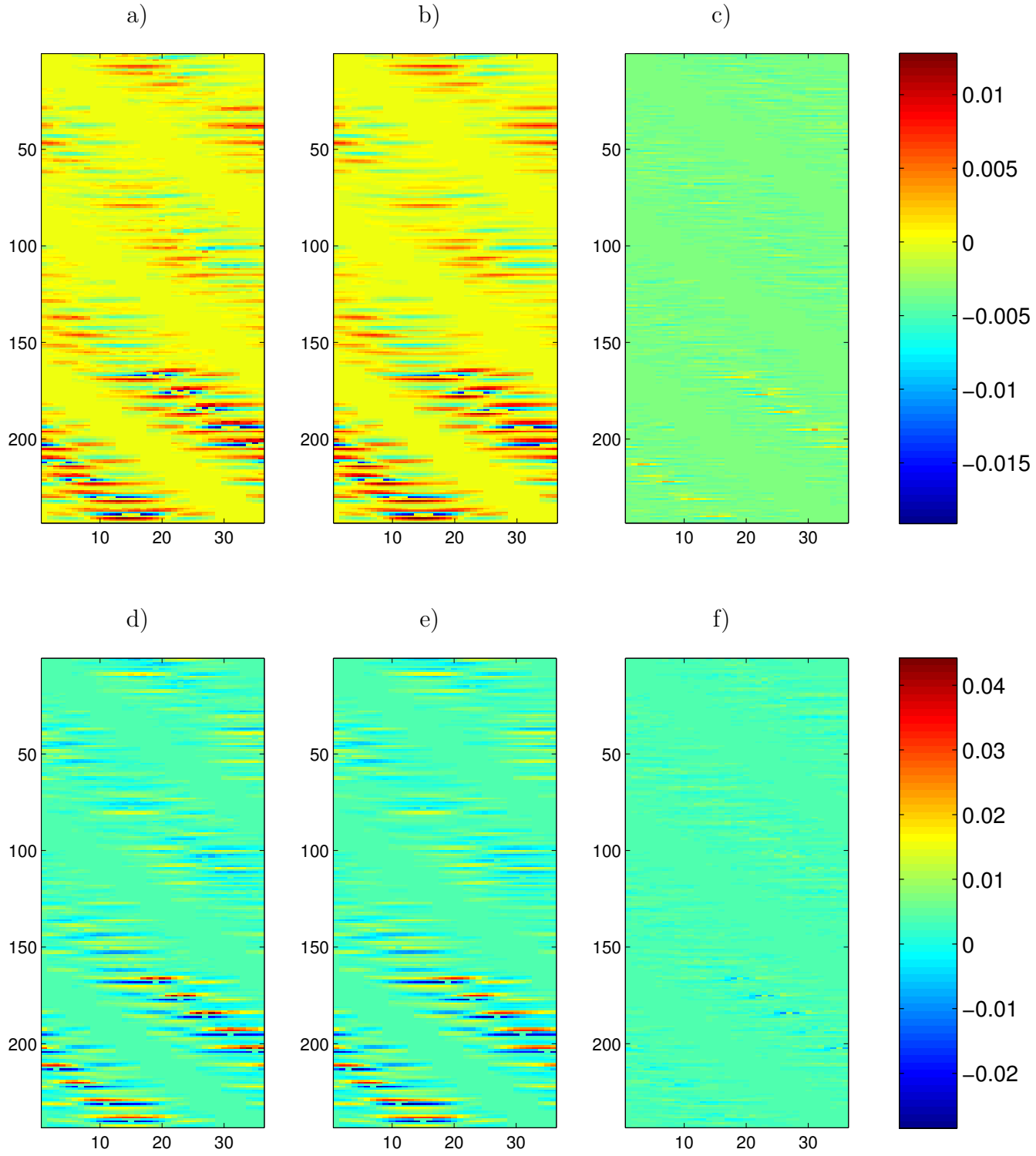


Figure 2.6 Comparaison des données réelles et simulées pour l'objet TwoCubes à 5 GHz. a) Partie réelle des données réelles ; b) partie réelle des données simulées ; c) différences entre les parties réelles ; d) partie imaginaire des données réelles ; e) partie imaginaire des données simulées ; f) différences entre les parties imaginaires.

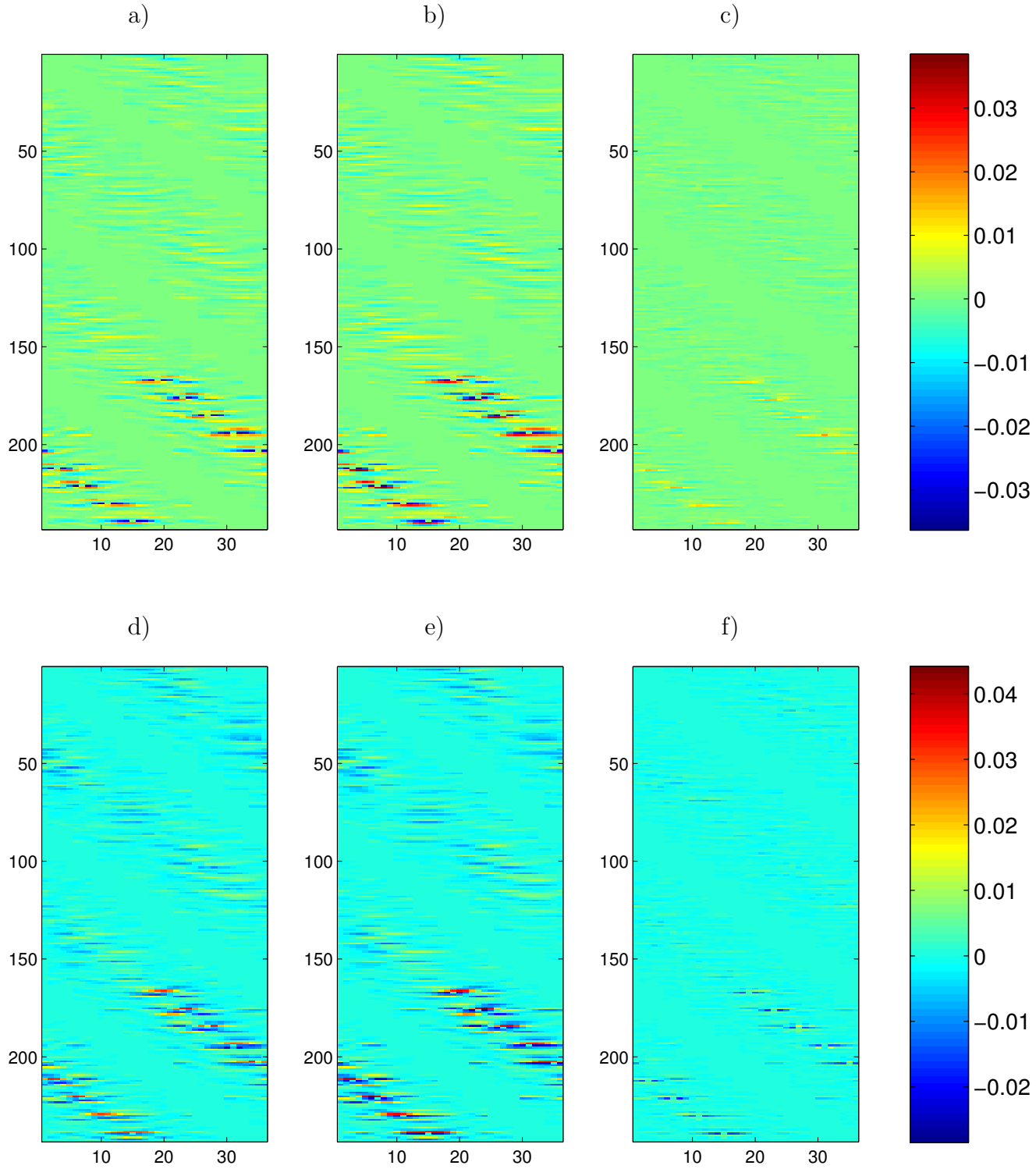


Figure 2.7 Comparaison des données réelles et simulées pour l'objet TwoCubes à 8 GHz. a) Partie réelle des données réelles ; b) partie réelle des données simulées ; c) différences entre les parties réelles ; d) partie imaginaire des données réelles ; e) partie imaginaire des données simulées ; f) différences entre les parties imaginaires.

2.5 Conclusion

Nous avons présenté dans ce chapitre la formulation du modèle direct que nous allons utiliser dans la suite de ce manuscrit. Celui-ci est basé sur la solution intégrale de l'équation d'onde et sur une discrétisation par la méthode des moments. L'expression matricielle du modèle direct (2.21), liant les mesures du champ électrique diffracté au contraste de l'objet, est le point de départ de la formulation du problème inverse que nous verrons dans le chapitre 3.

Nous avons également mis en évidence la principale difficulté du calcul du modèle direct qui réside dans la résolution d'un système linéaire de grande taille afin de calculer le champ total. Cette étape est la plus coûteuse dans le calcul des champs diffractés et nous verrons dans le chapitre suivant que ce calcul est récurrent pour la reconstruction en TMO. L'algorithme BiCGStab s'est imposé comme référence pour résoudre le problème direct. Nous verrons dans le chapitre 5 un nouveau moyen de résoudre efficacement ces systèmes linéaires intervenant dans le problème inverse.

Enfin, puisque ce modèle 3-D n'avait jamais été développé dans les laboratoires encadrant cette thèse, une étape de validation du modèle était indispensable. Le modèle numérique a été confronté aussi bien à la solution analytique de Mie qu'aux données expérimentales issues de l'Institut Fresnel. Les résultats obtenus permettent de conclure que notre modèle est suffisamment précis dans les cas d'utilisation en TMO. En revanche, on observe des erreurs élevées pour des objets de grande taille et fortement contrastés.

CHAPITRE 3 MÉTHODES D'INVERSION EN IMAGERIE MICRO-ONDES

Le chapitre précédent a présenté le modèle direct permettant de calculer le champ électrique diffracté par un objet diélectrique. Ce chapitre se concentre sur le problème inverse et sur les méthodes de reconstruction. Le problème inverse en TMO consiste en l'estimation des propriétés diélectriques d'un milieu inconnu à partir de mesures de champs diffractés. Cette estimation repose généralement sur la minimisation d'un critère d'adéquation aux données : on cherche les propriétés électriques pour lesquelles l'erreur entre les observations et le modèle numérique est la plus faible. Les problèmes inverses sont souvent mal posés au sens de Hadamard et c'est le cas du problème de TMO (Pastorino, 2010). Il est alors fréquent d'introduire un terme de régularisation permettant de stabiliser la solution. Enfin, un algorithme d'optimisation doit être choisi afin de trouver un optimum au critère à minimiser. Finalement, trois éléments sont primordiaux dans la définition de notre problème inverse en TMO : le critère, la régularisation et l'algorithme d'optimisation. Ce chapitre propose de comparer les différentes approches fréquentes dans la littérature pour ces trois éléments et nous présenterons notre méthode d'inversion en justifiant nos choix.

Ce chapitre s'articule autour de cinq sections :

- la première section présente le problème général d'imagerie micro-ondes ;
- la deuxième section traite de l'inversion vue comme un problème d'optimisation. Il s'agit de présenter les fonctions coût fréquemment utilisées dans la littérature et de comparer les avantages et inconvénients de chacune ;
- les différents types de régularisation sont discutés dans la troisième section et nous décrivons en détail le choix adopté pour notre problème ;
- la section 4 dépeint les différentes familles d'algorithmes d'optimisation utilisées pour la reconstruction d'images en TMO et compare les propriétés de chacune ;
- enfin, la dernière section présente le calcul du gradient de la fonction coût choisie. Ce calcul met en évidence les difficultés de la résolution du problème inverse.

3.1 L'imagerie micro-ondes

L'imagerie micro-ondes est un ensemble très vaste de techniques de détection, de localisation et d'estimation d'objets dissimulés à l'aide d'ondes électromagnétiques aux fréquences micro-ondes (300 MHz - 300 GHz). Cette science regroupe une gamme variée de problèmes et d'applications. Certaines méthodes d'imagerie ne se focalisent que sur la détection d'objets

particuliers cachés, c'est-à-dire de déceler la présence ou non d'un corps ; d'autres méthodes recherchent la position et/ou la forme de ces objets. Enfin, des méthodes plus complexes cherchent à estimer aussi bien la position, la forme et les propriétés diélectriques des objets, ce qui revient à réaliser une cartographie des propriétés diélectriques dans un volume d'intérêt. Dans ce manuscrit, nous nous intéressons uniquement à ces dernières, les méthodes dites quantitatives. Dans la littérature, l'expression *inverse scattering methods* est souvent utilisée pour nommer ce type d'approches.

Un grand nombre de méthodes de reconstruction a été développé depuis quelques décennies mais il n'est pas envisageable de mentionner toutes ces approches. Ici nous nous concentrerons sur les méthodes basées sur la minimisation d'un critère pénalisé par des algorithmes itératifs d'optimisation locale. Ces méthodes sont les plus courantes dans la littérature car elles donnent des résultats convenables pour un coût de calcul raisonnable. Plus d'informations sont données dans les sections suivantes. Nous pouvons citer malgré tout quelques méthodes "différentes" de celles que nous verrons par la suite, s'appuyant par exemple sur :

- différents modèles numériques. Nous avons parlé de la discrétisation des équations régissant la propagation d'onde et des méthodes numériques comme la méthode des moments, la méthode des éléments finis. Mais il existe des modèles bien différents comme celui basé sur la déformation d'un volume où un nombre fini de points de la surface fait office d'inconnues (Li et al., 2012), ou celui basé sur une représentation binaire du volume d'intérêt (Duchêne, 2001) ;
- différentes approximations. Certaines méthodes utilisent des approximations linéaires afin de contourner la non-linéarité du modèle direct et donc d'accélérer la reconstruction. C'est le cas de l'approximation de Born ou de Rytov, ou encore l'approche *Linear Sampling Method* (LSM) ;
- différents outils mathématiques, par exemple ceux permettant de résoudre un problème linéaire comme la troncature de la décomposition en valeurs singulières (TSVD).

Pour plus de références, Pastorino (2010) fournit de nombreuses descriptions d'algorithmes de reconstruction.

Pour la suite, on pose les notations suivantes : on suppose un volume homogène V dans lequel un ou plusieurs objets inconnus sont placés. Le volume est discrétisé en N voxels cubiques. Ce volume est illuminé successivement par N_S ondes incidentes notés $\boldsymbol{e}_i^{\text{inc}}$ $(i = 1, \ldots, N_S)$. N_R récepteurs sont situés autour de ce volume et mesurent le champ électrique diffracté noté $\widetilde{\boldsymbol{e}}_i^{\text{scat}}$ $(i = 1, \ldots, N_S)$. Le vecteur $\widetilde{\boldsymbol{e}}_i^{\text{scat}}$ de longueur N_R contient les mesures aux N_R récepteurs pour l'illumination i. On cherche alors à reconstruire la fonction de contraste sur les N voxels, représentée par le vecteur $\boldsymbol{x}$ de longueur N. On rappelle les équations matricielles

d'observation et du domaine issues du chapitre précédent :

$$\boldsymbol{e}^{\text{scat}} = \mathbf{G}_o \mathbf{X} \boldsymbol{e}^{\text{tot}} \qquad \text{(équation d'observation)} \tag{3.1}$$

$$\boldsymbol{e}^{\text{tot}} = \boldsymbol{e}^{\text{inc}} + \mathbf{G}_c \mathbf{X} \boldsymbol{e}^{\text{tot}} \qquad \text{(équation du domaine),} \tag{3.2}$$

ou de façon équivalente, l'équation non-linéaire :

$$\boldsymbol{e}^{\text{scat}} = \mathbf{G}_o \mathbf{X} \mathbf{L}_x^{-1} \boldsymbol{e}^{\text{inc}} \tag{3.3}$$

avec $\mathbf{L}_x = \mathbf{I} - \mathbf{G}_c \mathbf{X}$.

3.2 L'inversion vue comme un problème d'optimisation

Puisqu'à partir des mesures de champs diffractés, on ne peut pas directement obtenir une estimation du contraste, on considère le problème inverse comme la minimisation d'un critère permettant de réduire l'erreur entre les données mesurées et le modèle direct. En problème inverse, un critère des moindres carrés, c'est-à-dire l'erreur quadratique entre les mesures et le modèle, est souvent considéré. L'estimation du contraste correspond alors au minimiseur de ce critère qui est généralement obtenu par des algorithmes d'optimisation itératifs. S'appuyant sur le modèle numérique vu précédemment, deux formulations du critère ressortent fréquemment dans la littérature : une formulation bilinéaire et une formulation non-linéaire. Notons que ces formulations du critère ne sont pas réservées à une forme intégrale du modèle direct, elles peuvent être appliquées à une approche par différences finies ou éléments finis par exemple.

3.2.1 Formulation bilinéaire

On peut voir le problème inverse comme un problème d'optimisation de l'équation d'observation sous la contrainte de l'équation du domaine. Il est en général plus simple de se ramener à un problème sans contrainte où la contrainte d'égalité est introduite d'une certaine façon dans le critère. La formulation bilinéaire s'appuie justement sur l'équation d'observation pour le terme d'attache aux données pénalisée par l'erreur commise sur la contrainte. Dans cette équation, aussi bien le contraste $\boldsymbol{x}$ que les champs totaux $\boldsymbol{e}_i^{\text{tot}}$ sont inconnus : on cherche alors à les estimer conjointement. Dans la formulation bilinéaire, on ne cherche pas à résoudre exactement les champs totaux dans l'équation du domaine ; on relâche la contrainte d'égalité sur cette équation. Dans la littérature, ces méthodes prennent le nom MGM (*Modified Gradient Method*) (Kleinman et van den Berg, 1992) ou CSI (*Contrast Source Inversion*) (van den

Berg et Kleinman, 1995). Le critère à minimiser prend la forme générique suivante :

$$\mathcal{F}^{CSI}(\boldsymbol{x}, \boldsymbol{e}_1^{\text{tot}}, \ldots, \boldsymbol{e}_{N_S}^{\text{tot}}) = \frac{1}{2}\sum_{i=1}^{N_S} \|\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o \mathbf{X} \boldsymbol{e}_i^{\text{tot}}\|^2 + \frac{\gamma}{2}\sum_{i=1}^{N_S} \|\boldsymbol{e}_i^{\text{tot}} - \mathbf{G}_D \mathbf{X} \boldsymbol{e}_i^{\text{tot}} - \boldsymbol{e}_i^{\text{inc}}\|^2. \quad (3.4)$$

Le paramètre γ permet de contrôler la pénalisation que l'on souhaite attribuer à la contrainte sur les champs totaux : plus γ est élevé, plus l'erreur sur la contrainte dans le second terme va tendre vers zéro. Il serait nécessaire d'avoir un poids γ très grand afin de s'assurer que le champ total estimé correspond bien à l'équation $\boldsymbol{e}_i^{\text{tot}} = \mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}$. De nombreux auteurs ont proposé de faire varier γ au cours des itérations de l'algorithme d'optimisation (van den Berg et Kleinman, 1995). En revanche, Barrière et al. (2011) a montré qu'il était judicieux de fixer empiriquement la valeur de γ et de garder cette valeur au cours de la reconstruction afin d'assurer la convergence de l'algorithme d'inversion.

Ces approches bilinéaires sont fréquemment utilisées car elles ont l'avantage majeur que le critère est linéaire par rapport à $\boldsymbol{x}$ et par rapport aux champs totaux. De plus, il n'est pas nécessaire de résoudre le problème direct (inversion par $\mathbf{L}_x$) ; le coût de calcul par itération est considérablement réduit. En revanche, le nombre d'inconnues est très conséquent :$(N_S+1)\times N$ inconnues. A cause du nombre important d'inconnues, il est souvent nécessaire d'effectuer un grand nombre d'itérations pour converger.

Les algorithmes de reconstruction utilisant cette formulation bilinéaire du critère sont très fréquents notamment en 2-D car le nombre d'inconnues reste raisonnable (Abubakar et al., 2005; Barrière et al., 2011; van den Berg et Abubakar, 2001; Bozza et Pastorino, 2009; Gilmore et al., 2009b,a; Li et al., 2013; Mojabi et LoVetri, 2010; Ozdemir, 2014; Rubaek et al., 2011; Zakaria et LoVetri, 2012). En revanche, les méthodes d'inversion basées sur cette formulation pour des problèmes tridimensionnels sont bien plus rares. On peut tout de même en citer quelques-unes (Abubakar et van den Berg, 2004; Zhang et Liu, 2004; Catapano et al., 2006; Asefi et al., 2015).

3.2.2 Formulation non-linéaire

Contrairement à la formulation bilinéaire où la contrainte d'égalité est exprimée sous la forme d'un terme de pénalisation, la formulation non-linéaire s'appuie sur l'inclusion de cette contrainte dans le modèle direct. Cela revient donc à considérer l'équation non-linéaire (3.3). Le critère de cette formulation exprime directement l'erreur entre les données et le modèle

non-linéaire. Le critère d'adéquation aux données est alors défini par :

$$\mathcal{F}^D(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{N_S} \|\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}\|^2 \tag{3.5}$$

Les champs totaux ne sont plus des inconnues du problème, seule la fonction de contraste $\boldsymbol{x}$ en est une : le nombre d'inconnues est réduit à N. En revanche, le calcul du critère nécessite la résolution du problème direct, c'est-à-dire de la résolution des systèmes linéaires impliquant $\mathbf{L}_x$. Cela entraîne un coût de calcul conséquent à chaque itération. De plus, la forte non-linéarité du modèle direct peut engendrer la présence de multiples minima locaux et rendre la résolution du problème inverse difficile. En conclusion, le coût d'une itération du problème inverse est nettement plus important que celui pour le critère bilinéaire. En revanche, le nombre d'itérations est en général bien plus faible que pour l'approche bilinéaire puisque le nombre d'inconnues est bien plus faible.

En TMO 2-D, l'utilisation d'un critère non-linéaire a été utilisée (Roger, 1981; Chew et Wang, 1990; Remis et van den Berg, 2000; Franchois et Tijhuis, 2003; Fang et al., 2004; Hu et al., 2009). Pour des problèmes 3-D de grande taille, le fait que le nombre d'inconnues soit plus faible que pour les approches bilinéaires a incité les auteurs à se tourner davantage vers la formulation non-linéaire. Il s'agit de l'approche la plus fréquente dans la littérature ces dernières années (Harada et al., 2001; Soleimani et al., 2006; Yu et al., 2009; De Zaeytijd et Franchois, 2009; Chaumet et Belkebir, 2009; Winters et al., 2010; Grzegorczyk et al., 2012; Abubakar et al., 2012b,a; Estatico et al., 2013; Scapaticci et al., 2015). Notre choix s'est donc porté sur la formulation non-linéaire.

De nombreuses autres méthodes s'appuient sur différents critères. Ce manuscrit ne prétend pas comparer toutes ces méthodes. Nous pouvons citer tout de même les méthodes itératives basées sur des approximations du modèle direct, comme la méthode de Born itérative (Ali et Moghaddam, 2010; Zhang et Liu, 2015). Mudry et al. (2012) a proposé une méthode hybride mêlant les critères non-linéaire et bilinéaire. Des informations supplémentaires peuvent être obtenues dans Pastorino (2010).

3.3 La régularisation

Les problèmes inverses sont en général mal posés et c'est le cas en tomographie micro-ondes (Pastorino, 2010). Un problème bien posé respecte les trois conditions suivantes (Idier, 2001) :

— existence d'une solution,

— unicité de la solution,
— robustesse de la solution au bruit.

Obtenir ces trois conditions est très rare en pratique. Par exemple, la résolution d'un problème sous-déterminé donne une infinité de solutions possibles et l'unicité de la solution n'est donc pas vérifiée. C'est le cas en TMO, notamment en 3-D, où le nombre de voxels N est bien supérieur au nombre de données.

Pour éviter le caractère mal posé du problème en TMO, il est nécessaire d'ajouter des contraintes ou des informations *a priori* sur la solution. L'utilisation d'un terme de régularisation est une méthode très fréquente pour améliorer la stabilité de l'inversion et donc obtenir une solution correcte (Idier, 2001). Cette régularisation intervient comme un terme supplémentaire dans le critère à minimiser. Dans la suite, le terme de régularisation est noté $\mathcal{R}(\boldsymbol{x})$.

3.3.1 Régularisation "$\ell_2\ell_1$"

Différentes questions se posent quant au choix de la régularisation. Quelle connaissance *a priori* a-t-on sur l'objet ? Que cherche-t-on à pénaliser ? Quelle fonction de régularisation utiliser ? Comment intégrer le terme de régularisation dans le critère à minimiser ? Le choix de la régularisation reste ouvert et la quantité impressionnante d'approches de régularisation dans la littérature en est la preuve. Malgré tout, une régularisation conservant les discontinuités (dite "*edge-preserving*") semble mettre d'accord la plupart des auteurs. Ce type de régularisation fait l'hypothèse que l'objet à reconstruire est un objet homogène par morceaux, ce qui en effet représente la majorité des objets. Par exemple, pour une application biomédicale, on peut faire l'hypothèse que les propriétés diélectriques d'un tissu humain (os, muscle, etc.) sont constantes dans un même tissu.

L'idée est alors de pénaliser les variations de contraste, c'est-à-dire les différences premières : on pénalise fortement les faibles variations afin de favoriser des milieux constants et on pénalise dans une moindre mesure les fortes variations de contraste afin de permettre des discontinuités entre les milieux homogènes. De nombreuses fonctions de régularisation permettent ce type de pénalisation. Une liste est fournie dans Idier (2001). Nous avons fait le choix d'utiliser ce type de pénalisation dans cette thèse à l'aide des fonctions "$\ell_2\ell_1$" sur les différences premières. Ces fonctions ont l'avantage d'être continûment différentiables et donc d'avoir un gradient défini pour tout $\boldsymbol{x}$. De plus, ces fonctions sont convexes, ce qui permet d'assurer la robustesse à la solution et d'améliorer la convergence des algorithmes d'optimisation. Le nom "$\ell_2\ell_1$" provient du fait qu'elles ont un comportement quadratique à l'origine et asymptotiquement linéaire à l'infini. Elles conviennent donc pour pénaliser fortement les

faibles variations et préserver les grandes discontinuités. Il existe d'autres types de régularisation pour conserver les discontinuités comme les fonctions "$\ell_2\ell_0$" (asymptotiquement constantes à l'infini) mais celles-ci ne sont pas convexes, ce qui peut entraîner des instabilités lors de l'optimisation.

Le terme de régularisation utilisé dans nos travaux est le suivant :

$$\mathcal{R}(\boldsymbol{x}) = \sum_{n=1}^{N_x-1}\sum_{m=1}^{N_y-1}\sum_{p=1}^{N_z-1} \varphi(\chi_{n+1,m,p}-\chi_{n,m,p})+\varphi(\chi_{n,m+1,p}-\boldsymbol{x}_{n,m,p})+\varphi(\chi_{n,m,p+1}-\chi_{n,m,p}). \quad (3.6)$$

Rappelons que le terme $\chi_{n,m,p}$ correspond à la valeur du contraste χ sur le voxel d'indice (n, m, p) selon la représentation volumique des voxels et les entiers N_x, N_y et N_z représentent les dimensions du volume V discrétisé. On retrouve ici la pénalisation sur les différences premières à travers l'approximation au premier ordre : $\chi_{n+1,m,p} - \chi_{n,m,p}$ pour la direction x par exemple. La fonction $\varphi(u) = \sqrt{\delta^2 + |u|^2}$ est la fonction de pénalisation "$\ell_2\ell_1$" qui permet de préserver les fortes discontinuités. Dans cette fonction, le paramètre $\delta > 0$ doit être réglé. Dans ce manuscrit, nous avons fait le choix de fixer δ de façon empirique.

Ce type de régularisation semble faire l'unanimité en TMO bien que l'appellation ne soit pas partagée. Certains auteurs introduisent la variation totale comme terme de pénalisation (van den Berg et Kleinman, 1995; Abubakar et van den Berg, 2001) mais ce terme est équivalent à notre expression dans (3.6).

3.3.2 Régularisation additive ou multiplicative ?

Dans la majorité des méthodes d'inversion, les auteurs proposent d'intégrer le terme de régularisation dans le critère soit :

— sous forme additive. Le critère global prend la forme : $\mathcal{F} = \mathcal{F}^D + \lambda\mathcal{R}$

— sous forme multiplicative. Le critère global prend la forme $\mathcal{F} = \mathcal{F}^D \times \mathcal{R}$

La régularisation additive est très fréquente en problèmes inverses. Le paramètre de régularisation λ permet d'ajuster l'équilibre entre l'adéquation aux données $\mathcal{F}^D$ et l'information *a priori* de la régularisation $\mathcal{R}$. Le réglage de ce paramètre est important car il influence le résultat de la reconstruction. Choisir λ faible signifie que l'on fait plutôt confiance aux données et au modèle (par exemple, faible bruit sur les données). À l'inverse, choisir λ élevé signifie que l'on attribue plus de crédit à l'information *a priori*. Le choix de λ est donc fortement lié au niveau de bruit sur les données (Idier, 2001) : plus le niveau de bruit est élevé, plus λ sera élevé. Des techniques ont été développées pour ce réglage (Idier, 2001), on peut citer la méthode de la "courbe en L" par exemple. Ces techniques peuvent être coûteuses et le réglage de λ est alors réalisé empiriquement. Par exemple, Barrière et al. (2011) proposent

une procédure expérimentale pour ce réglage : une première étape consiste à choisir un objet de référence suffisamment grand et contrasté pour représenter le type d'objets à reconstruire pour l'application donnée, la seconde étape consiste à effectuer un certain nombre de reconstructions de l'objet référence en faisant varier les paramètres de la régularisation (ici, λ et δ) sur une grille logarithmique. Les valeurs des paramètres sont alors choisies en fonction de la qualité de la solution pour les différentes reconstructions de l'objet référence.

La régularisation multiplicative est apparue dans les algorithmes d'inversion en TMO depuis quelques années et semble avoir convaincu un grand nombre d'auteurs. L'avantage avancé pour une telle régularisation est le fait qu'il n'y ait pas besoin de paramètre de régularisation, ce qui évite la phase délicate du réglage de celui-ci. Barrière et al. (2011) montrent en revanche que ce type de régularisation présente des limites notamment pour la reconstruction d'objets fortement contrastés. D'autres approches de régularisation ont été étudiées dans la littérature, comme par exemple De Zaeytijd et al. (2007) qui mélange les régularisations additives et multiplicatives.

Dans la suite, nous avons choisi une régularisation additive. Pour le réglage du paramètre de régularisation λ, nous avons choisi de le fixer de façon empirique en suivant la procédure décrite dans Barrière et al. (2011).

3.4 Les algorithmes d'optimisation utilisés

Après avoir fait le choix d'un critère à minimiser et d'une régularisation qui convient au problème étudié, il est nécessaire de choisir un algorithme d'optimisation approprié. Comme le critère n'est pas convexe, des minima locaux peuvent être présents. La minimisation d'un tel critère nécessite généralement des méthodes d'optimisation globale, qu'elles soient déterministes ou stochastiques (Pastorino, 2007). Cependant, à cause du grand nombre d'inconnues et de la difficulté du problème, ces approches sont trop coûteuses. Malgré la non-linéarité du problème et le risque de tomber dans un minimum local non global, les approches d'optimisation locale sont tout de même robustes et permettent d'obtenir des résultats de reconstruction acceptables. Ce sont les algorithmes les plus fréquents pour minimiser le critère et nous énumérerons les principales méthodes ci-dessous. Pour assurer la convergence vers un minimum *acceptable*, il est nécessaire d'initialiser correctement ces algorithmes afin que l'initialisation appartienne au bassin d'attraction de ce minimum acceptable. Certaines techniques permettent d'obtenir une initialisation convenable. On peut citer deux d'entre elles :

— initialisation du contraste avec une solution approchée. Par exemple, la méthode de Born (linéarisation du modèle) permet d'obtenir une première idée de l'objet inconnu. Cette technique est surtout valable dans les conditions de validité de l'approximation

de Born. On peut également citer la méthode *Linear Sampling Method* (LSM) qui peut être utilisée pour déterminer une première approximation de la forme et de la position de l'objet (Catapano et al., 2009) ;

— la technique par saut de fréquence (*frequency hopping*). L'idée est de reconstruire l'objet à partir de mesures à différentes fréquences. Les basses fréquences permettent de résoudre un problème faiblement non-linéaire puisque l'objet est de taille électrique faible à ces fréquences. On obtient alors une solution lissée de l'objet inconnu. En reconstruisant l'objet à des fréquences de plus en plus élevées, il est possible d'obtenir une solution de plus en plus fidèle de l'objet. Cette technique est très utilisée, c'est par exemple le cas dans la plupart des méthodes reconstruisant les objets de la base Fresnel. Cette technique nécessite d'avoir fait l'acquisition des mesures à différentes fréquences en amont de la reconstruction.

Nous allons nous focaliser sur les méthodes d'optimisation locale qui couvrent la majeure partie des méthodes d'inversion en TMO. L'idée générale est d'estimer une direction de descente $\boldsymbol{d}$ et un pas de descente α à partir d'une estimation courante du contraste. Notons à l'itération k, l'itéré courant $\boldsymbol{x}^{(k)}$. L'itéré suivant $\boldsymbol{x}^{(k+1)}$ est défini sous la forme :

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}, \tag{3.7}$$

de façon à ce que $\mathcal{F}(\boldsymbol{x}^{(k+1)}) < \mathcal{F}(\boldsymbol{x}^{(k)})$. L'algorithme est arrêté quand l'itéré est proche d'un minimum local : en pratique, quand la norme du gradient $\nabla_x \mathcal{F}(\boldsymbol{x}^{(k)})$ est inférieure à un seuil défini par l'utilisateur. La différence principale entre les algorithmes itératifs est la façon dont la direction $\boldsymbol{d}$ est calculée. On peut classer les algorithmes en deux catégories : les algorithmes du premier ordre, c'est-à-dire n'utilisant que l'information du gradient pour calculer les directions de descente ; et les algorithmes du second ordre, utilisant le gradient et la matrice Hessienne afin de calculer une direction de descente plus efficace. Les méthodes du premier ordre sont généralement moins coûteuses pour calculer la direction de descente à chaque itération puisqu'il n'est pas nécessaire d'estimer la matrice Hessienne. En revanche, elles requièrent souvent un plus grand nombre d'itérations avant convergence. À l'inverse, le coût de calcul du Hessien pouvant être important, le coût d'une itération pour les méthodes du second ordre est souvent très élevé. Mais la convergence est atteinte en moins d'itérations, notamment quand le critère est proche d'une forme quadratique.

Concernant le pas de descente $\alpha^{(k)}$, trouver le pas optimal minimisant le critère selon la direction de descente est trop coûteux pour un critère non-linéaire. En général, le pas de descente est choisi de façon à respecter certaines conditions. Les conditions fortes de Wolfe peuvent par exemple être appliquées (Nocedal et Wright, 1999). Une condition moins stricte,

la condition d'Armijo, est souvent utilisée. Plusieurs procédures de recherche de pas ont été développées afin de trouver un pas respectant une de ces conditions. On peut citer la procédure de *backtracking* fréquemment utilisée.

Revenons aux algorithmes d'optimisation où nous présentons ceux qui sont largement appliqués à la reconstruction en imagerie micro-ondes. Les méthodes ne seront pas détaillées ici ; pour plus d'informations, le lecteur pourra se référer à Bertsekas (1999).

La méthode du gradient conjugué non-linéaire

La méthode du gradient conjugué non-linéaire est l'extension du réputé gradient conjugué pour des fonctions non-linéaires. On peut exprimer la direction de descente à l'itération k sous la forme :

$$\boldsymbol{d}^{(k)} = -\nabla_x \mathcal{F}(\boldsymbol{x}^{(k)}) + \beta_k \boldsymbol{d}^{(k-1)}, \tag{3.8}$$

où β_k est calculé selon une certaine formule. Les formules les plus connues sont celles de Fletcher-Reeves et celle de Polak-Ribière (Bertsekas, 1999).

La méthode de Newton

La méthode de Newton est une méthode du second degré basée sur la minimisation successive d'approximations quadratiques du critère. La direction de descente est donnée par :

$$\boldsymbol{d}^{(k)} = -\mathbf{H}_k^{-1} \nabla_x \mathcal{F}(\boldsymbol{x}^{(k)}), \tag{3.9}$$

où la matrice $\mathbf{H}_k$, de taille $N \times N$, est la matrice Hessienne. Pour des problèmes de grande taille, l'inversion directe de $\mathbf{H}_k$ est impossible ; il est nécessaire d'utiliser des méthodes itératives pour résoudre le système linéaire afin de déterminer la direction de descente $\boldsymbol{d}^{(k)}$. De plus, de par sa grande taille, la matrice $\mathbf{H}_k$ ne peut pas être stockée en mémoire directement. Tout comme la matrice $\mathbf{L}_x$ dans le modèle direct, seules quelques quantités matricielles exprimant une formulation implicite de $\mathbf{H}_k$ doivent être stockées.

La méthode originale de Newton consiste à choisir un pas de descente $\alpha^{(k)} = 1$. Si la matrice Hessienne n'est pas définie positive, ce choix peut entraîner non pas une direction de descente mais une direction de montée. Pour contourner ce problème, la variante relâchée de la méthode de Newton propose de choisir un pas de descente selon la même démarche que les autres algorithmes (par exemple par *backtracking*) qui peut être négatif pour assurer la descente de l'algorithme.

La méthode de Gauss-Newton

La méthode de Gauss-Newton (GN) est une méthode utilisable uniquement pour la minimisation de critères des moindres carrés. On peut voir cette approche comme une variante de la méthode de Newton dans laquelle la matrice Hessienne est approchée par $\mathbf{H}_k \approx \mathbf{J}^\dagger\mathbf{J}$, où $\mathbf{J}$ est le jacobien du modèle en $\boldsymbol{x}^{(k)}$. On obtient donc la direction de descente suivante :

$$\boldsymbol{d}^{(k)} = -(\mathbf{J}^\dagger\mathbf{J})^{-1}\,\nabla_x\mathcal{F}(\boldsymbol{x}^{(k)}). \tag{3.10}$$

La matrice $\mathbf{J}^\dagger\mathbf{J}$ étant semi-définie positive, il est assuré que la direction est une direction de descente. Tout comme la méthode originale de Newton, la méthode de Gauss-Newton propose un pas de descente $\alpha^{(k)} = 1$. Cela n'assure pas nécessairement que $\mathcal{F}(\boldsymbol{x}^{(k+1)}) < \mathcal{F}(\boldsymbol{x}^{(k)})$. On peut alors ajouter une recherche de pas efficace afin de satisfaire aux conditions d'Armijo par exemple.

Cet algorithme est très utilisé pour le problème inverse en TMO, on le retrouve notamment dans la littérature récente (Mojabi et LoVetri, 2009; Abubakar et al., 2012a; De Zaeytijd et Franchois, 2009; Ostadrahimi et al., 2013; Estatico et al., 2013).

Tout comme l'algorithme de Newton, il est impossible de stocker la matrice $\mathbf{J}^\dagger\mathbf{J}$ directement pour des problèmes de grande taille. Il est nécessaire d'utiliser d'autres solutions pour calculer la direction de descente. Une solution possible est la formulation implicite du jacobien décrite dans Abubakar et al. (2012a).

La méthode de Levenberg-Marquardt

La méthode de Levenberg-Marquardt (LM) est une amélioration de la méthode de Gauss-Newton qui permet de stabiliser la solution en améliorant le conditionnement de la matrice $\mathbf{J}^\dagger\mathbf{J}$. La direction de descente est :

$$\boldsymbol{d}^{(k)} = -(\mathbf{J}^\dagger\mathbf{J} + \lambda\mathbf{M})^{-1}\,\nabla_x\mathcal{F}(\boldsymbol{x}^{(k)}). \tag{3.11}$$

On peut voir la matrice $\mathbf{M}$ comme un terme de régularisation quadratique, autrement appelé régularisation de Tikhonov. Le paramètre λ peut aussi être considéré comme un facteur de régularisation. Le choix le plus courant pour $\mathbf{M}$ est de prendre la matrice identité.

Il a été prouvé que cette méthode est équivalente à la méthode *Distorted Born Iterative Method* (DBIM) avec régularisation quadratique (Chew et Wang, 1990; Joachimowicz et al., 1991; Franchois et Pichot, 1997). De nombreux auteurs utilisent le dénommé DBIM (Haddadin et Ebbini, 1998; Remis et van den Berg, 2000; Yu et al., 2009; Gilmore et al., 2009b;

Scapaticci et al., 2015). De plus, l'algorithme de Gauss-Newton est équivalent à la méthode DBIM pour laquelle aucune régularisation n'est utilisée (Carfantan, 1996; Azghani et al., 2015).

La méthode BFGS

La méthode Broyden-Fletcher-Goldfarb-Shanno (BFGS) est un algorithme de quasi-Newton, c'est-à-dire que l'inverse de la matrice Hessienne est approchée selon une formule donnée, uniquement à partir des gradients aux itérations précédentes. La direction de descente est définie par :

$$\boldsymbol{d}^{(k)} = -\mathbf{B}_k \nabla_x \mathcal{F}(\boldsymbol{x}^{(k)}), \tag{3.12}$$

où la matrice $\mathbf{B}_k$ est une approximation de l'inverse de la matrice Hessienne. Cette matrice est calculée à partir de $\mathbf{B}_{k-1}$, de la direction précédente et du gradient courant $\nabla_x \mathcal{F}(\boldsymbol{x}^{(k)})$. Les différents algorithmes de quasi-Newton se distinguent par la façon dont est calculée la matrice $\mathbf{B}_k$. L'intérêt de ces méthodes est qu'il n'y a pas de système linéaire à résoudre pour calculer la direction de descente, contrairement à la méthode de Newton, GN ou LM. Quelques auteurs ont utilisé cet algorithme pour la reconstruction en TMO (Franchois et Tijhuis, 2003; Hu et al., 2006; De Zaeytijd et al., 2007).

En revanche, comme les méthodes de Newton, GN et LM, la matrice $\mathbf{B}_k$ est de taille $N \times N$ et il n'est donc pas envisageable de la stocker directement en mémoire. Pour les problèmes de grande taille, une version permettant de limiter la mémoire nécessaire a été développée : L-BFGS (Limited-Memory BFGS) (Nocedal et Wright, 1999; Byrd et al., 1995). Cet algorithme permet de ne stocker qu'un certain nombre de vecteurs afin de représenter de façon implicite la matrice $\mathbf{B}_k$. Dans la suite de nos travaux, nous avons choisi d'utiliser l'algorithme L-BFGS.

3.5 Calcul du gradient pour le critère non-linéaire

Comme nous l'avons vu, la majorité des algorithmes itératifs requiert le calcul du gradient du critère afin d'évaluer une direction de descente. Dans notre cas, il est nécessaire d'obtenir une forme analytique du gradient du critère non-linéaire. Pour des raisons de simplicité, nous

omettons le terme de régularisation. Le gradient s'exprime sous la forme :

$$\nabla_x \mathcal{F} = -\sum_{i=1}^{N_S} \mathbf{J}_i^\dagger \left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o \mathbf{X} \mathbf{L}_x^{-1} \boldsymbol{e}_i^{\text{inc}} \right) \tag{3.13}$$

$$= -\mathbf{Q}^{\text{t}} \sum_{i=1}^{N_S} \text{diag}\left\{ \boldsymbol{e}_i^{\text{tot}} \right\}^\dagger \left(\overline{\mathbf{L}_x} \right)^{-1} \mathbf{G}_o^\dagger \left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o \mathbf{X} \mathbf{L}_x^{-1} \boldsymbol{e}_i^{\text{inc}} \right). \tag{3.14}$$

La matrice $\mathbf{J}_i = \mathbf{G}_o \left(\mathbf{L}_x^{\text{t}} \right)^{-1} \text{diag}\{\boldsymbol{e}_i^{\text{tot}}\} \mathbf{Q}$, de taille $N_R \times 3N$, dénote la matrice jacobienne du modèle direct pour la source i. La matrice $\mathbf{Q}$ de taille $3N \times N$ s'écrit sous la forme :

$$\mathbf{Q} = \begin{bmatrix} \mathbf{I}_N \\ \mathbf{I}_N \\ \mathbf{I}_N \end{bmatrix}, \tag{3.15}$$

où $\mathbf{I}_N$ est la matrice identité de taille $N \times N$. Le détail du calcul du gradient est donné en Annexe B. Notons que le calcul du gradient requiert plusieurs résolutions de systèmes faisant intervenir la matrice $\mathbf{L}_x$:

- N_S systèmes linéaires $\mathbf{L}_x^{-1} \boldsymbol{e}_i^{\text{inc}}$, avec $i = 1, \ldots, N_S$, afin de calculer les champs totaux $\boldsymbol{e}_i^{\text{tot}}$. Ces systèmes sont identiques à ceux apparaissant dans le critère ;
- des systèmes linéaires impliquant $\left(\overline{\mathbf{L}_x} \right)^{-1}$. Le nombre exact de systèmes à résoudre dépend de la mise en œuvre utilisée pour l'implémentation du calcul du gradient. Le chapitre 4 est consacré à ce calcul.

À chaque itération de l'algorithme de minimisation, le calcul du gradient est requis afin de déterminer une direction de descente. Nous venons de voir que le calcul du gradient pour un $\boldsymbol{x}$ donné nécessite la résolution d'un grand nombre de systèmes linéaires et ces opérations sont les plus coûteuses dans le calcul du gradient. De la même façon, la recherche de pas requiert l'évaluation du critère un certain nombre de fois ; la résolution des systèmes linéaires pour le calcul des champs totaux est aussi une opération coûteuse. Se concentrer sur ces résolutions de systèmes afin de réduire le coût de calcul de l'inversion semble donc naturel.

Dans un premier temps, la question du nombre de systèmes à résoudre se pose. Le calcul du critère nécessite la résolution de N_S systèmes. En revanche, nous venons de dire que le nombre de systèmes intervenant dans le gradient dépend de la mise en œuvre de celui-ci. Différentes approches permettent de calculer le jacobien et le gradient de façon efficace. Nous proposons dans le chapitre 4 une procédure générale permettant d'optimiser le nombre de systèmes linéaires à résoudre en fonction de la configuration du montage d'acquisition. Réduire le nombre de systèmes linéaires entraîne évidemment un gain de temps à chaque itération du problème inverse.

Dans un second temps, nous nous sommes focalisés sur l'accélération de la résolution de ces systèmes. Comme détaillé dans le chapitre 2, ces systèmes sont résolus par des méthodes itératives comme l'approche BiCGStab. Puisque tous ces systèmes font intervenir une même matrice $\mathbf{L}_x$, des approches par blocs semblent applicables. Ces approches par blocs ont pour intérêt de résoudre conjointement plusieurs systèmes linéaires impliquant une même matrice, ici $\mathbf{L}_x$. Le chapitre 5 propose une méthode efficace permettant d'appliquer une approche par blocs à la résolution de ces systèmes en TMO.

3.6 Conclusion

Pour résumer, nous avons fait le choix de voir le problème inverse comme la minimisation d'un critère pénalisé. Le terme d'adéquation aux données correspond à l'erreur quadratique entre le modèle direct non-linéaire et les mesures. La régularisation additive de type "$\ell_2\ell_1$" sur les différences premières a l'avantage de préserver les contours des objets de l'image. La fonction de contraste reconstruite est donnée par :

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \mathcal{F}(\boldsymbol{x}), \text{ avec } \mathcal{F}(\boldsymbol{x}) = \mathcal{F}^D(\boldsymbol{x}) + \mathcal{R}(\boldsymbol{x}) \tag{3.16}$$

Le terme d'adéquation aux données $\mathcal{F}^D$ est donné par :

$$\mathcal{F}^D(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{N_S} \|\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}\|^2 \tag{3.17}$$

et le terme de régularisation $\mathcal{R}$ par l'équation (3.6). La solution est obtenue par minimisation du critère par l'algorithme L-BFGS qui a l'avantage de ne pas utiliser le Hessien du critère.

Le calcul du critère et du gradient à chaque itération de l'algorithme d'optimisation requiert la résolution d'un grand nombre de systèmes linéaires impliquant la matrice $\mathbf{L}_x$. Les chapitres 4 et 5 proposent des outils permettant de réduire le nombre de ces systèmes et d'accélérer leurs résolutions.

