Approche parcimonieuse pour l'imagerie 3D haute résolution de surface équivalente radar.

Thomas Benoudiba-Campanini

▶ **To cite this version:**

Thomas Benoudiba-Campanini. Approche parcimonieuse pour l'imagerie 3D haute résolution de surface équivalente radar.. Traitement du signal et de l'image [eess.SP]. Université de Bordeaux, 2018. Français. NNT : 2018BORD0110 . tel-01944454

HAL Id: tel-01944454
https://theses.hal.science/tel-01944454

Submitted on 4 Dec 2018

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

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

université de BORDEAUX

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

École doctorale n° 209 : Sciences Physiques et de l'Ingénieur

Spécialité doctorale
« Automatique, Productique, Signal et Image, Ingénierie Cognitique »

présentée et soutenue publiquement par

Thomas BENOUDIBA--CAMPANINI

le 13 juillet 2018

Approche parcimonieuse pour l'imagerie 3D haute résolution de surface équivalente radar.

Directeur de thèse : **Jean-François GIOVANNELLI**
Co-encadrant de thèse : **Pierre MINVIELLE**

Jury

LITMAN	Amélie	Rapporteur
TUPIN	Florence	Rapporteur
CARFANTAN	Hervé	Examinateur
MALLEJAC	Nicolas	Examinateur

Titre

Approche parcimonieuse pour l'imagerie 3D haute résolution de surface équivalente radar.

Résumé

La SER (Surface Équivalente Radar) est une grandeur caractérisant le pouvoir rétro-diffuseur d'une cible soumise à un champ électromagnétique. Dans de nombreuses applications, il est capital d'analyser et de contrôler la SER. L'imagerie 3D est l'outil adapté pour localiser et caractériser en trois dimensions les principaux contributeurs à la SER. Cependant, ce traitement est un problème de synthèse de Fourier qui n'est pas inversible car il y a plus d'inconnues que de données. Les méthodes conventionnelles telles que le Polar Format Algorithm, consistant en un reformatage des données avec complétion de zéro suivi d'une transformée de Fourier inverse rapide, fournissent des résultats de qualité limitée.

Dans ce travail, nous proposons une nouvelle méthode haute résolution. Elle est dénommée SPRITE (pour SParse Radar Imaging TEchnique) et permet d'accroître considérablement la qualité des cartes de rétro-diffusion estimées. Elle repose sur une régularisation du problème consistant en la prise en compte d'informations *a priori* de parcimonie et d'une information de support. La solution est alors définie comme le minimiseur d'un critère pénalisé et contraint. L'optimisation est assurée par l'algorithme primal-dual ADMM (Alternating Direction Method of Multiplier) dont une adaptation aux spécificités du problème mène à des calculs efficaces à l'aide de transformées de Fourier rapides.

Finalement, la méthode est évaluée sur des données synthétiques et réelles. Comparativement à la méthode conventionnelle, la résolution est drastiquement accrue. Les images 3D produites sont alors un outil particulièrement adapté à l'analyse et au contrôle de SER.

Mots clés

SPRITE, Imagerie 3D HR, SER, Problème Inverse, Régularisation Parcimonieuse, ADMM

Title

Sparse approach for high resolution 3D radar cross section imaging.

Abstract

The RCS (Radar Cross Section) is a quantity which characterizes the scattering power of a target exposed to an electromagnetic field. Its analysis and control are important in many applications