CHAPITRE 4 ÉTUDE ALGÉBRIQUE DU GRADIENT ET RÉDUCTION DU NOMBRE DE SYSTÈMES LINÉAIRES

Nous avons vu précédemment que la reconstruction en tomographie micro-ondes passe généralement par la minimisation d'un critère où il est nécessaire de calculer le gradient afin d'estimer une direction de descente. Le calcul du critère et du gradient à chaque itération requiert la résolution d'un grand nombre de systèmes linéaires de grande taille. Ces multiples résolutions représentent la majeure partie du coût de calcul de la reconstruction. Plusieurs formulations différentes existent permettant de calculer le gradient et le nombre de systèmes linéaires à résoudre peut varier selon ces approches. De même, la configuration du montage d'acquisition peut influencer sur le nombre de systèmes intervenant dans le calcul du critère et du gradient. Nous proposons dans ce chapitre une procédure optimisant le nombre de systèmes à résoudre en s'appuyant sur deux outils : la mise en œuvre du calcul du gradient et le théorème de réciprocité de Lorentz. Il ne semble pas ou peu précisée dans la littérature le fait de pouvoir réduire le nombre de systèmes pour la résolution du problème inverse. C'est pourquoi nous présentons cette procédure qui peut s'appliquer à n'importe quel problème de reconstruction en TMO.

Dans un premier temps, nous allons nous concentrer sur la formulation matricielle du jacobien et du gradient, présentée dans le chapitre 3 et dans l'annexe B. Nous verrons que plusieurs mises en œuvre du calcul du gradient et du jacobien sont possibles permettant de réduire le nombre de systèmes linéaires à résoudre. Nous comparerons ensuite ce calcul aux approches présentes dans la littérature, notamment la formulation de l'état adjoint qui est une méthode très fréquente pour évaluer le gradient. Dans un second temps, nous présenterons le principe de réciprocité de Lorentz qui permet de considérer un montage d'acquisition adjoint. Nous verrons qu'une bonne utilisation de ce principe permet de réduire le nombre de systèmes linéaires intervenant dans le calcul du critère. Nous proposerons de s'appuyer sur quelques exemples concrets pour évaluer l'effet de l'utilisation de ce théorème. Enfin, à partir des deux outils présentés ci-dessus (mise en œuvre du gradient et principe de réciprocité de Lorentz), nous proposerons une procédure générale permettant d'optimiser le nombre de systèmes linéaires nécessaires au calcul du critère et du gradient. Nous étudierons quelques exemples de montage d'acquisition classiques dans la littérature et nous montrerons qu'il est possible d'optimiser le nombre de systèmes linéaires nécessaires.

Rappelons le critère non-linéaire d'adéquation aux données (ici, on omet le terme de régula-

risation afin d'alléger les formules) :

$$\mathcal{F}(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{N_S} \left\| \widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}} \right\|^2, \tag{4.1}$$

et le gradient correspondant :

$$\nabla_x \mathcal{F} = -\sum_{i=1}^{N_S} \mathbf{J}_i^{\dagger} \left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}} \right), \tag{4.2}$$

où les jacobiens $\mathbf{J}_i$ pour toutes les sources ($i = 1, \ldots, N_S$) valent :

$$\mathbf{J}_i = \mathbf{G}_o \left(\mathbf{L}_x^{\text{t}}\right)^{-1} \text{diag}\{\boldsymbol{e}_i^{\text{tot}}\}\, \mathbf{Q}. \tag{4.3}$$

4.1 Mise en œuvre du calcul du jacobien

Nous avons vu que de nombreux auteurs utilisent des algorithmes d'optimisation tels que Newton, Gauss-Newton et Levenberg-Marquardt ; ces approches nécessitent le calcul explicite du jacobien. Nous souhaitons dans cette section montrer qu'il est possible de calculer le jacobien de différentes façons.

En reprenant l'équation (4.3), le jacobien peut être exprimé sous la forme :

$$\mathbf{J}_i = \mathbf{G}_o \left(\mathbf{L}_x^{\text{t}}\right)^{-1} \Big(\text{diag}\{\boldsymbol{e}_i^{\text{tot}}\}\, \mathbf{Q}\Big). \tag{4.4}$$

Le calcul sous cette forme requiert le calcul de la matrice $\left(\mathbf{L}_x^{\text{t}}\right)^{-1} \Big(\text{diag}\{\boldsymbol{e}_i^{\text{tot}}\}\, \mathbf{Q}\Big)$ de taille $3N \times N$, soit la résolution de N systèmes linéaires de taille $3N \times 3N$. Pour des problèmes 3-D de grande taille, résoudre ces N systèmes à chaque itération du problème inverse est tout simplement inenvisageable. Par exemple, pour un volume discrétisé en $30 \times 30 \times 30$ voxels, le calcul du jacobien à chaque itération nécessiterait la résolution de 27 000 systèmes linéaires de taille $81\,000 \times 81\,000$!

Une astuce simple est de voir le calcul du jacobien sous une autre forme (accentuée par la présence des parenthèses) :

$$\mathbf{J}_i = \left(\mathbf{L}_x^{-1}\mathbf{G}_o^{\text{t}}\right)^{\text{t}} \text{diag}\{\boldsymbol{e}_i^{\text{tot}}\}\, \mathbf{Q}. \tag{4.5}$$

Avec cette mise en œuvre, le calcul de $\mathbf{L}_x^{-1}\mathbf{G}_o^{\text{t}}$ requiert la résolution de N_R systèmes linéaires de taille $3N \times 3N$. En pratique, le nombre de récepteurs N_R est bien plus faible que le nombre de voxels N : les problèmes en TMO prennent généralement en compte plusieurs dizaines voire centaines de milliers de voxels, alors que les montages d'acquisition utilisent quelques

dizaines voire centaines de récepteurs. Il est donc bien plus avantageux de mettre en œuvre le calcul du jacobien comme (4.5) plutôt que (4.4).

La mise en œuvre de l'équation (4.5) est appelé par certains la formulation de l'état adjoint (Belkebir et al., 2000; Abubakar et al., 2008b, 2012a). Cette notion d'état adjoint n'est pas clairement définie et deux visions sont établies dans la littérature :

1. pour certains auteurs, la formulation précédente du jacobien, donnée en (4.5), est appelée formulation de l'état adjoint : elle consiste à réduire le nombre de systèmes à résoudre dans le jacobien pour passer de N à N_R systèmes (Abubakar et al., 2012a) ;
2. pour d'autres il s'agit d'une mise en œuvre non pas au niveau de la matrice jacobienne, mais directement au niveau du gradient. Au lieu de résoudre N systèmes pour le calcul du gradient, la formulation de l'état adjoint propose de résoudre N_S systèmes (McGillivray et Oldenburg, 1990). Nous verrons cela dans la section 4.2 sur le calcul du gradient.

Dans cette section sur le calcul du jacobien, quand nous parlons d'état adjoint, il s'agit de la première définition. Nous allons maintenant voir comment cette formulation est présentée dans la littérature, puis nous allons faire le lien avec notre formulation matricielle du jacobien.

4.1.1 L'état adjoint pour le calcul du jacobien

Les auteurs utilisant la formulation de l'état adjoint pour le jacobien proposent une vision "physicienne" du problème. Ils présentent la variation du champ diffracté au niveau d'un récepteur comme le produit scalaire entre deux champs totaux : l'un provenant de la source et l'autre provenant du récepteur. La notion d'adjoint fait intervenir l'idée d'opérateur image allant de l'espace des données vers l'espace de l'objet, ce qui correspond au champ provenant du récepteur. Nous présentons dans un premier temps cette formulation tirée de Abubakar et al. (2012a) et nous établirons ensuite l'équivalence de cette formulation avec notre expression matricielle du jacobien.

Pour des raisons de simplicité dans la comparaison des formulations, nous travaillons ici avec la matrice jacobienne $\widetilde{\mathbf{J}}_i = \mathbf{G}_o \, (\mathbf{L}_x^{\mathrm{t}})^{-1} \, \mathrm{diag}\{\boldsymbol{e}_i^{\mathrm{tot}}\}$, de taille $N_R \times 3N$ (cf. Annexe B). La formulation de l'état adjoint, proposée par Abubakar et al. (2012a), est exprimé de façon continue :

$$\widetilde{\mathcal{J}}_{i,j,(p,\kappa)} = -j\omega\epsilon_b \int_{V_p} E_{i,\kappa}^{\mathrm{tot}}(\boldsymbol{r}) \cdot E_{j,\kappa}^{\mathrm{tot},R}(\boldsymbol{r}) \, d\boldsymbol{r}. \tag{4.6}$$

Le terme $\widetilde{\mathcal{J}}_{i,j,(p,\kappa)}$ dénote l'indice $(j,(p,\kappa))$ de la matrice jacobienne $\widetilde{\mathbf{J}}_i$ pour la source i. Pour rappel, $\widetilde{\mathbf{J}}_i$ est de taille $N_R \times 3N$: l'indice j représente le récepteur j et l'indice double (p,κ) représente la composante κ du champ sur le voxel p, soit $3N$ valeurs. κ vaut x, y ou z. V_p

représente le volume du voxel p. Le terme $E_{i,\kappa}^{\text{tot}}$ correspond à la composante κ du champ total $\overrightarrow{\boldsymbol{E}}_i^{\text{tot}}$ provenant de la source i. L'expression $E_{j,\kappa}^{\text{tot},R}$ correspond à la composante κ du champ total virtuel $\overrightarrow{\boldsymbol{E}}_j^{\text{tot},R}$ diffracté par l'objet provenant d'une source placée au récepteur j. L'exposant R permet d'accentuer le fait qu'il s'agisse d'un champ provenant d'un récepteur et non d'une source. La formulation de l'état adjoint décrit cette source virtuelle comme orientée dans la même direction que le récepteur, d'amplitude 1 et de phase nulle. Ces champs provenant des récepteurs sont généralement appelés champs adjoints.

Comme on le voit dans l'équation (4.6), le calcul du jacobien nécessite de calculer :

— des champs totaux $\overrightarrow{\boldsymbol{E}}_i^{\text{tot}}$ provenant des N_S sources. Dans une formulation discrète, ceux-ci correspondent aux N_S champs $\boldsymbol{e}_i^{\text{tot}} = \mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}$ intervenant dans le modèle direct,

— des champs totaux adjoints $\overrightarrow{\boldsymbol{E}}_j^{\text{tot},R}$ provenant des N_R récepteurs. Il est donc nécessaire de définir N_R champs incidents adjoints $\overrightarrow{\boldsymbol{E}}_j^{\text{inc},R}$ provenant des récepteurs, puis d'en estimer les champs totaux adjoints correspondants.

4.1.2 Équivalence entre l'état adjoint et la formulation matricielle du jacobien

Nous allons maintenant montrer que la formulation matricielle (4.5) et celle de l'état adjoint (4.6) sont équivalentes. Partons de la formulation de l'état adjoint définie dans l'équation (4.6). On suppose que le volume V est discrétisé en N voxels de volume ΔV et que les champs sont constants sur chacun des voxels. Ce choix de discrétisation est identique à celui défini dans la méthode des moments au chapitre 2. On exprime alors le jacobien sous la forme :

$$\widetilde{\mathcal{J}}_{i,j,(p,\kappa)} = -j\omega\epsilon_b\Delta V\; e_{i,(p,\kappa)}^{\text{tot}}\, e_{j,(p,\kappa)}^{\text{tot},R}. \tag{4.7}$$

Comme les champs totaux sont constants sur les voxels, l'intégrale dans (4.6) disparaît et les scalaires $e_{i,(p,\kappa)}^{\text{tot}}$ et $e_{j,(p,\kappa)}^{\text{tot},R}$ représentent respectivement les valeurs de la composante κ des champs totaux réel et adjoint dans le voxel p. En concaténant ce résultat pour tous les voxels du domaine et les trois composantes des champs, on obtient la j-ème ligne $\widetilde{\mathcal{J}}_{i,j}$ du jacobien $\widetilde{\mathbf{J}}_i$ qu'on peut écrire comme

$$\widetilde{\mathcal{J}}_{i,j} = -j\omega\epsilon_b\Delta V\; (\boldsymbol{e}_j^{\text{tot},R})^{\text{t}}\,\text{diag}\{\boldsymbol{e}_i^{\text{tot}}\}, \tag{4.8}$$

où $\boldsymbol{e}_i^{\text{tot}}$ et $\boldsymbol{e}_j^{\text{tot},R}$ sont des vecteurs colonnes de taille $3N$ contenant les trois composantes des champs totaux pour tous les voxels du domaine. On retrouve la même notation que celle utilisée dans les chapitres précédents.

Exprimons maintenant le champ total adjoint : le champ total $\boldsymbol{e}_j^{\text{tot},R}$ provient d'un champ

incident qu'on nomme $\boldsymbol{e}_j^{\mathrm{inc},R}$, version discrétisée du champ incident adjoint $\overrightarrow{\boldsymbol{E}}_j^{\mathrm{inc},R}$. Ce champ incident suit l'équation intégrale :

$$\boldsymbol{e}_j^{\mathrm{tot},R} = \mathbf{L}_x^{-1} \boldsymbol{e}_j^{\mathrm{inc},R}. \tag{4.9}$$

Comme nous l'avons vu dans la sous-section 4.1.1, la formulation de l'état adjoint énonce que le champ incident virtuel provient d'une source ponctuelle au niveau du récepteur j, orientée dans la même direction que celui-ci, d'amplitude unitaire et de phase nulle. On peut formaliser ce champ incident par l'équation intégrale suivante :

$$\overrightarrow{\boldsymbol{E}}_j^{\mathrm{inc},R}(\boldsymbol{r}) = j\omega\mu_b \, \overline{\overline{\mathbf{G}}}(\boldsymbol{r}, \boldsymbol{r}_j^R) \, \overrightarrow{\boldsymbol{J}}_j^R. \tag{4.10}$$

On rappelle que $\overline{\overline{\mathbf{G}}}$ est le tenseur dyadique de Green. On définit $\overrightarrow{\boldsymbol{J}}_j^R$ la densité de courant de la source virtuelle positionnée au niveau du récepteur j. Prenons par exemple un récepteur j faisant l'acquisition du champ selon l'axe z, la formulation de l'état adjoint établit que la densité de courant vaut $\overrightarrow{\boldsymbol{J}}_j^R = [0, 0, 1]^{\mathrm{t}}$, soit :

$$\overrightarrow{\boldsymbol{E}}_j^{\mathrm{inc},R}(\boldsymbol{r}) = j\omega\mu_b \, \overline{\overline{\mathbf{G}}}(\boldsymbol{r}, \boldsymbol{r}_j^R) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{4.11}$$

$$= j\omega\mu_b \, \overline{\mathbf{G}}_z(\boldsymbol{r}, \boldsymbol{r}_j^R), \tag{4.12}$$

où $\overline{\mathbf{G}}_z$ correspond à la troisième colonne du tenseur $\overline{\overline{\mathbf{G}}}$. En concaténant les valeurs de $\overrightarrow{\boldsymbol{E}}_j^{\mathrm{inc},R}(\boldsymbol{r})$ en tous les voxels, on obtient alors

$$\boldsymbol{e}_j^{\mathrm{inc},R} = j\omega\mu_b \, \boldsymbol{g}_j^R. \tag{4.13}$$

Le vecteur $\boldsymbol{g}_j^R$ de taille $3N$ provient de la discrétisation du tenseur de Green entre le récepteur positionné en $\boldsymbol{r}_j^R$ et tous les voxels du domaine V. Comme le tenseur de Green est symétrique, ce vecteur correspond exactement au vecteur $\boldsymbol{u}_j^R$ donné dans l'annexe A. Ce vecteur représente le tenseur de Green entre les voxels du domaine et la position du récepteur j. L'équation (4.9) du champ total virtuel devient alors :

$$\boldsymbol{e}_j^{\mathrm{tot},R} = j\omega\mu_b \, \mathbf{L}_x^{-1} \boldsymbol{u}_j^R \tag{4.14}$$

et la j-ème ligne $\widetilde{\mathcal{J}}_{i,j}$ du jacobien $\widetilde{\mathbf{J}}_i$ s'écrit finalement :

$$\widetilde{\mathcal{J}}_{i,j} = k_b^2 \Delta V \, (\mathbf{L}_x^{-1} \boldsymbol{u}_j^R)^{\mathrm{t}} \, \mathrm{diag}\{\boldsymbol{e}_i^{\mathrm{tot}}\}. \tag{4.15}$$

En concaténant toutes les lignes du jacobien $\widetilde{\mathbf{J}}_i$ correspondant aux récepteurs, revenons à l'expression matricielle du jacobien pour la source i :

$$\begin{aligned}\widetilde{\mathbf{J}}_i &= (\mathbf{L}_x^{-1}\mathbf{G}_o^{\mathrm{t}})^{\mathrm{t}}\ \mathrm{diag}\{\boldsymbol{e}_i^{\mathrm{tot}}\} \\ &= \mathbf{G}_o\ \left(\mathbf{L}_x^{\mathrm{t}}\right)^{-1}\mathrm{diag}\{\boldsymbol{e}_i^{\mathrm{tot}}\}. \end{aligned} \tag{4.16}$$

On retrouve bien la formulation matricielle du jacobien. Finalement, on peut voir la matrice $\mathbf{G}_o$ comme l'ensemble des champs incidents virtuels mesurés au niveau des N voxels et provenant des N_R récepteurs. Le terme $\mathbf{L}_x^{-1}\mathbf{G}_o^{\mathrm{t}}$ représente alors la version discrétisée des N_R champs adjoints $\overrightarrow{\boldsymbol{E}}_j^{\mathrm{tot},R}$.

En conclusion, la formulation de l'état adjoint pour le calcul du jacobien permet de réduire le nombre de systèmes linéaires à résoudre. Ce calcul nécessite donc la résolution de $N_S + N_R$ systèmes linéaires pour calculer les champs totaux réels et adjoints.

4.2 Mises en œuvre du calcul du gradient

Contrairement aux méthodes d'optimisation du second ordre, les méthodes du premier ordre ou les méthodes de quasi-Newton ne nécessitent pas forcément le calcul explicite du jacobien. Les auteurs qui utilisent des algorithmes nécessitant le calcul du gradient sans avoir besoin du jacobien mettent en œuvre celui-ci par des opérations algébriques moins coûteuses que pour le calcul du jacobien (McGillivray et Oldenburg, 1990). En revanche, selon la configuration du montage d'acquisition, nous verrons qu'il peut être avantageux de tout de même calculer le jacobien pour exprimer le gradient. On peut exprimer le gradient sous la forme suivante :

$$\nabla_x \mathcal{F} = -\mathbf{Q}^{\mathrm{t}} \sum_{i=1}^{N_S} \mathrm{diag}\left\{\boldsymbol{e}_i^{\mathrm{tot}}\right\}^{\dagger} (\overline{\mathbf{L}_x})^{-1}\mathbf{G}_o^{\dagger}\left(\widetilde{\boldsymbol{e}}_i^{\mathrm{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\mathrm{inc}}\right). \tag{4.17}$$

On distingue deux mises en œuvre que l'on détaille ci-dessous : une mise en œuvre que l'on appelle "mise en œuvre directe" et une mise en œuvre adjointe du jacobien.

4.2.1 Mise en œuvre directe

La première solution de mise en œuvre consiste à voir l'implémentation du calcul du gradient en plaçant les parenthèses comme suit :

$$\nabla_x \mathcal{F} = -\mathbf{Q}^{\mathrm{t}} \sum_{i=1}^{N_S} \mathrm{diag}\left\{\boldsymbol{e}_i^{\mathrm{tot}}\right\}^{\dagger} (\overline{\mathbf{L}_x})^{-1}\left[\mathbf{G}_o^{\dagger}\left(\widetilde{\boldsymbol{e}}_i^{\mathrm{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\mathrm{inc}}\right)\right]. \tag{4.18}$$

Dans ce cas, N_S systèmes linéaires $(\overline{\mathbf{L}_x})^{-1}\left[\mathbf{G}_o^\dagger\left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}\right)\right]$ sont à résoudre. Avec cette mise en oeuvre, le calcul du gradient requiert la résolution de $2N_S$ systèmes de taille $3N \times 3N$:

— N_S systèmes pour estimer les champs totaux $\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}$;

— N_S systèmes $(\overline{\mathbf{L}_x})^{-1}\left[\mathbf{G}_o^\dagger\left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}\right)\right]$.

Nous appelons cette mise en œuvre, la mise en œuvre directe.

Cette formulation est également appelée formulation de l'état adjoint par certains auteurs qui y voient l'intérêt de ne pas calculer le jacobien (McGillivray et Oldenburg, 1990; Plessix, 2006; Rekanos et al., 1999). Cette méthode est souvent utilisée pour calculer le gradient d'un critère des moindres carrés non-linéaire.

4.2.2 Mise en œuvre adjointe du jacobien

L'autre approche de mise en œuvre est de voir le gradient comme suit :

$$\nabla_x \mathcal{F} = -\mathbf{Q}^{\text{t}} \sum_{i=1}^{N_S} \text{diag}\left\{\boldsymbol{e}_i^{\text{tot}}\right\}^\dagger \overline{\left(\mathbf{L}_x^{-1}\mathbf{G}_o^{\text{t}}\right)} \left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}\right), \tag{4.19}$$

où les parenthèses placées autour de la matrice $\mathbf{L}_x^{-1}\mathbf{G}_o^{\text{t}}$ mettent l'accent sur la façon dont les systèmes sont résolus. Dans cette configuration, la résolution de N_R systèmes linéaires est nécessaire. On retrouve exactement l'approche de l'état adjoint pour le calcul du jacobien vue précédemment où la matrice $\mathbf{L}_x^{-1}\mathbf{G}_o^{\text{t}}$ représente les champs totaux adjoints. Le calcul du gradient nécessite au total la résolution de $(N_S + N_R)$ systèmes linéaires de taille $3N \times 3N$ pour estimer les N_S champs totaux $\boldsymbol{e}_i^{\text{tot}}$ et les N_R champs adjoints. Nous appelons cette mise en œuvre, la mise en œuvre adjointe du jacobien.

4.2.3 Choix de la formulation en fonction du montage d'imagerie

Nous avons vu qu'un certain nombre d'auteurs utilise la formulation adjointe du jacobien pour calculer la matrice jacobienne et le gradient. L'approche directe semble tout de même moins fréquente dans la littérature et la question de la mise en œuvre du calcul du gradient ne semble pas avoir fait l'objet d'études. Comme ces résolutions de systèmes concentrent la majeure partie du temps de calcul, il semble pertinent de prendre en compte cette mise en œuvre et de proposer une solution optimisée en fonction du montage d'acquisition.

Le choix de la mise en œuvre dépend du montage d'imagerie utilisé, plus précisément du nombre de sources et de récepteurs. Deux configurations sont possibles :

— $N_S < N_R$: le nombre de sources est inférieur au nombre de récepteurs. Dans ce cas, il est plus avantageux d'utiliser la mise en œuvre directe donnée par l'équation (4.18) où $2N_S$ systèmes doivent être résolus pour le calcul du gradient.

— $N_S > N_R$: le nombre de sources est supérieur au nombre de récepteurs. La mise en œuvre adjointe (4.19) est privilégiée. Au total, la résolution de $(N_S + N_R)$ systèmes linéaires est nécessaire pour calculer le gradient.

Dans le cas où $N_S = N_R$, il y autant de sources que de récepteurs. Quelle que soit l'approche utilisée, la résolution de $2N_S$ systèmes linéaires est requise. Finalement, en fonction de la configuration du montage d'acquisition, il est nécessaire de choisir la bonne formulation de façon à optimiser le nombre de systèmes à résoudre et donc réduire le coût de calcul du gradient. Nous verrons dans la dernière section de ce chapitre quelques montages utilisés dans la littérature et comment il est possible de réduire le nombre de systèmes à résoudre sur ces exemples précis.

4.3 Théorème de réciprocité de Lorentz

Nous avons vu qu'il est possible d'accélérer le calcul du gradient en choisissant la mise en œuvre du calcul du gradient la plus efficace selon le montage d'acquisition. En général, la reconstruction s'appuie directement sur le montage utilisé lors de l'acquisition, notamment sur la répartition des sources et des récepteurs : les champs incidents $\boldsymbol{e}_i^{\text{inc}}$ utilisés concordent avec les sources réelles du montage et les données concordent avec les mesures réalisées aux récepteurs réels. En revanche, certains auteurs utilisent le théorème de réciprocité de Lorentz, qui permet de considérer un problème adjoint où les sources physiques jouent le rôle de récepteurs virtuels et où les récepteurs physiques jouent le rôle de sources virtuelles. Nous verrons dans cette section qu'une bonne utilisation de ce théorème permet également de réduire le nombre de systèmes linéaires à résoudre pour le calcul du critère.

Le théorème de réciprocité de Lorentz énonce que la relation entre une source harmonique et un point de mesure de son champ électrique est inchangée si on intervertit les positions de la source et du point de mesure (Balanis, 2012). On comprend que ce principe de réciprocité est lié au concept d'état adjoint où les points de mesure jouent le rôle de sources virtuelles et inversement. Physiquement, cela revient à dire que si une densité de courant ponctuelle $\overrightarrow{\boldsymbol{J}}_1(\boldsymbol{r}_1)$ crée un champ électrique $\overrightarrow{\boldsymbol{E}}_1(\boldsymbol{r})$ et une autre densité de courant ponctuelle $\overrightarrow{\boldsymbol{J}}_2(\boldsymbol{r}_2)$

crée un champ électrique $\overrightarrow{\boldsymbol{E}}_2(\boldsymbol{r})$, on a alors la relation :

$$\overrightarrow{\boldsymbol{J}}_1(\boldsymbol{r}_1) \cdot \overrightarrow{\boldsymbol{E}}_2(\boldsymbol{r}_1) = \overrightarrow{\boldsymbol{J}}_2(\boldsymbol{r}_2) \cdot \overrightarrow{\boldsymbol{E}}_1(\boldsymbol{r}_2). \tag{4.20}$$

Si on applique ce théorème à un montage d'acquisition en tomographie micro-ondes, cela revient à inverser le rôle des sources et des récepteurs dans le modèle utilisé et les nouvelles données de champs diffractés avec le montage virtuel correspondent simplement aux données mesurées $\widetilde{\boldsymbol{e}}_i^{\text{scat}}$ à un coefficient près. Appliquer le théorème de réciprocité permet donc de considérer un nouveau montage d'acquisition dans lequel le nombre de sources et de récepteurs est inversé. C'est notamment utile dans le cas où le nombre de sources est plus élevé que le nombre de récepteurs. Prenons ce cas où $N_S > N_R$ et étudions le nombre de systèmes à résoudre pour évaluer le critère, le gradient et le jacobien :

- sans le théorème de réciprocité : d'après l'étude sur le calcul du gradient, comme $N_S > N_R$, la mise en œuvre la plus efficace est la mise en œuvre adjointe du jacobien (4.19), soit $N_S + N_R$ systèmes à résoudre. Si le jacobien est requis pour estimer une direction de descente, la formulation adjointe (4.5) est dans tous les cas nécessaire, soit $N_S + N_R$. Enfin, le calcul du critère pour estimer un pas de descente nécessite N_S résolutions ;
- avec le théorème de réciprocité : on se ramène à un montage adjoint de N_R sources virtuelles et N_S récepteurs virtuels. Le calcul du gradient nécessite au mieux $2N_R$ systèmes à résoudre en appliquant la mise en œuvre directe (4.18). Si le jacobien est requis, il est nécessaire de résoudre $N_R + N_S$ systèmes pour évaluer celui-ci. Enfin, la recherche de pas nécessite la résolution de N_R systèmes pour évaluer le critère.

L'utilisation du théorème de réciprocité quand $N_S > N_R$ permet donc de réduire aussi bien le coût de calcul du critère que celui du gradient. Notons que le nombre de systèmes à résoudre pour le calcul du jacobien est identique, avec ou sans l'utilisation du principe de Lorentz. Une procédure générale sera présentée dans la section 4.4.

Prenons maintenant l'exemple concret du montage utilisé pour l'acquisition de la base de données de l'institut Fresnel (Geffrin et Sabouroux (2009)). Le montage expérimental est composé de 81 sources qui émettent une onde incidente selon deux polarisations et 36 récepteurs récupèrent uniquement la composante verticale du champ électrique (selon z). Considérons une seule polarisation : $N_S = 81$. Comme nous sommes dans le cas $N_S > N_R$, il est possible d'appliquer le théorème de réciprocité de Lorentz afin de réduire le nombre de systèmes linéaires. En envisageant la configuration adjointe, cela revient à considérer 36 sources virtuelles placées au niveau des récepteurs émettant un champ polarisé selon z et 81 récepteurs virtuels au niveau des sources acquérant la composante correspondant à la polarisation du champ incident. Le nombre de sources est alors fortement réduit passant de 81 sources à

36 sources.

Le principe de réciprocité de Lorentz est donc une solution permettant de réduire le nombre de systèmes linéaires en profitant d'une représentation adjointe du montage d'acquisition. Comme nous l'avons vu, il peut être appliqué à n'importe quelle configuration d'imagerie dans laquelle le nombre de sources est plus grand que le nombre de récepteurs.

4.4 Procédure finale d'optimisation du nombre de systèmes

Cette section propose un condensé des deux techniques vues dans ce chapitre afin de réduire au mieux le nombre de systèmes linéaires à résoudre pour le calcul du critère et du gradient : l'utilisation ou non du théorème de réciprocité de Lorentz et le choix de la mise en œuvre du calcul du gradient. Ici, nous proposons un schéma pour lequel l'évaluation du jacobien n'est pas nécessaire. Nous verrons plus bas la stratégie similaire à adopter si le jacobien est requis pour estimer une direction de descente. Voici le schéma à suivre en fonction du nombre de sources et de récepteurs dans la configuration d'acquisition :

— $N_S < N_R$: comme le nombre de sources est inférieur au nombre de récepteurs, il n'est pas utile d'appliquer le théorème de réciprocité. Concernant le calcul du gradient, il est avantageux d'utiliser la mise en œuvre directe (4.18) où $2N_S$ systèmes sont à résoudre. Le calcul du critère nécessite quant à lui N_S résolutions.

— $N_S > N_R$: on applique le théorème de réciprocité de Lorentz afin d'intervertir les sources et les récepteurs. Dans la nouvelle configuration virtuelle, nous avons N_R sources et N_S récepteurs. Comme le nombre de sources virtuelles est inférieur au nombre de récepteurs virtuels, on utilise la mise en œuvre directe (4.18) afin d'avoir à résoudre $2N_R$ systèmes. Le calcul du critère nécessite N_R résolutions.

— $N_S = N_R$: dans cette configuration, toutes les possibilités reviennent à résoudre le même nombre de systèmes. L'utilisation du principe de réciprocité n'apporte donc aucun avantage. De même, le choix de la mise en œuvre du gradient n'influence pas le nombre de systèmes : $2N_S = 2N_R$ résolutions requises pour le gradient et $N_S = N_R$ systèmes pour le critère.

En conclusion, on remarque que quelle que soit la configuration du montage d'acquisition, la formulation de l'état adjoint n'est jamais utilisée si l'on souhaite minimiser le nombre de systèmes linéaires à résoudre. Bien que cette formulation soit souvent utilisée pour calculer

le gradient, aucune étude ne semble avoir été proposée dans la littérature. Nous avons donc montré avec cette étude algébrique du calcul du gradient que la formulation de l'état adjoint n'est pas efficace et qu'il est possible de réduire fortement le coût de calcul du critère et du gradient en choisissant correctement la configuration optimale.

En ce qui concerne les méthodes itératives du second ordre nécessitant le calcul du jacobien, la formulation adjointe (4.5) est obligatoire afin d'obtenir une forme explicite du jacobien. Le calcul du gradient et du jacobien requiert donc la résolution de $N_S + N_R$ systèmes, peu importe l'utilisation ou non du principe de réciprocité. En revanche, il est tout de même utile d'appliquer ce principe si $N_S > N_R$ car il permet de réduire le nombre de systèmes à résoudre pour le calcul du critère.

4.5 Exemple de configurations présentes dans la littérature

Dans cette section, différentes configurations présentes dans la littérature sont détaillées et nous illustrons sur ces exemples l'efficacité de la procédure proposée dans la section 4.4.

4.5.1 Montage générique d'acquisition

Il est très fréquent de voir, dans la littérature, des montages (simulés ou réels) comportant autant d'antennes émettrices que d'antennes réceptrices. C'est notamment le cas pour les montages utilisés en TMO pour la détection du cancer du sein : les antennes servent à tour de rôle d'émetteurs et de récepteurs. De nombreux auteurs présentent des résultats avec ce type de configuration (Meaney et al., 2007; Rubaek et al., 2009; Abubakar et al., 2012a; Grzegorczyk et al., 2012). On peut distinguer deux cas : le cas où les récepteurs mesurent une seule composante du champ électrique et le cas où les trois composantes du champ sont mesurées.

Dans le cas où seule une composante du champ électrique est acquise, la situation proposée est celle dans laquelle le nombre de sources est égale au nombre de récepteurs, soit $N_S = N_R$. D'après les conclusions des sections précédentes, il n'y a pas d'intérêt à utiliser le théorème de réciprocité de Lorentz. De plus, quelque soit la mise en œuvre, il faudra $2N_S$ résolutions de systèmes pour calculer le gradient.

Si maintenant les trois composantes du champ électrique sont mesurées au niveau des récepteurs, cela revient à acquérir trois fois plus de données. Augmenter le nombre de mesures est bien sûr un avantage pour la reconstruction. La configuration est alors de N_S sources et $N_R = 3N_S$ récepteurs (trois récepteurs par position d'antennes). Avec ce montage, N_S est plus faible que N_R, il est donc logique d'utiliser la mise en œuvre directe afin d'avoir $2N_S$

systèmes à résoudre, plutôt que $N_S + 3N_R = 4N_S$ systèmes.

4.5.2 Montage simulé de Zhang et Liu (2015)

Le montage d'acquisition utilisé par (Zhang et Liu, 2015) est un exemple simulé dans lequel le nombre de sources est supérieur au nombre de récepteurs : $N_S = 63$ et $N_R = 36$. L'algorithme utilisé par les auteurs est une variante de DBIM ; le jacobien est donc nécessaire. Le calcul du gradient et du jacobien requiert la formulation adjointe du jacobien et conduit à une résolution de $N_S + N_R = 99$ systèmes. Le calcul du critère nécessite $N_S = 63$ résolutions. En conservant l'algorithme d'optimisation que les auteurs utilisent, il est intéressant d'utiliser le principe de réciprocité afin de réduire le nombre de systèmes pour le calcul du critère, passant de 63 à 36. Le coût de calcul pour le jacobien reste en revanche inchangé : $N_S + N_R = 99$ systèmes.

Si nous souhaitons appliquer un algorithme d'optimisation où seul le gradient est requis (L-BFGS par exemple), appliquer la procédure définie dans la section précédente permet une réduction importante du nombre de systèmes à résoudre : le calcul du gradient requiert la résolution de $2N_R = 72$ systèmes et le calcul du critère en nécessite $N_R = 36$.

4.5.3 Configuration d'acquisition des données Fresnel

Reprenons le montage utilisé pour l'acquisition des données de la base Fresnel décrit dans Geffrin et Sabouroux (2009). Il s'agit d'un ensemble de 81 antennes émettrices émettant des ondes polarisées selon deux directions. Nous considérons ici les deux polarisations, soit $N_S = 162$ ondes incidentes. L'acquisition est réalisée sur un ensemble de 36 récepteurs orientés selon z, soit $N_R = 36$. Lors de la publication de ces données dans la communauté scientifique, six équipes de chercheurs ont proposé leurs méthodes et résultats de reconstruction dans la revue *Inverse Problems 25*. Parmi eux, certains ont considéré la configuration initiale du montage sans utiliser le principe de réciprocité (Eyraud et al., 2009; Li et al., 2009; Yu et al., 2009). Quel que soit l'algorithme d'inversion utilisé, le nombre de problèmes directs à résoudre est égal à $N_S = 162$. Pour réduire le coût de calcul, certains ont préféré ne considérer qu'une seule polarisation : le nombre de sources N_S est divisé par deux, soit $N_S = 81$. Évidemment, la quantité d'information est divisée par deux et la reconstruction devient plus difficile. Le calcul du gradient correspondant en utilisant la mise en œuvre adjointe conduit à la résolution de $N_S + N_R = 198$ systèmes.

En utilisant le théorème de réciprocité de Lorentz, on se ramène à une configuration adjointe composée de 36 sources virtuelles et 162 récepteurs virtuels. Parmi les six méthodes proposées

pour reconstruire les données Fresnel, uniquement deux ont utilisé le théorème de réciprocité (Catapano et al., 2009; Chaumet et Belkebir, 2009). Cela montre bien que l'utilisation de ce principe n'est pas systématique. En utilisant la mise en œuvre directe, cela revient à résoudre $2 \times 36 = 72$ systèmes pour calculer le gradient, quel que soit le nombre de récepteurs. On observe alors un gain de 63% par rapport à une non-utilisation du théorème de réciprocité.

Dans le cas de méthodes d'optimisation nécessitant le jacobien (De Zaeytijd et Franchois, 2009; Yu et al., 2009), nous avons vu que l'utilisation du principe de réciprocité n'est pas utile pour le calcul du jacobien et du gradient car celui-ci requiert la résolution de $N_S + N_R$ systèmes. En revanche, il est tout de même intéressant de l'appliquer car il permet de réduire le coût de calcul du critère, nécessaire à la recherche de pas.

4.6 Conclusion

Dans ce chapitre, nous avons étudié deux outils permettant de réduire le nombre de systèmes linéaires à résoudre dans l'inversion. Le premier est basé sur une étude algébrique du gradient et du jacobien. Différentes mises en œuvre ont été proposées. Cette étude a notamment montré que la formulation de l'état adjoint, largement utilisée en TMO, n'est pas optimale et qu'il est possible de réduire efficacement le nombre de systèmes selon la configuration du montage. Le second outil est l'utilisation du principe de réciprocité de Lorentz. Nous avons vu que si le nombre de sources est supérieur au nombre de récepteurs, l'application de ce principe permet de considérer une configuration adjointe qu mène à une réduction du nombre de systèmes à résoudre pour le calcul du critère et du gradient.

Finalement, nous avons proposé une procédure permettant d'optimiser le nombre de systèmes. Sur plusieurs exemples concrets comme les données expérimentales de l'institut Fresnel, nous avons démontré qu'il est possible de réduire de façon importante le nombre de systèmes à résoudre.

CHAPITRE 5 APPROCHES ITÉRATIVES PAR BLOCS POUR LA RÉSOLUTION DE SYSTÈMES MULTIPLES

Comme vu dans les chapitres précédents, les calculs du critère (3.5) (p. 37) et du gradient (3.14) (p. 45) sont nécessaires à l'évaluation d'une direction et d'un pas de descente dans les algorithmes itératifs d'optimisation locale. Ces calculs requièrent la résolution d'un grand nombre de systèmes linéaires impliquant la matrice $\mathbf{L}_x$ de taille $3N \times 3N$. À cause de la grande taille de cette matrice, en particulier en 3-D, ces résolutions sont très coûteuses et concentrent la majorité du temps de calcul dans l'inversion. Dans le chapitre 4, nous avons proposé une procédure permettant de réduire le nombre de systèmes à résoudre en choisissant la mise en œuvre adéquate pour le calcul du gradient et en considérant une configuration adjointe du montage d'origine.

Nous souhaitons maintenant étudier la possibilité d'accélérer ces résolutions de systèmes. Pour cela, l'idée mise en avant dans ce chapitre est de profiter de l'avantage que tous ces systèmes impliquent la même matrice $\mathbf{L}_x$. Les algorithmes itératifs par blocs sont des méthodes exploitant cette propriété de matrice commune pour résoudre conjointement plusieurs systèmes linéaires. Nous noterons dans la suite la résolution des multiples problèmes directs sous la forme :

$$\mathbf{L}_x \mathbf{E}^{\text{tot}} = \mathbf{E}^{\text{inc}}, \tag{5.1}$$

avec $\mathbf{E}^{\text{tot}} = \left[\boldsymbol{e}_1^{\text{tot}}, \ldots, \boldsymbol{e}_{N_S}^{\text{tot}}\right]$ et $\mathbf{E}^{\text{inc}} = [\boldsymbol{e}_1^{\text{inc}}, \ldots, \boldsymbol{e}_{N_S}^{\text{inc}}]$. Cette notation permet de mettre l'accent sur la résolution conjointe des systèmes. La matrice du système sera appelée *matrice opérateur* et les vecteurs connus $\mathbf{E}^{\text{inc}}$ les *vecteurs image* (en anglais, *Right-Hand Side (RHS) vectors*). On notera également $\mathbf{F}$ les solutions des systèmes supplémentaires intervenant dans le calcul du gradient. Comme nous l'avons vu dans le chapitre précédent, en fonction de la mise en œuvre du calcul du gradient, le nombre de systèmes et les systèmes peuvent varier (cf. équations (4.18) et (4.19), p. 52). Nous noterons tout de même de façon générale :

$$\mathbf{F} = \mathbf{L}_x^{-1} \mathbf{F}^0, \tag{5.2}$$

où $\mathbf{F}^0$ vaut soit $\mathbf{G}_o^{\text{t}}$ si la mise en œuvre adjointe du jacobien est utilisée, soit les vecteurs $\overline{\mathbf{G}_o^{\dagger}\left[\widetilde{\boldsymbol{e}_i}^{\text{scat}} - \mathbf{G}_o \mathbf{X} \boldsymbol{e}_i^{\text{tot}}\right]}$ si la mise en œuvre directe est utilisée.

Ce chapitre se décompose comme suit :

— la section 5.1 présente les approches par blocs existantes et leur intérêt. Le choix de travailler avec l'approche Block-BiCGStab est justifié ;

— la seconde section détaille l'algorithme Block-BiCGStab et propose de l'adapter à une utilisation efficace pour la résolution des problèmes directs en tomographie micro-ondes. Quelques résultats préliminaires permettent de valider le choix du réglage des paramètres ;
— la section 5.3 introduit une généralisation de l'algorithme Block-BiCGStab pour une résolution efficace sur une architecture multi-cœurs. Cette approche est appelée Partial-Block BiCGStab ;
— dans la section 5.4, nous proposons une mise en œuvre efficace de cette approche pour une utilisation dans l'algorithme d'optimisation réalisant l'inversion ;
— la section 5.5 propose de comparer les performances des algorithmes BiCGStab, Block-BiCGStab et Partial-Block BiCGStab sur la résolution des problèmes directs et d'étudier l'effet de certains paramètres sur le coût de calcul, comme le contraste ou la fréquence ;
— enfin, la dernière section présente des résultats de reconstruction et une comparaison des coûts de calcul entre l'utilisation du BiCGStab, Block-BiCGStab et du Partial-Block BiCGStab dans l'inversion.

5.1 Les approches de résolution de systèmes par blocs

Le principe des algorithmes itératifs par blocs est de résoudre conjointement plusieurs systèmes d'équations linéaires où la matrice opérateur est commune à tous les systèmes. La plupart des méthodes itératives connues ont leurs équivalents par blocs. Ces versions par blocs ont été développées dans le cadre des mathématiques appliquées. On peut notamment citer :

— le Block-CG, la version par blocs du gradient conjugué (O'Leary, 1980) ;
— le Block-BiCG, la version par blocs du gradient biconjugué (O'Leary, 1980) ;
— le Block-QMR, la version par blocs de l'algorithme Quasi-Minimal Residual (Freund et Malhotra, 1997; Simoncini, 1997) ;
— le Block-MINRES, la version par blocs de la méthode de minimisation du résidu (Soodhalter, 2015) ;
— le Block-GMRES, la version par blocs de la généralisation de la méthode de minimisation du résidu (Simoncini et Gallopoulos, 1996) ;
— le Block-BiCGStab, la version par blocs de l'algorithme du gradient biconjugué stabilisé (El Guennouni et al., 2003).

Ces approches par blocs ont été développées à l'origine comme un outil mathématique de résolution de systèmes multiples, sans avoir de réelles applications pratiques. Les résultats obtenus dans les publications citées précédemment sont généralement présentés sur des systèmes linéaires simples (matrices creuses, matrices tridiagonales, ...) avec des vecteurs image peu nombreux et aléatoires.

Pour ces différents algorithmes itératifs par blocs, l'intérêt de résoudre conjointement plusieurs systèmes est de profiter des directions de descente des autres systèmes à chaque itération. L'objectif est alors de réduire le nombre d'itérations avant convergence, ce qui revient à réduire le coût de calcul global de la résolution de ces systèmes. Les publications présentant ces approches par blocs mettent en avant les performances de ces résolutions conjointes comparées aux performances d'une résolution séquentielle, c'est-à-dire d'une résolution de chaque système de façon indépendante. Par exemple, O'Leary (1980) compare l'algorithme du gradient conjugué et l'approche Block-CG : sur un problème avec trois vecteurs image, l'approche Block-CG a convergé en 5 itérations et la version séquentielle CG en un total de 25 itérations, soit une moyenne de 8.3 itérations par système. Le gain en nombre d'itérations de l'algorithme Block-CG est donc de 40% par rapport à la version séquentielle. Mentionnons que le coût de calcul d'une itération pour les deux approches est du même ordre. Dans cet exemple, le gain de temps serait également de l'ordre de 40%.

En tomographie micro-ondes, les algorithmes par blocs n'ont jamais réellement été utilisés pour résoudre conjointement les multiples systèmes linéaires intervenant dans le calcul du critère et du gradient. Une seule référence en TMO mentionne l'utilisation d'un algorithme par blocs : Fang et al. (2004) proposent d'utiliser l'approche Block-QMR pour la résolution des problèmes directs. Un modèle numérique par éléments finis y est employé et les auteurs justifient le choix d'une approche QMR par les bonnes performances de cette approche avec une matrice opérateur creuse (provenant du modèle par éléments finis). Dans leur montage d'acquisition, seize sources sont utilisées. Leur exemple est relativement éloigné de notre problème : notre matrice opérateur $\mathbf{L}_x$ est pleine et plusieurs dizaines (voire centaines) de systèmes sont généralement à résoudre. De plus, les auteurs ne comparent malheureusement pas les performances de Block-QMR avec d'autres algorithmes.

L'utilisation très fréquente de l'algorithme BiCGStab pour résoudre le problème direct en TMO nous a poussés à étudier en détail la version par blocs de cet algorithme. El Guennouni et al. (2003) décrivent l'algorithme Block-BiCGStab et démontre son efficacité face à d'autres algorithmes par blocs, notamment Block-CG et Block-QMR. Les résultats présentés dans El Guennouni et al. (2003) sont tout de même éloignés de notre problème : les matrices utilisées sont des matrices creuses, les vecteurs image sont aléatoires et leur nombre

varie entre 5 et 10. Tout comme les résultats dans Fang et al. (2004), notre problème avec $\mathbf{L}_x$ pleine et un grand nombre de vecteurs image ne correspond pas à celui présenté dans El Guennouni et al. (2003). Il est donc primordial de s'assurer que l'algorithme Block-BiCGStab peut être utilisé pour résoudre les systèmes linéaires dans notre problème. Nous verrons dans la sous-section 5.2.2 qu'une utilisation naïve de l'algorithme telle que proposée dans El Guennouni et al. (2003) entraîne des difficultés de convergence et qu'il est nécessaire de régler correctement les paramètres afin de conserver les bonnes performances de celui-ci.

Dans la section suivante, nous verrons en détail l'algorithme Block-BiCGStab et nous verrons comment l'adapter au problème de TMO.

5.2 Adaptation de l'approche Block-BiCGStab pour une application en TMO

5.2.1 L'algorithme Block-BiCGStab

L'algorithme Block-BiCGStab, développé par El Guennouni et al. (2003), est une extension du BiCGStab pour résoudre plusieurs systèmes linéaires conjointement. L'Algorithme 3 présente cette méthode. L'approche BiCGStab pour la résolution de N_S systèmes est rappelée dans l'Algorithme 2 pour pouvoir comparer côte à côte avec Block-BiCGStab et visualiser facilement les différences. La notation $\langle \mathbf{T}, \mathbf{S} \rangle_F$ représente le produit scalaire de Frobenius qui est équivalent à Trace($\mathbf{T}^\dagger \mathbf{S}$). Les opérations de Block-BiCGStab sont similaires à celles du BiCGStab : les calculs des directions de descente vectorielles $\boldsymbol{s}$ et $\boldsymbol{t}$ et des pas α et β du BiCGStab sont analogues aux quantités matricielles $\mathbf{S}$, $\mathbf{T}$, $\mathbf{A}$ et $\mathbf{B}$. Si l'on prend un seul vecteur image ($N_S = 1$) dans l'algorithme Block-BiCGStab, on retombe exactement sur les opérations du BiCGStab.

Parmi ces opérations, on retrouve les deux multiplications par $\mathbf{L}_x$ aux lignes 5 et 8 qui sont les opérations les plus coûteuses. Si l'on compare une itération de l'algorithme, on note qu'il y a $2N_S$ multiplications par $\mathbf{L}_x$. L'algorithme BiCGStab en nécessite deux pour chaque système à chaque itération (lignes 5 et 8 de l'Agorithme 2). Il y a donc bien le même nombre de produits par $\mathbf{L}_x$ dans les deux méthodes : pour une itération de Block-BiCGStab, il y a autant de multiplications par $\mathbf{L}_x$ que pour une itération de N_S résolutions par BiCGStab.

On retrouve également le calcul des pas de descente aux lignes 6 et 12. Dans le cas du BiCGStab, cela concerne le calcul des scalaires α et β aux lignes 6 et 12. En revanche, pour le Block-BiCGStab, il s'agit de calculer les matrices $\mathbf{A}$ et $\mathbf{B}$ qui nécessitent l'inversion de la matrice $(\widetilde{\mathbf{R}}_0^\dagger \mathbf{V})$ de taille N_S. Puisqu'en pratique la taille de ces matrices est bien plus faible que la taille de $\mathbf{L}_x$, le coût de calcul de ces inversions est négligeable comparé au coût des produits par $\mathbf{L}_x$. La question de la résolution de ces systèmes de taille N_S se pose.

Cependant, comme il s'agit de la même matrice à inverser aux lignes 6 et 12, nous avons choisi de calculer directement la matrice inverse. Bien que des problèmes de conditionnement peuvent apparaître à l'inversion, nous verrons dans la suite que le conditionnement de la matrice $(\widetilde{\mathbf{R}}_0^\dagger \mathbf{V})$ à inverser sera suffisamment faible pour éviter cette difficulté.

En conclusion, les opérations de BiCGStab et Block-BiCGStab étant similaires, le coût d'une itération de Block-BiCGStab équivaut approximativement au coût de N_S itérations de BiCG-Stab. Il est donc important de comprendre que l'intérêt d'utiliser le Block-BiCGStab ne provient pas d'une réduction de coût des opérations réalisées à chaque itération mais du nombre d'itérations nécessaires à la convergence de la méthode. En effet, le calcul conjoint des directions de descente permet d'améliorer l'efficacité de l'algorithme et donc de converger plus rapidement que son homologue séquentiel BiCGStab. Dans la suite, quand nous comparerons les performances des différentes approches, nous parlerons aussi bien en termes de nombre d'itérations nécessaires pour converger que de temps de calcul.

Algorithme 2 : Résolution de N_S systèmes linéaires par BiCGStab

0: **for** $i = 1, \ldots, N_S$ **do**
1: Pour une initialisation $\boldsymbol{e}_i^{\mathrm{tot}(0)}$, $\boldsymbol{r}^{(0)} = \boldsymbol{e}_i^{\mathrm{inc}} - \mathbf{L}_x \boldsymbol{e}_i^{\mathrm{tot}(0)}$, $\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)}$
2: Choisir un vecteur arbitraire $\tilde{\boldsymbol{r}}_0$
3: $k = 0$
4: **repeat**
5: $\boldsymbol{v} = \mathbf{L}_x \boldsymbol{p}^{(k)}$
6: $\alpha = (\tilde{\boldsymbol{r}}_0^\dagger \boldsymbol{r}^{(k)})/(\tilde{\boldsymbol{r}}_0^\dagger \boldsymbol{v})$
7: $\boldsymbol{s} = \boldsymbol{r}^{(k)} - \alpha \boldsymbol{v}$
8: $\boldsymbol{t} = \mathbf{L}_x \boldsymbol{s}$
9: $\omega = \boldsymbol{t}^\dagger \boldsymbol{s} / \boldsymbol{t}^\dagger \boldsymbol{t}$
10: $\boldsymbol{e}_i^{\mathrm{tot}(k+1)} = \boldsymbol{e}_i^{\mathrm{tot}(k)} + \alpha \boldsymbol{p}^{(k)} + \omega \boldsymbol{s}$
11: $\boldsymbol{r}^{(k+1)} = \boldsymbol{s} - \omega \boldsymbol{t}$
12: $\beta = (\alpha \tilde{\boldsymbol{r}}_0^\dagger \boldsymbol{r}^{(k+1)})/(\omega \tilde{\boldsymbol{r}}_0^\dagger \boldsymbol{r}^{(k)})$
13: $\boldsymbol{p}^{(k+1)} = \boldsymbol{r}^{(k+1)} + \beta(\boldsymbol{p}^{(k)} - \omega \boldsymbol{v})$
14: $k \leftarrow k + 1$
15: **until** $\|\boldsymbol{r}^{(k)}\| / \|\boldsymbol{e}_i^{\mathrm{inc}}\| <$ tolérance,
16: **end for**

Algorithme 3 : Résolution de N_S systèmes linéaires par Block-BiCGStab

1: Pour une matrice d'initialisations $\mathbf{E}^{\mathrm{tot}(0)}$, $\mathbf{R}^{(0)} = \mathbf{E}^{\mathrm{inc}} - \mathbf{L}_x \mathbf{E}^{\mathrm{tot}(0)}$, $\mathbf{P}^{(0)} = \mathbf{R}^{(0)}$
2: Choisir une matrice $(3N, N_S)$ arbitraire $\widetilde{\mathbf{R}}_0$
3: $k = 0$
4: **repeat**
5: $\mathbf{V} = \mathbf{L}_x \mathbf{P}^{(k)}$
6: résoudre $(\widetilde{\mathbf{R}}_0^\dagger \mathbf{V})\mathbf{A} = \widetilde{\mathbf{R}}_0^\dagger \mathbf{R}^{(k)}$
7: $\mathbf{S} = \mathbf{R}^{(k)} - \mathbf{V}\mathbf{A}$
8: $\mathbf{T} = \mathbf{L}_x \mathbf{S}$
9: $\omega = \langle \mathbf{T}, \mathbf{S} \rangle_F / \langle \mathbf{T}, \mathbf{T} \rangle_F$
10: $\mathbf{E}^{\mathrm{tot}(k+1)} = \mathbf{E}^{\mathrm{tot}(k)} + \mathbf{P}^{(k)}\mathbf{A} + \omega \mathbf{S}$
11: $\mathbf{R}^{(k+1)} = \mathbf{S} - \omega \mathbf{T}$
12: résoudre $(\widetilde{\mathbf{R}}_0^\dagger \mathbf{V})\mathbf{B} = -\widetilde{\mathbf{R}}_0^\dagger \mathbf{T}$
13: $\mathbf{P}^{(k+1)} = \mathbf{R}^{(k+1)} + (\mathbf{P}^{(k)} - \omega \mathbf{V})\mathbf{B}$
14: $k \leftarrow k + 1$
15: **until** $\forall i, \|\boldsymbol{r}_i^{(k)}\| / \|\boldsymbol{e}_i^{\mathrm{inc}}\| <$ tolérance, où $\boldsymbol{r}_i^{(k)}$ est la i-ème colonne de $\mathbf{R}^{(k)}$

5.2.2 Réglage des paramètres du Block-BiCGStab pour une application en TMO

Bien que l'algorithme Block-BiCGStab semble adapté à la résolution des multiples problèmes directs en imagerie micro-ondes, nous avons vu précédemment que les exemples d'utilisation de cette méthode dans El Guennouni et al. (2003) sont très différents de notre problème (matrice creuse, vecteurs image aléatoires et en faible nombre). Nous montrons ici que l'application naïve de la méthode par blocs entraîne des problèmes de convergence et nous proposons un réglage des paramètres du Block-BiCGStab pour une résolution efficace des systèmes linéaires en TMO.

Réglage de l'initialisation

L'algorithme Block-BiCGStab a montré de bonnes performances sur des problèmes génériques et avec des vecteurs image aléatoires. En revanche, dans une application comme la TMO, les vecteurs image ne sont pas aléatoires et sont au contraire corrélés. Cette corrélation est d'autant plus importante quand deux sources sont proches car les champs incidents sont similaires. Cela peut entraîner des performances dégradées dans la résolution et nous le montrons ci-après.

Prenons par exemple la résolution de deux systèmes $[\boldsymbol{e}_1^{\text{tot}}, \boldsymbol{e}_2^{\text{tot}}] = \mathbf{L}_x^{-1} [\boldsymbol{e}_1^{\text{inc}}, \boldsymbol{e}_2^{\text{inc}}]$ où les deux champs incidents $\boldsymbol{e}_1^{\text{inc}}$ et $\boldsymbol{e}_2^{\text{inc}}$ proviennent de deux sources proches l'une de l'autre. Les deux vecteurs $\boldsymbol{e}_1^{\text{inc}}$ et $\boldsymbol{e}_2^{\text{inc}}$ sont alors fortement corrélés. Une façon usuelle d'initialiser ces systèmes est de choisir soit des vecteurs nuls, soit les champs incidents. Dans les deux cas, les initialisations sont proches et donc corrélées. En reprenant les étapes de l'Algorithme 3, on a :

- ligne 1 : l'initialisation $\mathbf{E}^{\text{tot}(0)}$ contient deux colonnes fortement corrélées. Les deux colonnes de $\mathbf{R}^{(0)}$ et donc de $\mathbf{P}^{(0)}$ sont alors également corrélées.
- ligne 5 : la matrice $\mathbf{V}$ contient deux colonnes corrélées.
- ligne 6 : la matrice à inverser $(\widetilde{\mathbf{R}}_0^\dagger \mathbf{V})$ contient aussi deux colonnes fortement corrélées. La matrice est donc mal conditionnée et la résolution du système de l'étape 6 est instable.
- ligne 12 : tout comme la résolution du système à la ligne 6, le système linéaire est mal conditionné.

On comprend donc que si deux vecteurs image de $\mathbf{E}^{\text{inc}}$ sont fortement corrélées (ce qui est le cas pour deux sources proches) et que les initialisations sont proches également, cela entraîne des problèmes de conditionnement dans la matrice $(\widetilde{\mathbf{R}}_0^\dagger \mathbf{V})$ aux lignes 6 et 12 et donc des erreurs numériques dans le calcul des matrices $\mathbf{A}$ et $\mathbf{B}$. En pratique, l'accumulation de ces erreurs mène à des difficultés de convergence de l'algorithme Block-BiCGStab. On rappelle

que les vecteurs image dans El Guennouni et al. (2003) ont été pris aléatoirement ; le problème de conditionnement n'apparaît donc pas dans cet article.

L'initialisation de l'algorithme s'est avérée être le réglage primordial afin d'éviter les instabilités dues au mauvais conditionnement : le choix habituel d'initialiser les systèmes avec des vecteurs nuls ou avec les champs incidents ne permet pas une résolution efficace des systèmes linéaires. Pour contourner cette difficulté, nous proposons d'initialiser la résolution des systèmes linéaires par des vecteurs contenant une composante aléatoire. Cela permet de réduire la corrélation entre les colonnes de la matrice d'initialisation $\mathbf{E}^{\mathrm{tot}\,(0)}$. La corrélation est donc faible pour les colonnes de $\mathbf{R}^{(0)}$ et $\mathbf{V}$, ce qui permet d'améliorer le conditionnement de la matrice $(\widetilde{\mathbf{R}}_0^{\dagger}\mathbf{V})$ à inverser. Dans la suite, nous distinguerons deux cas : le cas où aucune information n'est connue sur la solution du système linéaire et le cas où l'on possède une connaissance *a priori* de la solution. On retrouve le second cas en TMO quand à chaque itération du processus d'inversion, les champs totaux calculés à l'itération précédente de l'inversion (disons $k-1$) peuvent être utilisés comme information initiale pour résoudre les problèmes directs à l'itération k.

Quelque soit l'initialisation choisie (champs incidents ou information *a priori* sur les champs totaux), il sera judicieux d'initialiser les résolutions de systèmes de cette manière en y ajoutant une composante aléatoire (en pratique, un bruit gaussien complexe circulaire) afin de décorréler les initialisations. La variance du bruit à ajouter est un compromis entre l'initialisation et la décorrélation : un bruit faible entraînera une faible altération de l'information issue de l'initialisation et une faible décorrélation des vecteurs initiaux, tandis qu'un bruit important entraînera une forte altération de l'initialisation et une forte décorrélation des vecteurs initiaux. Dans le cas où aucun *a priori* n'est fourni sur les champs totaux, il n'est pas gênant d'altérer l'initialisation. En revanche, si une information fiable sur les champs totaux est donnée (ce sera le cas lorsque l'algorithme de reconstruction s'approche de la convergence), il n'est pas judicieux de trop bruiter cette initialisation afin de conserver une initialisation proche de la solution. Les deux cas de figure seront présentés dans les résultats préliminaires en sous-section 5.2.3 afin de mettre en évidence l'effet de l'initialisation sur la convergence de l'approche Block-BiCGStab. Nous montrerons plus tard qu'ajouter du bruit de façon à ce que le rapport signal sur bruit (RSB) soit égal à 50 dB permet de résoudre efficacement les systèmes linéaires par Block-BiCGStab, quelle que soit l'initialisation.

Réglage des autres paramètres

Le choix de la matrice $\widetilde{\mathbf{R}}_0$ (ligne 2 de l'Algorithme 3) de taille $3N \times N_S$ est un paramètre supplémentaire à régler. Dans l'algorithme BiCGStab, il est également nécessaire de régler

le vecteur $\widetilde{\boldsymbol{r}}_0$ (ligne 2 de l'Algorithme 2). Van der Vorst (1992); Gutknecht (1993); Xu et al. (2002) proposent de choisir $\widetilde{\boldsymbol{r}}_0 = \boldsymbol{r}^{(0)}$, ce qui semble fournir de bons résultats de convergence. Pour l'algorithme Block-BiCGStab, El Guennouni et al. (2003) proposent de prendre $\widetilde{\mathbf{R}}_0 = \mathbf{R}^{(0)}$ pour résoudre les systèmes génériques avec des vecteurs image aléatoires. Pour notre problème de TMO, nous avons testé différentes formes pour le choix $\widetilde{\mathbf{R}}_0$: matrice aléatoire, matrice identité rectangulaire (des 1 sur la diagonale principale), matrice égale à $\mathbf{R}^{(0)}$ ou encore matrice égale à $\mathbf{L}_x\mathbf{R}^{(0)}$. Après avoir comparé les performances sur différents problèmes, le choix de $\widetilde{\mathbf{R}}_0 = \mathbf{R}^{(0)}$ est celui qui a montré les meilleures performances. Nous avons donc fixé $\widetilde{\mathbf{R}}_0 = \mathbf{R}^{(0)}$ pour les résultats présents dans cette thèse.

Pour le choix du critère d'arrêt, la tolérance a été fixée de la même façon que pour l'algorithme BiCGStab afin de pouvoir comparer correctement les deux méthodes : l'algorithme est arrêté quand le résidu relatif $\|\boldsymbol{r}^{(k)}\|/\|\boldsymbol{e}^{\text{inc}}\|$ de chacun des systèmes est inférieur à un certain seuil fixé par l'opérateur. La tolérance est la même pour les N_S systèmes. Concernant l'algorithme Block-BiCGStab, il est arrêté lorsque tous les résidus relatifs sont inférieurs à la tolérance. Cela signifie que même si certains résidus relatifs parmi les N_S résidus sont inférieurs au seuil, l'algorithme continue jusqu'à que tous les résidus aient atteint le seuil (cf. ligne 15 de l'Algorithme 3). Pour une même tolérance, on obtient donc des résidus généralement plus faibles par une résolution Block-BiCGStab que par des résolutions BiCGStab.

5.2.3 Résultats justifiant le réglage de l'initialisation

Nous allons étudier l'influence de l'initialisation sur la convergence de l'algorithme Block-BiCGStab et comparer les performances avec l'approche BiCGStab. Dans ce chapitre, tous les résultats seront présentés sur un objet simulé décrit ci-dessous. Les tests seront réalisés sur des résolutions de problèmes directs $\mathbf{L}_x\mathbf{E}^{\text{tot}} = \mathbf{E}^{\text{inc}}$. Rappelons que la résolution de tels systèmes est également nécessaire pour le calcul du gradient.

Description de l'objet simulé et du montage d'acquisition

L'objet simulé qui sera utilisé dans ce chapitre représente deux cubes imbriqués placés dans l'air. Cet exemple est similaire à celui exploité dans Abubakar et al. (2012a) mais le contraste de l'objet a été augmenté pour que le problème soit plus difficile à reconstruire.

La figure 5.1 représente le volume d'étude et l'emplacement des sources. L'objet est illuminé par 160 sources à la fréquence $f = 3$ GHz (la longueur d'onde λ est d'environ 10 cm). Les sources sont des dipôles électriques placés verticalement et situés sur cinq anneaux parallèles au plan (xOy). Le diamètre des anneaux est de 6λ, soit 60 cm et les hauteurs sont

$z = -2\lambda, -\lambda, 0, \lambda, 2\lambda$. On considère que l'objet est placé dans un domaine d'étude V cubique de côté $3\lambda = 30$ cm, discrétisé en $30 \times 30 \times 30$ voxels. La résolution est donc de 1 cm. L'objet d'étude représente deux cubes imbriqués concentriques. Le gros cube est de taille $2\lambda = 20$ cm de côté avec un contraste constant $\chi_1 = 0.6 + j0.8$ et le petit cube à l'intérieur est de taille $\lambda = 10$ cm de contraste constant $\chi_2 = 1.2 + j0.4$. La figure 5.2 donne une représentation de trois coupes de l'objet.

Dans cette section, les simulations de problèmes directs pour calculer les champs totaux seront basées sur la description précédente de l'objet et du montage. Dans la suite du chapitre, nous chercherons à reconstruire cet objet. Pour cela, nous complétons le montage avec 160 récepteurs placés autour de l'objet aux mêmes positions que les sources. Les récepteurs ne mesurent que la composante verticale du champ électrique. On a donc $30^3 = 27\,000$ inconnues complexes et $160 \times 160 = 25\,600$ mesures du champ diffracté. Comme l'objet est fortement contrasté, il est utile d'avoir un bon *a priori* de l'objet pour initialiser l'algorithme d'inversion. Pour cela, on utilise une procédure par saut de fréquence. L'idée de cette procédure est de reconstruire partiellement l'objet avec des acquisitions successives aux basses fréquences afin d'obtenir une initialisation convenable pour la fréquence finale (la plus élevée). Dans cet exemple simulé, on cherche à reconstruire l'objet à 3 GHz. On réalise alors l'acquisition de mesures à 1 GHz, puis à 2 GHz avant de reconstruire l'objet à 3 GHz en initialisant avec la dernière solution obtenue, c'est-à-dire celle obtenue à 2 GHz. Les données aux récepteurs sont générées sur une grille plus fine que celle uilisée pour la reconstruction, soit $60 \times 60 \times 60$ voxels, et un bruit gaussien avec un rapport signal sur bruit de 20 dB est ajouté aux données générées. Cela permet d'éviter le crime inverse, c'est-à-dire de reconstruire l'objet avec exactement le même modèle que celui qui a permis de générer les données.

Effet de l'initialisation

Nous souhaitons mettre en évidence l'effet de l'initialisation sur la convergence de l'approche Block-BiCGStab. Pour cela, nous allons comparer le comportement de BiCGStab et Block-BiCGStab pour différentes initialisations.

Les simulations ont été exécutées avec un processeur Intel Core i7-5960X (8 cœurs) cadencé à 3 GHz. Le langage utilisé est MATLAB. Pour éviter les effets potentiels dus à la parallélisation des opérations algébriques élémentaires qui peuvent dépendre du langage utilisé, les simulations dans cette section sont réalisées sur un seul cœur.

La figure 5.3 présente l'évolution des 160 résidus relatifs $\|\boldsymbol{e}_i^{\text{inc}} - \mathbf{L}_x \boldsymbol{e}_i^{\text{tot}(k)}\| / \|\boldsymbol{e}_i^{\text{inc}}\|$ en fonction des itérations k de BiCGStab et Block-BiCGStab. Sur cette figure, l'initialisation des deux approches est prise égale aux champs incidents : $\boldsymbol{e}_i^{\text{tot}(0)} = \boldsymbol{e}_i^{\text{inc}}$. La tolérance est fixée

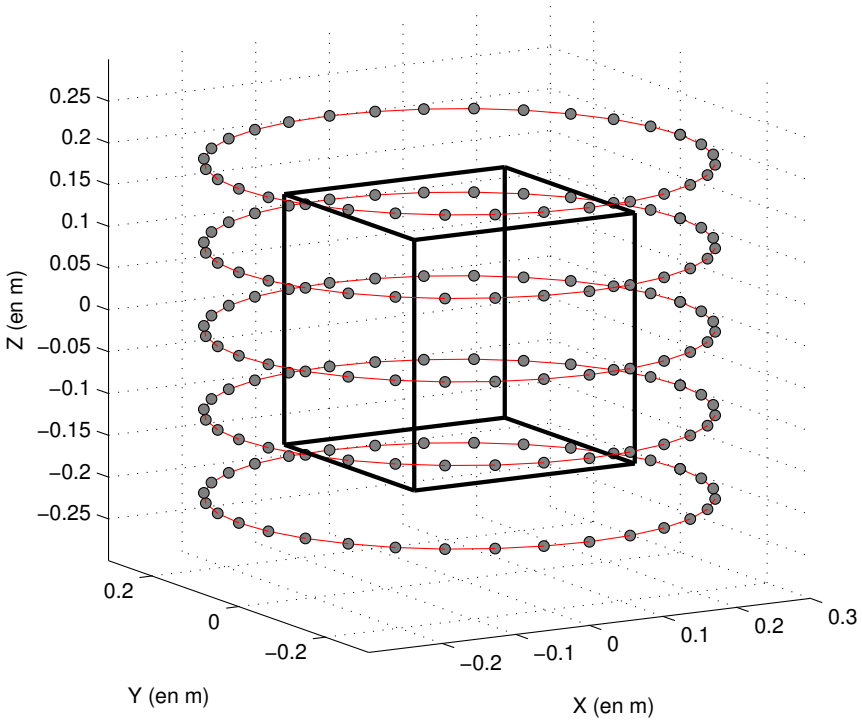


Figure 5.1 Schéma du montage d'acquisition pour l'objet simulé. Le cube noir représente le volume V dans lequel l'objet est placé. Les ronds gris représentent les sources placées sur cinq cercles autour du volume d'étude. Les récepteurs sont placés aux mêmes positions que les sources.

ici à 10^{-6}. Les 160 résolutions par BiCGStab convergent en moyenne en 19.5 itérations ; la résolution la plus rapide atteint la tolérance en 18 itérations et la plus lente en 21 itérations. Au contraire, les résidus relatifs par Block-BiCGStab n'atteignent pas la tolérance et la résolution a un comportement divergent à partir d'une quarantaine d'itérations. Cela met en évidence l'effet du mauvais conditionnement des matrices $\mathbf{A}$ et $\mathbf{B}$ sur la convergence des systèmes. La figure 5.4 présente l'évolution des 160 résidus relatifs pour une initialisation aux champs incidents avec une composante aléatoire de façon à ce que le RSB soit égal à 50 dB. Les résidus relatifs par BiCGStab atteignent la tolérance en moyenne en 19.6 itérations : l'ajout de la composante aléatoire ne change pas significativement la convergence de l'approche BiCGStab. La résolution par Block-BiCGStab converge en 12 itérations. L'ajout de la composante aléatoire permet donc à l'algorithme Block-BiCGStab de converger. Quantitativement, l'approche Block-BiCGStab nécessite environ 40% d'itérations en moins que BiCGStab. L'ajout d'une composante aléatoire a donc l'effet escompté.

Nous souhaitons maintenant connaître l'influence du niveau de bruit à ajouter à l'initialisa-

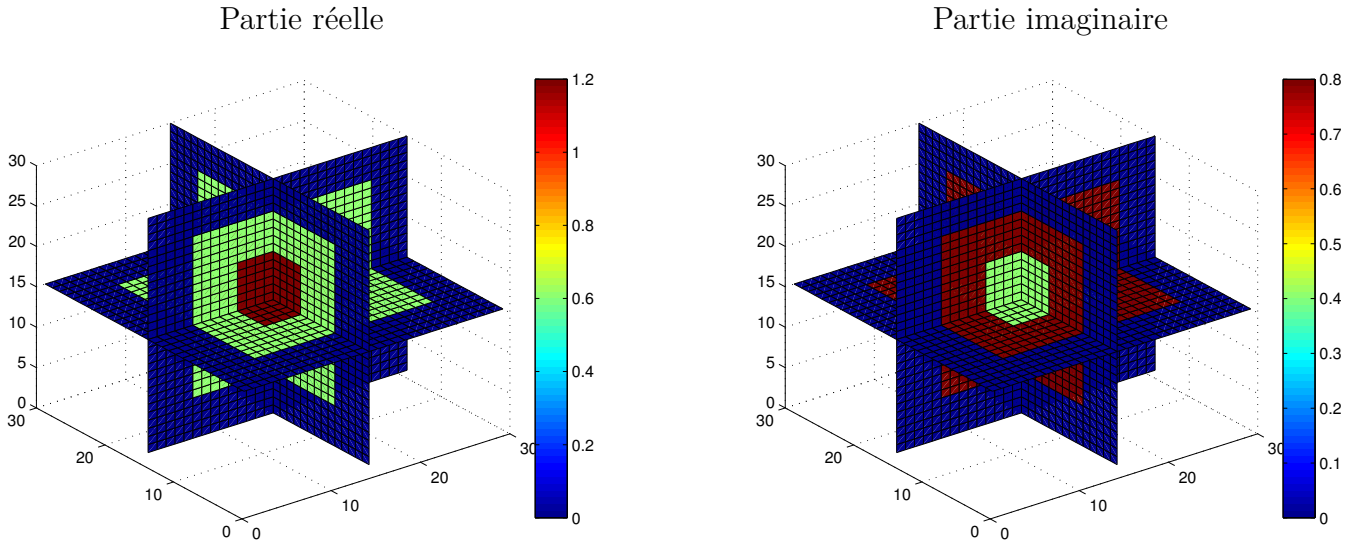


Figure 5.2 Représentation 3-D des parties réelle et imaginaire du contraste de l'objet simulé. Le contraste du cube externe est $\chi_1 = 0.6 + j0.8$ et celui du cube interne est $\chi_2 = 1.2 + j0.4$. Les axes représentent le nombre de voxels, soit $30 \times 30 \times 30$ voxels.

tion sur la convergence de l'approche Block-BiCGStab. Le tableau 5.1 présente le nombre d'itérations pour atteindre la convergence pour différents niveaux de bruit ajoutés aux champs incidents. On remarque que pour de très faibles niveaux de bruit, c'est-à-dire pour de grandes valeurs de rapports signal sur bruit (RSB), la résolution par Block-BiCGStab diverge. On retrouve ce qu'on observe sur la figure 5.3 où les résidus du Block-BiCGStab n'atteignent pas la tolérance quand aucun bruit n'est ajouté (RSB = Inf). En revanche, quand on ajoute une composante aléatoire un peu plus conséquente, la résolution converge. On note d'ailleurs, que quelque soit le RSB entre -10 dB et 80 dB, la résolution converge en à peu près le même nombre d'itérations, soit 12 ou 13 itérations. Cette grande plage de niveaux de bruit est due au fait que l'initialisation n'est pas proche de la solution exacte. L'initialisation et l'initialisation bruitée sont donc tout aussi éloignées de la solution exacte. On observe tout de même des valeurs minimales du nombre d'itérations qui sont obtenues pour un RSB entre 20 dB et 60 dB.

On réalise maintenant la même étude sur l'influence de l'initialisation quand une information *a priori* sur les champs à calculer est connue. Ici on suppose que l'on connaît les champs totaux pour un objet très proche de l'objet simulé : on connaît les champs totaux $\mathbf{E}_{\widetilde{x}}^{\text{tot}} = \mathbf{L}_{\widetilde{x}}^{-1}\mathbf{E}^{\text{inc}}$ pour le contraste $\widetilde{\boldsymbol{x}} = 0.99\boldsymbol{x}$. On initialise alors la résolution des systèmes $\mathbf{E}^{\text{tot}} = \mathbf{L}_x^{-1}\mathbf{E}^{\text{inc}}$ par les champs $\mathbf{E}_{\widetilde{x}}^{\text{tot}}$ perturbés par une composante aléatoire.

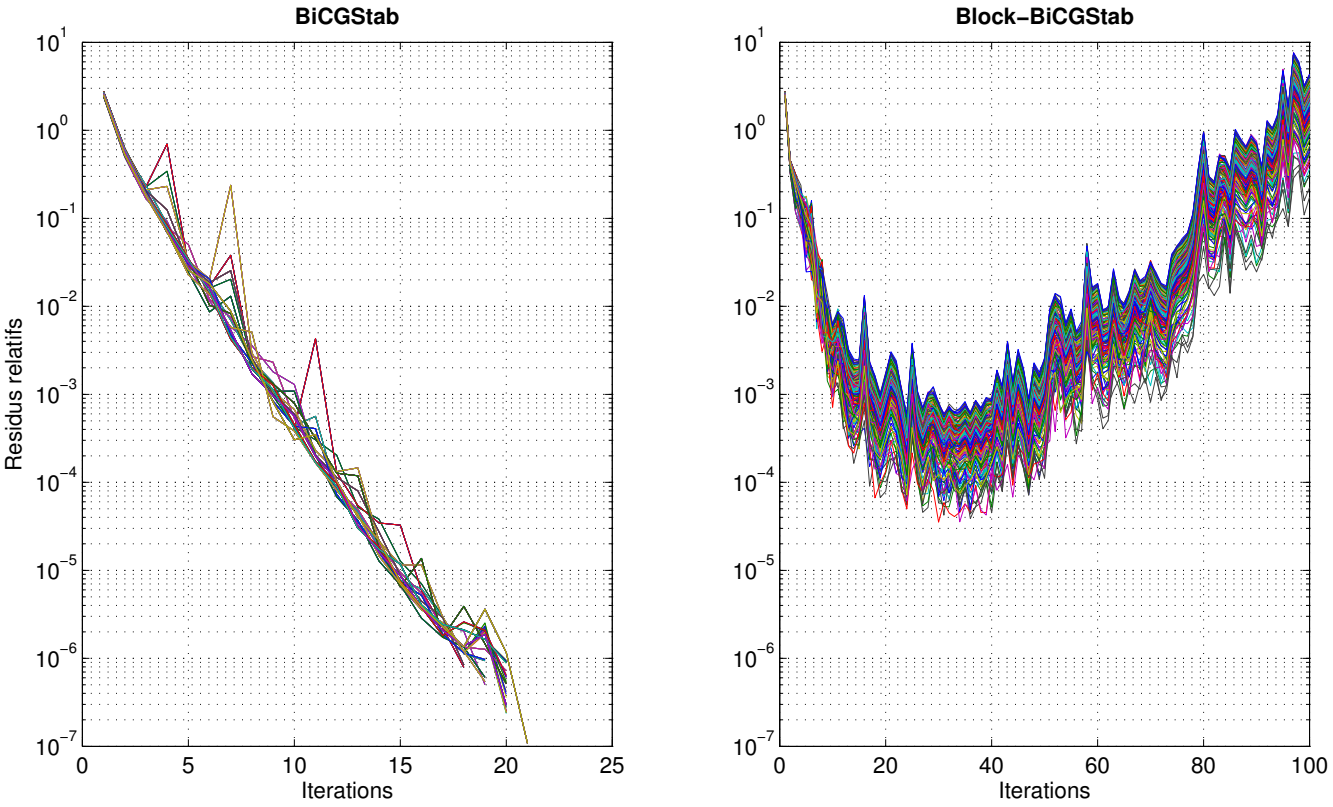


Figure 5.3 Évolution des résidus relatifs des 160 résolutions de systèmes par BiCGStab (à gauche) et par Block-BiCGStab (à droite) en fonction des itérations. Les initialisations des systèmes sont choisies égales aux champs incidents $\boldsymbol{e}_i^{\text{inc}}$. La tolérance est de 10^{-6}.

Le tableau 5.2 représente le nombre d'itérations en fonction du niveau de bruit ajouté à l'initialisation des champs totaux $\mathbf{E}_{\widetilde{x}}^{\text{tot}}$. Ici, on note que quelque soit le niveau de bruit ajouté, la tolérance est atteinte. L'ajout d'aucun bruit ne permet pas de converger et l'ajout d'un bruit faible (entre un RSB de 40 dB et de 100 dB) entraîne une convergence en environ 10 itérations. Enfin, pour l'ajout d'un bruit trop fort élevé, la composante aléatoire prend le dessus sur l'information *a priori* et la résolution nécessite environ 13 itérations, comme si aucun *a priori* n'est connu. Finalement, on choisit de fixer un bruit dont le RSB est de 50 dB. Pour cette initialisation, l'algorithme BiCGStab converge en moyenne en 12.2 itérations. L'approche Block-BiCGStab converge donc en environ 20% d'itérations de moins que l'approche BiCGStab.

En conclusion, on a vu expérimentalement qu'un ajout d'un bruit de RSB de 50 dB sur l'initialisation donne de bonnes performances du Block-BiCGStab, quelle que soit l'initialisation.

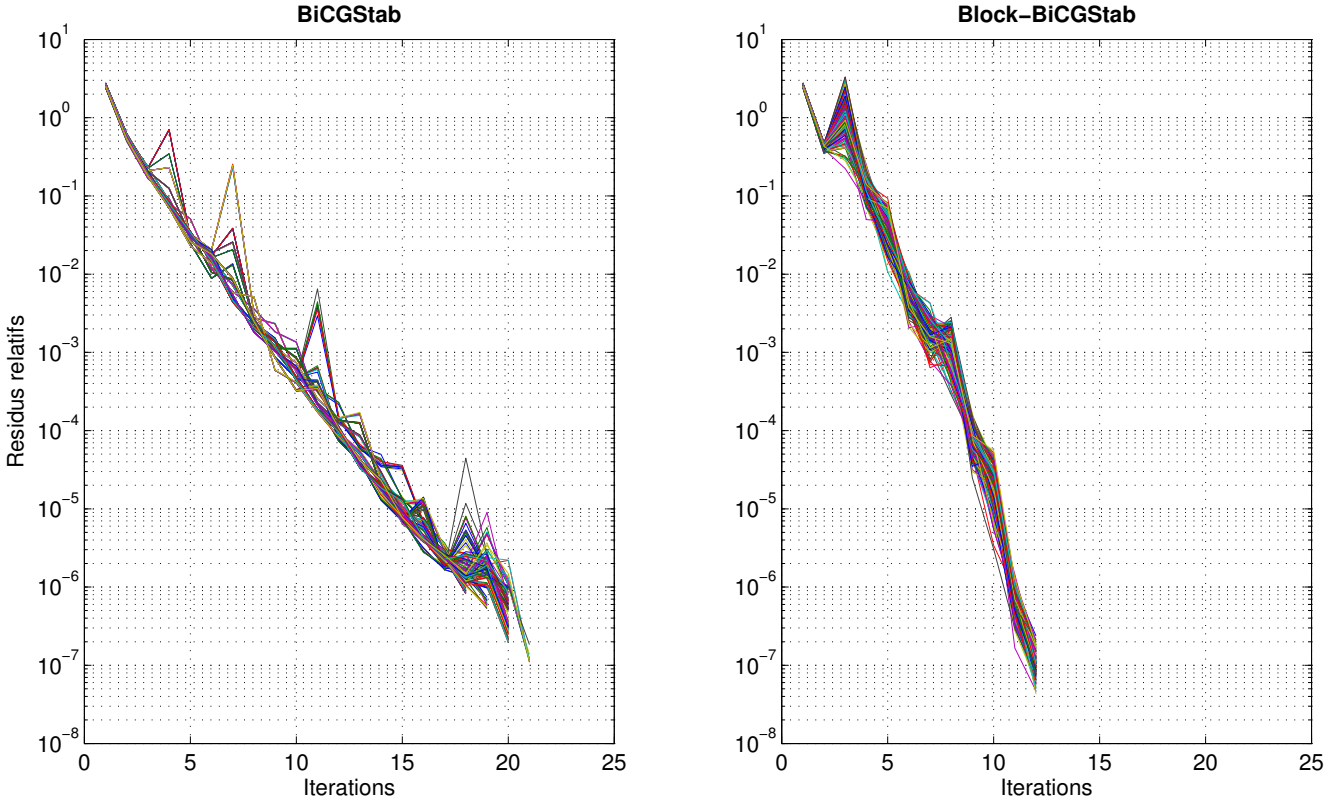


Figure 5.4 Évolution des résidus relatifs des 160 résolutions de systèmes par BiCGStab (à gauche) et par Block-BiCGStab (à droite) en fonction des itérations. Les initialisations des systèmes sont choisies égales aux champs incidents $\boldsymbol{e}_i^{\text{inc}}$ avec une composante aléatoire de 50 dB. La tolérance est de 10^{-6}.

RSB	Nb iter. Block
-10 dB	12.95
0 dB	12.6
10 dB	12.3
20 dB	12.2
30 dB	12
40 dB	12.25
50 dB	12.2
60 dB	12.25
70 dB	12.3
80 dB	13.15
90 dB	diverge
100 dB	diverge
Inf	diverge

Tableau 5.1 Nombre d'itérations du Block-BiCGStab en fonction du niveau de bruit ajouté à l'initialisation. L'initialisation est choisie égale aux champs incidents. Le nombre d'itérations correspond à une valeur moyenne calculée sur 20 réalisations. Pour les mêmes problèmes, BiCGStab converge en moyenne en 19.5 itérations.

RSB	Nb iter. Block
-10 dB	13.04
0 dB	12.1
10 dB	11.4
20 dB	11
30 dB	10.25
40 dB	9.85
50 dB	9.65
60 dB	10
70 dB	9.85
80 dB	9.75
90 dB	9.9
100 dB	9.8
Inf	diverge

Tableau 5.2 Nombre d'itérations du Block-BiCGStab en fonction du niveau de bruit ajouté à l'initialisation. L'initialisation est ici égale aux champs totaux calculés pour le contraste 0.99χ, soit proche de la solution exacte. Moyenne sur 20 réalisations. Pour les mêmes problèmes, BiCGStab converge en moyenne en 12.2 itérations.

5.3 Amélioration du Block-BiCGStab pour une architecture multi-cœurs

Avec le développement moderne des architectures multi-cœurs, il devient facile de profiter de cette puissance de calcul pour répartir les opérations sur les différents cœurs et donc accélérer le calcul global. La parallélisation de processus est une science à part entière en pleine expansion. De nombreuses problématiques concrètes s'appuient désormais sur cette science pour réduire le temps de calcul de processus chronophages. Un exemple simple de processus parallélisable est le calcul d'une boucle `FOR` où les itérations sont indépendantes les unes des autres. Chaque itération peut être considérée comme un processus indépendant exécutable sur un cœur. Un autre type d'utilisation fréquente des architectures multi-cœurs est l'accélération d'opérations élémentaires, comme une multiplication matricielle, une inversion matricielle ou encore une transformée de Fourier rapide (FFT). Les opérations algébriques de ces opérations élémentaires sont découplées et envoyées sur les différents cœurs. Cette accélération est souvent implémentée directement dans le langage et est donc dépendante du langage utilisé. C'est par exemple le cas du logiciel MATLAB dans lequel la parallélisation de nombreuses opérations élémentaires est implémentée par défaut (option *multithreading*). Dans cette thèse, le travail porte uniquement sur des approches algorithmiques. La parallélisation des opérations élémentaires ne sera pas utilisée dans la suite de ce chapitre afin de quantifier les gains obtenus uniquement grâce aux différentes structures d'algorithmes.

Les N_S problèmes directs en TMO sont indépendants les uns des autres. La résolution de ces N_S systèmes par BiCGStab peut donc être facilement parallélisée en envoyant les résolutions sur les cœurs disponibles (Abubakar et al., 2012a). On s'attend d'ailleurs à une accélération due à la parallélisation théoriquement linéaire, c'est-à-dire que le temps de calcul global serait divisé par le nombre de cœurs disponibles. En pratique, nous verrons dans la suite que l'accélération n'est pas tout à fait linéaire mais que l'on s'en rapproche. En revanche, l'algorithme Block-BiCGStab ne peut pas être parallélisé directement puisque les systèmes linéaires sont résolus de façon conjointe. Nous proposons dans cette section une solution efficace pour adapter l'approche Block-BiCGStab à une architecture multi-cœurs. Cette approche, que nous nommons Partial-Block BiCGStab est un compromis entre BiCGStab séquentiel et Block-BiCGStab.

5.3.1 L'approche Partial-Block BiCGStab

Pour profiter des atouts du Block-BiCGStab et d'une machine multi-cœurs, l'idée est de séparer les N_S systèmes linéaires en plusieurs groupes de taille moindre et d'appliquer l'algorithme Block-BiCGStab sur ces différents groupes. Les résolutions par Block-BiCGStab de

ces groupes de systèmes sont indépendantes et peuvent être envoyées sur les différents cœurs. Un argument supplémentaire de cette division par groupes est d'améliorer le conditionnement dans chacun des groupes en découpant efficacement l'ensemble des vecteurs image afin que la convergence des résolutions soit la meilleure possible. Cette approche est appelée Partial-Block BiCGStab et est décrite dans l'Algorithme 4. Le nombre entier P définit le nombre de groupes choisi par l'utilisateur. Les N_S sources sont alors réparties en P groupes et les matrices $\mathbf{E}_p^{\text{inc}}$ et $\mathbf{E}_p^{\text{tot}}$ $(p = 1 \dots P)$ définissent respectivement des sous-matrices de $\mathbf{E}^{\text{inc}}$ et $\mathbf{E}^{\text{tot}}$ représentant les P groupes choisis.

Algorithme 4 : Résolution de N_S systèmes par Partial-Block BiCGStab

1: Créer P groupes $\mathbf{E}_p^{\text{inc}}$ $(p = 1 \dots P)$ en partitionnant les N_S sources.
2: **for** $p = 1 \dots P$ **do**
3: Exécuter l'algorithme Block-BiCGStab avec la matrice $\mathbf{E}_p^{\text{inc}}$ pour calculer $\mathbf{E}_p^{\text{tot}}$.
4: **end for**

5.3.2 Réglages des paramètres de l'algorithme Partial-Block BiCGStab

Le principe de l'algorithme Partial-Block BiCGStab est élémentaire mais, tout comme Block-BiCGStab, plusieurs paramètres doivent être ajustés pour que la résolution des systèmes linéaires en TMO ait un comportement stable. Les paramètres à régler sont le nombre de groupes et la répartition des sources dans ces groupes. De plus, comme pour Block-BiCGStab, l'initialisation joue un rôle important. Nous verrons ici comment ajuster ces paramètres.

Nous ne parlerons pas des autres paramètres (matrice $\widetilde{\mathbf{R}}_0$, tolérance) puisque ceux-ci restent les mêmes que pour la résolution par Block-BiCGStab.

Réglage du nombre de groupes

On peut voir l'algorithme Partial-Block BiCGStab comme une généralisation des approches BiCGStab et Block-BiCGStab. Le nombre de groupes P est le paramètre qui permet de contrôler la balance entre les deux : si l'on prend $P = 1$, cela signifie qu'on résout tous les systèmes linéaires conjointement, ce qui correspond au Block-BiCGStab. À l'opposé, si $P = N_S$, chaque groupe contient un seul système ; cela revient à résoudre tous les systèmes indépendamment, ce qui correspond au BiCGStab. Le choix du nombre de groupes P est donc un compromis entre les atouts du Block-BiCGStab et les capacités de parallélisation du BiCGStab en fonction du nombre de cœurs disponibles. Pour bénéficier des avantages d'une machine multi-cœurs, nous montrerons dans la sous-section 5.3.3 qu'il est nécessaire de séparer les systèmes en un nombre de groupes égal au nombre de cœurs disponibles : si par

exemple N_c cœurs sont disponibles sur une machine, nous choisissons de diviser les systèmes en $P = N_c$ groupes.

Il est logique d'affirmer que plus le nombre de cœurs disponibles est élevé, plus la résolution des N_S systèmes sera rapide. C'est par exemple le cas dans Abubakar et al. (2012a) où les auteurs utilisent 128 cœurs pour résoudre les nombreux systèmes. Idéalement, avoir autant de cœurs que de systèmes à résoudre serait la configuration la plus efficace : chaque système serait résolu par BiCGStab sur un cœur différent (Golnabi et al., 2011). En revanche, l'utilisation d'un ordinateur avec suffisamment de cœurs (voire un supercalculateur) nécessite des coûts importants et n'est pas à la portée de tous les utilisateurs. C'est pourquoi l'approche Partial-Block BiCGStab est la solution adaptée à des machines courantes de l'ordre de quelques cœurs, voire une dizaine de cœurs.

Répartition des sources dans les groupes

Une fois le nombre de groupes P fixé, nous proposons de répartir astucieusement les systèmes dans les groupes. Comme il a été vu dans la section 5.2.2, deux sources proches peuvent entraîner des difficultés dans la résolution dues aux matrices $\mathbf{A}$ et $\mathbf{B}$ mal conditionnées. Pour éviter cela, nous proposons que la répartition des sources dans les groupes soit réalisée de façon à ce que les sources d'un même groupe soient le plus éloignées les unes des autres. Cela permet d'avoir une corrélation faible entre les vecteurs des champs incidents $\boldsymbol{e}_i^{\text{inc}}$ pour un même groupe.

Déterminer la répartition optimale des sources en P groupes de façon à avoir la corrélation intra-groupe la plus faible est un problème d'optimisation combinatoire complexe. Cependant en pratique, les sources sont réparties selon une structure régulière : sur des cercles ou des sphères autour du domaine à imager par exemple. Dans le cas de sources réparties sur des cercles, une solution simple et suffisante est de créer des groupes en prenant une source toutes les P sources sur ces cercles. Bien que la répartition ne soit pas garantie d'être optimale, en pratique la distance entre les sources d'un même groupe est suffisante pour avoir un faible conditionnement dans les P groupes. Par exemple, on reprend le montage de l'objet simulé, présenté en 5.2.3, où 160 antennes sont réparties sur cinq cercles. La figure 5.5 présente la répartition des sources en $P = 4$ groupes. Les sources d'un même groupe sont suffisamment éloignées les unes des autres ce qui permet d'obtenir une corrélation relativement faible entre elles. Dans le cas de sources réparties sur une sphère, il est possible de reprendre la même répartition afin d'éviter d'avoir deux sources proches dans le même groupe. Selon l'espacement des sources sur la sphère, il peut être plus judicieux de suivre l'ordre selon les méridiens ou les parallèles. Ces deux types de montage (cercles, sphère) représentent la majorité des montages

en TMO.

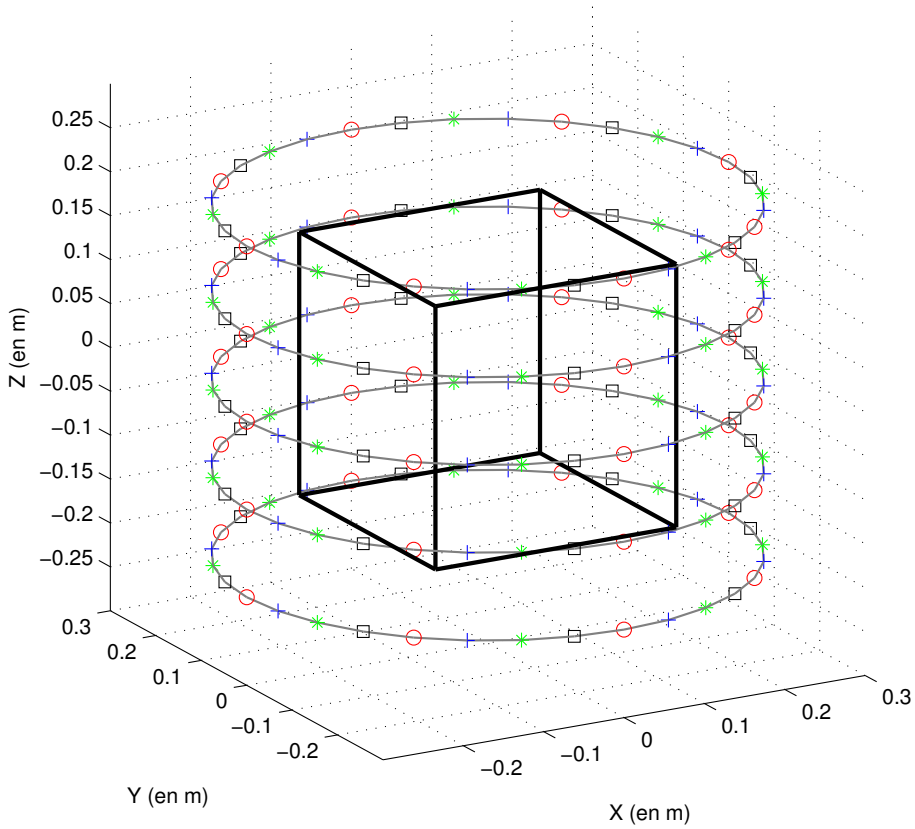


Figure 5.5 Schéma de représentation du montage simulé pour l'application de l'algorithme Partial-Block BiCGStab où les sources sont partitionnées en $P = 4$ groupes. Chaque symbole représente un groupe.

Réglage de l'initialisation

On rappelle que dans le cas du Block-BiCGStab ($P = 1$), la corrélation de l'initialisation est corrigée en ajoutant une composante aléatoire. A l'opposé, pour le BiCGStab ($P = N_S$), il n'est pas nécessaire de décorréler les initialisations car les résolutions sont indépendantes. La question de l'initialisation pour Partial-Block BiCGStab se pose alors.

Plus le nombre de groupes P est grand, plus les sources dans un même groupe sont ainsi éloignées les unes des autres et donc plus la corrélation est faible entre les vecteurs champs incidents $\boldsymbol{e}_i^{\text{inc}}$ d'un même groupe. Il semble donc naturel de se dire que plus P est grand, moins il est nécessaire de décorréler les initialisations en ajoutant une composante aléatoire

tel que nous l'avons proposé en 5.2.2. Il est difficile de fournir une solution "clé en main" pour l'initialisation du Partial-Block BiCGStab car cela dépend du nombre de groupes, de la répartition des sources et de la géométrie du montage d'acquisition. Afin d'assurer la convergence des P résolutions, nous avons fait le choix d'initialiser de la même façon que dans le cas de Block-BiCGStab, c'est-à-dire d'ajouter un bruit complexe gaussien de façon à avoir un RSB de 50 dB. Ce choix reste empirique mais a montré de bonnes performances avec l'algorithme Partial-Block BiCGStab.

5.3.3 Effet des réglages du Partial-Block BiCGStab sur la convergence

Nous cherchons maintenant à étudier l'effet du nombre de groupes et de la répartition des sources dans ces groupes sur le comportement de l'algorithme Partial-Block BiCGStab.

Reprenons l'objet simulé décrit dans le paragraphe 5.2.3. Afin d'étudier la corrélation qu'il y a entre les 160 champs incidents $\boldsymbol{e}_i^{\text{inc}}$, nous pouvons calculer la matrice de corrélation de ces champs. La figure 5.6 représente la valeur absolue des coefficients de corrélation des champs incidents. Notons que les sources sont rangées selon un certain ordre : les 32 premières sources sont celles du premier cercle, les 32 suivantes sont celles du deuxième cercle, et ainsi de suite. On remarque que les champs incidents de deux sources consécutives sur un même cercle sont fortement corrélés. Cela correspond aux valeurs sur les premières diagonales supérieure et inférieure de la matrice. On remarque également une corrélation moyenne pour deux sources placées l'une au-dessus de l'autre, c'est-à-dire sur deux cercles différents mais ayant les mêmes coordonnées x et y. Cela s'observe sur les différentes diagonales décalées de 32 sources. Cette analyse justifie le choix de répartir les sources selon le schéma décrit dans le paragraphe 5.3.2.

Nous étudions maintenant la question du réglage du nombre de groupes en fonction de l'architecture multi-cœurs disponible. L'idée est de comparer pour différentes valeurs de P le temps de calcul sur une architecture donnée. La figure 5.7 présente le temps de calcul pour résoudre les 160 systèmes linéaires par Partial-Block BiCGStab en fonction du nombre de groupes P. Ici, nous supposons qu'aucune connaissance *a priori* n'est utilisée : les initialisations sont choisies égales aux champs incidents pour BiCGStab ($P = N_S$) et égales aux champs incidents bruités pour les autres ($P < N_S$). Quatre courbes sont tracées, chacune représentant une architecture différente, c'est-à-dire un nombre de cœurs N_c différent variant entre 1 et 8. Pour réaliser ces tests en pratique, nous avons utilisé le même processeur à 8 cœurs mais nous avons limité le nombre de cœurs pour l'exécution du code. Il faut noter que les points à gauche ($P = 1$) représente l'algorithme Block-BiCGStab et les points à droite ($P = 160$) représente BiCGStab. Pour la courbe $N_c = 1$, la méthode Block-BiCGStab est la plus rapide et l'approche BiCGStab quasiment la plus lente. Puisque $N_c = 1$, les résolutions par BiCGStab

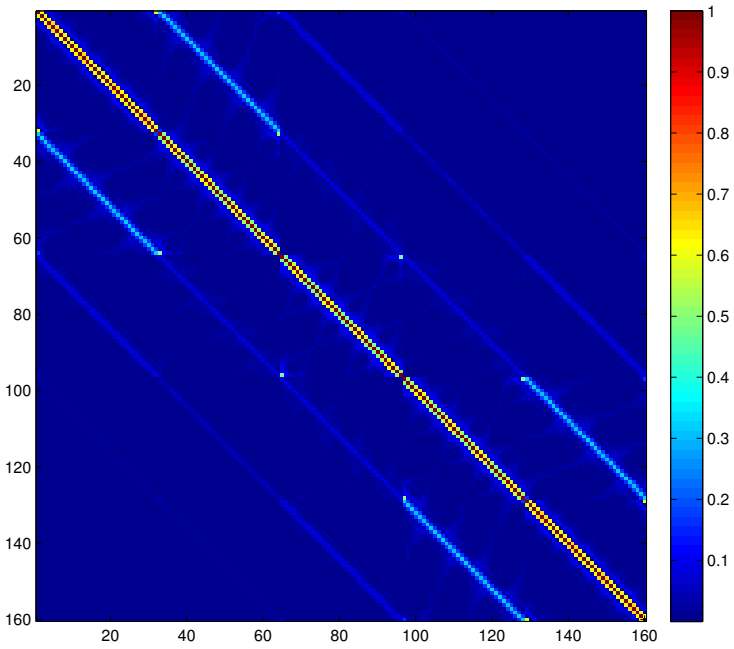


Figure 5.6 Matrice de corrélation des champs incidents provenant des 160 sources du montage simulé

et Partial-Block BiCGStab n'ont pas été parallélisées et ont été exécutés séquentiellement.

Quand le nombre de cœurs augmente, les méthodes pouvant être parallélisées le sont. On remarque par exemple pour $N_c = 2$ que le temps de calcul du Block-BiCGStab ne change pas tandis que le temps de calcul du BiCGStab est approximativement divisé par deux. Dans ce cas, la méthode la plus efficace est l'algorithme Partial-Block BiCGStab avec $P = 2$. Pour les deux dernières courbes ($N_c = 4$ et $N_c = 8$), la méthode la plus rapide est le Partial-Block BiCGStab avec $P = N_c$. Il semble donc préférable de choisir l'algorithme Partial-Block BiCGStab avec $P = N_c$ afin d'obtenir les meilleures performances. Ce résultat confirme notre choix naturel de prendre $P = N_c$. Cette figure permet aussi de supposer que si $N_c = N_S$ cœurs sont disponibles, la méthode la plus efficace est BiCGStab où chaque résolution est envoyée sur un cœur différent.

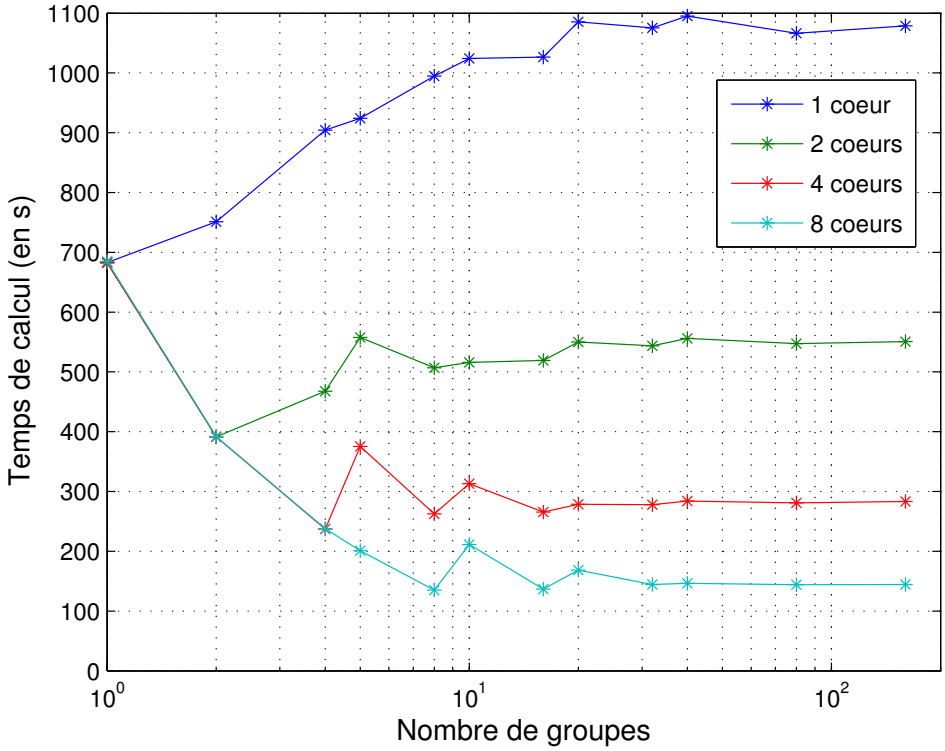


Figure 5.7 Temps de calcul pour la résolution des 160 problèmes directs de l'objet simulé par Partial-Block BiCGStab en fonction du nombre de groupes P. Quatre architectures multi-cœurs sont présentées : $N_c = 1, 2, 4$ et 8.

5.4 Exploitation des approches par blocs dans le processus d'inversion

Comme nous l'avons présenté dans le chapitre 3, la reconstruction en TMO requiert la minimisation itérative d'un critère. A chaque itération de la reconstruction, le calcul du gradient permet de déterminer une direction de descente et le pas de descente est déduit à partir d'une ou plusieurs évaluations du critère. On rappelle que le calcul du gradient requiert la résolution d'un certain nombre de systèmes selon la mise en œuvre choisie (voir chapitre 4) et que le calcul du critère nécessite la résolution de N_S systèmes linéaires. Tous ces systèmes impliquent la même matrice opérateur $\mathbf{L}_x$. Pour résoudre ces nombreux systèmes, les méthodes de reconstruction actuelles utilisent des algorithmes itératifs, notamment le BiCGStab (Abubakar et al., 2012a; Zhang et Liu, 2015). Nous proposons dans cette section d'appliquer les méthodes Block-BiCGStab et Partial-Block BiCGStab pour résoudre ces systèmes afin d'accélérer le coût global des méthodes de reconstruction.

Le réglage de l'approche Partial-Block BiCGStab a été décrit dans les sections précédentes.

L'initialisation, paramètre majeur du réglage, doit être choisie judicieusement pour conserver l'efficacité des méthodes. Dans le cas de l'inversion, il est possible d'avoir une information *a priori* sur les solutions des systèmes à partir des résolutions effectuées aux itérations précédentes de l'algorithme de reconstruction. Prenons l'exemple d'un algorithme de reconstruction qui converge vers un minimum local. Le contraste estimé à l'itération k sera proche du contraste estimé à l'itération précédente $k-1$. Cela signifie que la matrice $\mathbf{L}_x$ varie peu entre ces deux itérations et les solutions des systèmes linéaires seront également proches. L'idée est donc d'initialiser la résolution des systèmes à l'itération k de façon à profiter de la solution obtenue à l'itération $k-1$. Les solutions à l'itération $k-1$ pouvant être fortement corrélées entre elles, il est cependant nécessaire d'ajouter un terme aléatoire afin de décorréler l'initialisation à l'itération k, comme étudié au paragraphe 5.2.2.

Pour résumer le processus d'inversion, nous notons $\widehat{\mathbf{E}}^{(k)}$ les champs totaux estimés par résolution des problèmes directs à l'itération k de l'algorithme de reconstruction, c'est-à-dire pour le contraste $\boldsymbol{x}^{(k)}$. De même, nous notons $\widehat{\mathbf{F}}^{(k)}$ les solutions des systèmes intervenant dans le calcul du gradient. On résume alors une itération de l'inversion par la structure suivante :

- on connaît l'estimation du contraste $\boldsymbol{x}^{(k)}$ à l'itération k et les solutions $\widehat{\mathbf{E}}^{(k-1)}$ et $\widehat{\mathbf{F}}^{(k-1)}$ des systèmes linéaires, calculées à l'itération précédente ;
- on résout les systèmes donnant $\widehat{\mathbf{E}}^{(k)}$ et $\widehat{\mathbf{F}}^{(k)}$ pour le contraste $\boldsymbol{x}^{(k)}$. Les résolutions de ces systèmes sont initialisées respectivement par $\widehat{\mathbf{E}}^{(k-1)}$ et $\widehat{\mathbf{F}}^{(k-1)}$, perturbés par un bruit gaussien ;
- on en déduit le gradient $\nabla_x \mathcal{F}$, donné en (3.14), évalué en $\boldsymbol{x}^{(k)}$ et on calcule la direction de descente $\boldsymbol{d}^{(k)}$ fourni par l'algorithme de minimisation (dans notre cas, il s'agit de L-BFGS) ;
- on détermine le pas de descente par évaluations successives du critère. Pour calculer le critère, les résolutions de problèmes directs pour calculer les champs totaux sont initialisées avec la dernière estimation des champs totaux (ici $\widehat{\mathbf{E}}^{(k)}$) perturbés par du bruit ;
- on obtient une nouvelle estimation du contraste $\boldsymbol{x}^{(k+1)}$.

L'utilisation d'une telle initialisation à chaque itération est bien plus efficace qu'une initialisation avec les champs incidents puisque l'information précise fournie à l'itération précédente est prise en compte et permet de réduire considérablement le nombre d'itérations. Cette stratégie est appliquée aussi bien au BiCGStab qu'au Partial-Block BiCGStab, c'est-à-dire que les résolutions par BiCGStab à l'itération k sont initialisées avec $\widehat{\mathbf{E}}^{(k-1)}$ et $\widehat{\mathbf{F}}^{(k-1)}$ calculées à l'itération précédente.

Enfin, il est important de préciser que l'utilisation de l'approche Block-BiCGStab ou Partial-

Block BiCGStab en remplacement de BiCGStab ne dépend pas de l'algorithme d'inversion utilisé. Dès que le calcul de systèmes linéaires multiples dans le critère et/ou le gradient est requis, il est possible d'utiliser ces approches par blocs quelle que soit la formulation du critère, du gradient ou l'algorithme de minimisation utilisé. Par exemple, les méthodes développées dans Eyraud et al. (2009) (gradient conjugué non-linéaire), dans Abubakar et al. (2012a) (Gauss-Newton) ou encore dans Yu et al. (2009) (*Distorted Born Iterative Method*) utilisent BiCGStab pour la résolution des systèmes. Il est donc possible d'accélérer la reconstruction grâce à l'utilisation de ces approches par blocs.

5.5 Évaluation des coûts de calcul des différents algorithmes pour la résolution de problèmes directs

Dans cette section, nous allons comparer les performances des algorithmes Block-BiCGStab et Partial-Block BiCGStab et celles de l'approche BiCGStab sur des résolutions de problèmes directs. Nous allons étudier ces performances selon différents paramètres du problème, notamment le contraste de l'objet et la fréquence de l'onde incidente.

5.5.1 Influence du contraste

Dans un premier temps, nous souhaitons étudier l'influence du contraste sur la vitesse de convergence du Partial-Block BiCGStab et comparer ses performances par rapport à l'algorithme BiCGStab. Nous nous plaçons dans le cas où nous n'avons qu'un seul cœur disponible. La méthode Partial-Block BiCGStab correspond alors à Block-BiCGStab.

Pour étudier l'effet du contraste, nous allons résoudre les $N_S = 160$ problèmes directs pour l'objet simulé, présenté en 5.2.3, pour lequel nous faisons varier le contraste original $\chi_1 = 0.6 + j0.8$ et $\chi_2 = 1.2 + j0.4$. On définit un facteur de contraste η de façon à ce que le contraste du nouvel objet $\chi_\eta = \eta\chi$. Par exemple, pour un facteur de contraste de $\eta = 0.5$, cela signifie que le nouvel objet à un contraste $\chi_{\eta,1} = 0.3 + j0.4$ et $\chi_{\eta,2} = 0.6 + j0.2$. Plus le facteur de contraste est élevé, plus l'objet a une permittivité et une conductivité élevées. Le tableau 5.3 présente le nombre d'itérations et le temps de calcul pour atteindre la convergence pour différents facteurs de contraste. Deux méthodes sont comparées : BiCGStab et Block-BiCGStab. Ici, l'initialisation est choisie égale aux champs incidents. Afin d'améliorer la convergence de l'approche BiCGStab, nous avons ajouté une procédure dite "marching-on-in-source" permettant d'initialiser encore mieux les résolutions de systèmes (De Zaeytijd et al., 2007). Cette procédure propose d'initaliser le calcul du champ total $\boldsymbol{e}_i^{\text{tot}}$ en s'appuyant sur une combinaison linéaire des champs totaux estimées aux sources précédentes $\boldsymbol{e}_{i-1}^{\text{tot}}, \boldsymbol{e}_{i-2}^{\text{tot}}, \ldots$.

Nous avons choisi d'utiliser les trois sources précédentes pour proposer une meilleure initialisation pour les résolutions BiCGStab. Pour les initialisations Block-BiCGStab, comme les résolutions sont conjointes, la procédure "marching-on-in-source" n'est pas applicable. Comme le préconise notre étude au paragraphe 5.2.3, nous prenons une initialisation égale aux champs incidents bruités avec un RSB de 50 dB. D'après le tableau 5.3, on note immédiatement que plus le contraste est élevé, plus le gain entre BiCGStab et Block-BiCGStab est important. Ainsi, pour un contraste élevé ($\eta = 2.5$), la résolution par Block-BiCGStab est quasiment deux fois plus rapide que par BiCGStab. À l'inverse, pour des contrastes très faibles, l'approche Block-BiCGStab converge en quasiment le même nombre d'itérations que BiCGStab. Le gain est donc négligeable a très faible contraste.

Dans l'optique de reconstruire des objets fortement contrastés, ces résultats justifient l'utilisation de l'approche par blocs pour la résolution des systèmes linéaires.

Facteur de	Temps CPU		Nombre d'itérations		Ratio du temps
contraste η	BiCGStab	Block	BiCGStab	Block	Block / BiCGStab
1/4	416 s	405 s	[7 7.01 8]	[7]	0.97
1/2	585 s	514 s	[10 10.3 12]	[9]	0.88
1	989 s	689 s	[17 17.9 21]	[13]	0.70
1.5	1468 s	897 s	[25 27.1 30]	[16]	0.61
2	1922 s	1109 s	[34 36.0 39]	[20]	0.58
2.5	2436 s	1326 s	[44 45.9 50]	[24]	0.54

Tableau 5.3 Temps de calcul et nombre d'itérations pour la résolution de 160 problèmes directs avec BiCGStab et Block-BiCGStab. Différents contrastes de l'objet simulé présenté en 5.2.3 sont testés. L'initialisation est choisie égale aux champs incidents avec procédure de marching-on-in-source pour BiCGStab, et égale aux champs incidents bruités (RSB = 50 dB) pour Block-BiCGStab. Le nombre d'itérations pour les 160 systèmes par BiCGStab entre crochets correspond à [minimum moyenne maximum].

5.5.2 Influence de la fréquence

Nous étudions maintenant l'effet de la fréquence sur le temps de calcul. Notons que plus la fréquence est élevée, plus la taille électrique de l'objet, c'est-à-dire, sa taille en fonction de la longueur d'onde, est grande. Étudier l'effet de la fréquence revient donc à étudier l'effet de la taille de l'objet. Bien que plus difficile, la reconstruction à hautes fréquences permet une meilleure résolution de reconstruction et est donc un enjeu en tomographie micro-ondes.

Pour cela, nous choisissons de résoudre les problèmes directs pour différentes fréquences : 1 GHz, 2 GHz et 3 GHz. Cela représente respectivement une taille d'objet de $\frac{2}{3}\lambda$, $\frac{4}{3}\lambda$ et 2λ. On suppose qu'aucun *a priori* n'est connu sur les champs totaux. Les initialisations

sont choisies égales aux champs incidents (avec procédure “marching-on-in-source”) pour BiCGStab et égales aux champs incidents bruités pour Block-BiCGStab. Le tableau 5.4 compare les temps de calcul et le nombre d'itérations des deux méthodes pour les différentes fréquences. L'algorithme Block-BiCGStab est d'autant plus performant face à BiCGStab que la fréquence est élevée (ou de façon que la taille de l'objet est grande). En revanche, pour de faibles fréquences, les deux approches atteignent la tolérance en quasiment le même nombre d'itérations.

Le gain de Block-BiCGStab est important face à BiCGStab dans des situations complexes (forts contrastes, grands objets) ; ce sont ces situations pour lesquelles les résolutions de systèmes et l'inversion sont les plus coûteuses. L'utilisation de l'algorithme Block-BiCGStab est donc vivement justifiée pour la reconstruction d'objets fortement contrastés et de grande taille.

Fréquence	Temps CPU		Nombre d'itérations		Ratio du temps
	BiCGStab	Block	BiCGStab	Block	Block / BiCGStab
1 GHz	498 s	499 s	[7 7.97 10]	[8]	1.00
2 GHz	809 s	607 s	[13 13.37 15]	[10]	0.75
3 GHz	989 s	689 s	[17 17.9 21]	[13]	0.70

Tableau 5.4 Temps de calcul et nombre d'itérations pour la résolution de 160 problèmes directs avec BiCGStab et Block-BiCGStab sur l'objet simulé présenté en 5.2.3. Différentes fréquences d'ondes incidentes sont testées. L'initialisation est choisie égale aux champs incidents avec procédure de marching-on-in-source pour BiCGStab, et égale aux champs incidents bruités (RSB = 50 dB) pour Block-BiCGStab. Le nombre d'itérations pour les 160 systèmes par BiCGStab entre crochets correspond à [minimum moyenne maximum].

5.5.3 Influence du nombre de systèmes à résoudre

Nous allons maintenant analyser l'effet du nombre de systèmes N_S à résoudre sur le gain de Block-BiCGStab face à BiCGStab. Pour cela, nous allons limiter le nombre d'illuminations dans la configuration du montage simulé présenté en 5.2.3.

Cinq simulations des problèmes directs ont été réalisées pour cinq valeurs différentes de N_S : 32, 64, 96, 128, 160. En pratique, nous sélectionnons respectivement un cercle de sources, deux cercles de sources, et ainsi de suite jusqu'à l'ensemble des cinq cercles. Le tableau 5.5 donne les temps de calcul et les nombres d'itérations pour BiCGStab et Block-BiCGStab pour différentes valeurs de N_S.

On remarque que l'approche Block-BiCGStab est d'autant meilleure face à BiCGStab que le nombre de systèmes linéaires à résoudre augmente. On rappelle que les algorithmes par blocs

dans la littérature sont souvent appliqués à un faible nombre de vecteurs image. Ce résultat permet de valider le fait que les réglages proposés dans la section 5.2 permettent de conserver les bonnes performances de Block-BiCGStab pour un grand nombre de systèmes linéaires.

Ce résultat met également en lumière le fait que diviser les résolutions en groupes dans Partial-Block BiCGStab ne permet pas de conserver l'efficacité de la méthode Block-BiCGStab et que l'intérêt de cette division réside essentiellement dans la possibilité de parallélisation.

Nombre N_S de systèmes linéaires	Temps CPU		Nombre d'itérations		Ratio du temps Block / BiCGStab
	BiCGStab	Block	BiCGStab	Block	
32	206 s	204 s	[17 17.7 21]	[18]	0.99
64	408 s	346 s	[17 17.8 21]	[15]	0.85
96	615 s	470 s	[17 18 21]	[13]	0.76
128	823 s	611 s	[17 18 21]	[13]	0.74
160	989 s	689 s	[17 17.9 21]	[13]	0.70

Tableau 5.5 Temps de calcul et nombre d'itérations pour la résolution des 160 problèmes directs avec BiCGStab et Block-BiCGStab sur l'objet simulé présenté en 5.2.3. Différentes valeurs de N_S sont testées. L'initialisation est choisie égale aux champs incidents avec procédure de "marching-on-in-source" pour BiCGStab, et égale aux champs incidents bruités (RSB = 50 dB) pour Block-BiCGStab. Le nombre d'itérations pour les 160 systèmes par BiCGStab entre crochets correspond à [minimum moyenne maximum].

5.5.4 Effet de l'architecture multi-cœurs

Nous étudions enfin l'effet du nombre de cœurs N_c sur le comportement de Partial-Block BiCGStab. On a vu que pour une architecture donnée, l'algorithme le plus efficace est Partial-Block BiCGStab où le nombre de groupes P est égale au nombre de cœurs disponibles. Nous comparons ici les performances de cet algorithme et celles de BiCGStab. Le tableau 5.6 donne les temps de calcul des méthodes BiCGStab et Partial-Block BiCGStab pour différentes architectures. Ces architectures sont simulées en limitant le nombre de cœurs octroyées aux résolutions. On remarque tout d'abord que, quelle que soit l'architecture, l'approche Partial-Block BiCGStab est toujours plus efficace que BiCGStab. Logiquement, plus le nombre de cœurs disponibles N_c augmente, plus l'écart entre les deux approches est réduit. On retrouve le comportement observé dans la figure 5.7. En pratique, le nombre de cœurs disponibles est souvent limité ; l'algorithme Partial-Block BiCGStab est donc très performant face à BiCGStab.

Nombre de cœurs N_c	Temps CPU Partial		Ratio du temps Partial / BiCGStab
	BiCGStab	$P = N_c$	
1 cœur	1079 s	683 s	0.63
2 cœurs	551 s	391 s	0.71
4 cœurs	283 s	237 s	0.84
8 cœurs	145 s	135 s	0.93

Tableau 5.6 Temps de calcul et nombre d'itérations pour la résolution de 160 problèmes directs avec BiCGStab et Partial-Block BiCGStab sur l'objet simulé présenté en 5.2.3. Différentes architectures matérielles sont testées. L'initialisation est choisie égale aux champs incidents pour BiCGStab, et égale aux champs incidents bruités (RSB = 50 dB) pour Partial-Block BiCGStab.

5.6 Évaluation des approches pour la reconstruction

Nous présentons maintenant les résultats de reconstruction de l'objet simulé ainsi qu'une comparaison des inversions utilisant les approches BiCGStab et Partial-Block BiCGStab.

5.6.1 Procédure de reconstruction

Afin d'effectuer la reconstruction de l'objet simulé défini en 5.2.3 à partir des données mesurées à 3 GHz, nous avons vu précédemment qu'il était nécessaire d'obtenir une bonne initialisation du contraste. Nous avons fait le choix de réaliser l'acquisition à des fréquences plus basses (1 et 2 GHz) et d'appliquer une procédure par saut de fréquence. Nous avons alors défini la procédure de reconstruction suivante :

— on initialise la reconstruction avec un contraste nul en tout point ;
— on minimise le critère pénalisé (3.5) pour les données à 1 GHz par L-BFGS pendant 20 itérations. La solution obtenue est nommée $\hat{\boldsymbol{x}}_1$;
— on minimise le critère pour les données à 2 GHz en initialisant avec $\hat{\boldsymbol{x}}_1$ et on effectue 20 itérations. La solution obtenue est nommée $\hat{\boldsymbol{x}}_2$;
— on minimise le critère pour les données à 3 GHz en initialisant avec $\hat{\boldsymbol{x}}_2$. On arrête l'algorithme L-BFGS quand la norme du gradient (3.14) du critère atteint la tolérance 10^{-6}.

Le fait d'effectuer quelques itérations avec les données à basses fréquences permet d'obtenir une solution à basse résolution et donc d'initialiser correctement la reconstruction à 3 GHz afin que la minimisation converge vers un minimum local convenable. Concernant la régularisation, les paramètres sont fixés empiriquement : $\lambda = 10^{-6}$ et $\delta = 10^{-2}$.

Cette procédure de reconstruction est la même quelle que soit la méthode utilisée pour résoudre les systèmes linéaires intervenant dans le critère et le gradient. La tolérance dans la résolution de ces systèmes linéaires est fixée à 10^{-6}. Comme cette tolérance est choisie identique pour BiCGStab et Partial-Block BiCGStab et qu'elle est relativement stricte, les évaluations du critère et du gradient par BiCGStab et Partial-Block BiCGStab fournissent quasiment les mêmes valeurs. Cela signifie que le comportement de l'algorithme de reconstruction avec utilisation de BiCGStab ou Partial-Block BiCGStab est similaire. La différence entre les deux reconstructions n'intervient donc pas au niveau de la solution ou du nombre d'itérations de l'optimisation par l'algorithme L-BFGS, mais au niveau du temps de calcul global.

5.6.2 Résultats de reconstruction

Nous présentons ici les résultats de reconstruction de l'objet simulé. Les figures 5.8 et 5.10 représentent respectivement les parties réelle et imaginaire de l'objet exact. Les figures 5.9 et 5.11 représentent respectivement les parties réelle et imaginaire de la solution de la reconstruction. Dans chaque figure, les 30 tranches du volume d'étude sont présentées. Cette représentation par tranches permet de visualiser tous les voxels du domaine d'étude V. Certains auteurs ne présentent que quelques tranches mais il semble pertinent de visualiser l'ensemble afin de ne manquer aucun détail de l'objet reconstruit, notamment les discontinuités.

Tout d'abord, notons que la position et la forme de l'objet sont relativement bien estimées. On retrouve correctement le cube externe de taille 2λ. Le cube interne est plus difficile à reconstruire, on remarque que les discontinuités du cube interne ne sont pas aussi franches et aussi droites que celle du cube externe. La raison principale vient du fait que, le cube externe étant fortement contrasté, l'onde électromagnétique pénètre plus difficilement dans l'objet et il est donc plus compliqué de reconstruire l'intérieur du cube. Cependant, le cube interne est tout de même détecté et la taille de celui-ci est également bien déterminée. De plus, on remarque que les discontinuités dans le plan horizontal (tranches 6 et 25) ne sont pas aussi franches que les discontinuités des faces verticales. Cela vient du fait que seules les composantes verticales du champ sont acquises ce qui entraîne une résolution dégradée dans cette direction.

Concernant l'aspect quantitatif de la reconstruction, on peut voir que les valeurs du contraste sont légèrement mal estimées. Cela provient probablement de la régularisation qui a tendance à réduire les différences au niveau des discontinuités. Par exemple, on observe que la partie imaginaire du contraste du cube interne est sur-estimée et que la différence entre le cube interne et externe est plus faible que pour l'objet exact.

Nous pouvons comparer nos résultats de reconstruction avec Abubakar et al. (2012a) où le même objet est simulé avec un contraste moindre. Comme les méthodes utilisées ne sont pas comparables (critère, régularisation et algorithme d'optimisation différents), nous ne faisons une comparaison uniquement en termes de qualité d'image. Notre reconstruction semble tout à fait convenable, étant donné le contraste plus élevé de notre problème. les discontinuités semblent même plus franches que celles estimées dans Abubakar et al. (2012a). Bien que l'objectif de ce chapitre soit de présenter une accélération de la reconstruction, ces résultats permettent tout de même de valider les bonnes performances de notre algorithme d'inversion en termes de qualité de reconstruction.

5.6.3 Comparaison des coûts de reconstruction selon les approches

Nous proposons ici de comparer les coûts de calcul de la reconstruction de l'objet simulé en fonction de l'approche utilisée pour la résolution des nombreux systèmes linéaires.

Dans un premier temps, nous limitons le nombre de cœurs disponibles à 1 et nous comparons les reconstructions où les systèmes sont résolus avec les approches BiCGStab et Block-BiCGStab. Le tableau 5.7 présente ces coûts de calcul. Le fort contraste de l'objet rend la reconstruction difficile, les temps de calcul sont donc conséquents. Sur un unique cœur, la reconstruction avec l'utilisation de BiCGStab met presque 80 heures et la reconstruction avec Block-BiCGStab met environ 65 heures, soit un gain de 18%. Regardons plus en détail le coût de calcul de l'inversion :

- les 20 itérations de L-BFGS avec les données à 1 GHz ont pris 5 h avec BiCGStab et 4.7 h avec Block-BiCGStab, soit un ratio de 0.94. Ce faible écart est attendu puisqu'à cette fréquence, l'objet est de petite taille électrique. Le gain de Block-BiCGStab face à BiCGStab est donc relativement faible ;
- les 20 itérations de L-BFGS avec les données à 2 GHz ont pris respectivement 7.9 h et 6.9 h avec BiCGStab et Block-BiCGStab, soit un ratio de 0.88. Tout d'abord, on observe que réaliser le même nombre d'itérations avec les données à 2 GHz est bien plus long qu'avec les données à 1 GHz, cela vient du fait que les résolutions de systèmes sont plus lentes puisque la fréquence augmente. De plus, le gain relatif de Block-BiCGStab par rapport à BiCGStab est plus important que précédemment ;
- enfin les 175 itérations avant convergence avec les données à 3 GHz ont pris respectivement 66.7 h et 53.8 h avec BiCGStab et Block-BiCGStab, soit un ratio d'environ 0.81.

Nous avons effectué la même reconstruction avec une architecture à quatre cœurs. Une première reconstruction a été réalisée où les systèmes sont résolus par BiCGStab et une seconde

où les systèmes sont résolus par Partial-Block BiCGStab avec quatre groupes. Les résolutions, aussi bien par BiCGStab que Partial-Block BiCGStab, ont été parallélisées sur les quatre coeurs disponibles. Le tableau 5.7 donne ces résultats. Comme attendu, la reconstruction avec Partial-Block BiCGStab est plus rapide que celle avec BiCGStab mais le gain entre les deux reconstructions est réduit : on observe pour cet objet un gain de 5%.

Nombre de cœurs N_c	Temps		Ratio du temps Partial / BiCGStab
	BiCGStab	$P = N_c$	
1 cœur	79.6 h	65.4 h	0.82
4 cœurs	21.9 h	20.9 h	0.95

Tableau 5.7 Temps de reconstruction pour l'objet simulé avec résolutions des systèmes par BiCGStab et Block-BiCGStab ($N_c = 1$) ou Partial-Block BiCGStab ($N_c = 4$).

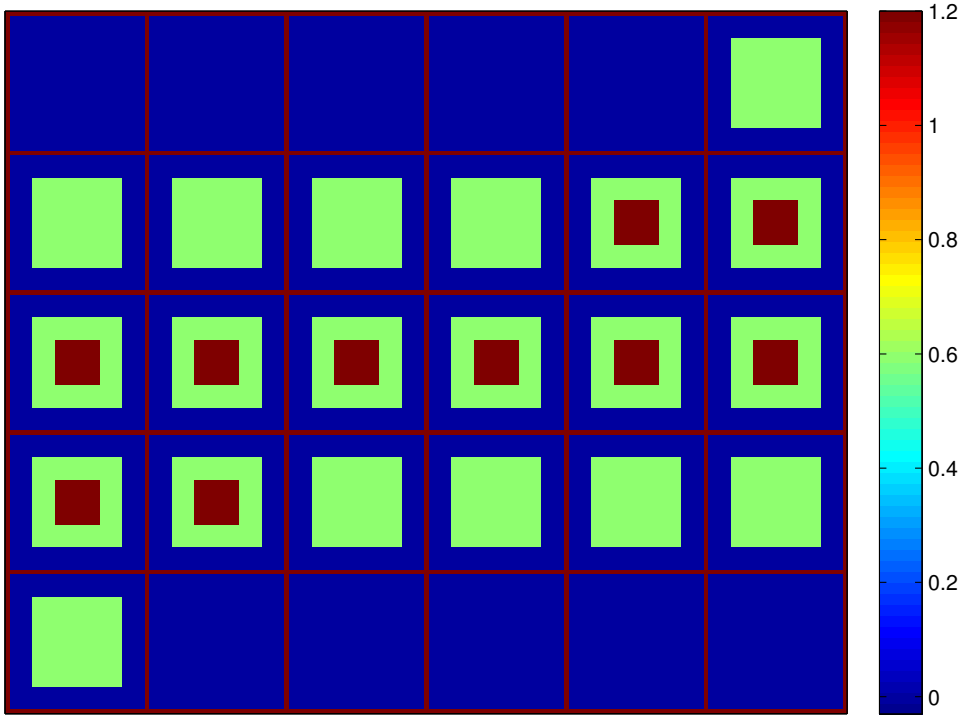


Figure 5.8 30 tranches horizontales de la partie réelle de l'objet exact.

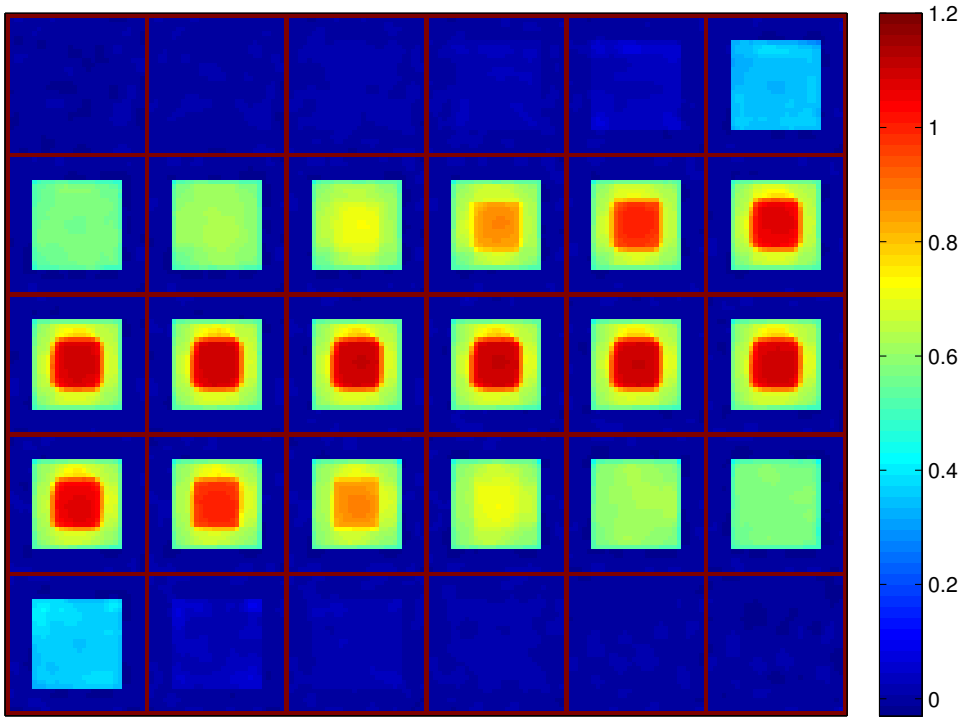


Figure 5.9 30 tranches horizontales de la partie réelle de l'objet reconstruit.

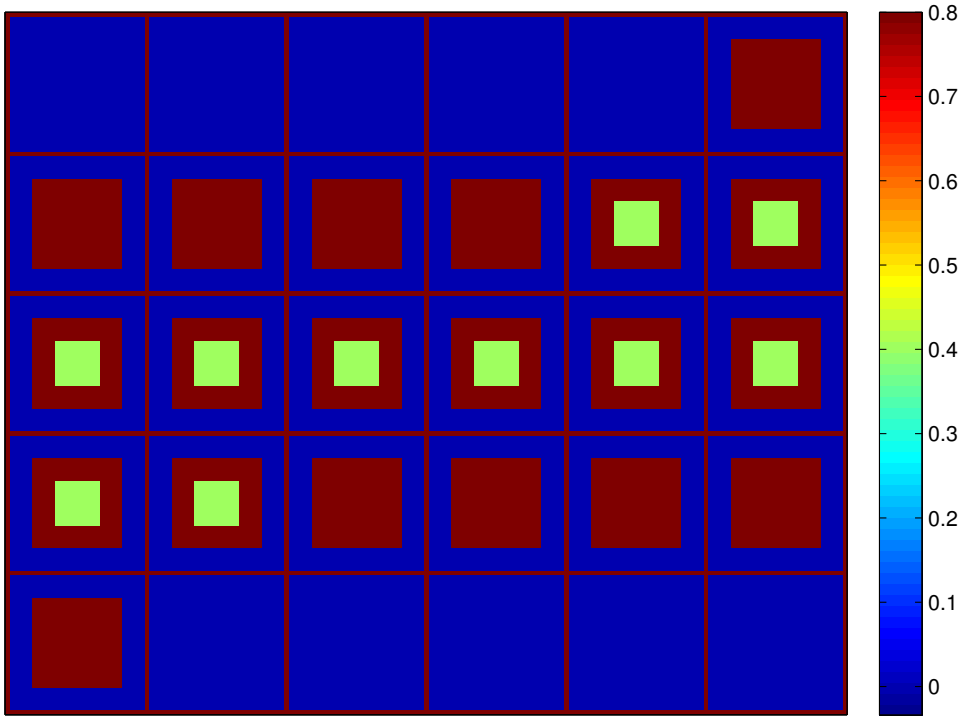


Figure 5.10 30 tranches horizontales de la partie imaginaire de l'objet exact.

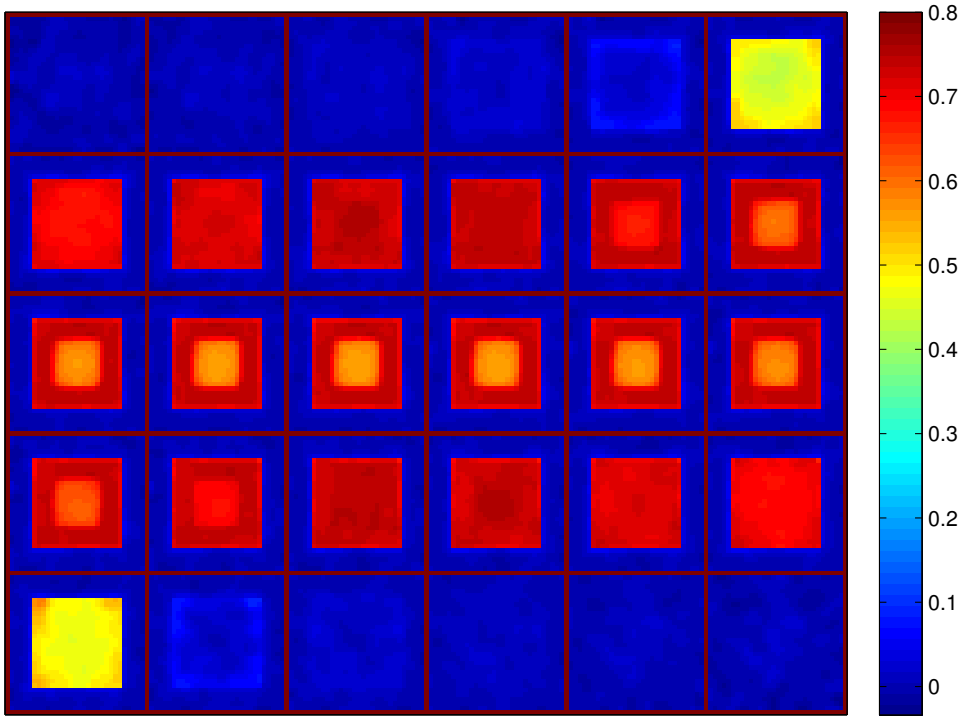


Figure 5.11 30 tranches horizontales de la partie imaginaire de l'objet reconstruit.

5.7 Conclusion

Dans ce chapitre, nous avons cherché à accélérer la résolution des nombreux systèmes linéaires intervenant dans le calcul du critère et du gradient. Pour cela, nous avons tout d'abord proposé une adaptation de l'algorithme Block-BiCGStab pour la résolution des systèmes linéaires en TMO. Cette méthode permet de résoudre conjointement plusieurs systèmes linéaires impliquant la même matrice opérateur. L'adaptation de cet algorithme à la TMO a nécessité de régler correctement l'initialisation, paramètre critique dans le comportement de la méthode. Pour éviter des difficultés de convergence dues à un mauvais conditionnement, nous avons proposé d'ajouter une composante aléatoire à l'initialisation, ce qui a pour effet de décorréler les vecteurs initiaux et donc de conserver un bon conditionnement au cours de la résolution.

Comme l'algorithme Block-BiCGStab n'est pas adapté pour une résolution sur une architecture multi-cœurs, nous avons proposé une généralisation de Block-BiCGStab, nommée Partial-Block BiCGStab, afin de pouvoir profiter de cette architecture. L'idée est de séparer l'ensemble des systèmes linéaires en sous-groupes et d'appliquer la méthode Block-BiCGStab sur chacun de ces groupes. Il est donc possible d'envoyer chaque résolution par Block-BiCGStab sur les cœurs disponibles. L'approche Partial-Block BiCGStab est donc un

compromis entre une résolution séquentielle par BiCGStab et une résolution conjointe par Block-BiCGStab. Nous avons notamment montré que la méthode la plus efficace est Partial-Block BiCGStab pour laquelle le nombre de groupes est équivalent au nombre de cœurs disponibles.

Nous avons présenté des résultats et comparaisons entre l'approche BiCGStab et l'approche Partial-Block BiCGStab. Tout d'abord, nous avons étudié l'influence du contraste de l'objet et de la fréquence des champs incidents sur le comportement de ces méthodes. On a observé que plus le contraste était élevé, plus l'approche Block-BiCGStab est avantageuse comparée à BiCGStab. De même, pour des incidences à hautes fréquences, ou de façon équivalente pour des objets de grande taille électrique, les performances de Block-BiCGStab sont accrues, comparées à celles de BiCGStab. Finalement, les approches proposées permettent de réduire d'autant plus le coût de calcul par rapport à BiCGStab que le problème est difficile, c'est-à-dire quand le contraste et la taille de l'objet sont élevés.

Nous avons finalement comparé l'utilisation de Partial-Block BiCGStab et BiCGStab dans l'inversion. Tout d'abord, la reconstruction a permis de confirmer que notre méthode d'inversion fournit des résultats très acceptables en termes de qualité d'image, même pour des objets fortement contrastés. De plus, nous observons un gain de temps de quasiment 20% quand les reconstructions sont réalisées avec l'utilisation de Block-BiCGStab plutôt que BiCGStab.

CHAPITRE 6 VALIDATION DE NOTRE ALGORITHME D'INVERSION SUR DONNÉES RÉELLES

Ce chapitre se concentre sur la reconstruction d'objets à partir de données réelles. L'objectif est double :

- évaluer les performances de notre algorithme d'inversion en termes de qualité de reconstruction sur des données expérimentales ;
- comparer l'utilisation de l'approche Block-BiCGStab par rapport au BiCGStab dans le cadre de reconstruction des données réelles.

Il n'existe que peu de données expérimentales accessibles librement pour la reconstruction en imagerie micro-ondes. La base de données [1] mise à disposition par l'Institut Fresnel (Geffrin et Sabouroux, 2009) permettent à quiconque n'ayant pas de moyens matériels, financiers ou techniques de pouvoir appliquer des méthodes de reconstruction sur des données réelles. Au laboratoire de Génie Électrique à l'École Polytechnique de Montréal, des configurations d'acquisition 2-D ont été développées pour l'imagerie du sein (Diaz Bolado, 2012). En revanche, dans le cadre de nos travaux uniquement tournés vers la résolution du problème inverse, il nous a semblé plus judicieux de se concentrer sur les données expérimentales de l'Institut Fresnel plutôt que de travailler sur un nouveau montage d'acquisition 3-D.

Plusieurs objets purement diélectriques de formes et de propriétés différentes ont été utilisés afin de générer les données expérimentales de la base Fresnel. Le montage d'acquisition et les objets sont décrits dans Geffrin et Sabouroux (2009). Le post-traitement des données y est également présenté.

Ces données ont fait l'objet d'un *concours* où plusieurs équipes scientifiques ont proposé leur algorithme d'inversion et leurs résultats de reconstruction. Le volume 25 d'*Inverse Problems* contient les différents papiers résultant de ce concours. Nous nous appuierons sur leurs résultats afin de comparer la qualité de reconstruction de notre méthode. Litman et Crocco (2009) résument les contributions des équipes ayant participé à ce concours et dressent un comparatif des méthodes selon différents critères.

Ce chapitre se décompose en trois sections :

- la première section propose un descriptif du montage d'acquisition et des objets utilisés. Une première analyse met en avant les propriétés de ces objets rendant la reconstruction difficile ;

1. Données disponibles à l'adresse : `http://www.fresnel.fr/3Ddatabase/`

— la deuxième section présente la mise en œuvre de notre algorithme d'inversion pour la reconstruction des données réelles. L'utilisation d'une procédure par saut de fréquence est notamment décrite ;
— la troisième section montre nos résultats de reconstruction obtenus avec la méthode décrite dans la section précédente. Ces résultats sont comparés à ceux de la littérature ;
— la dernière section compare les temps de calcul des reconstructions en fonction de la résolution des systèmes par BiCGStab ou Block-BiCGStab.

6.1 Présentation de la base de données de l'Institut Fresnel

Les données de l'Institut Fresnel forment un ensemble de fichiers contenant les valeurs de champs électriques diffractés pour différents objets, différentes fréquences et polarisations. Bien que le montage d'acquisition ait été brièvement présenté dans les chapitres 2 et 4, nous allons donner plus d'informations ici, issues de Geffrin et Sabouroux (2009).

6.1.1 La configuration du montage d'acquisition

Les mesures ont été réalisées dans une chambre anéchoïque, ce qui permet de faire l'hypothèse que les ondes électromagnétiques sont rayonnées à l'infini. Cette condition de rayonnement est essentielle pour notre modèle direct dans lequel l'hypothèse d'un milieu homogène infini est faite. 81 sources sont placées sur une sphère de rayon $r = 1.796$ m centrée sur l'origine du repère. Les ondes émises par ces sources peuvent être considérées comme des ondes planes. Des ondes incidentes avec deux polarisations sont émises : une co-polarisation et une cross-polarisation.

36 récepteurs sont placés sur un cercle dans le plan (xOy), centré sur l'origine et de même rayon que la sphère des sources. Les récepteurs ne mesurent que la composante verticale du champ diffracté. Des zones d'occlusion sont considérés dans le montage : pour chaque position de source, le champ n'est pas mesuré aux 36 récepteurs mais uniquement à un certain nombre positionnés suffisamment loin de la source. De plus, certaines mesures de champs ont saturé et ont donc été supprimées des données.

Les données mesurées sont ensuite corrigées et calibrées de façon à ce que les champs incidents soient considérés comme des ondes planes d'amplitude 1 V/m et de phase nulle au centre du repère.

Pour chaque objet, des mesures ont été réalisées pour 21 fréquences différentes : de 3 GHz à 8 GHz avec un pas de 0.25 GHz.

6.1.2 Les objets de la base

Cinq objets purement diélectriques ont été utilisés pour acquérir les données. Nous les présentons ici.

L'objet TwoCubes

L'objet TwoCubes est composé de deux cubes de 25 mm de côté et de contraste $\chi = 1.35$. Ces cubes sont en contact par un angle. La figure 6.1 présente cet objet.

L'objet TwoCubes est le plus facile à reconstruire pour différentes raisons :

- la taille des cubes est relativement faible par rapport à la longueur d'onde : 25 mm correspond à un quart de la longueur d'onde à 3 GHz et deux tiers à 8 GHz ;
- le contraste est relativement faible par rapport aux autres objets ;
- la forme de l'objet n'est pas trop complexe, contrairement aux deux derniers objets.

Le fait que l'objet soit de petite taille électrique et de faible contraste rend le critère moins non-linéaire et donc plus facile à minimiser.

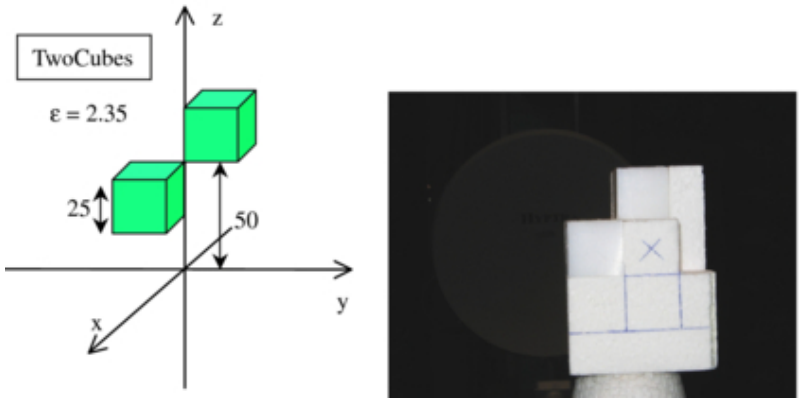


Figure 6.1 L'objet TwoCubes de la base Fresnel. Figure tirée de Geffrin et Sabouroux (2009).

L'objet TwoSpheres

L'objet TwoSpheres est composé de deux sphères identiques en contact de 50 mm de diamètre et de contraste $\chi = 1.6$. La figure 6.2 présente cet objet.

Par rapport aux deux cubes, l'objet TwoSpheres est plus difficile à reconstruire à cause d'une taille plus conséquente et d'un contraste légèrement plus élevé. De plus, la discrétisation cartésienne rend plus difficile la reconstruction d'objets sphériques. Il est possible de raffiner le maillage afin de mieux représenter les sphères mais cela se fait au détriment d'un coût

de calcul plus important. Ici, nous ne raffinerons pas le maillage et conserverons le même maillage pour tous les objets.

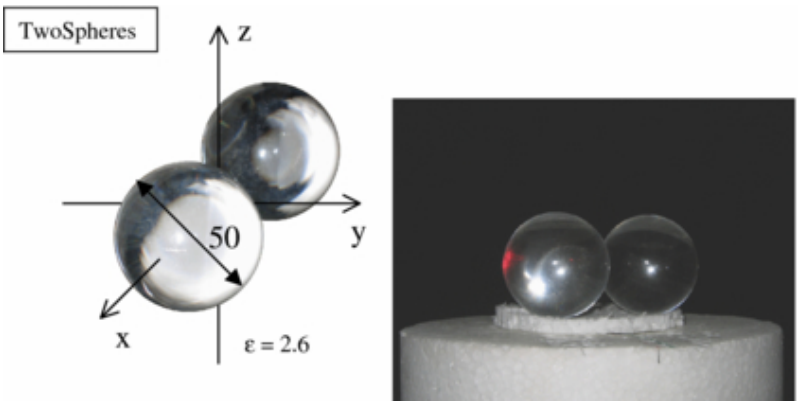


Figure 6.2 L'objet TwoSpheres de la base Fresnel. Figure tirée de Geffrin et Sabouroux (2009).

L'objet Cylinder

L'objet Cylinder, présenté dans la figure 6.3, est composé d'un cylindre de 80 mm de longueur, de 80 mm de diamètre et d'un contraste $\chi = 2.05$.

Cet objet est le plus volumineux et le plus contrasté de l'ensemble de la base. À 8 GHz, l'objet est plus grand que deux longueurs d'onde. Il est donc le plus difficile à reconstruire. Quelques auteurs ont essayé d'appliquer leurs méthodes de reconstruction sur cet objet et les résultats sont globalement médiocres (Yu et al., 2009; Li et al., 2009).

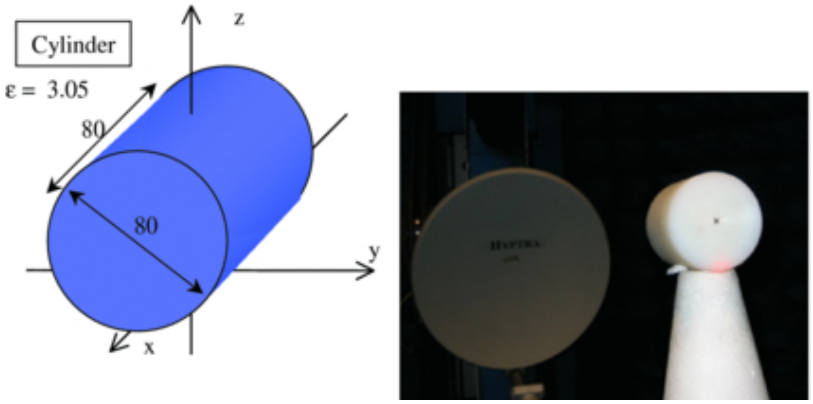


Figure 6.3 L'objet Cylinder de la base Fresnel. Figure tirée de Geffrin et Sabouroux (2009).

L'objet CubeSpheres

L'objet CubeSpheres est composé de 27 sphères identiques formant un cube de $3 \times 3 \times 3$ sphères. Les sphères ont un diamètre de 15.9 mm et un contraste $\chi = 1.6$. L'enjeu de cet objet réside non pas dans la taille et le contraste de l'ensemble mais dans le détail des formes. En effet, les sphères sont de petite taille ; la difficulté est donc de discerner chacune des sphères composant l'ensemble. L'objet est présenté dans la figure 6.4.

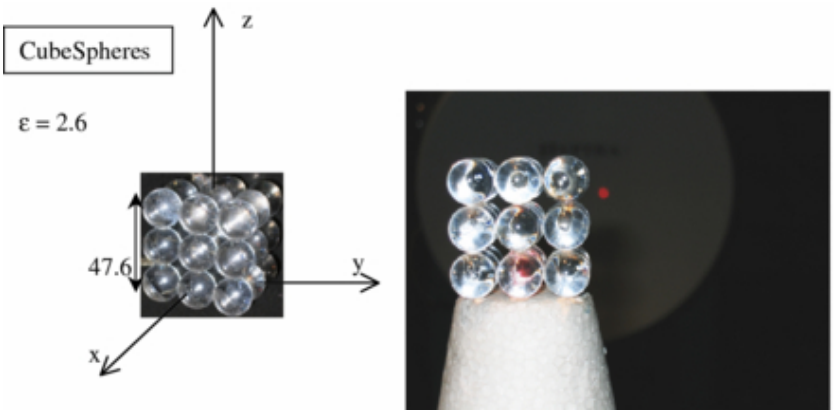


Figure 6.4 L'objet CubeSpheres de la base Fresnel. Figure tirée de Geffrin et Sabouroux (2009).

L'objet IsocaSphere

L'objet IsocaSphere est composé de 12 sphères où chacune des sphères est placées à un sommet d'un isocahèdre. Les sphères ont un diamètre de 23.8 mm et un contraste $\chi = 1.6$. Cet objet était à l'origine un objet *mystère* pour les équipes participant au concours. Les chercheurs n'avaient donc aucun renseignement sur cet objet ; seul un volume d'étude était donné. Pour nos résultats, nous avons également considéré cet objet inconnu et nous avons repris la même procédure de reconstruction que pour les autres objets. L'objet est présenté dans la figure 6.5.

6.2 Mise en œuvre de notre algorithme de reconstruction

Dans cette section, nous présentons la procédure utilisée pour réaliser la reconstruction des données expérimentales.

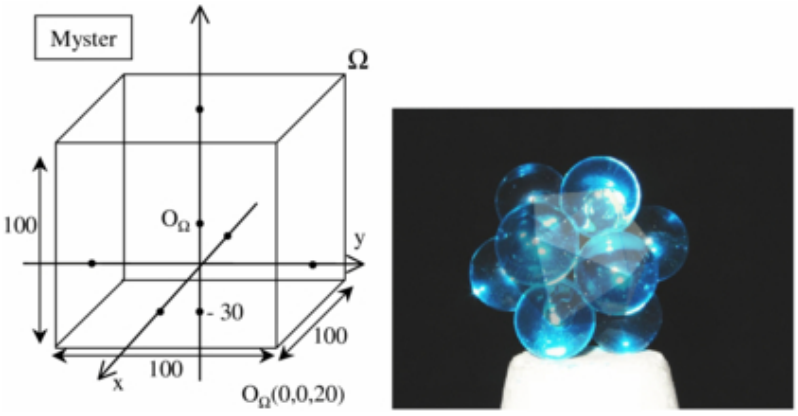


Figure 6.5 L'objet IsocaSphere de la base Fresnel. Figure tirée de Geffrin et Sabouroux (2009).

6.2.1 Réciprocité de Lorentz et polarisations utilisées

Comme nous l'avons vu dans la section 4.5 (page 57), la configuration du montage d'acquisition utilisé est propice à l'utilisation du théorème de réciprocité de Lorentz afin de considérer un montage adjoint dans lequel 36 sources virtuelles sont considérées. Pour rappel, parmi les six équipes ayant participé, seules deux d'entre elles ont appliqué le principe de réciprocité (Catapano et al., 2009; Chaumet et Belkebir, 2009). Les autres considèrent le montage d'acquisition réel directement dans leur modélisation, soit 81 sources.

Concernant les polarisations utilisées, nous avons pris en compte les deux types de polarisation simultanément. Avec le théorème de réciprocité, cela revient à considérer qu'à chaque récepteur virtuel, la mesure du champ est vectorielle. Nous pouvons donc supposer que pour chaque récepteur, les trois composantes spatiales du champ diffracté $\overrightarrow{\boldsymbol{E}}^{\text{scat}}$ sont mesurées.

La configuration adjointe peut être résumée comme suit :

- 36 sources virtuelles (aux positions des récepteurs réels) émettent une onde plane incidente polarisée verticalement.
- 81 récepteurs virtuels mesurent les trois composantes du champ diffracté pour chaque illumination. Pour simplifier le problème, cette interprétation ne prend pas en compte les zones d'occlusion et les données saturées. Il est évidemment nécessaire de les prendre en compte dans le modèle direct et l'inversion. On suppose donc $N_R = 81 \times 3 = 243$ mesures par source.

Comme détaillé dans le chapitre 4, la mise en œuvre optimale pour le calcul du gradient avec cette configuration est la mise en œuvre directe (4.18), soit $2N_S = 72$ systèmes à résoudre. Comparons l'utilisation de L-BFGS et de Gauss-Newton pour la reconstruction. L'algorithme L-BFGS, nécessitant uniquement le gradient pour estimer une direction de descente, est

bien adapté à cette configuration puisque seulement $2N_S$ systèmes doivent être résolus. En revanche, l'algorithme Gauss-Newton nécessite le calcul du jacobien, ce qui revient à résoudre $N_S + N_R = 279$ systèmes pour évaluer une direction de descente, soit quasiment 8 fois plus que pour L-BFGS. Il faut ajouter à cela le coût de calcul supplémentaire de la résolution des N systèmes linéaires pour calculer la direction de descente, où N est le nombre de voxels. Ceci n'est qu'une étude du nombre de systèmes à résoudre par itération du problème inverse et cela ne prend évidemmment pas en compte le nombre d'itérations global avant convergence. Notre travail ne met pas l'accent sur cette comparaison entre les différents algorithmes d'inversion.

6.2.2 Volume d'étude et discrétisation

Tous les objets sont contenus dans un volume cubique de taille $10 \times 10 \times 10$ cm. Les objets ne sont pas forcément centrés sur l'origine ; l'emplacement du volume est donc choisi différemment pour chaque objet de façon à ce que l'objet soit dans le volume et ne soit pas en contact avec les limites du volume d'étude. Seul l'objet TwoSpheres atteint les bords de ce volume car les deux sphères en contact mesurent chacune 5 cm.

Nous avons décidé de discrétiser le volume en $30 \times 30 \times 30$ voxels, ce qui correspond à une résolution d'environ 3.3 mm. Pour la fréquence la plus faible (3 GHz), il y a 30 voxels par longueur d'onde ; pour la fréquence la plus élevée (8 GHz), il y a environ 11.2 voxels par longueur d'onde. Ces valeurs restent convenables pour se placer dans le domaine de validité du modèle numérique. Ces valeurs sont dans la gamme de celles utilisées par les participants au défi de reconstruction. Certains ont utilisé une résolution allant jusqu'à 2 mm, d'autres ont utilisé des résolutions de 5 mm.

De même, on peut calculer la longueur d'onde à 8 GHz dans l'objet le plus contrasté, soit l'objet Cylinder. Celle-ci vaut 21.5 mm, soit environ 6.4 voxels par longueur d'onde. On arrive ici aux limites du modèle. Nous verrons tout de même dans la section suivante que la reconstruction de l'objet Cylinder avec cette résolution conduit à des résultats convenables.

6.2.3 Procédure par saut de fréquence

Pour reconstruire au mieux les objets de la base, les données pour la plus haute fréquence, c'est-à-dire 8 GHz, sont utilisées dans l'algorithme d'inversion. Puisqu'à cette fréquence, le problème peut être fortement non-linéaire, nous avons choisi d'appliquer une procédure par saut de fréquence. L'intérêt est d'utiliser successivement les données aux fréquences plus basses afin de déterminer une initialisation pour les données à 8 GHz permettant d'obtenir une reconstruction convenable. Bien que des mesures à 21 fréquences différentes ont été

réalisées par l'Institut Fresnel, nous n'utiliserons que les fréquences entières : 3, 4, 5, 6, 7 et 8 GHz.

Afin de choisir une procédure acceptable pour la reconstruction des objets Fresnel, nous avons défini la procédure suivante :

- aucun *a priori* n'est connu sur l'objet à reconstruire. On initialise alors la reconstruction avec un contraste nul en tout voxel du domaine d'étude V ;
- avec les données à 3 GHz, on réalise dix itérations de descente par L-BFGS sur le critère $\mathcal{F}$. La solution après ces dix itérations est nommée $\hat{\boldsymbol{x}}_3$;
- avec les données à 4 GHz, on initialise l'algorithme L-BFGS avec $\hat{\boldsymbol{x}}_3$ et on réalise dix itérations pour minimiser le critère. La solution est nommée $\hat{\boldsymbol{x}}_4$;
- on répète l'étape précédente avec les données à 5 GHz, 6 GHz et 7 GHz. Après dix itérations, on obtient respectivement les solutions $\hat{\boldsymbol{x}}_5$, $\hat{\boldsymbol{x}}_6$ et $\hat{\boldsymbol{x}}_7$;
- on reconstruit finalement avec les données à 8 GHz en initialisant avec $\hat{\boldsymbol{x}}_7$. L'algorithme L-BFGS est arrêté quand la norme du gradient atteint la tolérance fixée à 10^{-6}.

Cette procédure est identique pour tous les objets sauf pour l'objet Cylinder. En effet, à cause du volume important et du contraste plus élevé que les autres objets, le critère est fortement non-linéaire et la procédure précédente conduit à une initialisation $\hat{\boldsymbol{x}}_7$ qui ne permet pas de faire converger l'algorithme vers un minimum local convenable. Nous avons donc fait le choix d'effectuer plus d'itérations par fréquence afin d'obtenir une reconstruction acceptable. Pour l'objet Cylinder, nous réalisons 20 itérations à chaque fréquence plutôt que 10 itérations pour les autres objets.

Les paramètres du terme de régularisation (3.6) sont fixés empiriquement et sont identiques pour toutes les reconstructions. On a posé le coefficient de régularisation $\lambda = 10^{-5}$ et le paramètre "$\ell_2\ell_1$" $\delta = 10^{-2}$. De même, la tolérance pour la résolution des systèmes par BiCGStab ou Partial-Block BiCGStab est identique pour tous les objets ; elle est fixée à 10^{-6}.

On notera enfin que, mis à part le terme de régularisation, aucune information *a priori* n'est utilisée sur le contraste des objets : aucune hypothèse de non-négativité sur le contraste ou de binarité (c'est-à-dire que le contraste ne peut prendre que deux valeurs différentes) n'est faite.

6.3 Résultats de reconstruction

Nous présentons dans cette section les résultats de la reconstruction des différents objets de la base Fresnel suivant la procédure définie précédemment. Les images obtenues sont analysées et

comparées à celles présentées dans la littérature. Il est évident que les chercheurs ont appliqué chacun leurs méthodes de reconstruction et les procédures sont toutes différentes : mono-fréquence, multi-fréquence, procédure par saut de fréquence, polarisation simple ou multiple, etc. Nous nous restreindrons à une comparaison en termes de qualité de reconstruction et non à une comparaison algorithmique avec les méthodes utilisées dans la littérature.

Les résultats suivants ne représentent que la partie réelle du contraste des objets. Comme les objets sont purement diélectriques, la partie imaginaire du contraste est nulle. Lors de la reconstruction, le contraste estimé n'est évidemment pas purement réel mais la partie imaginaire estimée est relativement faible. Elle n'est donc pas représentée dans les résultats de cette section.

Objet TwoCubes

La figure 6.6 présente la partie réelle du contraste exact et du contraste reconstruit pour l'objet TwoCubes. Les 30 tranches horizontales du volume d'étude discrétisé sont représentées. Il est important de voir que les tranches de l'objet "exact" ne sont qu'une représentation discrétisée de l'objet réel : si le centre d'un voxel appartient à l'objet, alors tout le voxel prend la valeur du contraste de l'objet. On voit bien qu'un cube sur la figure 6.6 est inscrit sur 7 tranches, soit une taille de $7 \times 3.3 = 23.3$ mm plutôt que les 25 mm du cube réel.

La reconstruction de l'objet TwoCubes permet de distinguer clairement les deux cubes ; les faces et angles sont relativement bien marqués. Le point de contact est également visible, bien que les valeurs du contraste à cet endroit soient inférieures au reste de l'objet. Quantitativement, le contraste est légèrement sous-estimé : entre 1 et 1.1 au lieu d'un contraste réel de 1.35. C'est également le cas pour les autres objets et pour les reconstructions présentes dans la littérature.

La figure 6.7 donne une représentation 3-D d'une isosurface de la reconstruction. Le seuil de l'isosurface a été fixé à un contraste de 0.6, soit environ la moitié du contraste réel. Cette représentation confirme l'analyse proposée ci-dessus. On note également la difficulté de la reconstruction au niveau du point de contact.

Enfin, la figure 6.8 montre l'initialisation de l'algorithme pour la reconstruction à 8 GHz. On rappelle que cette initialisation est la solution de reconstruction à 7 GHz où l'algorithme L-BFGS est tronqué à 10 itérations. L'initialisation est une solution lissée et faiblement contrastée de l'objet réel. Cela permet d'avoir une idée de la position et de la forme de l'objet afin de converger vers un minimum local acceptable.

Objet TwoSpheres

La figure 6.9 représente les 30 tranches de l'objet exact et reconstruit TwoSpheres. La solution obtenue correspond bien aux deux sphères. Tout comme l'objet TwoCubes, on observe que le point de contact entre les deux sphères semble difficile à détecter correctement. Quantitativement, le contraste est aussi légèrement sous-estimé mais on observe tout de même des valeurs très fortes au point de contact, valeurs étonnamment supérieures au contraste réel qui est de 1.6.

La figure 6.10 présente une vue 3-D de la reconstruction. Le seuil de l'isosurface est fixé à 0.6. Les deux sphères sont bien représentées. On observe l'effet de la discrétisation cartésienne du volume sur la forme des sphères : la forme arrondie ne peut pas être représentée fidèlement. Enfin, comme nous l'avons dit précédemment, l'objet TwoSpheres "touche" les bords du volume d'étude. On observe que les extrémités de l'objet sont mal estimées (forme légèrement cubique). Cela peut être corrigé en choisissant un volume légèrement plus grand. La figure 6.11 montre justement le résultat de reconstruction en utilisant un volume de taille $11 \times 11 \times 11$ cm ; les extrémités sont cette fois-ci bien estimées.

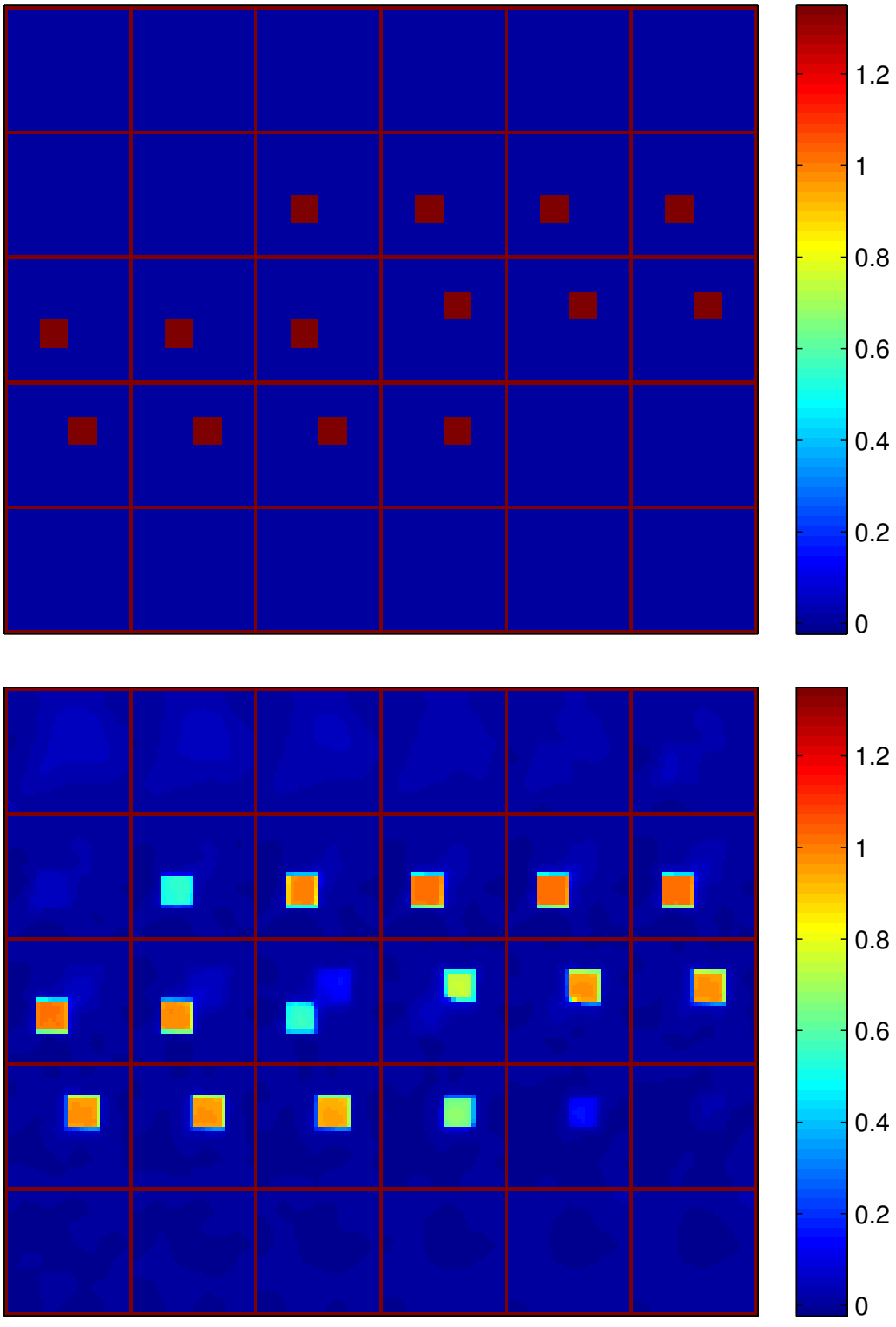


Figure 6.6 Reconstruction de l'objet TwoCubes avec procédure par saut de fréquence. En haut, les 30 tranches horizontales de l'objet exact. En bas, les 30 tranches de la solution de la reconstruction.

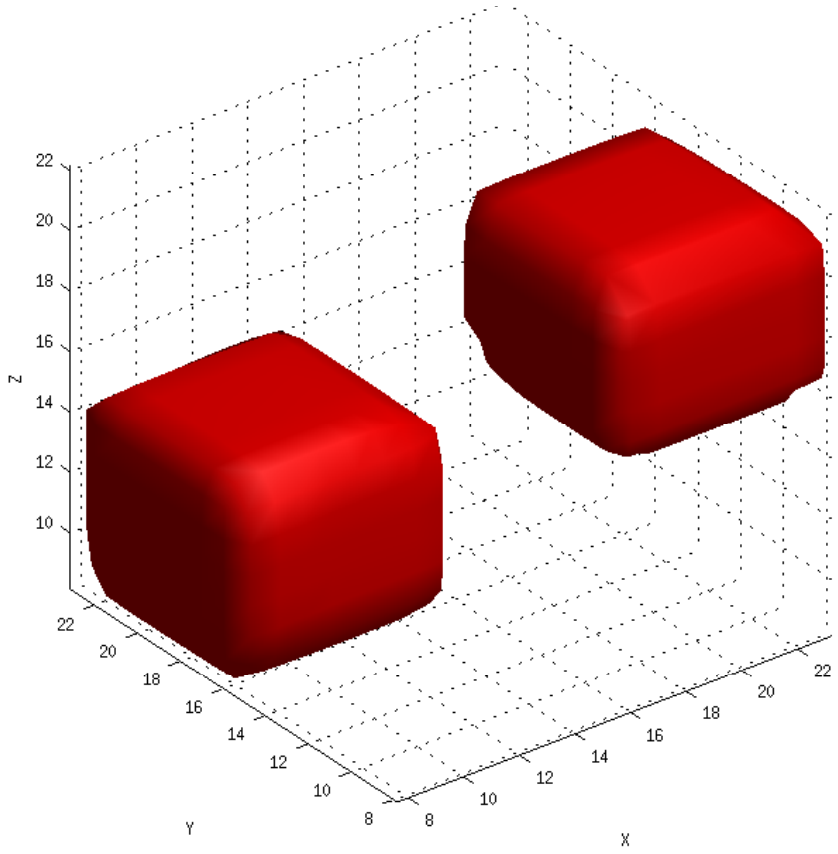


Figure 6.7 Isosurface de la solution de la reconstruction TwoCubes. Le seuil de l'isosurface est fixé à un contraste de 0.6.

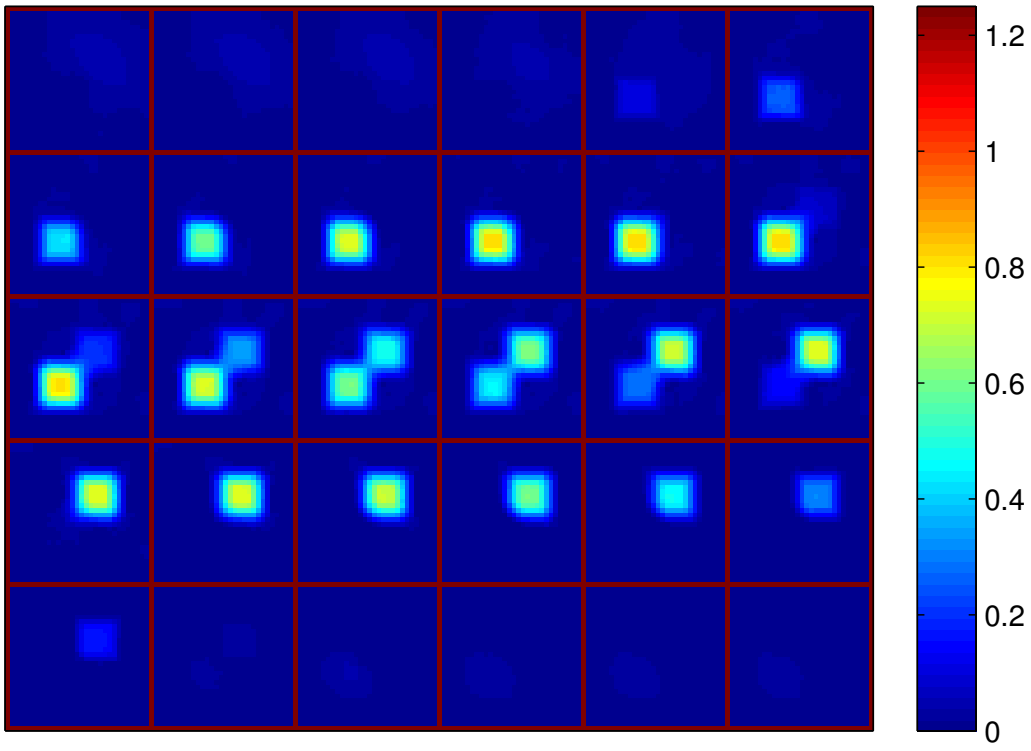


Figure 6.8 Initialisation de la reconstruction de l'objet TwoCubes pour la dernière fréquence (soit 8 GHz). Il s'agit de la solution obtenue après dix itérations de la reconstruction avec les données à 7 GHz.

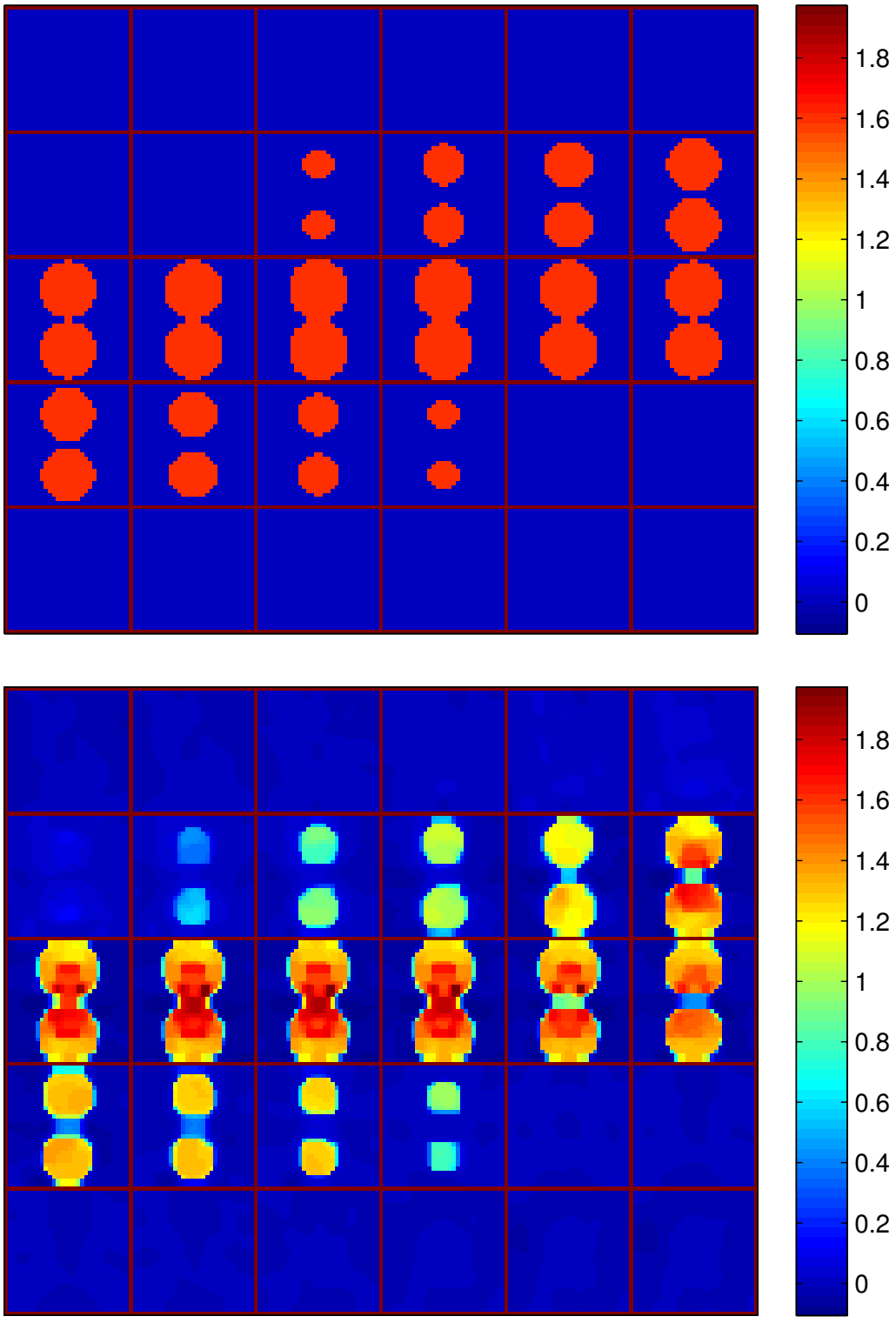


Figure 6.9 Reconstruction de l'objet TwoSpheres avec procédure par saut de fréquence. En haut, les 30 tranches horizontales de l'objet exact. En bas, les 30 tranches de la solution de la reconstruction.

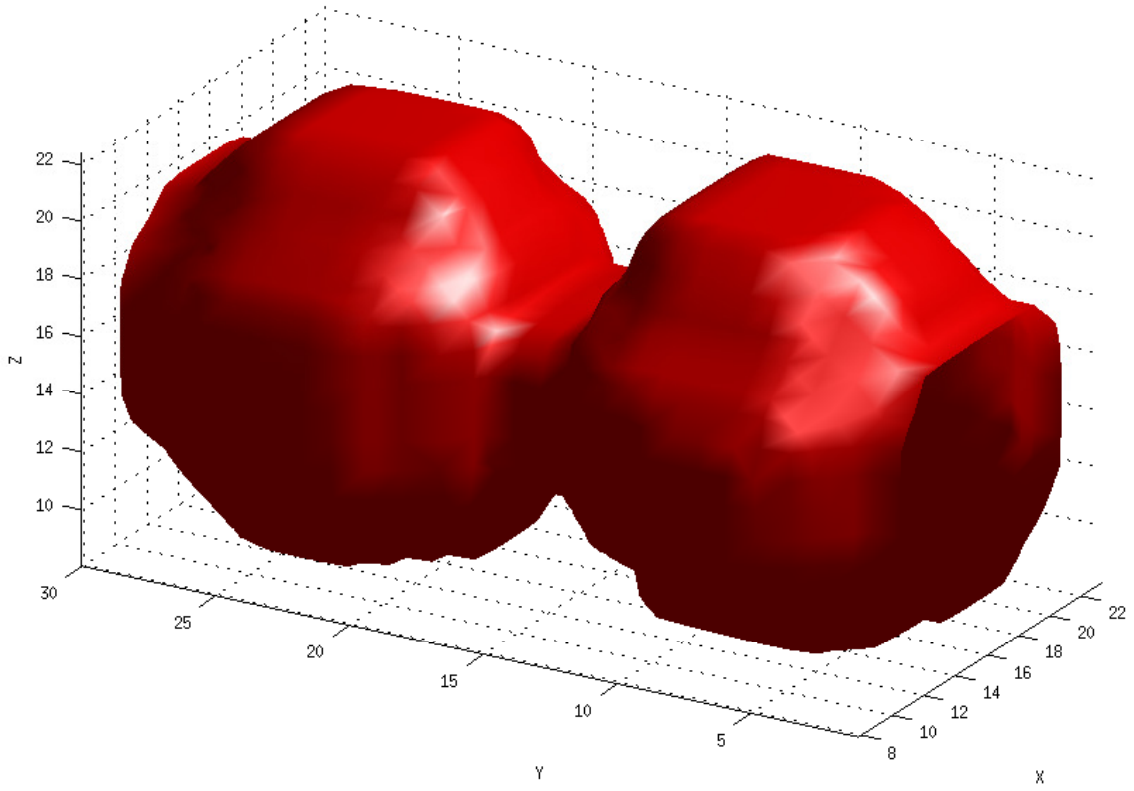


Figure 6.10 Isosurface de la solution de la reconstruction TwoSpheres. Le seuil de l'isosurface est fixé à un contraste de 0.6.

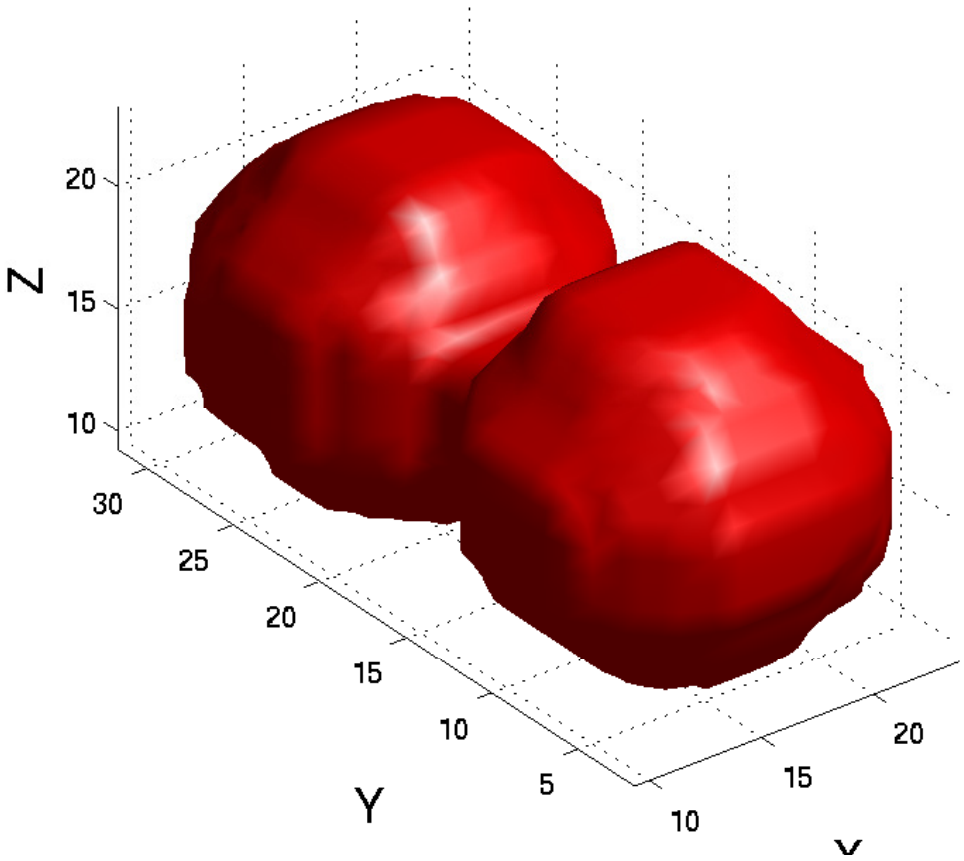


Figure 6.11 Isosurface de la solution de la reconstruction TwoSpheres pour un volume d'étude plus grand. Le seuil de l'isosurface est fixé à un contraste de 0.6.

Objet Cylinder

L'objet Cylinder est le plus volumineux et le plus contrasté de la base Fresnel. L'application de la procédure par saut de fréquence décrite en 6.2.3 n'a pas fourni de résultats satisfaisants. En revanche, en réalisant 20 itérations par fréquence plutôt que 10, la solution obtenue à la fréquence 7 GHz permet d'initialiser correctement l'inversion à 8 GHz, et donc de converger vers un minimum local acceptable.

La figure 6.12 présente les 30 tranches de la reconstruction de l'objet Cylinder et la figure 6.13 en montre une représentation 3-D. Ici, le seuil de l'isosurface est fixé à un contraste de 1, soit environ la moitié du contraste réel. Le cylindre est bien reconstruit : la forme est bien détectée et le contraste est légèrement sous-estimé (entre 1.8 et 2 au lieu de 2.05).

Objet CubeSpheres

L'objet CubeSpheres est difficile à reconstruire à cause des détails de l'objet et de la faible taille des sphères. Les figures 6.14 et 6.15 représentent la reconstruction de cet objet. On voit bien que la solution est éloignée de l'objet exact. La représentation 3-D montre 9 "colonnes", au lieu des 27 sphères. Les sphères sont donc estimées correctement dans le plan (xOy) mais invisibles selon la direction z. Mudry et al. (2012) estiment d'ailleurs la résolution de la reconstruction à $\lambda/4$ dans le plan (xOy) et $\lambda/1.8$ selon z. À 8 GHz, cela revient à une résolution de reconstruction verticale d'environ 21 mm alors que les sphères ont un diamètre d'environ 16 mm. Cette analyse de la résolution explique la qualité de la reconstruction. Les auteurs s'étant attaqué à cet objet ont des solutions semblables à celle que nous avons obtenue (De Zaeytijd et Franchois, 2009; Yu et al., 2009).

Objet IsocaSphere

Les sphères de l'objet IsocaSphere sont plus grandes que celles de l'objet CubeSpheres. Les sphères sont donc plus faciles à détecter car la taille de celles-ci est plus élevée que la résolution de la reconstruction. Les figures 6.16 et 6.17 montrent la reconstruction de cet objet. La forme de l'objet est bien déterminée mais le contraste est légèrement sous-estimé, entre 1 et 1.4 au lieu de 1.6. L'isosurface montre la géométrie de l'objet et les sphères sont clairement visibles. En revanche, la reconstruction à l'intérieur de l'objet est moins nette, tout comme les points de contact entre les sphères qui restent flous. Les résultats des autres équipes sur cet objet ne sont pas meilleurs.

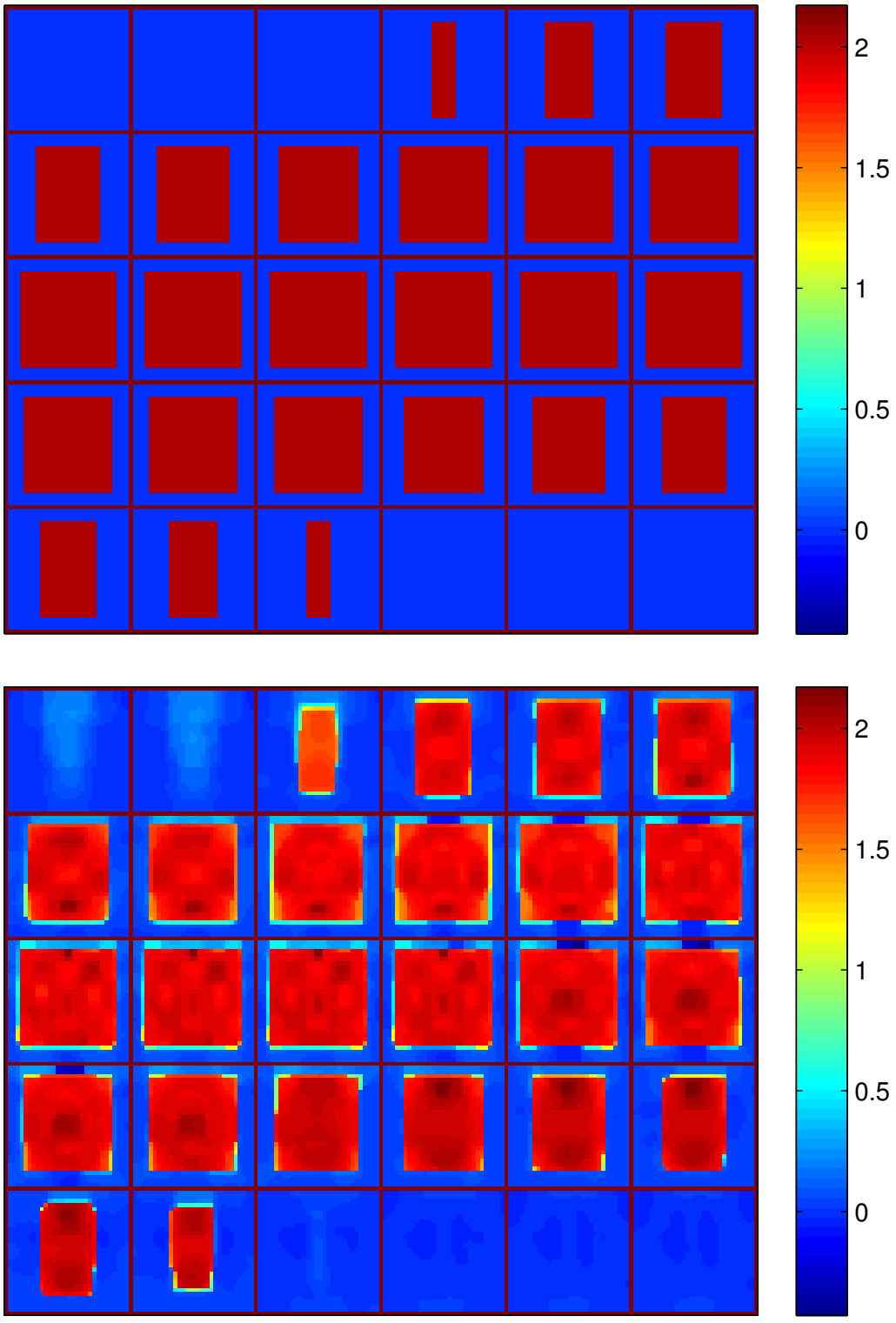


Figure 6.12 Reconstruction de l'objet Cylinder avec procédure par saut de fréquence. En haut, les 30 tranches horizontales de l'objet exact. En bas, les 30 tranches de la solution de la reconstruction.

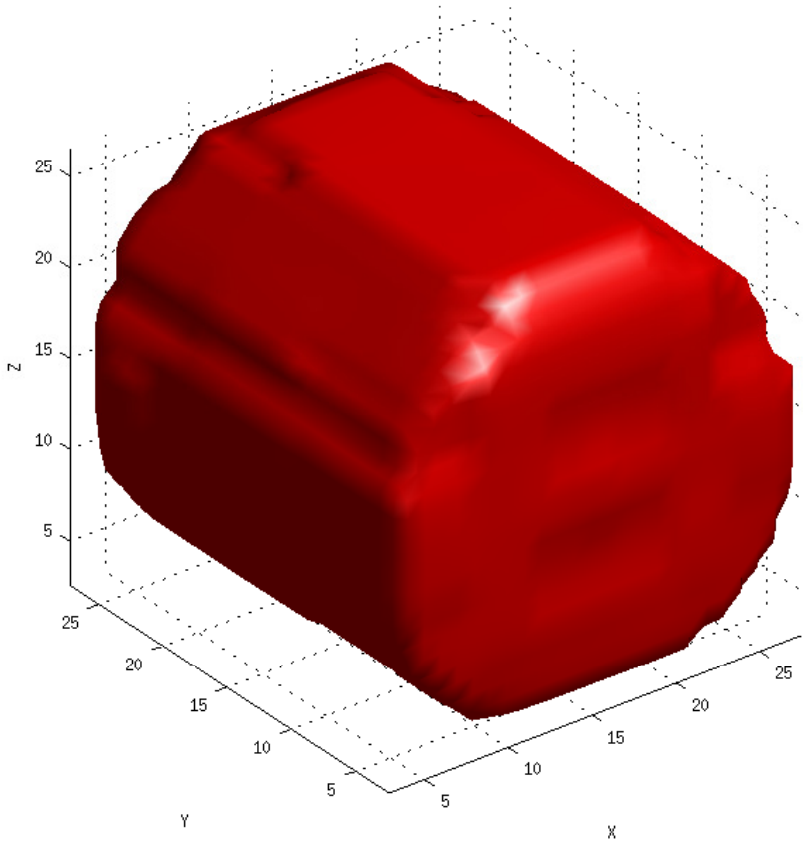


Figure 6.13 Isosurface de la solution de la reconstruction Cylinder. Le seuil de l'isosurface est fixé à un contraste de 1.

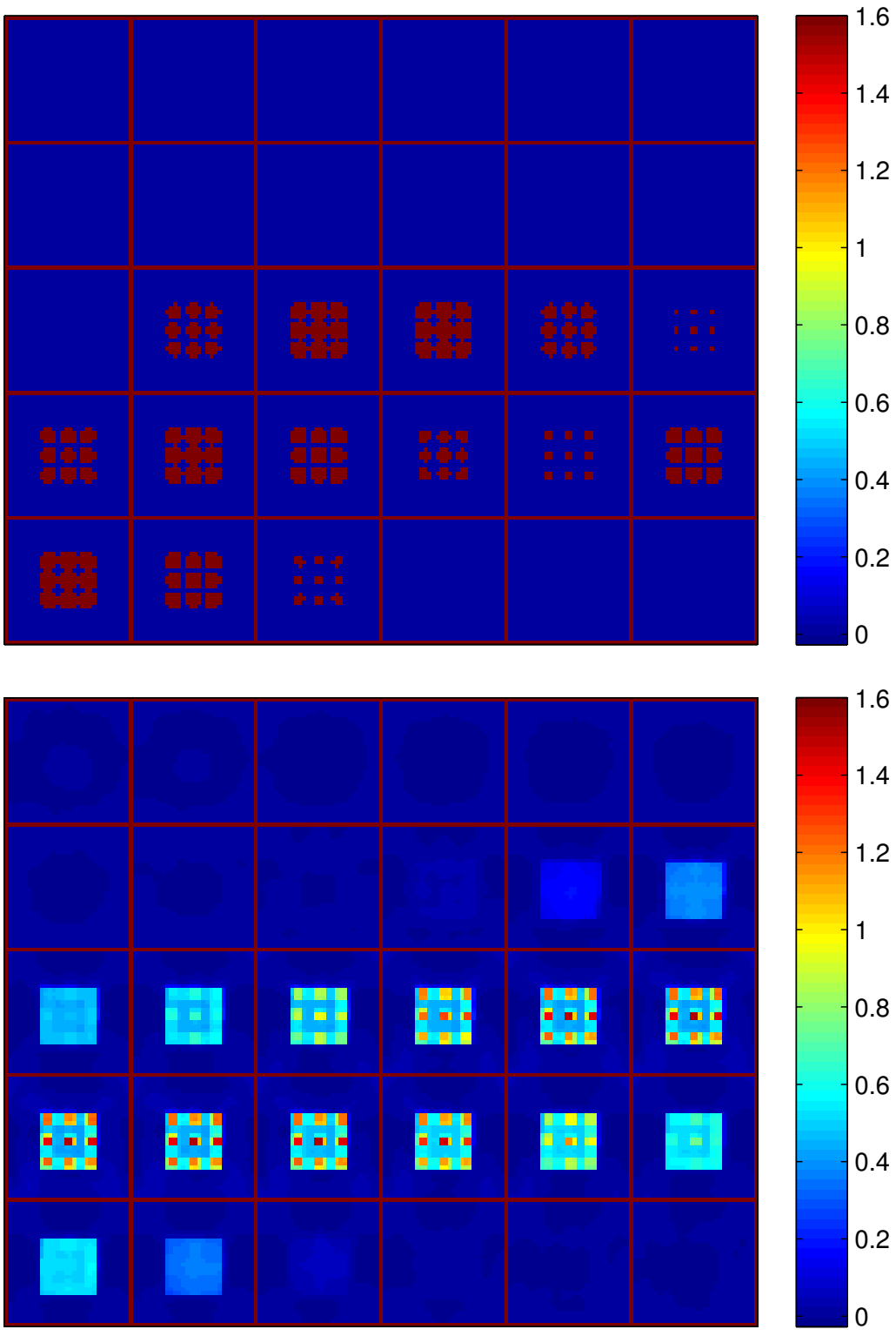


Figure 6.14 Reconstruction de l'objet CubeSpheres avec procédure par saut de fréquence. En haut, les 30 tranches horizontales de l'objet exact. En bas, les 30 tranches de la solution de la reconstruction.

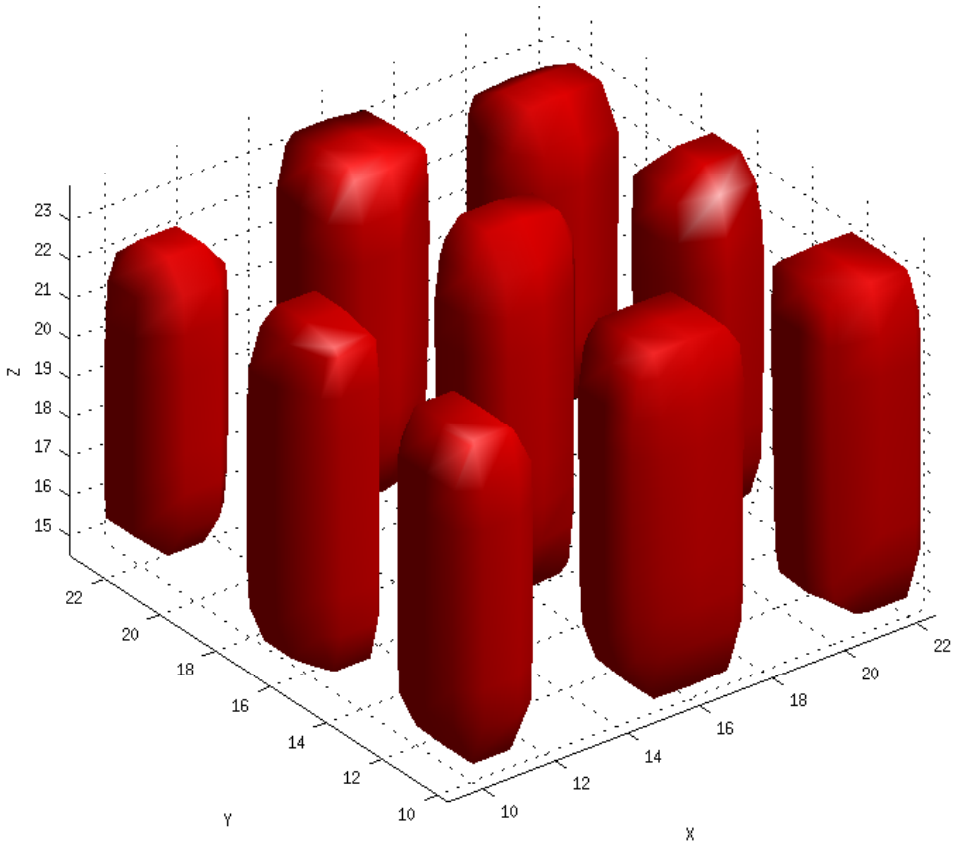


Figure 6.15 Isosurface de la solution de la reconstruction CubeSpheres. Le seuil de l'isosurface est fixé à un contraste de 0.8.

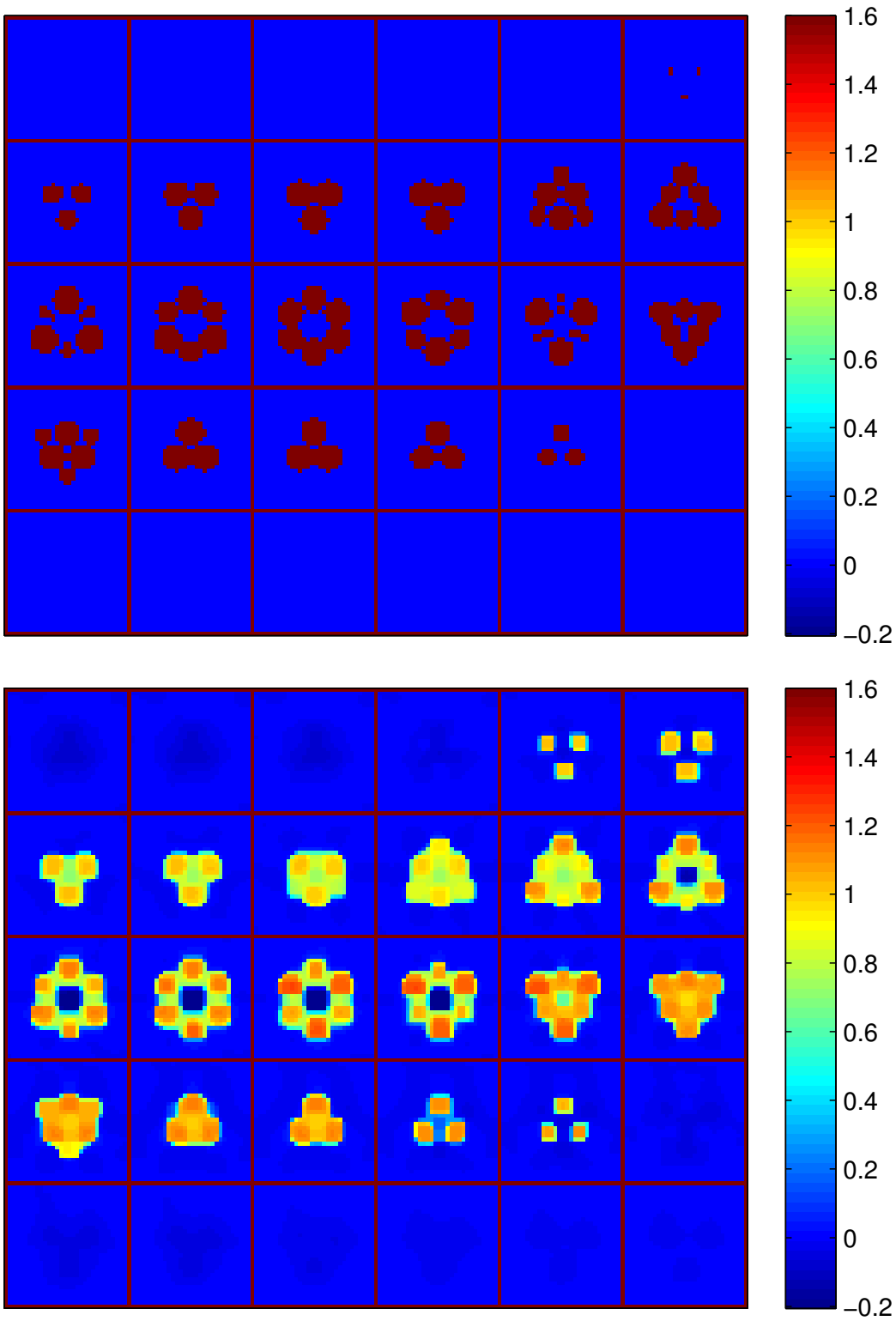


Figure 6.16 Reconstruction de l'objet IsocaSphere avec procédure par saut de fréquence. En haut, les 30 tranches horizontales de l'objet exact. En bas, les 30 tranches de la solution de la reconstruction.

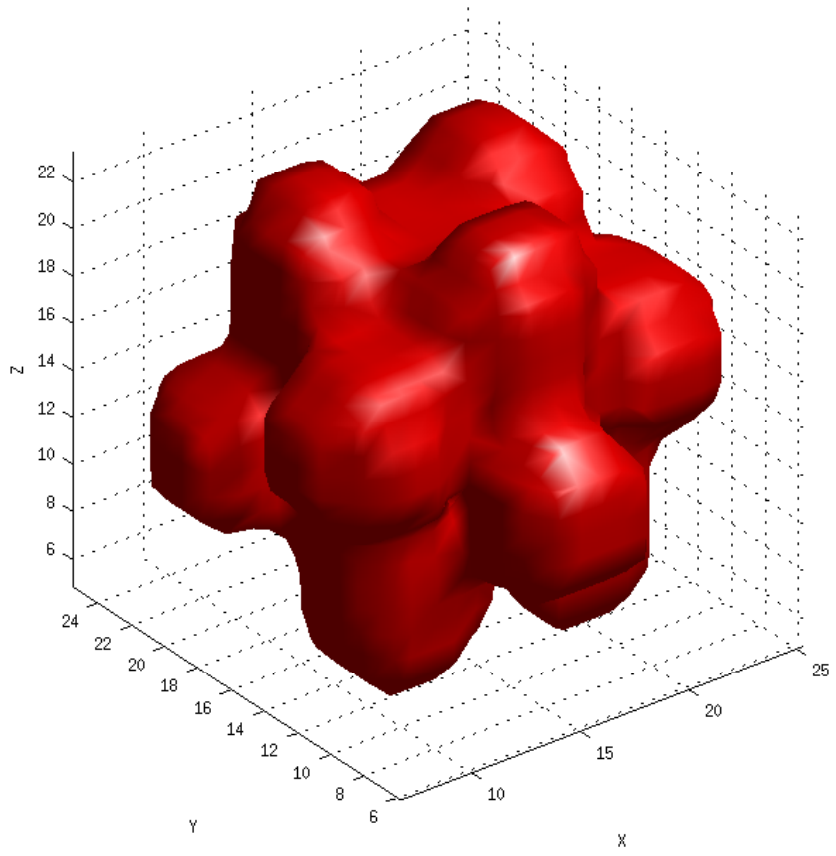


Figure 6.17 Isosurface de la solution de la reconstruction IsocaSphere. Le seuil de l'isosurface est fixé à un contraste de 0.8.

6.4 Comparaison des coûts de calcul des différentes approches

Nous cherchons maintenant à comparer le coût de calcul de la reconstruction avec l'utilisation de BiCGStab et Block-BiCGStab pour la résolution des systèmes. Dans le chapitre précédent, nous nous étions limités à l'utilisation de l'algorithme Block-BiCGStab sur un cœur afin de mettre en avant uniquement les contributions algorithmiques sans considérer les possibilités de parallélisation des opérations élémentaires. Nous rappelons que ces travaux de thèse n'ont pas pour objectif d'étudier des solutions de parallélisation plus complexes des algorithmes présentés. Nous utilisons le logiciel MATLAB pour réaliser les reconstructions. MATLAB propose nativement l'utilisation implicite de la parallélisation des fonctions intégrées telles que le produit matriciel ou la `FFT`.

Dans ce chapitre, nous avons donc effectué les reconstructions non pas en limitant les opérations sur un unique cœur mais en profitant de la parallélisation des opérations élémentaires proposée par MATLAB. Comme les algorithmes BiCGStab et Block-BiCGStab font appel aux mêmes opérations, les différences de coût de calcul avec cette solution sont similaires aux différences obtenues avec une reconstruction sur un cœur. L'intérêt majeur est de profiter de l'architecture multi-cœurs de la machine et de la parallélisation des opérations intégrées de MATLAB pour utiliser l'algorithme Block-BiCGStab dans une version "parallélisée". On parlera de résolution avec *multithreading*.

Les résultats de coût de calcul présentés dans cette section permettent de comparer le temps de reconstruction avec BiCGStab multithreadé et Block-BiCGStab multithreadé. Le tableau 6.1 présente les coûts de calcul de la reconstruction avec BiCGStab et Block-BiCGStab pour les différents objets. Tout d'abord, notons que la reconstruction avec Block-BiCGStab est plus rapide que celle avec BiCGStab, quel que soit l'objet. Mais surtout, plus l'objet est volumineux et contrasté, plus la reconstruction est lente et plus l'approche Block-BiCGStab est avantageuse face à BiCGStab. Par exemple, pour l'objet TwoCubes, petit et faiblement contrasté, la reconstruction dure environ deux heures et l'approche Block-BiCGStab dure 10% de moins que BiCGStab. À l'opposé, pour l'objet Cylinder, la reconstruction est 30 fois plus longue pour atteindre la même tolérance sur la norme du gradient ; l'approche Block-BiCGStab est deux fois plus rapide que l'approche BiCGStab. Cette première analyse met à nouveau en avant le fait que Block-BiCGStab est bien plus performant pour des objets de grande taille électrique et de contraste élevé.

Pour mieux comprendre l'écart de temps entre BiCGStab et Block-BiCGStab, nous étudions les résolutions de systèmes intervenant à chaque itération de l'algorithme L-BFGS. Prenons l'objet Cylinder pour lequel nous observons le plus gros écart entre BiCGStab et Block-

Objet de la base	Temps BiCGStab	Temps Block-BiCGStab	Rapport temps Block/BiCGStab
TwoCubes	2.33 h	2.07 h	0.89
TwoSpheres	8.60 h	6.41 h	0.75
Cylinder	60.0 h	28.3 h	0.47
CubeSpheres	3.21 h	2.60 h	0.81
IsocaSphere	4.38 h	3.58 h	0.82

Tableau 6.1 Temps de reconstruction pour les objets de la base Fresnel. Les coûts de calcul des reconstructions avec BiCGStab et Block-BiCGStab sont comparées.

BiCGStab. Sur les dix dernières itérations de la reconstruction, la moyenne du temps pour une itération de L-BFGS avec l'utilisation de BiCGStab est d'environ 1250 s : environ 130 itérations par système pour la résolution des champs totaux $\mathbf{E}^{\text{tot}}$ et environ 165 itérations pour résoudre les systèmes $\mathbf{F}$ intervenant dans le gradient. Pour la reconstruction avec l'utilisation de Block-BiCGStab, sur les dix dernières itérations, le temps moyen d'une itération L-BFGS est de 490 s, soit 60% de moins que le temps moyen d'une itération avec l'utilisation de BiCGStab. Sur ces dix dernières itérations de L-BFGS, le nombre d'itérations moyen pour résoudre les champs totaux par Block-BiCGStab est d'environ 58 itérations et 54 itérations pour le gradient. On voit donc bien que les résolutions par Block-BiCGStab pour un tel objet sont bien plus efficaces que par BiCGStab.

6.5 Conclusion

Dans ce chapitre, nous avons présenté les résultats de reconstruction sur l'ensemble des objets de la base Fresnel. Une procédure par saut de fréquence a été utilisée, identique pour tous les objets : 10 itérations de L-BFGS aux fréquences entières de 3 GHz à 7 GHz afin d'initialiser la reconstruction finale à 8 GHz. Seul l'objet Cylinder, plus difficile à cause d'une taille plus conséquente et d'un contraste plus élevé, a nécessité de réaliser plus d'itérations (20 au lieu de 10) à chaque fréquence afin d'obtenir une bonne initialisation pour la reconstruction finale.

Notre algorithme d'inversion, basé sur la minimisation d'un critère non-linéaire avec régularisation additive et sur la méthode L-BFGS, fournit des solutions de reconstruction aussi bonnes voire meilleures que celles présentées dans la littérature. L'objet difficile Cylinder a été reconstruit correctement, comme les autres objets de la base. Seul l'objet CubeSpheres, composé de petites sphères, n'a pas été reconstruit correctement. Cela est dû à une limite physique : les sphères étant petites devant la longueur d'onde, la résolution de la reconstruction selon l'axe d'acquisition des données (ici, l'axe vertical) ne permet pas de distinguer clairement ces sphères selon l'axe vertical.

Finalement, nous avons comparé notre algorithme de reconstruction avec l'utilisation de BiCGStab et Block-BiCGStab pour résoudre les systèmes linéaires dans le critère et le gradient. La reconstruction avec l'approche par blocs a toujours été plus rapide que celle avec BiCGStab. Nous obtenons même un facteur 2 pour la reconstruction de l'objet Cylinder. Ces résultats confirment l'efficacité de la méthode Block-BiCGStab pour des objets de grande taille électrique et fortement contrastés.

CHAPITRE 7 VERS UNE GESTION EFFICACE DES EFFETS DE LA TRONCATURE SUR LA COHÉRENCE ENTRE LE CRITÈRE ET LE GRADIENT

Comme nous l'avons vu dans les chapitres 3 et 5, les résolutions des nombreux systèmes linéaires intervenant dans le calcul du critère et du gradient sont en pratique effectuées par des méthodes itératives, notamment par BiCGStab ou Block-BiCGStab, et sont les opérations les plus coûteuses dans l'inversion. Pour réduire le coût de calcul, les résolutions de ces systèmes sont généralement tronquées : les itérations sont arrêtées quand le résidu relatif atteint un seuil fixé par l'utilisateur (Zhang et al., 2003a). Les quantités calculées ne sont donc pas exactes et les erreurs d'approximation entraînent naturellement des évaluations inexactes du critère et du gradient. En pratique, un compromis doit être obtenu entre :

- une résolution avec une tolérance stricte mais coûteuse : le critère et le gradient calculés sont proches des valeurs exactes ;
- une résolution avec une tolérance grossière mais plus rapide : le critère et le gradient sont fortement altérés par les approximations dues à la troncature.

Non seulement les quantités calculées (critère et gradient) sont approchées et ne correspondent pas aux valeurs exactes, mais de plus, il y a une incohérence entre ces deux quantités, c'est-à-dire que le gradient évalué n'est pas exactement celui du critère calculé. Si la troncature est grossière, cette incohérence entre le critère et le gradient calculés est accentuée et peut conduire à des difficultés dans la convergence de l'algorithme de reconstruction. En effet, le calcul de la direction de descente et de la recherche de pas est étroitement lié à la cohérence qu'il y a entre le critère et le gradient. S'il y a une incohérence entre le critère et le gradient, la direction de descente calculée peut ne pas en être une et la recherche de pas peut être inefficace. Ces approximations peuvent alors mener à un arrêt prématuré de l'algorithme d'optimisation.

Dans ce chapitre, nous montrons tout d'abord que la troncature peut avoir un effet néfaste sur la convergence de l'algorithme d'inversion. Puis nous proposons une nouvelle formulation du critère et du gradient afin d'assurer la cohérence entre ces deux quantités, même en présence d'effets de la troncature. L'idée est donc d'améliorer la convergence de l'algorithme d'optimisation. Pour cela, nous définissons un critère approché où les champs totaux sont remplacés par une projection de ceux-ci dans un sous-espace, résultant par exemple des directions de descente calculées par BiCGStab. Cette approximation peut être explicitée analytiquement, ce qui permet de définir une formulation exacte du gradient de ce critère approché.

Pour des raisons de simplification, nous supposerons dans ce chapitre qu'une seule source éclaire le volume inconnu et nous négligerons le terme de régularisation. Cela revient à travailler avec le critère suivant :

$$\mathcal{F}^D(\boldsymbol{x}) = \frac{1}{2}\left\|\widetilde{\boldsymbol{e}}^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}^{\text{inc}}\right\|^2. \tag{7.1}$$

Le gradient pour une source vaut :

$$\nabla_x\mathcal{F}^D = -\mathbf{Q}^{\text{t}}\,\text{diag}\left\{\boldsymbol{e}^{\text{tot}}\right\}^{\dagger}\left(\overline{\mathbf{L}_x}\right)^{-1}\mathbf{G}_o^{\dagger}\left(\widetilde{\boldsymbol{e}}^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}^{\text{inc}}\right). \tag{7.2}$$

Ce chapitre se décompose comme suit :

- la première section 7.1 met en évidence sur un exemple concret l'effet de la troncature sur les approximations du critère et du gradient ;
- la section 7.2 présente l'approximation du champ total projeté dans un sous-espace ;
- la dernière section définit le nouveau critère approché et le gradient correspondant à ce critère.

7.1 Mise en évidence de l'effet de la troncature sur la convergence de l'algorithme d'optimisation

Dans cette section, nous souhaitons mettre en évidence l'incohérence qu'il y a entre le critère et le gradient quand les résolutions des systèmes linéaires sont tronquées. L'idée est de montrer que le gradient calculé n'est pas forcément une direction de descente pour le critère calculé. Si le critère et le gradient sont calculés de façon exacte, nous devrions avoir pour tout $\boldsymbol{x}$ et pour tout α positif et suffisamment petit :

$$\mathcal{F}^D(\boldsymbol{x} - \alpha\nabla_x\mathcal{F}^D) < \mathcal{F}^D(\boldsymbol{x}), \tag{7.3}$$

ce qui signifie que la direction opposée au gradient est une direction de descente pour le critère $\mathcal{F}^D$. Nous allons montrer empiriquement que pour une évaluation approchée du critère et du gradient, la relation (7.3) peut ne pas être vérifiée. Pour cela, nous nous plaçons dans le cas où le critère et le gradient exacts sont estimés de façon approchée : les systèmes dans le critère $\mathcal{F}^D(\boldsymbol{x})$ en (7.1) et son gradient défini en (7.2) sont résolus par BiCGStab avec une tolérance, par exemple 10^{-3} (Zhang et al., 2003a). Ce niveau de tolérance n'est pas si grossier mais est suffisant pour mettre en évidence l'incohérence entre le critère et le gradient. Rappelons que dans les chapitres précédents, nous avons fait le choix de fixer la tolérance à 10^{-6}.

Prenons un montage d'acquisition quelconque et un objet de grande taille électrique et de contraste relativement fort. On peut prendre par exemple une fonction de contraste $\boldsymbol{x}$ aléatoire avec un objet dont la taille est équivalente à quelques longueurs d'onde. La troncature des résolutions des systèmes avec une tolérance de 10^{-3} entraîne une approximation relativement grossière des solutions. Par exemple, le calcul tronqué du champ total $\boldsymbol{e}^{\text{tot}}$ donne une approximation $\hat{\boldsymbol{e}}^{\text{tot}}$. L'évaluation approchée du critère en $\boldsymbol{x}$ est notée f_x :

$$f_x = \frac{1}{2}\left\|\tilde{\boldsymbol{e}}^{\text{scat}} - \mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}^{\text{tot}}\right\|^2. \tag{7.4}$$

De même, l'estimation approchée du gradient $\nabla_x\mathcal{F}^D$ est notée $\boldsymbol{g}_x$. Évidemment, plus la tolérance est stricte, plus f_x tend vers le critère exact $\mathcal{F}^D(\boldsymbol{x})$ et plus $\boldsymbol{g}_x$ tend vers le gradient exact $\nabla_x\mathcal{F}^D(\boldsymbol{x})$.

Nous cherchons maintenant à savoir si le gradient approché $\boldsymbol{g}_x$ évalué en $\boldsymbol{x}$ est une direction de descente pour le critère approché. Pour cela, nous calculons le critère en un point proche de $\boldsymbol{x}$ dans la direction opposée au gradient, soit $\mathcal{F}^D(\boldsymbol{x}-\alpha\boldsymbol{g}_x)$, avec α positif et suffisamment petit. Cette évaluation du critère en $\boldsymbol{x}-\alpha\boldsymbol{g}_x$ est également réalisée de façon approchée avec la même tolérance que pour le calcul des autres quantités, à savoir 10^{-3} dans notre exemple. Nous notons cette évaluation tronquée $f_{x-\alpha g}$. Si $f_{x-\alpha g} < f_x$, alors le gradient approché $\boldsymbol{g}_x$ permet de proposer de réduire la fonction coût lors d'une itération de l'algorithme d'optimisation. Au contraire, si $f_{x-\alpha g} > f_x$, alors on est dans l'incapacité d'assurer une réduction du critère.

En pratique, pour mettre en évidence cette incohérence, un grand nombre d'évaluations de critère et gradient a été réalisé pour de nombreux contrastes aléatoires $\boldsymbol{x}$. Pour chaque $\boldsymbol{x}$, le critère f_x, le gradient $\boldsymbol{g}_x$ et le critère en $\boldsymbol{x} - \alpha\boldsymbol{g}_x$, $f_{x-\alpha g}$, sont évalués. Parmi toutes ces évaluations, le cas où $f_{x-\alpha g} > f_x$ a été rencontré plusieurs fois. Cela signifie que lorsqu'on suit la direction opposée au gradient, l'évaluation du critère est plus élevée qu'en $\boldsymbol{x}$ alors qu'on devrait observer l'inverse. Cela met en évidence le fait que si la tolérance sur les systèmes est grossière, le calcul approché du critère et du gradient n'assure pas que l'on puisse réduire le critère lors d'une itération de l'algorithme d'inversion. Dans le cas où des contrastes $\boldsymbol{x}$ sont choisis encore plus élevés, le cas où $f_{x-\alpha g} > f_x$ apparaît d'autant plus. De même, si la tolérance est encore plus grossière (par exemple 10^{-1}), le cas où $f_{x-\alpha g} > f_x$ apparaît plus régulièrement.

Maintenant que nous avons vu qu'il est possible de mettre en défaut la relation (7.3) pour une troncature grossière, on peut vérifier que pour un même contraste $\boldsymbol{x}$, la relation est vérifiée si la troncature est plus stricte. Prenons un contraste $\boldsymbol{x}$ pour lequel $f_{x-\alpha g} > f_x$, on peut recalculer f_x, $\boldsymbol{g}_x$ et $f_{x-\alpha g}$ pour une tolérance plus stricte, par exemple 10^{-6}. Cette fois-ci,

le résultat obtenu est $f_{x-\alpha g} < f_x$. Ce résultat semble naturel : les résolutions ne sont pas tronquées grossièrement et les solutions obtenues sont alors proches des solutions exactes des systèmes linéaires. Les évaluations approchées du critère et du gradient sont donc plus proches du critère et gradient exacts. Pour visualiser ces résultats, nous avons tracé la variation du critère $f_{x-\alpha g} - f_x$ en fonction de α pour différentes valeurs de tolérance : 10^{-3}, 10^{-6} et 10^{-15}. La figure 7.1 représente ces variations. On note tout d'abord que le pas α, compris entre 0 et 10^{-9}, est tellement faible que la décroissance du critère exact est infime. On observe cela avec la courbe de tolérance de 10^{-15} : la tolérance est très stricte, les systèmes sont alors résolus de façon quasiment exacte et le critère évalué est quasiment égale au critère exacte. De même pour la courbe de tolérance égale à 10^{-6}, les systèmes sont résolus avec des approximations relativement faibles. En revanche, pour des résolutions de systèmes tronquées à 10^{-3}, on observe que pour certaines valeurs de α, le critère évalué en $\boldsymbol{x} - \alpha \boldsymbol{g}_x$ peut être supérieur à f_x. Cette figure met bien en évidence l'effet nuisible des approximations dans le calcul du critère et du gradient.

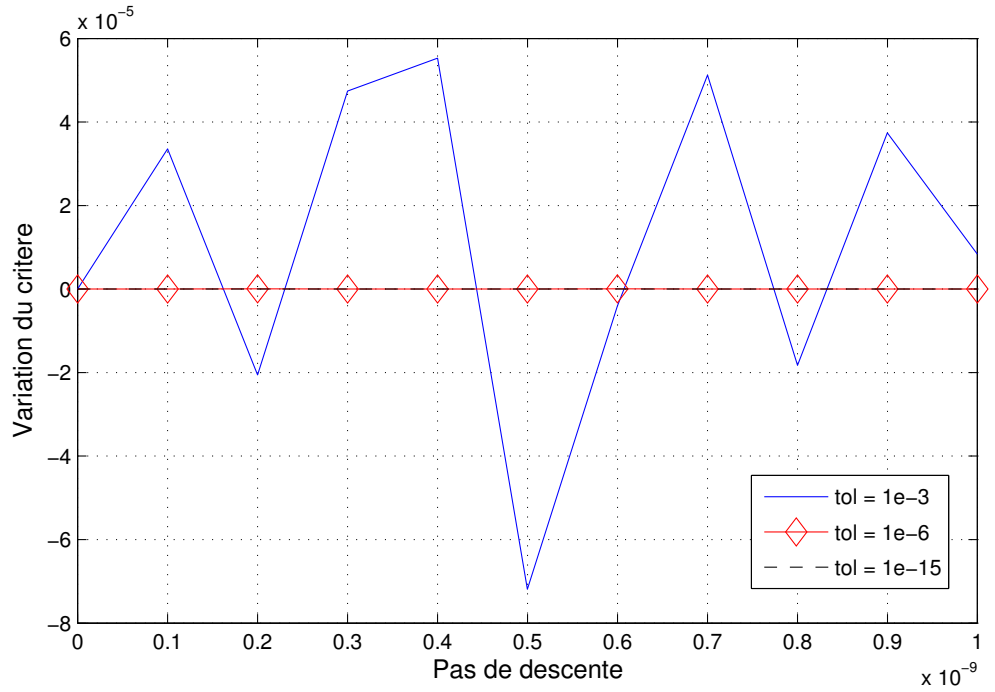


Figure 7.1 Variation du critère approché dans la direction du gradient approché. Trois valeurs de tolérance sur la résolution des systèmes sont données : 10^{-3}, 10^{-6} et 10^{-15}.

Dans la littérature, les auteurs utilisant des algorithmes itératifs pour résoudre les systèmes linéaires ne mentionnent que rarement les tolérances utilisées. Zhang et al. (2003a) choisit par exemple une tolérance de 10^{-3}. Il ne serait pas étonnant que certains appliquent des

tolérances grossières afin de réduire le coût de calcul du critère et du gradient. Comme nous venons de le voir, cette troncature a un effet néfaste sur la convergence de l'algorithme de reconstruction puisque la cohérence entre le critère et le gradient calculés de façon approchée n'est plus assurée. On peut alors observer un arrêt prématuré de l'algorithme d'optimisation et obtenir une solution médiocre. C'est pour cette raison que nous proposons dans ce chapitre un moyen d'assurer la cohérence entre le critère et le gradient quelle que soit la tolérance utilisée.

7.2 Expression du champ total dans un sous-espace

Maintenant que nous avons mis en évidence l'effet de la troncature sur la convergence de la reconstruction, nous allons proposer une nouvelle formulation du critère et du gradient afin d'assurer la cohérence. Avant tout, dans cette section, nous donnons une formulation algébrique de l'approximation du champ total dans un sous-espace constitué des directions de descente calculées par BiCGStab. Cette nouvelle définition du champ total est à l'origine de la nouvelle formulation du critère et du gradient.

7.2.1 Expression du champ

Supposons que l'on recherche l'approximation du champ total dans un sous-espace donné. Nous verrons en 7.2.2 que BiCGStab fournit la construction d'un tel sous-espace. Le champ total $\boldsymbol{e}^{\text{tot}} = \mathbf{L}_x^{-1}\boldsymbol{e}^{\text{inc}}$ est approché par une projection dans ce sous-espace. Notons $\hat{\boldsymbol{e}}^{\text{tot}}$ cette projection. On suppose que $\hat{\boldsymbol{e}}^{\text{tot}}$ est la projection dans un sous-espace de dimension m. On peut alors écrire $\hat{\boldsymbol{e}}^{\text{tot}} = \mathbf{B}\widehat{\boldsymbol{u}}$ où $\mathbf{B}$ est une matrice de taille $3N \times m$ contenant les m vecteurs engendrant le sous-espace et $\widehat{\boldsymbol{u}}$ contient les coefficients de la projection. Le vecteur $\widehat{\boldsymbol{u}}$ minimise l'erreur dans le sous-espace $\mathbf{B}$:

$$\widehat{\boldsymbol{u}} = \arg\min_{\boldsymbol{u}} \|\boldsymbol{e}^{\text{inc}} - \mathbf{L}_x\mathbf{B}\boldsymbol{u}\|^2. \tag{7.5}$$

La solution exacte de ce problème est

$$\begin{aligned}\widehat{\boldsymbol{u}} &= \left(\mathbf{B}^\dagger\mathbf{L}_x^\dagger\mathbf{L}_x\mathbf{B}\right)^{-1}\mathbf{B}^\dagger\mathbf{L}_x^\dagger\boldsymbol{e}^{\text{inc}} \\ &= \left(\mathbf{C}_x^\dagger\mathbf{C}_x\right)^{-1}\mathbf{C}_x^\dagger\boldsymbol{e}^{\text{inc}},\end{aligned} \tag{7.6}$$

avec $\mathbf{C}_x = \mathbf{L}_x\mathbf{B}$, matrice de taille $3N \times m$. En remplaçant $\widehat{\boldsymbol{u}}$ dans l'expression du champ total, on obtient

$$\hat{\boldsymbol{e}}^{\text{tot}} = \mathbf{B}\left(\mathbf{C}_x^\dagger\mathbf{C}_x\right)^{-1}\mathbf{C}_x^\dagger\boldsymbol{e}^{\text{inc}}. \tag{7.7}$$

On remarque qu'avec cette formulation, il n'est plus nécessaire d'inverser la matrice $\mathbf{L}_x$ de taille $3N \times 3N$ mais simplement la matrice $\mathbf{C}_x^\dagger \mathbf{C}_x$ de taille $m \times m$. Évidemment, cela ne prend pas en compte le coût de calcul pour déterminer les vecteurs de $\mathbf{B}$ engendrant le sous-espace. De plus, le calcul des vecteurs de $\mathbf{C}_x$ requiert m produits matrice-vecteur avec $\mathbf{L}_x$ et $\mathbf{B}$.

7.2.2 Choix d'un sous-espace

Nous proposons maintenant de choisir un sous-espace pertinent afin que la projection $\hat{\boldsymbol{e}}^{\text{tot}}$ du champ total dans ce sous-espace soit le plus proche possible du champ total exact $\boldsymbol{e}^{\text{tot}} = \mathbf{L}_x^{-1}\boldsymbol{e}^{\text{inc}}$.

Il est possible par exemple de prendre m vecteurs de la base canonique pour créer $\mathbf{B}$. Ce choix n'est évidemment pas judicieux mais il a l'avantage de ne nécessiter aucun calcul pour calculer la famille de vecteurs $\mathbf{B}$.

Un choix naturel consiste à prendre pour vecteurs engendrant le sous-espace les directions de descente issues de BiCGStab (ou Partial-Block BiCGStab). L'algorithme est utilisé car il converge rapidement, ce qui siginifie qu'il nécessite peu d'itérations pour réduire significativement le résidu. En prenant les directions de descente calculées par cet algorithme, nous devrions avoir un sous-espace convenable pour un faible nombre m de vecteurs. En reprenant l'Algorithme 1 (page 20), les directions de descente à chaque itération sont les vecteurs $\boldsymbol{p}$ et $\boldsymbol{s}$. Pour d itérations de BiCGStab, on peut construire la matrice $\mathbf{B}$ sous la forme :

$$\mathbf{B} = \left[\boldsymbol{e}^{\text{tot}(0)}, \boldsymbol{p}_1, \boldsymbol{s}_1, \ldots, \boldsymbol{p}_d, \boldsymbol{s}_d\right]. \tag{7.8}$$

Ici, la taille du sous-espace vaut $m = 2d + 1$. Le calcul de la matrice $\mathbf{C}_x$ requiert les m produits des vecteurs de $\mathbf{B}$ par $\mathbf{L}_x$. Or, dans BiCGStab, les produits $\mathbf{L}_x\boldsymbol{p}$ et $\mathbf{L}_x\boldsymbol{s}$ sont réalisés à chaque itération. Il suffit donc de conserver ces vecteurs lors de la résolution par BiCGStab.

Une fois les matrices $\mathbf{B}$ et $\mathbf{C}_x$ créées, il est nécessaire de résoudre l'équation (7.7) afin de calculer la projection du champ total. Notons que l'algorithme BiCGStab est une méthode utilisant les sous-espaces de Krylov, c'est-à-dire les sous-espaces engendrés par les vecteurs $\boldsymbol{e}^{\text{inc}}$, $\mathbf{L}_x\boldsymbol{e}^{\text{inc}}$, $\mathbf{L}_x^2\boldsymbol{e}^{\text{inc}}$, ... Tout comme la méthode GMRES (Saad et Schultz, 1986), notre approche basée sur la projection du champ total permet de minimiser le résidu dans le sous-espace engendré par les vecteurs de $\mathbf{B}$. Notre expression du champ total (7.7) est donc plus proche du champ exact que la solution par BiCGStab.

7.3 Définition du nouveau critère et calcul du gradient

7.3.1 Critère approché

Maintenant que la projection du champ total dans un sous-espace a été formalisée, nous définissons le nouveau critère approché $\widehat{\mathcal{F}}^D$. On rappelle qu'on se limite à une seule source et sans régularisation. Le critère approché prend la forme :

$$\widehat{\mathcal{F}}^D(\boldsymbol{x}) = \frac{1}{2}\left\|\boldsymbol{y} - \mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}^{\text{tot}}\right\|^2 \tag{7.9}$$

$$= \frac{1}{2}\left\|\boldsymbol{y} - \mathbf{G}_o\mathbf{X}\mathbf{B}\left(\mathbf{B}^\dagger\mathbf{L}_x^\dagger\mathbf{L}_x\mathbf{B}\right)^{-1}\mathbf{B}^\dagger\mathbf{L}_x^\dagger\boldsymbol{e}^{\text{inc}}\right\|^2 \tag{7.10}$$

$$= \frac{1}{2}\left\|\boldsymbol{y} - \mathbf{G}_o\mathbf{X}\mathbf{B}\left(\mathbf{C}_x^\dagger\mathbf{C}_x\right)^{-1}\mathbf{C}_x^\dagger\boldsymbol{e}^{\text{inc}}\right\|^2. \tag{7.11}$$

Le cas où le champ total est calculé exactement peut s'écrire en considérant la matrice $\mathbf{B}$ égale à l'identité de taille $3N \times 3N$ et le champ total approché $\hat{\boldsymbol{e}}^{\text{tot}} = \boldsymbol{e}^{\text{tot}} = \widehat{\boldsymbol{u}}$. Cela revient à dire que la solution exacte $\boldsymbol{e}^{\text{tot}}$ appartient au sous-espace. On retombe alors sur le critère exact : $\widehat{\mathcal{F}}^D(\boldsymbol{x}) = \mathcal{F}^D(\boldsymbol{x})$.

Le coût de calcul pour évaluer ce nouveau critère peut être décomposé en plusieurs étapes :

- la première étape est de définir la famille de vecteurs formant $\mathbf{B}$ et engendrant le sous-espace. Si $\mathbf{B}$ est composée des directions de descente de BiCGStab, il est nécessaire d'appliquer l'algorithme à la résolution du système $\mathbf{L}_x\boldsymbol{e}^{\text{tot}} = \boldsymbol{e}^{\text{inc}}$;
- l'étape suivante est le calcul des vecteurs $\mathbf{C}_x = \mathbf{L}_x\mathbf{B}$;
- enfin les dernières étapes correspondent au calcul de $(\mathbf{C}_x^\dagger\mathbf{C}_x)^{-1}$ et des produits matriciels intervenant dans (7.11).

Notons que le coût pour évaluer $\widehat{\mathcal{F}}^D(\boldsymbol{x})$ peut être légèrement plus élevé que celui pour évaluer $\mathcal{F}^D(\boldsymbol{x})$ car l'inversion supplémentaire de $\mathbf{C}_x^\dagger\mathbf{C}_x$ est requise. Cette matrice est tout de même de petite taille. L'intérêt ne réside pas dans l'accélération du coût de calcul mais dans la cohérence du critère et du gradient. Le gain est attendu du point de vue de la convergence qui peut être améliorée.

7.3.2 Gradient correspondant au critère approché

Pour calculer le gradient, on fait l'hypothèse que les vecteurs de $\mathbf{B}$ ne dépendent pas de la valeur du contraste $\boldsymbol{x}$. Seules les quantités $\mathbf{X}$ et $\mathbf{L}_x$ (et donc $\mathbf{C}_x$) dépendent du contraste. L'annexe C (p. 159) développe les opérations afin d'obtenir le gradient correspondant au

critère approché $\widehat{\mathcal{F}}^D$. On obtient :

$$\begin{aligned}\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}}^D = &-\mathbf{Q}^{\mathrm{t}}\left[\mathbf{G}_o\left(\mathbf{I}+\mathbf{X}\mathbf{B}(\mathbf{C}_x^\dagger\mathbf{C}_x)^{-1}\mathbf{C}_x^\dagger\mathbf{G}_c\right)\mathrm{diag}\{\hat{\boldsymbol{e}}^{\mathrm{tot}}\}\right]^\dagger\left(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}^{\mathrm{tot}}\right)\\ &-\mathbf{Q}^{\mathrm{t}}\,\mathrm{diag}\left\{(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}^{\mathrm{tot}})^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}_x^\dagger\mathbf{C}_x)^{-1}\mathbf{B}^\dagger\right\}\mathbf{G}_c^\dagger\left(\mathbf{C}_x(\mathbf{C}_x^\dagger\mathbf{C}_x)^{-1}\mathbf{C}_x^\dagger-\mathbf{I}\right)\boldsymbol{e}^{\mathrm{inc}}.\end{aligned} \tag{7.12}$$

On observe tout d'abord que le calcul du gradient ne nécessite plus la résolution de systèmes linéaires avec la matrice opérateur $\mathbf{L}_x$ mais il requiert l'inversion de la matrice $\mathbf{C}_x^\dagger\mathbf{C}_x$, de taille $m \times m$, qui intervient trois fois dans le gradient. On rappelle que la matrice $\mathbf{C}_x$ provient du produit de $\mathbf{L}_x$ avec les vecteurs engendrant le sous-espace. Ces inversions, qu'elles soient réalisées directement, par factorisation ou par méthode itérative, sont d'autant moins coûteuses que la dimension m du sous-espace est faible. Prenons par exemple un problème où le nombre de voxels N vaut $30^3 = 27\,000$ et où la dimension m du sous-espace vaut 101 (soit 50 itérations de BiCGStab). Ces valeurs sont représentatives d'un problème de TMO 3-D. La résolution d'un système impliquant $\mathbf{L}_x$ est de taille $81\,000 \times 81\,000$, tandis que la résolution des trois systèmes impliquant $\mathbf{C}_x^\dagger\mathbf{C}_x$ est de taille 100×100.

Finalement, deux points sont à noter sur cette nouvelle formulation du gradient :

- il s'agit de la formulation *exacte* du gradient du critère *approché*. Assurer la cohérence entre le critère et son gradient permet d'éviter des difficultés de recherche de pas et d'améliorer la convergence de l'algorithme d'optimisation vers un minimum local ;
- le calcul de ce gradient ne nécessite plus la résolution d'un système de grande taille impliquant $\mathbf{L}_x$ mais la résolution de trois systèmes de plus petite taille impliquant la matrice $\mathbf{C}_x^\dagger\mathbf{C}_x$.

7.3.3 Vérification du gradient pour B inversible

Nous proposons dans cette section de valider le fait que le gradient (exact) (7.12) du critère approché et le gradient (exact) (3.14) du critère exact sont équivalents dans le cas où $\mathbf{B}$ est inversible. Supposons que le champ total exact $\boldsymbol{e}^{\mathrm{tot}}$ appartient au sous-espace engendré par les vecteurs de $\mathbf{B}$. Cela signifie que $\hat{\boldsymbol{e}}^{\mathrm{tot}} = \boldsymbol{e}^{\mathrm{tot}}$. On peut donc prendre $\mathbf{B}$ égale à la matrice identité et $\boldsymbol{u} = \boldsymbol{e}^{\mathrm{tot}}$. Dans la nouvelle formulation du gradient (7.12), le terme $\mathbf{C}_x(\mathbf{C}_x^\dagger\mathbf{C}_x)^{-1}\mathbf{C}_x^\dagger - \mathbf{I}$ se

simplifie :

$$\begin{aligned}\mathbf{C}_x(\mathbf{C}_x^\dagger\mathbf{C}_x)^{-1}\mathbf{C}_x^\dagger - \mathbf{I} &= \mathbf{L}_x\mathbf{B}\left(\mathbf{B}^\dagger\mathbf{L}_x^\dagger\mathbf{L}_x\mathbf{B}\right)^{-1}\mathbf{B}^\dagger\mathbf{L}_x^\dagger - \mathbf{I}\\ &= \mathbf{L}_x\left(\mathbf{L}_x^\dagger\mathbf{L}_x\right)^{-1}\mathbf{L}_x^\dagger - \mathbf{I}\\ &= \mathbf{I} - \mathbf{I} = \mathbf{0}.\end{aligned}$$

Le second terme du gradient dans (7.12) s'annule. De même, on simplifie le premier terme du gradient :

$$\begin{aligned}\mathbf{I} + \mathbf{X}\mathbf{B}(\mathbf{C}_x^\dagger\mathbf{C}_x)^{-1}\mathbf{C}_x^\dagger\mathbf{G}_c &= \mathbf{I} + \mathbf{X}\mathbf{L}_x^{-1}\mathbf{G}_c\\ &= (\mathbf{I} - \mathbf{X}\mathbf{G}_c)^{-1}\\ &= (\mathbf{L}_x^{\mathrm{t}})^{-1}.\end{aligned}$$

Finalement, on simplifie la formulation (7.12) dans le cas où le champ total exact $\boldsymbol{e}^{\mathrm{tot}}$ appartient au sous-espace :

$$\begin{aligned}\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}} &= -\mathbf{Q}^{\mathrm{t}}\left[\mathbf{G}_o\left(\mathbf{L}_x^{\mathrm{t}}\right)^{-1}\mathrm{diag}\{\hat{\boldsymbol{e}}^{\mathrm{tot}}\}\right]^\dagger\left(\boldsymbol{y} - \mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}^{\mathrm{tot}}\right)\\ \nabla_{\boldsymbol{x}}\widehat{\mathcal{F}} &= \nabla_{\boldsymbol{x}}\mathcal{F}\end{aligned}$$

On retrouve donc bien le gradient exact du critère exact.

7.3.4 Étude de la variation du critère selon la direction du gradient

Tout comme la figure 7.1, nous cherchons à tracer le critère selon la direction opposée au gradient afin de vérifier que la nouvelle formulation du critère et du gradient assure la cohérence entre ces deux quantités.

Pour cela, nous calculons $\widehat{\mathcal{F}}(\boldsymbol{x})$ où les vecteurs de $\mathbf{B}$ proviennent des directions de descente de BiCGStab. La tolérance de ces résolutions est fixée à 10^{-3}, soit la même que dans la figure 7.1 où l'incohérence a été mis en valeur. Le gradient $\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}}$ est calculé et on évalue le critère $\widehat{\mathcal{F}}$ en $\boldsymbol{x} - \alpha\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}}$, où le même sous-espace est utilisé, c'est-à-dire la même matrice $\mathbf{B}$ est utilisée. La figure 7.2 présente la variation du critère $\widehat{\mathcal{F}}(\boldsymbol{x}-\alpha\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}}) - \widehat{\mathcal{F}}(\boldsymbol{x})$ en fonction de α. On remarque en premier lieu que, malgré la troncature à 10^{-3}, le fait d'assurer la cohérence entre le critère et le gradient permet d'éviter que les approximations dans le critère et le gradient mènent aux variations de la courbe bleue (critère et gradient incohérents). La figure 7.3 permet de zoomer sur l'échelle des ordonnées afin de mieux visualiser les variations du critère pour la

cohérence du gradient. On voit bien que le critère $\widehat{\mathcal{F}}(\boldsymbol{x} - \alpha\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}})$ décroît en fonction de α et est toujours inférieur à $\widehat{\mathcal{F}}(\boldsymbol{x})$. Cette figure confirme que la cohérence entre le critère et le gradient permet d'assurer que le gradient est une direction de descente, malgré la troncature grossière des résolutions de systèmes.

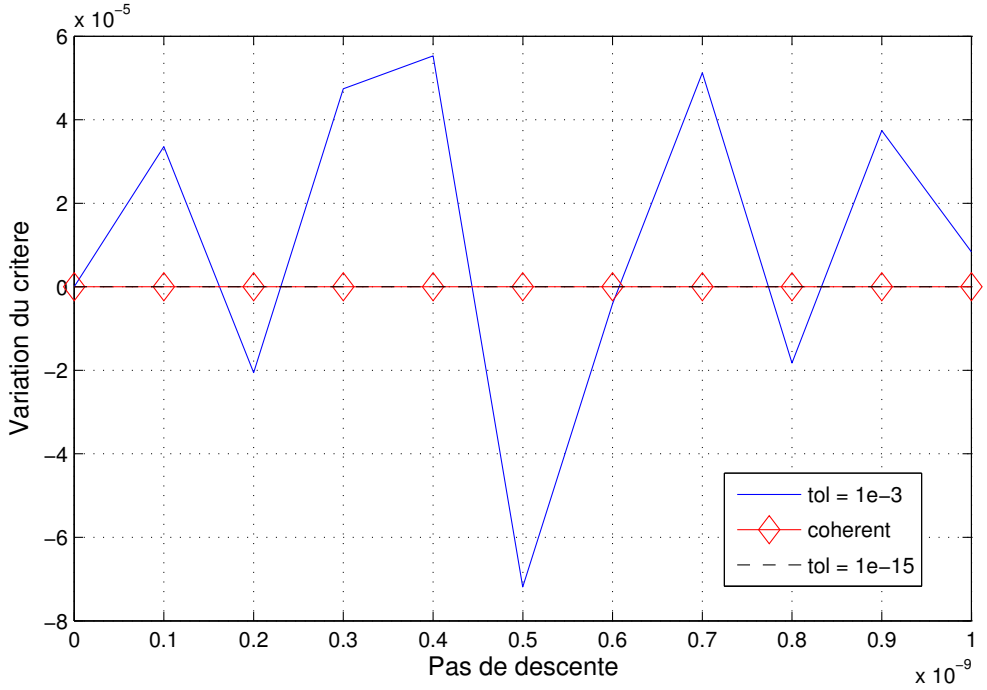


Figure 7.2 Variation du critère approché dans la direction du gradient exact du critère approché. Deux valeurs de tolérance sur la résolution des systèmes sont testées : 10^{-3} et 10^{-15}. La cohérence est assurée avec un sous-espace obtenu à partir des directions de descente de BiCGStab pour une tolérance de 10^{-3}.

7.4 Conclusion

Nous avons mis en évidence dans ce chapitre le fait que la troncature des systèmes linéaires à résoudre entraîne une incohérence entre le critère et le gradient. Nous avons vu sur un exemple simple que cette incohérence peut engendrer une difficulté de convergence à travers le fait que le gradient approché ne correspond pas à une direction de descente pour le critère. La résolution du problème inverse peut alors stagner et la recherche de pas dans l'algorithme d'optimisation peut être rendue impossible à cause de cette incohérence.

Nous avons donc proposé une nouvelle formulation des calculs d'une itération de l'algorithme d'optimisation en définissant un nouveau critère et son gradient exact. Cette formulation

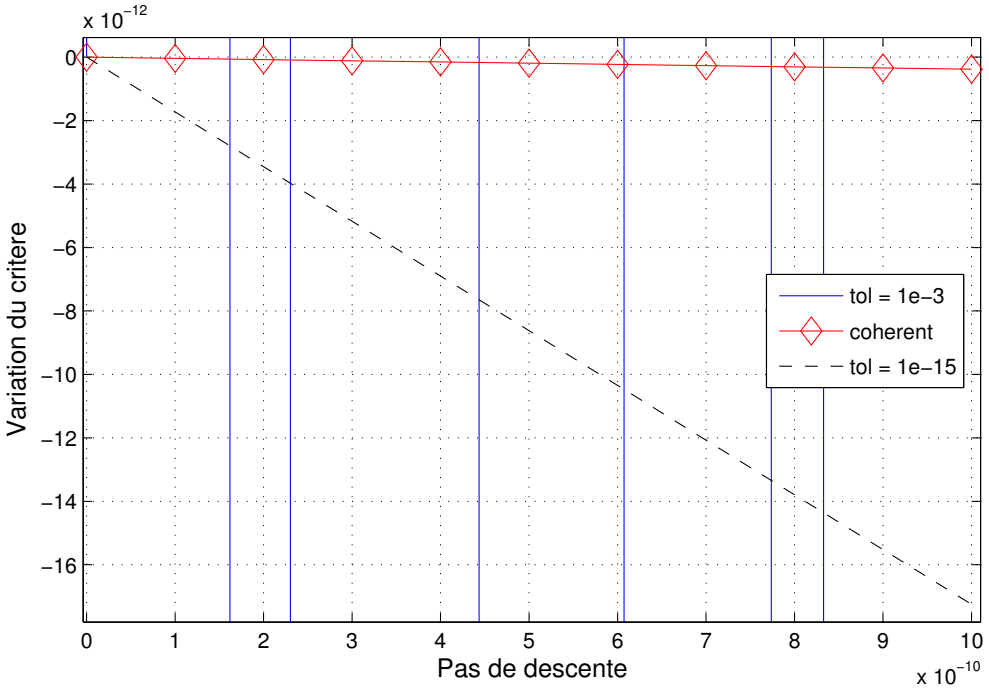


Figure 7.3 Zoom de la figure 7.2.

s'appuie sur la projection du champ total dans un sous-espace connu, en particulier pouvant résulter d'un algorithme itératif comme BiCGStab. La cohérence entre le critère et le gradient est assurée, ce qui permet d'assurer la décroissance de la fonction coût à chaque itération de l'algorithme de reconstruction.

Nous n'avons pas eu le temps de tester cette approche sur des exemples concrets et de comparer le comportement de l'inversion avec une approche classique par troncature de l'algorithme BiCGStab. Nous avons vu à travers l'analyse des coûts de calcul que cette approche peut être plus coûteuse que l'approche standard par résolutions tronquées. En revanche, il est probable que l'assurance de la cohérence entre le critère et le gradient permettra de réduire le nombre d'itérations global avant convergence.

Quelques tests préliminaires ont été réalisés et ils ont mis en avant un problème de conditionnement dans la famille engendrant le sous-espace. En effet, les directions de descente issues de BiCGStab peuvent être fortement corrélées et cela entraîne des erreurs numériques importantes dans l'inversion du système impliquant $\mathbf{C}_x$. Il peut être donc judicieux de chercher une nouvelle famille engendrant le même sous-espace mais dont les vecteurs sont décorrélées. Une solution naturelle consiste à appliquer une orthogonalisation QR sur la matrice $\mathbf{B}$ afin de déterminer une famille orthogonale. Cela entraîne évidemment un coût supplémentaire et

est à prendre en compte dans la perspective d'étude du coût de calcul.

Finalement, l'intérêt d'une telle formulation est de pouvoir tronquer fortement les résolutions de systèmes tout en assurant la cohérence entre le critère approché et le gradient. À cause des problèmes de conditionnement de $\mathbf{B}$, il n'est pas imaginable de fixer une tolérance stricte (soit un sous-espace de grande taille), ce qui entraînerait un coût de calcul important pour orthogonaliser cette famille de vecteurs. La perspective est donc d'utiliser cette formulation avec une troncature grossière afin de converger rapidement vers une première solution grossière. On peut par exemple réaliser les premières itérations d'une reconstruction ou les premières étapes d'une procédure par saut de fréquence avec cette formulation afin d'obtenir une estimation rapide de la solution. Ensuite, lorsque l'algorithme est proche d'un minimum, il est nécessaire de résoudre les systèmes de façon la plus exacte possible afin d'évaluer au mieux le critère et le gradient. De cette façon, la convergence de l'algorithme d'optimisation est assurée.

CHAPITRE 8 CONCLUSION ET PERSPECTIVES

8.1 Synthèse des travaux

Dans le cadre de ces travaux, nous nous sommes intéressés à l'accélération des méthodes de reconstruction en tomographie micro-ondes. Nous avons fait le choix dès le départ de considérer le problème inverse comme un problème d'optimisation où l'on cherche à minimiser les erreurs entre les mesures du champ diffracté et un modèle direct non-linéaire. Un modèle numérique basé sur la formulation intégrale de l'équation d'onde a été choisi et la méthode des moments a permis d'obtenir une forme algébrique des équations du modèle direct. Le problème inverse a, quant à lui, été pris comme la minimisation d'un critère non-linéaire régularisé. L'algorithme L-BFGS a été choisi afin de minimiser efficacement la fonction coût. Dans L-BFGS, seul le gradient est requis pour calculer une direction de descente, contrairement aux méthodes courantes de Gauss-Newton, Levenberg-Marquardt et DBIM qui nécessitent l'inversion d'un système de grande taille pour calculer la direction de descente. Malgré tout, nous avons vu que l'évaluation du critère et du gradient nécessaire à chaque itération de la reconstruction mène à la résolution de nombreux systèmes linéaires de grande taille. Les contributions de cette thèse concernent essentiellement ces multiples résolutions de systèmes afin de réduire le coût de calcul des algorithmes d'inversion.

8.1.1 Réduction du nombre de systèmes linéaires

Dans un premier temps, nous avons proposé une procédure permettant de réduire le nombre de systèmes linéaires à résoudre pour évaluer le critère et le gradient. Pour cela, nous nous sommes appuyés sur le théorème de réciprocité de Lorentz qui considère une configuration adjointe du montage d'acquisition, permettant notamment d'inverser le rôle des sources et des récepteurs. Comme le calcul du critère requiert la résolution d'un nombre de systèmes égal au nombre de sources, l'application de ce principe conduit à la réduction du nombre de systèmes pour l'évaluation du critère, et donc pour la recherche de pas dans l'inversion.

Il est également possible de réduire le nombre de systèmes à résoudre pour le calcul du gradient en s'appuyant sur différentes mises en œuvre de ce calcul. En fonction du montage d'acquisition, il est possible de résoudre un nombre de systèmes soit égal au nombre de sources, soit égal au nombre de récepteurs. Certains auteurs appellent ces mises en œuvre, la formulation de l'état adjoint.

Finalement, la combinaison de ces deux techniques permet de réduire le nombre de systèmes

pour le calcul du critère et du gradient en fonction de la configuration du montage. Aucune analyse n'a été effectuée dans la littérature et les auteurs utilisent souvent une approche qui n'est pas forcément la plus efficace. C'est pour cette raison que nous avons proposé une telle procédure. L'application sur la configuration d'acquisition des données 3-D de l'Institut Fresnel permet de réduire considérablement le nombre de systèmes à résoudre pour le critère et le gradient.

8.1.2 Accélération des résolutions de systèmes

Dans un second temps, nous avons proposé d'adapter efficacement un algorithme par blocs afin d'accélérer la résolution de ces systèmes. L'algorithme Block-BiCGStab permet en effet de résoudre conjointement plusieurs systèmes linéaires avec une matrice opérateur commune. Nous avons proposé un réglage afin de rendre efficace l'application de cet algorithme pour la résolution des problèmes directs en TMO. L'effet de certains paramètres du problème sur la vitesse de convergence a été étudié : pour des contrastes élevés, des fréquences élevées ou un grand nombre de systèmes, l'approche Block-BiCGStab converge bien plus rapidement que BiCGStab. On peut obtenir par exemple un facteur d'accélération de trois sur certains exemples.

Nous avons généralisé l'approche Block-BiCGStab, nommée Partial-Block BiCGStab, pour une résolution efficace sur des ordinateurs multi-cœurs. L'idée est de diviser l'ensemble des systèmes à résoudre en plusieurs groupes et d'appliquer l'algorithme Block-BiCGStab sur ces différents groupes. Ces résolutions sont finalement envoyées indépendamment sur les cœurs disponibles.

Finalement, l'utilisation de l'algorithme Partial-Block BiCGStab pour la résolution des systèmes dans l'inversion permet de réduire le coût de calcul de la reconstruction, comparée à une utilisation de l'approche BiCGStab. Ces méthodes sont utilisables dans n'importe quel algorithme d'inversion, dès lors que l'inversion requiert la résolution de plusieurs systèmes linéaires impliquant la même matrice opérateur. D'autres applications peuvent être également intéressées par cette approche ; nous en discutons dans les perspectives.

8.1.3 Validation de l'algorithme de reconstruction sur des données expérimentales

Bien que le cœur de ces travaux concerne la réduction du coût de calcul des méthodes de reconstruction, notre procédure d'optimisation est nouvelle et nous avons souhaité montrer ses performances en termes de qualité de reconstruction sur la reconstruction d'objets réels.

La base de données de l'Institut Fresnel a été utilisée afin de tester notre méthode et de comparer les objets reconstruits aux résultats publiés la littérature.

Notre algorithme d'optimisation est basé simplement sur un terme d'adéquation aux données non-linéaire et un terme de régularisation additive préservant les discontinuités. Cette fonction coût est minimisée à l'aide de l'algorithme L-BFGS, méthode de quasi-Newton. Nos résultats sont très concluants et permettent même de reconstruire un objet volumineux et fortement contrasté (objet Cylinder de la base de l'Institut Fresnel) qui a causé des difficultés de reconstruction pour la plupart des chercheurs.

La reconstruction de ces objets a également permis de comparer les performances de l'algorithme Block-BiCGStab et celles de BiCGStab pour la résolution des systèmes linéaires dans le critère et le gradient. Le gain relatif de l'utilisation de l'approche Block-BiCGStab varie de 10 à 50% selon les objets. En particulier, la reconstruction de l'objet Cylinder avec l'utilisation de Block-BiCGStab est deux fois plus rapide que cette même reconstruction avec BiCGStab.

8.1.4 Cohérence entre le critère et le gradient

Enfin, nous avons montré dans le chapitre 7 que la troncature grossière des résolutions des systèmes linéaires introduit des approximations importantes dans l'évaluation du critère et du gradient. Ces approximations entraînent l'incohérence entre le critère et le gradient calculés : le gradient approché ne correspond pas à une direction de descente pour le critère. Cela peut conduire à des difficultés pour la détermination d'une direction de descente et pour la recherche de pas dans l'algorithme d'inversion.

Nous avons donc proposé une nouvelle formulation du critère et du gradient permettant d'assurer la cohérence entre ces deux quantités, quel que soit le degré de troncature des systèmes. Cela permet de garantir que la direction de descente en est réellement une. La convergence de l'algorithme est donc améliorée.

Ce travail sur la cohérence ne représente pour l'instant qu'une étude méthodologique. Des perspectives d'étude de l'amélioration de la convergence et des comparaisons aux autres approches sont données en 8.3.1.

8.2 Limitations des solutions proposées

Quelques limites des solutions proposées sont à mentionner. Tout d'abord, la procédure présentée dans le chapitre 4 et permettant de réduire le nombre de systèmes linéaires pour

calculer le critère et le gradient s'applique uniquement à des méthodes de reconstruction basées sur le gradient. Les approches fréquentes de Gauss-Newton, Levenberg-Marquardt et DBIM requièrent une expression du jacobien. Cela signifie qu'il n'y a plus le choix de la mise en œuvre et qu'il est obligatoire de calculer les champs totaux adjoints selon la formulation de l'état adjoint (**??**). En revanche, l'utilisation du principe de réciprocité détaillé en 4.3 permet tout de même de réduire le coût de calcul du critère et donc de la recherche de pas.

Concernant l'algorithme Block-BiCGStab et Partial-Block BiCGStab, le gain relatif par rapport à BiCGStab est réduit dans certains cas, comme pour un faible contraste ou une petite taille électrique de l'objet diffractant. Il peut même être moins efficace que la méthode classique BiCGStab. Il est donc nécessaire d'établir un moyen de connaître la méthode la plus efficace sur un problème donné.

Dans le même contexte, quand le nombre de cœurs disponibles est important, la méthode Partial-Block BiCGStab se rapproche d'une version massivement parallèle de BiCGStab et le gain relatif entre les deux méthodes est également faible. En revanche, si la puissance de calcul continue à croître, il sera possible d'acquérir des données avec un nombre toujours plus grand d'illuminations. Ce type de configuration permettra donc aux approches par blocs de conserver leur intérêt.

La nouvelle formulation du critère et du gradient cohérent semble prometteuse du fait de l'amélioration de la convergence. En revanche, nous avons vu que le coût de calcul pour évaluer ce nouveau critère peut être très élevé. Il est donc nécessaire de s'assurer que le coût supplémentaire pour calculer le critère est compensé par la convergence plus rapide de l'inversion.

8.3 Améliorations futures et pistes de recherche

Nous étudions enfin quelques perspectives possibles, soit dans la continuité de ces travaux, soit dans l'objectif d'élargir les solutions proposées à un cadre plus vaste que celui posé dans cette thèse.

8.3.1 Étude de la formulation cohérente du gradient

Une première étape est de poursuivre le travail au sujet de la cohérence du critère et du gradient en réalisant des tests concrets de la nouvelle formulation du problème. Le travail à réaliser peut être décomposé comme suit :

— implémenter la résolution du problème inverse avec cette nouvelle formulation et étudier la convergence, notamment avec une troncature très grossière. Valider d'abord

sur des problèmes simples et de petite taille ;
— optimiser les opérations nécessaires au calcul du critère et du gradient afin de réduire les coûts de calcul (par exemple : orthogonalisation, inversion matricielle) ;
— comparer les performances en coût de calcul et en amélioration de la convergence de cette formulation et celles de la formulation “classique” où les résolutions des systèmes sont réalisées par BiCGStab et tronquées.

8.3.2 Études complémentaires de l’approche Partial-Block BiCGStab

L’algorithme Partial-Block BiCGStab a été étudié en fonction de plusieurs paramètres de l’objet et du montage. Pour l’instant, seuls deux montages ont été étudiés pour l’application de l’approche Partial-Block BiCGStab : un montage simulé et un montage réel. Bien que ces deux configurations soient différentes, les performances de l’approche par blocs ne peuvent pas être généralisées à tout type de configuration. Il serait nécessaire d’approfondir cette étude en caractérisant les performances de cette méthode en fonction d’autres paramètres, par exemple d’autres géométries de montage d’acquisition (plus ou moins de sources, plus ou moins éloignées) ou d’autres objets (formes complexes, objets fortement inhomogènes, etc.). Finalement, l’idée serait de pouvoir définir une procédure permettant de choisir s’il est plus avantageux d’appliquer une résolution par blocs ou une résolution séquentielle, selon les différents paramètres de l’objet, du montage ou même de l’ordinateur à disposition.

8.3.3 Approches de parallélisation de l’algorithme Block-BiCGStab

Comme nous l’avons vu lors de la reconstruction des objets de la base Fresnel, le logiciel MATLAB parallélise implicitement les opérations algébriques élémentaires (option *multi-threading* du logiciel). Ce schéma de parallélisation est un schéma différent de celui proposé par Partial-Block BiCGStab où la parallélisation agit au niveau de la répartition des nombreuses résolutions de systèmes. Bien que les équipes impliquées dans ce travail n’aient pas d’expertise pointue en termes de parallélisation, nous sommes conscients que les opérations présentes dans la reconstruction peuvent être parallélisées à différents niveaux : calcul du critère et gradient, calcul de la direction de descente de L-BFGS, résolutions des systèmes, calcul des multiples `FFT`, opérations élémentaires, etc. Un travail approfondi concernant la parallélisation de l’approche Block-BiCGStab pourrait alors être réalisé afin d’accélérer d’un point de vue matériel la résolution des systèmes.

8.3.4 Comparer notre algorithme de reconstruction à d'autres algorithmes fréquents en TMO

Bien que notre algorithme ait fourni de bons résultats de reconstruction sur les objets de la base Fresnel, il serait pertinent de le comparer aux algorithmes usuels utilisés en TMO. La difficulté réside dans la notion de *comparaison* : compare-t-on les résultats de reconstruction, compare-t-on les coûts de calcul de l'inversion, le nombre d'itérations avant convergence ? Doit-on comparer deux méthodes basées sur un même critère, une même régularisation, un même algorithme d'optimisation, ou les trois simultanément ? Définir un schéma de comparaison fiable est une véritable difficulté.

De plus, les auteurs ne détaillent jamais complètement leurs algorithmes. Par exemple, de nombreux auteurs utilisent l'approche BiCGStab pour résoudre les systèmes mais aucun ne donne la tolérance choisie pour ces résolutions. De même, pour l'algorithme de Gauss-Newton ou de Levenberg-Marquardt, le calcul de la direction de descente nécessite la résolution d'un système avec l'approximation de la matrice Hessienne. Aucun auteur ne détaille précisément la façon dont est réalisée et tronquée cette résolution. Il est donc impossible de reproduire précisément les méthodes observables dans la littérature. Dans ce manuscrit, nous avons donné toutes les informations nécessaires à la reproduction exacte des procédures de reconstruction.

8.3.5 Élargir le champ des applications

Reconstruction de données médicales

Une perspective intéressante est d'élargir le champ d'applications afin d'étudier la viabilité de l'algorithme d'inversion proposé sur d'autres problèmes. Dans ces travaux, nous nous sommes principalement concentrés à la reconstruction d'objets simulés ou d'objets réels de la base Fresnel. Il serait pertinent de travailler par exemple à la reconstruction de tissus humains qui ont des formes plus complexes et des inhomogénéités plus importantes que les objets utilisées dans cette thèse. Des bases de données existent issues d'essais cliniques, comme par exemple, des mesures sur l'imagerie du sein (Zastrow et al., 2008). Cela permettrait également de voir si les algorithmes Block-BiCGStab et Partial-Block BiCGStab sont aussi efficaces pour résoudre les problèmes directs sur des objets aussi particuliers qu'un corps humain.

Modèle numérique

On rappelle que l'utilisation d'un modèle numérique basé sur la méthode des moments est pertinente pour la reconstruction des objets de la base Fresnel car la configuration d'acquisi-

tion correspond bien à ce choix. En revanche, pour des applications biomédicales, il peut être judicieux d'utiliser des modèles différentiels (différences finies ou éléments finis). Les performances de l'approche par blocs avec un modèle différentiel ne sont pas intuitives et il serait intéressant d'étudier son utilisation pour des modèles numériques autres que la méthode des moments.

Certains matériaux sont très dispersifs, c'est-à-dire que les propriétés diélectriques du matériau sont fonction de la fréquence. Dans les reconstructions présentées dans nos travaux, l'hypothèse de matériaux non dispersifs est utilisée. En revanche, ce n'est pas le cas pour des tissus humains par exemple. Il serait donc nécessaire de prendre en compte la dispersion dans le modèle numérique. Le modèle de Debye permet de prendre en compte cette dispersion mais nécessite de considérer des variables inconnues supplémentaires.

Applications autres que la tomographie micro-ondes

Les approches par blocs en TMO trouvent leur intérêt dans le fait que de nombreux systèmes avec une même matrice opérateur doivent être résolus. Des applications autres que la TMO ont le besoin similaire de résoudre de multiples systèmes impliquant la même matrice. L'idée est donc d'ouvrir l'utilisation des approches par blocs à ces applications. On peut citer entre autres les techniques d'imagerie basées sur la diffraction de quantités physiques ; ces techniques peuvent être regroupées sous le nom de *tomographie de diffraction*. On peut citer la tomographie ultrasonore par diffraction ou l'imagerie sismique (Vautrin, 2011). Un travail en cours sur la tomographie par induction magnétique fait d'ailleurs l'objet d'une maîtrise à l'École Polytechnique de Montréal. On peut également citer le calcul de surfaces équivalentes radar (RCS). D'autres applications totalement différentes peuvent également être intéressées par l'utilisation d'approches par blocs, comme la chromodynamique quantique (*Lattice QCD*) (Nakamura et al., 2012).

RÉFÉRENCES

A. Abubakar, “Three-dimensional nonlinear inversion of electrical conductivity”, Thèse de doctorat, TU Delft, Delft, The Netherlands, 2000.

A. Abubakar et P. M. van den Berg, “Total variation as a multiplicative constraint for solving inverse problems”, *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1384–1392, 2001.

——, “Iterative forward and inverse algorithms based on domain integral equations for three-dimensional electric and magnetic objects”, *J. Comput. Phys.*, vol. 195, no. 1, pp. 236–262, October 2004.

A. Abubakar, P. M. van den Berg, et T. M. Habashy, “Application of the multiplicative regularized contrast source inversion method on TM- and TE-polarized experimental Fresnel data”, *Inverse Probl.*, vol. 21, no. 6, pp. S5–S13, November 2005.

A. Abubakar, T. M. Habashy, V. L. Druskin, L. Knizhnerman, et D. L. Alumbaugh, “2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements”, *Geophysics*, vol. 73, no. 4, pp. F165–F177, 2008.

A. Abubakar, W. Hu, P. M. van den Berg, et T. M. Habashy, “A finite-difference contrast source inversion method”, *Inverse Probl.*, vol. 24, no. 6, p. 065004, 2008.

A. Abubakar, T. M. Habashy, et G. Pan, “Microwave data inversions using the source-receiver compression scheme”, *IEEE Trans. Antennas Propag.*, vol. 60, no. 6, pp. 2853–2864, June 2012.

A. Abubakar, T. M. Habashy, G. Pan, et M.-K. Li, “Application of the multiplicative regularized Gauss-Newton algorithm for three-dimensional microwave imaging”, *IEEE Trans. Antennas Propag.*, vol. 60, no. 5, pp. 2431–2441, May 2012.

S. A. Akhmanov et S. Y. Nikitin, *Physical Optics*. Clarendon Press, 1997.

M. Al Sharkawy et H. El-Ocla, “Electromagnetic scattering from 3-D targets in a random medium using finite difference frequency domain”, *Antennas and Propagation, IEEE Transactions on*, vol. 61, no. 11, pp. 5621–5626, 2013.

M. A. Ali et M. Moghaddam, “3D nonlinear super-resolution microwave inversion technique using time-domain data”, *IEEE Trans. Antennas Propag.*, vol. 58, no. 7, pp. 2327–2336,

July 2010.

M. Asefi, I. Jeffrey, J. LoVetri, C. Gilmore, P. Card, et J. Paliwal, “Grain bin monitoring via electromagnetic imaging”, *Computers and Electronics in Agriculture*, vol. 119, pp. 133–141, 2015.

M. Azghani, P. Kosmas, et F. Marvasti, “Microwave medical imaging based on sparsity and an iterative method with adaptive thresholding”, *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 357–365, 2015.

C. A. Balanis, *Advanced Engineering Electromagnetics ; 2nd Ed.* John Wiley & Sons, 2012.

P.-A. Barrière, “Développement d’algorithmes d’inversion rapides et propositions relatives à la configuration du montage de mesures dans un contexte de tomographie micro-ondes appliquée à la détection du cancer du sein”, Thèse de doctorat, École polytechnique de Montréal, École Centrale de Nantes (IRCCyN), 2008.

P.-A. Barrière, J. Idier, J.-J. Laurin, et Y. Goussard, “Contrast source inversion method applied to relatively high contrast objects”, *Inverse Probl.*, vol. 27, no. 7, p. 075012, 2011.

K. Belkebir, S. Bonnard, F. Pezin, P. Sabouroux, et M. Saillard, “Validation of 2D inverse scattering algorithms from multi-frequency experimental data”, *Journal of Electromagnetic Waves and Applications*, vol. 14, no. 12, pp. 1637–1667, 2000.

J.-P. Berenger, “A perfectly matched layer for the absorption of electromagnetic waves”, *J. Comput. Phys.*, vol. 114, no. 2, pp. 185–200, 1994.

D. P. Bertsekas, *Nonlinear Programming.* Athena Scientific, 1999.

M. Bjelogrlic, M. Mattes, I. Koufogiannis, S. Capdevila, et J. R. Mosig, “Volume integral equation formulation for medical applications”, dans *Proc. 9th European Conf. Antennas Propag.*, April 2015.

G. Bozza et M. Pastorino, “An inexact Newton-based approach to microwave imaging within the contrast source formulation”, *IEEE Trans. Antennas Propag.*, vol. 57, no. 4, pp. 1122–1132, April 2009.

R. H. Byrd, P. Lu, J. Nocedal, et C. Zhu, “A limited memory algorithm for bound constrained optimization”, *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, 1995.

H. Carfantan, “Approche bayésienne pour un problème inverse non linéaire en imagerie à ondes diffractées”, Thèse de doctorat, Université de Paris-Sud, Centre d’Orsay, 1996.

I. Catapano, L. Crocco, M. D’Urso, et T. Isernia, “A novel effective model for solving 3-D nonlinear inverse scattering problems in lossy scenarios”, *Geoscience and Remote Sensing Letters, IEEE*, vol. 3, no. 3, pp. 302–306, July 2006.

——, “3D microwave imaging via preliminary support reconstruction : Testing on the Fresnel 2008 database”, *Inverse Problems*, vol. 25, no. 2, p. 024002, 2009.

P. C. Chaumet et K. Belkebir, “Three-dimensional reconstruction from real data using a conjugate gradient-coupled dipole method”, *Inverse Problems*, vol. 25, no. 2, p. 024003, 2009.

P. C. Chaumet et A. Rahmani, “Efficient iterative solution of the discrete dipole approximation for magnetodielectric scatterers”, *Optics letters*, vol. 34, no. 7, pp. 917–919, 2009.

P. C. Chaumet, A. Sentenac, et A. Rahmani, “Coupled dipole method for scatterers with large permittivity”, *Physical Review E*, vol. 70, no. 3, p. 036606, 2004.

P. C. Chaumet, T. Zhang, A. Rahmani, B. Gralak, et K. Belkebir, “Discrete dipole approximation in time domain through the Laplace transform”, *Physical Review E*, vol. 88, no. 6, p. 063303, 2013.

W. C. Chew et Y. M. Wang, “Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method”, *Medical Imaging, IEEE Transactions on*, vol. 9, no. 2, pp. 218–225, 1990.

E. Darve, “The fast multipole method : Numerical implementation”, *Journal of Computational Physics*, vol. 160, no. 1, pp. 195–240, 2000.

D. B. Davidson, *Computational Electromagnetics for RF and Microwave Engineering ; 2nd Ed.* Cambridge University Press, 2010.

J. De Zaeytijd et A. Franchois, “Three-dimensional quantitative microwave imaging from measured data with multiplicative smoothing and value picking regularization”, *Inverse Problems*, vol. 25, no. 2, p. 024004, 2009.

J. De Zaeytijd, A. Franchois, C. Eyraud, et J.-M. Geffrin, “Full-wave three-dimensional microwave imaging with a regularized Gauss-Newton method - theory and experiment”, *IEEE Trans. Antennas Propag.*, vol. 55, no. 11, pp. 3279–3292, November 2007.

J. De Zaeytijd, I. Bogaert, et A. Franchois, “An efficient hybrid MLFMA–FFT solver for the volume integral equation in case of sparse 3D inhomogeneous dielectric scatterers”, *Journal of Computational Physics*, vol. 227, no. 14, pp. 7052–7068, 2008.

A. Diaz Bolado, “Contribution to the design and implementation of a microwave tomography system for breast cancer detection”, Thèse de doctorat, École Polytechnique de Montréal, December 2012.

B. T. Draine et P. J. Flatau, “Discrete-dipole approximation for scattering calculations”, *J. Opt. Soc. Am.*, vol. 11, no. 4, 1994.

D. G. Drogoudis, G. A. Kyriacou, et J. N. Sahalos, “Microwave tomography employing an adjoint network based sensitivity matrix”, *Progress In Electromagnetics Research*, vol. 94, pp. 213–242, 2009.

B. Duchêne, “Inversion of experimental data using linearized and binary specialized nonlinear inversion schemes”, *Inverse Probl.*, vol. 17, no. 6, pp. 1623–1634, 2001.

A. El Guennouni, K. Jbilou, et H. Sadok, “A block version of BiCGSTAB for linear systems with multiple right-hand sides”, *Electronic Transactions on Numerical Analysis*, vol. 16, no. 129-142, p. 2, 2003.

R. G. Ellis, “Electromagnetic inversion using the QMR-FFT fast integral equation method”, dans *2002 SEG Annual Meeting*, 2002.

C. Estatico, M. Pastorino, et A. Randazzo, “Microwave imaging of three-dimensional targets by means of an inexact-Newton-based inversion algorithm”, *International Journal of Antennas and Propagation*, vol. 2013, 2013.

C. Eyraud, A. Litman, A. Hérique, et W. Kofman, “Microwave imaging from experimental data within a Bayesian framework with realistic random noise”, *Inverse problems*, vol. 25, no. 2, p. 024005, 2009.

Q. Fang, P. M. Meaney, S. D. Geimer, A. V. Streltsov, et K. D. Paulsen, “Microwave image reconstruction from 3-D fields coupled to 2-D parameter estimation”, *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 475–484, April 2004.

Q. Fang, P. M. Meaney, et K. D. Paulsen, “Viable three-dimensional medical microwave tomography : Theory and numerical experiments”, *IEEE Trans. Antennas Propag.*, vol. 58, no. 2, pp. 449–458, February 2010.

A. Fhager, S. K. Padhi, et J. Howard, “3D image reconstruction in microwave tomography using an efficient FDTD model”, *IEEE Antennas Wirel. Propag. Lett.*, vol. 8, pp. 1353–1356, 2009.

P. J. Flatau et B. T. Draine, “Light scattering by hexagonal columns in the discrete dipole approximation”, *Optics express*, vol. 22, no. 18, pp. 21 834–21 846, 2014.

E. Forte, M. Pipan, D. Casabianca, R. Di Cuia, et A. Riva, “Imaging and characterization of a carbonate hydrocarbon reservoir analogue using GPR attributes”, *Journal of Applied Geophysics*, vol. 81, pp. 76–87, 2012.

G. Franceschini, D. Franceschini, et A. Massa, “Full-vectorial three-dimensional microwave imaging through the iterative multiscaling strategy - a preliminary assessment”, *Geoscience and Remote Sensing Letters, IEEE*, vol. 2, no. 4, pp. 428–432, October 2005.

A. Franchois et C. Pichot, “Microwave imaging-complex permittivity reconstruction with a Levenberg-Marquardt method”, *IEEE Trans. Antennas Propag.*, vol. 45, no. 2, pp. 203–215, February 1997.

A. Franchois et A. G. Tijhuis, “A quasi-newton reconstruction algorithm for a complex microwave imaging scanner environment”, *Radio Science*, vol. 38, no. 2, pp. VIC–12, 2003.

R. W. Freund et M. Malhotra, “A block QMR algorithm for non-Hermitian linear systems with multiple right-hand sides”, *Linear Algebra and its Applications*, vol. 254, no. 1, pp. 119–157, 1997.

S. Fujino, “GPBiCG (m, l) : A hybrid of BiCGSTAB and GPBiCG methods with efficiency and robustness”, *Applied Numerical Mathematics*, vol. 41, no. 1, pp. 107–117, 2002.

H. Gan et W. C. Chew, “A discrete BCG-FFT algorithm for solving 3D inhomogeneous scatterer problems”, *Journal of Electromagnetic Waves and Applications*, vol. 9, no. 10, pp. 1339–1357, 1995.

J.-M. Geffrin et P. Sabouroux, “Continuing with the Fresnel database : Experimental setup and improvements in 3D scattering measurements”, *Inverse Probl.*, vol. 25, no. 2, p. 024001, 2009.

N. Geng, A. Sullivan, et L. Carin, “Fast multipole method for scattering from 3-D PEC targets situated in a half-space environment”, *Microwave and Optical Technology Letters*, vol. 21, no. 6, pp. 399–405, 1999.

C. Gilmore, A. Abubakar, W. Hu, T. M. Habashy, et P. M. van den Berg, “Microwave biomedical data inversion using the Finite-Difference contrast source inversion method”, *IEEE Trans. Antennas Propag.*, vol. 57, no. 5, pp. 1528–1538, May 2009.

C. Gilmore, P. Mojabi, et J. LoVetri, “Comparison of an enhanced distorted Born iterative method and the multiplicative-regularized contrast source inversion method”, *IEEE Trans. Antennas Propag.*, vol. 57, no. 8, pp. 2341–2351, August 2009.

A. H. Golnabi, P. M. Meaney, N. R. Epstein, et K. D. Paulsen, “Microwave imaging for breast cancer detection : Advances in three—dimensional image reconstruction”, dans *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011, pp. 5730–5733.

J. J. Goodman, B. T. Draine, et P. J. Flatau, “Application of fast-Fourier-transform techniques to the discrete-dipole approximation”, *Opt. Lett*, vol. 16, no. 15, pp. 1198–1200, 1991.

T. M. Grzegorczyk, P. M. Meaney, P. A. Kaufman, R. M. diFlorio Alexander, et K. D. Paulsen, “Fast 3-D tomographic microwave imaging for breast cancer detection”, *IEEE Trans. Med. Imaging*, vol. 31, no. 8, pp. 1584–1592, August 2012.

M. H. Gutknecht, “Variants of BICGSTAB for matrices with complex spectrum”, *SIAM Journal on Scientific Computing*, vol. 14, no. 5, pp. 1020–1033, 1993.

O. S. Haddadin et E. S. Ebbini, “Imaging strongly scattering media using a multiple frequency distorted Born iterative method”, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 45, no. 6, pp. 1485–1496, 1998.

H. Harada, M. Tanaka, et T. Takenaka, “Image reconstruction of a three-dimensional dielectric object using a gradient-based optimization”, *Microwave and Optical Technology Letters*, vol. 29, no. 5, pp. 332–336, 2001.

R. F. Harrington et J. L. Harrington, *Field Computation by Moment Methods*. Oxford University Press, 1996.

J.-L. Hu, Z. Wu, H. McCann, L. E. Davis, et C.-G. Xie, “Quasi-three-dimensional method of moments for analyzing electromagnetic wave scattering in microwave tomography systems”, *IEEE Sens. J.*, vol. 5, no. 2, pp. 216–223, April 2005.

——, “BFGS quasi-Newton method for solving electromagnetic inverse problems”, dans

Microwaves, Antennas and Propagation, IEE Proceedings, vol. 153, no. 2. IET, 2006, pp. 199–204.

W. Hu, A. Abubakar, et T. M. Habashy, "Simultaneous multifrequency inversion of full-waveform seismic data", *Geophysics*, vol. 74, no. 2, pp. R1–R14, 2009.

X. Hu, D. Z. Ding, et R. S. Chen, "A novel flexible GMRES with deflated restarted for efficient solution of electromagnetic problems", *Microwave and Optical Technology Letters*, vol. 53, no. 6, pp. 1360–1364, 2011.

G. Hursan et M. S. Zhdanov, "Contraction integral equation method in three-dimensional electromagnetic modeling", *Radio Science*, vol. 37, no. 6, pp. 1–1–1–13, 2002.

J. Idier, éd., *Approche Bayésienne pour les Problèmes Inverses*, série Traitement du Signal et de l'Image. Hermes Science – Lavoisier, 2001.

N. Joachimowicz, C. Pichot, et J.-P. Hugonin, "Inverse scattering : An iterative numerical method for electromagnetic imaging", *IEEE Trans. Antennas Propag.*, vol. 39, no. 12, pp. 1742–1753, December 1991.

Z. Ke-Ming et S. Yu-Fa, "Fast calculation of wide-band RCS for three-dimensional PEC objects based on preconditioned AWE technique", dans *Signals Systems and Electronics (ISSSE), 2010 International Symposium on*, vol. 2. IEEE, 2010, pp. 1–4.

R. E. Kleinman et P. M. van den Berg, "A modified gradient method for two-dimensional problems in tomography", *Journal of Computational and Applied Mathematics*, vol. 42, no. 1, pp. 17–35, 1992.

R. G. Kouyoumjian et P. H. Pathak, "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface", *Proceedings of the IEEE*, vol. 62, no. 11, pp. 1448–1461, 1974.

G. Li, L. Zhang, et B. Han, "Stable electromagnetic modeling using a multigrid solver on stretching grids : The magnetotelluric example", *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 3, pp. 334–338, 2016.

M. Li, A. Abubakar, et P. M. van den Berg, "Application of the multiplicative regularized contrast source inversion method on 3D experimental Fresnel data", *Inverse Problems*, vol. 25, no. 2, p. 024006, 2009.

M. Li, A. Abubakar, et T. M. Habashy, “A three-dimensional model-based inversion algorithm using radial basis functions for microwave data”, *IEEE Trans. Antennas Propag.*, vol. 60, no. 7, pp. 3361–3372, July 2012.

M. Li, O. Semerci, et A. Abubakar, “A contrast source inversion method in the wavelet domain”, *Inverse Probl.*, vol. 29, no. 2, p. 025015, 2013.

X. Li, E. J. Bond, B. D. Van Veen, et S. C. Hagness, “An overview of ultra-wideband microwave imaging via space-time beamforming for early-stage breast-cancer detection”, *IEEE Antennas and Propagation Magazine*, vol. 47, no. 1, pp. 19–34, 2005.

A. Litman et L. Crocco, “Testing inversion algorithms against experimental data : 3D targets”, *Inverse Problems*, vol. 25, no. 2, p. 020201, 2009.

J.-X. Liu, P.-f. Jiang, X.-z. Tong, L.-h. Xu, W. Xie, et H. Wang, “Application of BICGSTAB algorithm with incomplete LU decomposition preconditioning to two-dimensional magnetotelluric forward modeling [j]”, *Journal of Central South University (Science and Technology)*, vol. 2, p. 042, 2009.

N. Maaref, P. Millot, C. Pichot, et O. Picon, “FMCW ultra-wideband radar for through-the-wall detection of human beings”, dans *2009 International Radar Conference" Surveillance for a Safer World"(RADAR 2009)*. IEEE, 2009, pp. 1–5.

T. A. Maniatis, K. S. Nikita, et N. K. Uzunoglu, “Two-dimensional dielectric profile reconstruction based on spectral-domain moment method and nonlinear optimization”, *Microwave Theory and Techniques, IEEE Transactions on*, vol. 48, no. 11, pp. 1831–1840, 2000.

P. R. McGillivray et D. W. Oldenburg, “Methods for calculating Fréchet derivatives and sensitivities for then non-linear inverse problem : A comparative study”, *Geophysical Prospecting*, vol. 38, no. 5, pp. 499–524, 1990.

P. M. Meaney, Q. Fang, T. Rubaek, E. Demidenko, et K. D. Paulsen, “Log transformation benefits parameter estimation in microwave tomographic imaging”, *Med. Phys.*, vol. 34, no. 6, pp. 2014–2023, June 2007.

P. M. Meaney, T. M. Grzegorczyk, E. Attardo, et K. D. Paulsen, “Ultrafast 3D microwave tomography utilizing the direct dipole approximation”, dans *2012 International Conference on Electromagnetics in Advanced Applications*, September 2012, pp. 838–839.

P. Mojabi et J. LoVetri, “Microwave biomedical imaging using the multiplicative regularized Gauss-Newton inversion”, *IEEE Antennas Wirel. Propag. Lett.*, vol. 8, pp. 645–648, 2009.

——, “Eigenfunction contrast source inversion for circular metallic enclosures”, *Inverse Probl.*, vol. 26, no. 2, p. 025010, 2010.

E. Mudry, P. C. Chaumet, K. Belkebir, et A. Sentenac, “Electromagnetic wave imaging of three-dimensional targets using a hybrid iterative inversion method”, *Inverse Probl.*, vol. 28, no. 6, p. 065007, 2012.

Y. Nakamura, K.-I. Ishikawa, Y. Kuramashi, T. Sakurai, et H. Tadano, “Modified block BiCGSTAB for lattice QCD”, *Computer Physics Communications*, vol. 183, no. 1, pp. 34–37, 2012.

J. Nocedal et S. J. Wright, *Numerical Optimization*, J. Nocedal et S. J. Wright, éds. Springer Verlag, 1999.

D. P. O’Leary, “The block conjugate gradient algorithm and related methods”, *Linear algebra and its applications*, vol. 29, pp. 293–322, 1980.

G. Oliveri, L. Lizzi, M. Pastorino, et A. Massa, “A nested multi-scaling inexact-Newton iterative approach for microwave imaging”, *Antennas and Propagation, IEEE Transactions on*, vol. 60, no. 2, pp. 971–983, 2012.

M. Ostadrahimi, P. Mojabi, A. Zakaria, J. LoVetri, et L. Shafai, “Enhancement of Gauss-Newton inversion method for biological tissue imaging”, *Microwave Theory and Techniques, IEEE Transactions on*, vol. 61, no. 9, pp. 3424–3434, 2013.

O. Ozdemir, “Cauchy data contrast source inversion method”, *Geoscience and Remote Sensing Letters, IEEE*, vol. 11, no. 4, pp. 858–862, April 2014.

M. Pastorino, “Stochastic optimization methods applied to microwave imaging : A review”, *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 538–548, March 2007.

M. Pastorino, éd., *Microwave imaging*. John Wiley & Sons, 2010, vol. 208.

R.-E. Plessix, “A review of the adjoint-state method for computing the gradient of a functional with geophysical applications”, *Geophysical Journal International*, vol. 167, no. 2, pp. 495–503, 2006.

I. T. Rekanos et T. D. Tsiboukis, “An iterative numerical method for inverse scattering problems”, *Radio Science*, vol. 34, no. 6, pp. 1401–1412, 1999.

I. T. Rekanos, S. M. Panas, et T. D. Tsiboukis, “Microwave imaging using the finite-element method and a sensitivity analysis approach”, *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, pp. 1108–1114, 1999.

R. F. Remis et P. M. van den Berg, “On the equivalence of the Newton-Kantorovich and distorted Born methods”, *Inverse Problems*, vol. 16, no. 1, pp. L1–L4, 2000.

A. Roger, “Newton-Kantorovitch algorithm applied to an electromagnetic inverse problem”, *Antennas and Propagation, IEEE Transactions on*, vol. 29, no. 2, pp. 232–238, 1981.

T. Rubaek, O. S. Kim, et P. Meincke, “Computational validation of a 3-D microwave imaging system for breast-cancer screening”, *IEEE Trans. Antennas Propag.*, vol. 57, no. 7, pp. 2105–2115, July 2009.

T. Rubaek, P. M. Meaney, et K. D. Paulsen, “A contrast source inversion algorithm formulated using the log-phase formulation”, *Int. J. Antennas Propag.*, vol. 2011, pp. 1–10, 2011.

Y. Saad et M. H. Schultz, “GMRES : A generalized minimal residual algorithm for solving nonsymmetric linear systems”, *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 3, pp. 856–869, 1986.

T. K. Sarkar, E. Arvas, et S. M. Rao, “Application of FFT and the conjugate gradient method for the solution of electromagnetic radiation from electrically large and small conducting bodies”, *Antennas and Propagation, IEEE Transactions on*, vol. 34, no. 5, pp. 635–640, 1986.

R. Scapaticci, L. Di Donato, I. Catapano, et L. Crocco, “A feasibility study on microwave imaging for brain stroke monitoring”, *Progress In Electromagnetics Research B*, vol. 40, pp. 305–324, 2012.

R. Scapaticci, P. Kosmas, et L. Crocco, “Wavelet-based regularization for robust microwave imaging in medical applications”, *Biomedical Engineering, IEEE Transactions on*, vol. 62, no. 4, pp. 1195–1202, April 2015.

V. Simoncini, “A stabilized QMR version of block BICG”, *SIAM Journal on Matrix Analysis and Applications*, vol. 18, no. 2, pp. 419–434, 1997.

V. Simoncini et E. Gallopoulos, “Convergence properties of block GMRES and matrix polynomials”, *Linear Algebra and its Applications*, vol. 247, pp. 97–119, 1996.

M. Soleimani, W. R. B. Lionheart, A. J. Peyton, X. Ma, et S. R. Higson, “A three-dimensional inverse finite-element method applied to experimental eddy-current imaging data”, *Magnetics, IEEE Transactions on*, vol. 42, no. 5, pp. 1560–1567, May 2006.

K. M. Soodhalter, “A block MINRES algorithm based on the band Lanczos method”, *Numerical Algorithms*, vol. 69, no. 3, pp. 473–494, May 2015.

T. Takenaka, H. Zhou, et T. Tanaka, “Inverse scattering for a three-dimensional object in the time domain”, *JOSA A*, vol. 20, no. 10, pp. 1867–1874, 2003.

A. Trillon, “Eddy current tomography with a differential forward model”, Thèse de doctorat, Ecole Centrale de Nantes, Nantes, France, 2010.

P. M. van den Berg et A. Abubakar, “Contrast source inversion method : State of art”, *Progress in Electromagnetics Research*, vol. 34, no. 11, pp. 189–218, 2001.

P. M. van den Berg et R. E. Kleinman, “A total variation enhanced modified gradient algorithm for profile reconstruction”, *Inverse Probl.*, vol. 11, no. 3, pp. L5–L10, 1995.

——, “A contrast source inversion method”, *Inverse Probl.*, vol. 13, no. 6, p. 1607, 1997.

H. A. Van der Vorst, “Bi-CGSTAB : A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems”, *SIAM Journal on scientific and Statistical Computing*, vol. 13, no. 2, pp. 631–644, 1992.

D. Vautrin, “Régularisation et optimisation pour l’imagerie sismique des fondations de pylônes”, Thèse de doctorat, École Centrale de Nantes, Nantes, France, 2011.

C. F. Wang et J. M. Jin, “Simple and efficient computation of electromagnetic fields in arbitrarily shaped inhomogeneous dielectric bodies using transpose-free QMR and FFT”, *Microwave Theory and Techniques, IEEE Transactions on*, vol. 46, no. 5, pp. 553–558, 1998.

D. W. Winters, B. D. Van Veen, et S. C. Hagness, “A sparsity regularization approach to the electromagnetic inverse scattering problem”, *IEEE Trans. Antennas Propag.*, vol. 58, no. 1, pp. 145–154, January 2010.

F. Wu, Y. Zhang, Z. Z. Oo, et E. Li, “Parallel multilevel fast multipole method for solving large-scale problems”, *Antennas and Propagation Magazine, IEEE*, vol. 47, no. 4, pp. 110–118, 2005.

X. Xu, Q. H. Liu, et Z. Q. Zhang, “The stabilized biconjugate gradient fast Fourier transform method for electromagnetic scattering”, dans *Antennas and Propagation Society International Symposium, 2002. IEEE*, vol. 2. IEEE, 2002, pp. 614–617.

C. Yu, M. Yuan, et Q. H. Liu, “Reconstruction of 3D objects from multi-frequency experimental data with a fast DBIM-BCGS method”, *Inverse Problems*, vol. 25, no. 2, p. 024007, 2009.

M. A. Yurkin et A. G. Hoekstra, “The discrete dipole approximation : An overview and recent developments”, *J. Quant. Spectrosc. Radiat. Transf.*, vol. 106, pp. 558–589, 2007.

S. H. Zainud-Deen, W. M. Hassen, E. M. Ali, K. H. Awadalla, et H. A. Sharshar, “Breast cancer detection using a hybrid finite difference frequency domain and particle swarm optimization techniques”, dans *Radio Science Conference, 2008. NRSC 2008. National.* IEEE, 2008, pp. 1–8.

A. Zakaria et J. LoVetri, “Application of multiplicative regularization to the finite-element contrast source inversion method”, *IEEE Trans. Antennas Propag.*, vol. 59, no. 9, pp. 3495–3498, September 2011.

——, “The finite-element method contrast source inversion algorithm for 2D transverse electric vectorial problems”, *IEEE Trans. Antennas Propag.*, vol. 60, no. 10, pp. 4757–4765, October 2012.

A. Zakaria, C. Gilmore, et J. LoVetri, “Finite-element contrast source inversion method for microwave imaging”, *Inverse Problems*, vol. 26, no. 11, p. 115010, 2010.

E. Zastrow, S. Davis, M. Lazebnik, F. Kelcz, B. Van Veem, et S. C. Hagness, “Database of 3D grid-based numerical breast phantoms for use in computational electromagnetics simulations”, *Department of Electrical and Computer Engineering University of Wisconsin-Madison*, 2008.

S.-L. Zhang, “GPBi-CG : Generalized product-type methods based on Bi-CG for solving nonsymmetric linear systems”, *SIAM Journal on Scientific Computing*, vol. 18, no. 2, pp. 537–551, 1997.

W. Zhang et Q. H. Liu, “Three-dimensional scattering and inverse scattering from objects with simultaneous permittivity and permeability contrasts”, *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, no. 1, pp. 429–439, January 2015.

Z. Q. Zhang et Q. H. Liu, “Three-dimensional nonlinear image reconstruction for microwave biomedical imaging”, *IEEE Trans. Biomed. Eng.*, vol. 51, no. 3, pp. 544–548, March 2004.

Z. Q. Zhang, Q. H. Liu, C. Xiao, E. Ward, G. Ybarra, et W. T. Joines, “Microwave breast imaging : 3-D forward scattering simulation”, *IEEE Trans. Biomed. Eng.*, vol. 50, no. 10, pp. 1180–1189, October 2003.

Z. Q. Zhang, Q. H. Liu, et X. M. Xu, “RCS computation of large inhomogeneous objects using a fast integral equation solver”, *Antennas and Propagation, IEEE Transactions on*, vol. 51, no. 3, pp. 613–618, 2003.

P. Zwamborn et P. M. van den Berg, “The three-dimensional weak form of the conjugate gradient FFT method for solving scattering problems”, *IEEE Trans. Microw. Theory Tech.*, vol. 40, no. 9, pp. 1757–1766, September 1992.

ANNEXE A DISCRÉTISATION DE L'ÉQUATION INTÉGRALE ET FORMULATION MATRICIELLE

A.1 Discrétisation de l'équation du domaine

On rappelle l'équation intégrale du domaine reliant implicitement le champ total dans le domaine V de l'objet au contraste χ :

$$\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}) = \overrightarrow{\boldsymbol{E}}^{\text{inc}}(\boldsymbol{r}) + k_b^2 \int_V \overline{\overline{\mathbf{G}}}(\boldsymbol{r}, \boldsymbol{r}')\, \chi(\boldsymbol{r}')\, \overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}') d\boldsymbol{r}' \tag{A.1}$$

avec le tenseur de Green

$$\overline{\overline{\mathbf{G}}} = (\mathbf{I} + \frac{1}{k^2}\nabla\nabla)\, g(\boldsymbol{r}, \boldsymbol{r}')$$

et $g(\boldsymbol{r}, \boldsymbol{r}')$ la fonction scalaire de Green. Il est possible de séparer le tenseur de Green afin de réécrire l'équation intégrale :

$$\overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}) = \overrightarrow{\boldsymbol{E}}^{\text{inc}}(\boldsymbol{r}) + k_b^2 \overrightarrow{\boldsymbol{A}}(\boldsymbol{r}) + \nabla\nabla \cdot \overrightarrow{\boldsymbol{A}}(\boldsymbol{r}) \tag{A.2}$$

avec

$$\overrightarrow{\boldsymbol{A}}(\boldsymbol{r}) = \int_V g(\boldsymbol{r}, \boldsymbol{r}')\, \chi(\boldsymbol{r}')\, \overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}') d\boldsymbol{r}' \tag{A.3}$$

En pratique, il est possible d'obtenir une expression analytique du tenseur de Green $\overline{\overline{\mathbf{G}}}$ et de travailler directement avec les noyaux de convolution qui en résultent. Mais pour des raisons de simplicité et de réduction du temps de calul, nous avons préféré décrire l'équation intégrale sous la forme (A.2) et calculer d'abord l'intégrale $\overrightarrow{\boldsymbol{A}}$ faisant intervenir uniquement la fonction scalaire g et non le tenseur. On posera $\overrightarrow{\boldsymbol{B}} = \nabla\nabla \cdot \overrightarrow{\boldsymbol{A}}$.

Le domaine d'étude est discrétisé en voxels cubiques de côté Δx et de volume $\Delta V = \Delta x^3$. On suppose que le volume discrétisé contient N_x voxels selon l'axe x, N_y voxels selon l'axe y et N_z voxels selon l'axe z. La discrétisation par méthode des moments de cette expression donne pour tout $1 \leq m \leq N_x$, $1 \leq n \leq N_y$, $1 \leq p \leq N_z$:

$$E^{\text{tot}}_{\kappa;m,n,p} = E^{\text{inc}}_{\kappa;m,n,p} + k_b^2 A_{\kappa;m,n,p} + B_{\kappa;m,n,p} \tag{A.4}$$

où l'indice κ représente la composante spatiale des vecteurs (x,y ou z) et les indices m, n, p

représentent les voxels du domaine d'étude discrétisé. Le vecteur $\overrightarrow{\boldsymbol{A}}$ discrétisé sur les voxels

$$A_{\kappa;m,n,p} = \Delta V \sum_{m'=1}^{N_x} \sum_{n'=1}^{N_y} \sum_{p'=1}^{N_z} g_{m-m',n-n',p-p'}\ \chi_{m',n',p'}\ E^{\text{tot}}_{\kappa;m',n',p'} \tag{A.5}$$

pour $0 \leq m \leq N_x + 1$, $0 \leq n \leq N_y + 1$, $0 \leq p \leq N_z + 1$. $g_{m-m',n-n',p-p'}$ correspond à la fonction de Green discrétisée définie par : Afin de gérer la singularité de la fonction de Green, la formulation faible est utilisée (Abubakar et van den Berg, 2004) :

$$g_{m-m',n-n',p-p'} = \begin{cases} \dfrac{\left(1 - \frac{1}{2} j k_b \Delta x\right) \exp(\frac{1}{2} j k_b \Delta x) - 1}{\frac{1}{6}\pi k_b^2 \Delta x^3} & \text{si } m-m', n-n', p-p' = 0 \\ g(\boldsymbol{r}_{m,n,p}, \boldsymbol{r}_{m',n',p'}) & \text{otherwise} \end{cases} \tag{A.6}$$

On souhaite exprimer cette relation sous forme matricielle. On stocke les trois composantes du champ total pour les N voxels dans un vecteur de longeur $3N$ de la façon suivante :

$$\boldsymbol{e}^{\text{tot}} = \begin{pmatrix} \boldsymbol{e}_x^{\text{tot}} \\ \boldsymbol{e}_y^{\text{tot}} \\ \boldsymbol{e}_z^{\text{tot}} \end{pmatrix} \tag{A.7}$$

où les vecteurs $\boldsymbol{e}_\kappa^{\text{tot}}$ ($\kappa = x, y$ ou z) contiennent le champ total $\overrightarrow{\boldsymbol{E}}^{\text{tot}}$ des N voxels pour la composante κ. Le même arrangement est réalisé pour le vecteur $\boldsymbol{e}^{\text{inc}}$ qui contient les composantes du champ incident $\overrightarrow{\boldsymbol{E}}^{\text{inc}}$ sur les N voxels et pour les vecteurs $\boldsymbol{A}$ et $\boldsymbol{B}$ qui représentent la discrétisation de $\overrightarrow{\boldsymbol{A}}$ et $\overrightarrow{\boldsymbol{B}}$. De même pour le contraste χ, les valeurs sur les voxels sont stockées dans le vecteur $\boldsymbol{x}$ de taille N. On peut alors écrire (A.5) comme

$$\begin{pmatrix} \boldsymbol{A}_x \\ \boldsymbol{A}_y \\ \boldsymbol{A}_z \end{pmatrix} = \Delta V \begin{bmatrix} \widehat{\mathbf{G}_A} & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \widehat{\mathbf{G}_A} & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & \widehat{\mathbf{G}_A} \end{bmatrix} \begin{bmatrix} \text{diag}(\boldsymbol{x}) & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \text{diag}(\boldsymbol{x}) & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & \text{diag}(\boldsymbol{x}) \end{bmatrix} \begin{pmatrix} \boldsymbol{e}_x^{\text{tot}} \\ \boldsymbol{e}_y^{\text{tot}} \\ \boldsymbol{e}_z^{\text{tot}} \end{pmatrix} \tag{A.8}$$

ou de façon équivalente

$$\boldsymbol{A} = \Delta V \mathbf{G}_A \mathbf{X} \boldsymbol{e}^{\text{tot}} \tag{A.9}$$

La matrice $\widehat{\mathbf{G}_A}$ correspond à la matrice de convolution de l'équation (A.5). En pratique, ce calcul est fait par transformées de Fourier sur chacune des trois composantes x, y et z.

Enfin, le terme $\overrightarrow{\boldsymbol{B}} = \nabla\nabla \cdot \overrightarrow{\boldsymbol{A}}$ est approché par différences finies. On obtient par exemple pour

la composante x :

$$\begin{aligned} B_{x;m,n,p} = & \frac{A_{x;m-1,n,p} - 2A_{x;m,n,p} + A_{x;m+1,n,p}}{\Delta x^2} \\ & + \frac{A_{y;m-1,n-1,p} - A_{y;m-1,n+1,p} - A_{y;m+1,n-1,p} + A_{y;m+1,n+1,p}}{4\Delta x^2} \\ & + \frac{A_{z;m-1,n,p-1} - A_{z;m-1,n,p+1} - A_{z;m+1,n,p-1} + A_{z;m+1,n,p+1}}{4\Delta x^2} \end{aligned} \tag{A.10}$$

En pratique, le calcul du vecteur $\boldsymbol{B}$ sera réalisé comme (A.10). Pour l'expression matricielle, on peut exprimer $\boldsymbol{B}$ par :

$$\boldsymbol{B} = \mathbf{D}\boldsymbol{A} \tag{A.11}$$

où la matrice $\mathbf{D}$ de taille $(3N, 3N)$ correspond à la matrice des différences finies exprimées dans (A.10).

Finalement, la relation (A.4) peut être réécrite de façon matricielle :

$$\boldsymbol{e}^{\text{tot}} = \boldsymbol{e}^{\text{inc}} \; + \; k_b^2 \, \Delta V \, \mathbf{G}_A \, \mathbf{X} \, \boldsymbol{e}^{\text{tot}} \; + \; \Delta V \, \mathbf{D} \, \mathbf{G}_A \, \mathbf{X} \, \boldsymbol{e}^{\text{tot}} \tag{A.12}$$

$$= \boldsymbol{e}^{\text{inc}} \; + \; \mathbf{G}_D \, \mathbf{X} \, \boldsymbol{e}^{\text{tot}} \tag{A.13}$$

avec $\mathbf{G}_D = k_b^2 \, \Delta V \, \mathbf{G}_A + \Delta V \, \mathbf{D} \, \mathbf{G}_A$.

Pour le calcul de $\mathbf{G}_o$, repartons de l'équation intégrale d'observation :

$$\overrightarrow{\boldsymbol{E}}^{\text{scat}}(\boldsymbol{r}^R) = k_b^2 \int_V \overline{\overline{\mathbf{G}}}(\boldsymbol{r}^R, \boldsymbol{r}') \, \chi(\boldsymbol{r}') \, \overrightarrow{\boldsymbol{E}}^{\text{tot}}(\boldsymbol{r}') d\boldsymbol{r}' \; , \quad \forall \boldsymbol{r}^R \notin V. \tag{A.14}$$

Comme les points de mesure $\boldsymbol{r}^R$ ne coïncident pas avec les voxels du domaine V, le calcul du produit de convolution ne peut pas être effectué par des routines `FFT`. De plus, pour l'équation du domaine, il est possible de faire l'approximation des dérivées secondes du tenseur par différences finies puisqu'on connaît les quantités sur une grille régulière. Mais cela n'est pas le cas pour le calcul au niveau des récepteurs. Pour l'équation d'observation, nous exprimons analytiquement le tenseur de Green, ce qui permettra de discrétiser l'équation d'observation en considérant directement le tenseur. On obtient

$$\overline{\overline{\mathbf{G}}}_{\nu\kappa}(\boldsymbol{r}, \boldsymbol{r}') = \left[k_b^2 \delta_{\nu\kappa} + \frac{j k_b \delta_{\nu\kappa}}{|\boldsymbol{r} - \boldsymbol{r}'|} - \frac{k_b^2 r_\nu r_\kappa + \delta_{\nu\kappa}}{|\boldsymbol{r} - \boldsymbol{r}'|^2} - \frac{3 j k_b r_\nu r_\kappa}{|\boldsymbol{r} - \boldsymbol{r}'|^3} + \frac{3 r_\nu r_\kappa}{|\boldsymbol{r} - \boldsymbol{r}'|^4} \right] g(\boldsymbol{r}, \boldsymbol{r}'), \; \forall \boldsymbol{r} \neq \boldsymbol{r}', \tag{A.15}$$

où $\overline{\overline{\mathbf{G}}}_{\nu\kappa}$ représente l'élément (ν, κ) du tenseur de Green (par exemple $\overline{\overline{\mathbf{G}}}_{xx}$ ou $\overline{\overline{\mathbf{G}}}_{xy}$). Le terme $\delta_{\nu\kappa}$ correspond à l'indice de Kronecker qui vaut 1 si $\nu = \kappa$ et 0 sinon (par exemple $\delta_{xx} = 1$ et $\delta_{xy} = 0$). Enfin, r_ν (ou de façon équivalente r_κ) correspond à la composante ν du vecteur $\boldsymbol{r} - \boldsymbol{r}'$.

A.2 Discrétisation de l'équation d'observation

La discrétisation de l'équation d'observation est similaire à l'équation du domaine mis à part le fait que le tenseur de Green n'est pas divisé en deux termes $\overrightarrow{\boldsymbol{A}}$ et $\overrightarrow{\boldsymbol{B}}$. On obtient alors la discrétisation vectorielle suivante :

$$\overrightarrow{\boldsymbol{E}}^{\text{scat}}_{m,n,p} = k_b^2 \Delta V \sum_{m'=1}^{N_x} \sum_{n'=1}^{N_y} \sum_{p'=1}^{N_z} \overline{\overline{\mathbf{G}}}_{m-m',n-n',p-p'} \; \chi_{m',n',p'} \; \overrightarrow{\boldsymbol{E}}^{\text{tot}}_{m',n',p'}. \tag{A.16}$$

Les indices m,n et p représentent ici les coordonnées du récepteur. Le tenseur $\overline{\overline{\mathbf{G}}}_{m-m',n-n',p-p'}$ dénote $\overline{\overline{\mathbf{G}}}(\boldsymbol{r}^R, \boldsymbol{r'})$ entre chaque voxel et le récepteur. Pour obtenir une composante κ particulière du champ diffracté, on obtient finalement la relation :

$$E^{\text{scat}}_{\kappa;m,n,p} = k_b^2 \Delta V \sum_{\nu=x}^{z} \sum_{m'=1}^{N_x} \sum_{n'=1}^{N_y} \sum_{p'=1}^{N_z} \overline{\overline{\mathbf{G}}}_{\nu\kappa;m-m',n-n',p-p'} \; \chi_{m',n',p'} \; E^{\text{tot}}_{\nu;m',n',p'}. \tag{A.17}$$

En gardant les mêmes notations que précédemment, la représentation matricielle de cette équation donne :

$$E^{\text{scat}}_{\kappa;m,n,p} = k_b^2 \Delta V (\boldsymbol{u}^R)^{\text{t}} \mathbf{X} \boldsymbol{e}^{\text{tot}}, \tag{A.18}$$

où le vecteur $\boldsymbol{u}^R$ de taille $3N$ correspond aux valeurs du tenseur de Green pour les N couples voxel/récepteur et cela pour les trois composantes vectorielles.

Supposons maintenant que nous avons N_R récepteurs. Nous pouvons concaténer les champs diffractés mesurés à ces récepteurs dans un vecteur $\boldsymbol{e}^{\text{scat}}$. Les vecteurs $\boldsymbol{u}^R$ pour chaque récepteur sont notés $\boldsymbol{u}^R_j$, $j = 1, \ldots, N_R$.

$$\boldsymbol{e}^{\text{scat}} = k_b^2 \Delta V \begin{bmatrix} (\boldsymbol{u}^R_1)^{\text{t}} \\ \vdots \\ (\boldsymbol{u}^R_{N_R})^{\text{t}} \end{bmatrix} \mathbf{X} \boldsymbol{e}^{\text{tot}}. \tag{A.19}$$

Pour simplifier les notations, on pose la matrice $\mathbf{G}_o$ de taille $N_R \times 3N$:

$$\mathbf{G}_o = k_b^2 \Delta V \begin{bmatrix} (\boldsymbol{u}^R_1)^{\text{t}} \\ \vdots \\ (\boldsymbol{u}^R_{N_R})^{\text{t}} \end{bmatrix}. \tag{A.20}$$

On a finalement l'équation d'observation discrétisée pour N_R récepteurs :

$$\boldsymbol{e}^{\text{scat}} = \mathbf{G}_o \mathbf{X} \boldsymbol{e}^{\text{tot}}. \tag{A.21}$$

ANNEXE B CALCUL DU GRADIENT DU CRITÈRE NON-LINÉAIRE

B.1 Expression algébrique du gradient

Rappelons le critère non-linéaire d'adéquation aux données :

$$\mathcal{F}^D(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{N_S} \|\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}\|^2. \tag{B.1}$$

Dans la suite de cette annexe, posons la fonction du modèle direct $\boldsymbol{f}_i(\boldsymbol{x}) = \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}$ et posons le vecteur $\boldsymbol{y}$ de taille $3N$:

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{x} \\ \boldsymbol{x} \end{pmatrix}. \tag{B.2}$$

On définit alors la fonction $\boldsymbol{h}_i(\boldsymbol{y}) = \mathbf{G}_o\mathbf{Y}\mathbf{L}_y^{-1}\boldsymbol{e}_i^{\text{inc}}$ où $\mathbf{Y} = \text{diag}\{\boldsymbol{y}\}$ et $\mathbf{L}_y = \mathbf{I} - \mathbf{G}_D\mathbf{Y}$.

Afin de calculer le gradient du critère, nous allons calculer les jacobiens $\mathbf{J}_i = \nabla_x \boldsymbol{f}_i$, de taille $N_R \times N$. Celui-ci s'exprime sous la forme :

$$\nabla_x \boldsymbol{f}_i = \nabla_y \boldsymbol{h}_i \cdot \nabla_x \boldsymbol{y}, \tag{B.3}$$

avec

$$\nabla_x \boldsymbol{y} = \mathbf{Q} = \begin{bmatrix} \mathbf{I}_N \\ \mathbf{I}_N \\ \mathbf{I}_N \end{bmatrix}. \tag{B.4}$$

Pour rappel, la matrice $\mathbf{I}_N$ représente la matrice identité de taille N. Pour des raisons de simplicité, nous notons $\widetilde{\mathbf{J}}_i$ le jacobien de la fonction $\boldsymbol{h}_i$ de taille $N_R \times 3N$: $\widetilde{\mathbf{J}}_i = \nabla_y \boldsymbol{h}_i$ et $\mathbf{J}_i = \widetilde{\mathbf{J}}_i\,\mathbf{Q}$. Le gradient sera alors exprimé comme :

$$\nabla_x \mathcal{F} = -\mathbf{Q}^{\text{t}} \sum_{i=1}^{N_S} \widetilde{\mathbf{J}}_i^{\dagger} \left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o\mathbf{X}\mathbf{L}_x^{-1}\boldsymbol{e}_i^{\text{inc}}\right). \tag{B.5}$$

B.2 Expression du jacobien

Exprimons maintenant le jacobien $\mathbf{J}_i$. Pour cela, nous effectuons le développement de Taylor au premier ordre de $\boldsymbol{h}_i$:

$$\boldsymbol{h}_i(\boldsymbol{y}+\delta\boldsymbol{y}) \;=\; \boldsymbol{h}_i(\boldsymbol{y}) \;+\; \mathbf{J}_i\,\delta\boldsymbol{y} \;+\; \mathcal{O}\Big(\|\delta\boldsymbol{y}\|^2\Big). \tag{B.6}$$

Calculons $\boldsymbol{h}_i(\boldsymbol{y}+\delta\boldsymbol{y})$:

$$\boldsymbol{h}_i(\boldsymbol{y}+\delta\boldsymbol{y}) = \mathbf{G}_o\Big(\mathbf{Y}+\delta\mathbf{Y}\Big)\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}-\mathbf{G}_D\delta\mathbf{Y}\Big)^{-1}\boldsymbol{e}_i^{\text{inc}}, \tag{B.7}$$

où $\delta\mathbf{Y}$ est la matrice diagonale dont la diagonale est $\delta\boldsymbol{y}$. Pour de faibles variations de $\delta\boldsymbol{y}$, on peut faire l'approximation de Taylor au premier ordre de la matrice $\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}-\mathbf{G}_D\delta\mathbf{Y}\Big)^{-1}$:

$$\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}-\mathbf{G}_D\delta\mathbf{Y}\Big)^{-1} = \Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1} + \Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\mathbf{G}_D\delta\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1} + \;+\mathcal{O}\Big(\|\delta\boldsymbol{y}\|^2\Big). \tag{B.8}$$

En remplaçant dans (B.7), on obtient :

$$\begin{aligned}
\boldsymbol{h}_i(\boldsymbol{y}+\delta\boldsymbol{y}) &= \mathbf{G}_o\Big(\mathbf{Y}+\delta\mathbf{Y}\Big)\Big[\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1} + \Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\mathbf{G}_D\delta\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\Big]\boldsymbol{e}_i^{\text{inc}} \;+\; \mathcal{O}\Big(\|\delta\boldsymbol{y}\|^2\Big)\\
&= \mathbf{G}_o\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\boldsymbol{e}_i^{\text{inc}} \;+\; \mathbf{G}_o\delta\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\boldsymbol{e}_i^{\text{inc}}\\
&\quad + \mathbf{G}_o\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\mathbf{G}_D\delta\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\boldsymbol{e}_i^{\text{inc}} \;+\; \mathcal{O}\Big(\|\delta\boldsymbol{y}\|^2\Big)\\
&= \boldsymbol{h}_i(\boldsymbol{y}) + \mathbf{G}_o\delta\mathbf{Y}\boldsymbol{e}_i^{\text{tot}} + \mathbf{G}_o\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\mathbf{G}_D\delta\mathbf{Y}\boldsymbol{e}_i^{\text{tot}} \;+\; \mathcal{O}\Big(\|\delta\boldsymbol{y}\|^2\Big)\\
&= \boldsymbol{h}_i(\boldsymbol{y}) + \mathbf{G}_o\Big[\mathbf{I}+\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\mathbf{G}_D\Big]\delta\mathbf{Y}\boldsymbol{e}_i^{\text{tot}} \;+\; \mathcal{O}\Big(\|\delta\boldsymbol{y}\|^2\Big).
\end{aligned}$$

Comme $\delta\mathbf{Y}$ est diagonale, on arrange $\delta\mathbf{Y}\boldsymbol{e}_i^{\text{tot}} = \text{diag}\{\boldsymbol{e}_i^{\text{tot}}\}\delta\boldsymbol{y}$. De plus, on montre aisément que $\mathbf{I}+\mathbf{Y}\Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{-1}\mathbf{G}_D = \Big(\mathbf{I}-\mathbf{Y}\mathbf{G}_c\Big)^{-1}$ et que $\mathbf{I}-\mathbf{Y}\mathbf{G}_D = \Big(\mathbf{I}-\mathbf{G}_D\mathbf{Y}\Big)^{\text{t}} = \mathbf{L}_y^{\text{t}}$. On a alors :

$$\boldsymbol{h}_i(\boldsymbol{y}+\delta\boldsymbol{y}) \;=\; \boldsymbol{h}_i(\boldsymbol{y}) \;+\; \mathbf{G}_o(\mathbf{L}_y^{\text{t}})^{-1}\text{diag}\{\boldsymbol{e}_i^{\text{tot}}\}\,\delta\boldsymbol{y} \;+\; \mathcal{O}\Big(\|\delta\boldsymbol{y}\|^2\Big). \tag{B.9}$$

Le jacobien s'exprime finalement sous la forme :

$$\widetilde{\mathbf{J}}_i \;=\; \mathbf{G}_o(\mathbf{L}_y^{\text{t}})^{-1}\text{diag}\{\boldsymbol{e}_i^{\text{tot}}\}. \tag{B.10}$$

Pour se ramener aux inconnues $\boldsymbol{x}$, on remplace $\mathbf{L}_y$ par $\mathbf{L}_x$ et on se ramène à l'expression du gradient :

$$\nabla_x \mathcal{F} = -\sum_{i=1}^{N_S} \mathbf{J}_i^\dagger \left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o \mathbf{X} \mathbf{L}_x^{-1} \boldsymbol{e}_i^{\text{inc}} \right) \tag{B.11}$$

$$= -\mathbf{Q}^{\text{t}} \sum_{i=1}^{N_S} \text{diag}\left\{ \boldsymbol{e}^{\text{tot}} \right\}^\dagger \left(\overline{\mathbf{L}_x} \right)^{-1} \mathbf{G}_o^\dagger \left(\widetilde{\boldsymbol{e}}_i^{\text{scat}} - \mathbf{G}_o \mathbf{X} \mathbf{L}_x^{-1} \boldsymbol{e}_i^{\text{inc}} \right) \tag{B.12}$$

ANNEXE C CALCUL DU GRADIENT COHÉRENT AU CRITÈRE APPROCHÉ

C.1 Rappel du critère approché

Rappelons l'expression du critère approché où les champs totaux sont projetés dans des sous-espaces. Nous faisons l'hypothèse simplificatrice qu'une seule source éclaire le volume et nous omettons le terme de régularisation :

$$\widehat{\mathcal{F}}^D(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - \mathbf{G}_o\mathbf{X}\left(\mathbf{B}^\dagger\mathbf{L}_x^\dagger\mathbf{L}_x\mathbf{B}\right)^{-1}\mathbf{B}^\dagger\mathbf{L}_x^\dagger\boldsymbol{e}^{\text{inc}}\|^2, \tag{C.1}$$

où la matrice $\mathbf{B}$ de taille $3N \times m$ contient les m vecteurs engendrant le sous-espace. Pour la suite, on pose $\mathbf{C} = \mathbf{L}_x\mathbf{B}$.

C.2 Calcul du gradient cohérent

Le gradient correspondant nécessite de dériver les quantités dépendantes de $\boldsymbol{x}$, soit $\mathbf{L}_x$. Pour cela, on passe par la série de Taylor du critère au premier ordre et on cherche les variations du critère :

$$\begin{aligned}
\widehat{\mathcal{F}}(\boldsymbol{x}+\delta\boldsymbol{x}) &= \frac{1}{2}\|\boldsymbol{y} - \mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x})\|^2 \\
&= \frac{1}{2}\Big[\boldsymbol{y}^\dagger\boldsymbol{y} \;-\; \boldsymbol{y}^\dagger\mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x}) \;-\; \hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x})^\dagger(\mathbf{X}+\delta\mathbf{X})^\dagger\mathbf{G}_o^\dagger\boldsymbol{y} \\
&\quad + \;\hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x})^\dagger(\mathbf{X}+\delta\mathbf{X})^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x})\Big] \\
&= \frac{1}{2}\big[\boldsymbol{y}^\dagger\boldsymbol{y} - u - \bar{u} + v\big],
\end{aligned} \tag{C.2}$$

avec $u = \boldsymbol{y}^\dagger\mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x})$ et $v = \hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x})^\dagger(\mathbf{X}+\delta\mathbf{X})^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x})$. En utilisant le lemme d'inversion et en négligeant les termes en $\delta\mathbf{X}$ d'ordre supérieur ou égal à

2, on décompose u pour faire ressortir les termes en $\delta\mathbf{X}$ et $\delta\mathbf{X}^\dagger$:

$$
\begin{aligned}
u &= \boldsymbol{y}^\dagger \mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\,\hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x}) \\
&= \boldsymbol{y}^\dagger \mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\,\mathbf{B}\left[\mathbf{B}^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})\mathbf{B}\right]^{-1}\mathbf{B}^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})^\dagger\boldsymbol{e}^{\text{inc}} \\
&= \boldsymbol{y}^\dagger \mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\,\mathbf{B}\left[\mathbf{C}^\dagger\mathbf{C}-\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C}-\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\mathbf{B}+\mathcal{O}(\|\delta\mathbf{X}\|^2)\right]^{-1}\mathbf{B}^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})^\dagger\boldsymbol{e}^{\text{inc}} \\
&= \boldsymbol{y}^\dagger \mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\,\mathbf{B}\left[(\mathbf{C}^\dagger\mathbf{C})^{-1}+(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C}+\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\mathbf{B})(\mathbf{C}^\dagger\mathbf{C})^{-1}+\mathcal{O}(\|\delta\mathbf{X}\|^2)\right] \\
&\qquad \mathbf{B}^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})^\dagger\boldsymbol{e}^{\text{inc}}
\end{aligned}
$$

$$
\begin{aligned}
u = {} & \boldsymbol{y}^\dagger\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}} \;+\; \boldsymbol{y}^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\boldsymbol{e}^{\text{inc}} \\
& + \boldsymbol{y}^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\hat{\boldsymbol{e}} \;-\; \boldsymbol{y}^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\boldsymbol{e}^{\text{inc}} \\
& + \boldsymbol{y}^\dagger\mathbf{G}_o\delta\mathbf{X}\hat{\boldsymbol{e}} \;+\; \mathcal{O}(\|\delta\mathbf{X}\|^2)
\end{aligned}
$$

De même, on décompose l'expression de v pour faire ressortir les termes en $\delta\mathbf{X}$:

$$
\begin{aligned}
v &= \hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x})^\dagger(\mathbf{X}+\delta\mathbf{X})^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o(\mathbf{X}+dX)\hat{\boldsymbol{e}}(\boldsymbol{x}+\delta\boldsymbol{x}) \\
&= (\boldsymbol{e}^{\text{inc}})^\dagger\,(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})\,\mathbf{B}\left[\mathbf{B}^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})\mathbf{B}\right]^{-1}\mathbf{B}^\dagger\,(\mathbf{X}+\delta\mathbf{X})^\dagger\mathbf{G}_o^\dagger \\
&\quad \mathbf{G}_o(\mathbf{X}+\delta\mathbf{X})\,\mathbf{B}\left[\mathbf{B}^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})\mathbf{B}\right]^{-1}\mathbf{B}^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})^\dagger\boldsymbol{e}^{\text{inc}} \\
&= (\boldsymbol{e}^{\text{inc}})^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})\mathbf{B}\left[(\mathbf{C}^\dagger\mathbf{C})^{-1}+(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C}+\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\mathbf{B})(\mathbf{C}^\dagger\mathbf{C})^{-1}+\mathcal{O}(\|\delta\mathbf{X}\|^2)\right] \\
&\qquad \mathbf{B}^\dagger\,(\mathbf{X}+\delta\mathbf{X})^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\,(\mathbf{X}+\delta\mathbf{X})\,\mathbf{B} \\
&\quad \left[(\mathbf{C}^\dagger\mathbf{C})^{-1}+(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C}+\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\mathbf{B})(\mathbf{C}^\dagger\mathbf{C})^{-1}+\mathcal{O}(\|\delta\mathbf{X}\|^2)\right]\mathbf{B}^\dagger(\mathbf{L}_x-\mathbf{G}_c\delta\mathbf{X})^\dagger\boldsymbol{e}^{\text{inc}}
\end{aligned}
$$

$$
\begin{aligned}
v = {} & \hat{\boldsymbol{e}}^\dagger\,\mathbf{X}^\dagger\,\mathbf{G}_o^\dagger\,\mathbf{G}_o\,\mathbf{X}\,\hat{\boldsymbol{e}} \\
& - (\boldsymbol{e}^{\text{inc}})^\dagger\,\mathbf{G}_c\delta\mathbf{X}\,\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}} \\
& + \hat{\boldsymbol{e}}^\dagger\,(\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o\,\mathbf{X}\,\hat{\boldsymbol{e}} \\
& + (\boldsymbol{e}^{\text{inc}})^\dagger\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\,\mathbf{B})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o\,\mathbf{X}\,\hat{\boldsymbol{e}} \\
& + \hat{\boldsymbol{e}}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\,\mathbf{X}\,\hat{\boldsymbol{e}} \\
& + \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\delta\mathbf{X}\hat{\boldsymbol{e}} \\
& + \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\boldsymbol{e}^{\text{inc}} \\
& + \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X})\hat{\boldsymbol{e}} \\
& - \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\boldsymbol{e}^{\text{inc}} \;+\; \mathcal{O}(\|\delta\mathbf{X}\|^2)
\end{aligned}
$$

Revenons au calcul de $\widehat{\mathcal{F}}(\boldsymbol{x}+\delta\boldsymbol{x})$ et regroupons les termes indépendants de $\delta\mathbf{X}$, les termes en $\delta\mathbf{X}$ puis en $\delta\mathbf{X}^\dagger$:

$$\widehat{\mathcal{F}}(\boldsymbol{x}+\delta\boldsymbol{x}) = \boldsymbol{y}^\dagger\boldsymbol{y} \;-\; \boldsymbol{y}^\dagger\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}} \;-\; \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\boldsymbol{y} \;+\; \hat{\boldsymbol{e}}^\dagger\,\mathbf{X}^\dagger\,\mathbf{G}_o^\dagger\,\mathbf{G}_o\,\mathbf{X}\,\hat{\boldsymbol{e}}$$

$$\begin{aligned}
&- \boldsymbol{y}^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\hat{\boldsymbol{e}} \;-\; \boldsymbol{y}^\dagger\mathbf{G}_o\delta\mathbf{X}\hat{\boldsymbol{e}}\\
&- (\boldsymbol{e}^{\text{inc}})^\dagger\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\boldsymbol{y}\\
&+ (\boldsymbol{e}^{\text{inc}})^\dagger\mathbf{G}_c\delta\mathbf{X}\,\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\boldsymbol{y}\\
&- (\boldsymbol{e}^{\text{inc}})^\dagger\,\mathbf{G}_c\delta\mathbf{X}\,\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}\\
&+ (\boldsymbol{e}^{\text{inc}})^\dagger\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\,\mathbf{B})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o\,\mathbf{X}\,\hat{\boldsymbol{e}}\\
&+ \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\delta\mathbf{X}\hat{\boldsymbol{e}}\\
&+ \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X})\hat{\boldsymbol{e}}
\end{aligned}$$

$$\begin{aligned}
&- \boldsymbol{y}^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\boldsymbol{e}^{\text{inc}}\\
&+ \boldsymbol{y}^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\boldsymbol{e}^{\text{inc}}\\
&- \hat{\boldsymbol{e}}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\boldsymbol{y} - \hat{\boldsymbol{e}}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_o^\dagger\boldsymbol{y}\\
&+ \hat{\boldsymbol{e}}^\dagger\,(\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\mathbf{G}_o\,\mathbf{X}\,\hat{\boldsymbol{e}}\\
&+ \hat{\boldsymbol{e}}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\,\mathbf{X}\,\hat{\boldsymbol{e}}\\
&+ \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\boldsymbol{e}^{\text{inc}}\\
&- \hat{\boldsymbol{e}}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger\,\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\boldsymbol{e}^{\text{inc}}
\end{aligned}$$

$$+\; \mathcal{O}(\|\delta\mathbf{X}\|^2)$$

Au final, en combinant les termes en $\delta\mathbf{X}$ (ou de façon équivalente les termes en $\delta\mathbf{X}^\dagger$), on obtient la forme suivante :

$$
\begin{aligned}
\widehat{\mathcal{F}}(\boldsymbol{x}+\delta\boldsymbol{x}) = \quad & \widehat{\mathcal{F}}(\boldsymbol{x}) \\
& - \ (\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\hat{\boldsymbol{e}} \\
& - \ (\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})^\dagger\mathbf{G}_o\delta\mathbf{X}\hat{\boldsymbol{e}} \\
& - \ (\boldsymbol{e}^{\text{inc}})^\dagger\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\mathbf{G}_c\delta\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}) \\
& + \ (\boldsymbol{e}^{\text{inc}})^\dagger\mathbf{G}_c\delta\mathbf{X}\,\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}) \\
& - \hat{\boldsymbol{e}}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}) \\
& - \hat{\boldsymbol{e}}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_o^\dagger(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}) \\
& - (\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}(\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\mathbf{C})(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\boldsymbol{e}^{\text{inc}} \\
& + (\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\delta\mathbf{X}^\dagger\mathbf{G}_c^\dagger\boldsymbol{e}^{\text{inc}} \\
& + \ \mathcal{O}(\|\delta\mathbf{X}\|^2)
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\mathcal{F}}(\boldsymbol{x}+\delta\boldsymbol{x}) = \quad & \widehat{\mathcal{F}}(\boldsymbol{x}) \\
& - \ (\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})^\dagger\mathbf{G}_o\left(\mathbf{I}+\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\mathbf{G}_c\right)\text{diag}\{\hat{\boldsymbol{e}}\}\,\delta\boldsymbol{x} \\
& - \ (\boldsymbol{e}^{\text{inc}})^\dagger\left(\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger-\mathbf{I}\right)\mathbf{G}_c\,\text{diag}\left\{\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\mathbf{G}_o^\dagger(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})\right\}\,\delta\boldsymbol{x} \\
& - \delta\boldsymbol{x}^\dagger\,\text{diag}\{\hat{\boldsymbol{e}}^\dagger\}\left(\mathbf{I}+\mathbf{G}_c^\dagger\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\mathbf{X}^\dagger\right)\mathbf{G}_o^\dagger(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}}) \\
& - \delta\boldsymbol{x}^\dagger\,\text{diag}\left\{(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\right\}\mathbf{G}_c^\dagger\left(\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger-\mathbf{I}\right)\boldsymbol{e}^{\text{inc}} \\
& + \ \mathcal{O}(\|\delta\mathbf{X}\|^2)
\end{aligned}
$$

$$
\widehat{\mathcal{F}}(\boldsymbol{x}+\delta\boldsymbol{x}) = \quad \widehat{\mathcal{F}}(\boldsymbol{x}) \ + \ \frac{1}{2}(\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}})^\dagger\,\delta\boldsymbol{x} \ + \ \frac{1}{2}\delta\boldsymbol{x}^\dagger\,\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}} \ + \ \mathcal{O}(\|\delta\mathbf{X}\|^2) \tag{C.3}
$$

On en déduit l'expression du gradient qui correspond au critère $\widehat{\mathcal{F}}$:

$$\begin{aligned}\nabla_{\boldsymbol{x}}\widehat{\mathcal{F}} = &-2\left[\mathbf{G}_o\left(\mathbf{I}+\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger\mathbf{G}_c\right)\mathrm{diag}\{\hat{\boldsymbol{e}}\}\right]^\dagger(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})\\ &-2\,\mathrm{diag}\left\{(\boldsymbol{y}-\mathbf{G}_o\mathbf{X}\hat{\boldsymbol{e}})^\dagger\mathbf{G}_o\mathbf{X}\mathbf{B}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{B}^\dagger\right\}\mathbf{G}_c^\dagger\left(\mathbf{C}(\mathbf{C}^\dagger\mathbf{C})^{-1}\mathbf{C}^\dagger-\mathbf{I}\right)\boldsymbol{e}^{\mathrm{inc}}\end{aligned} \tag{C.4}$$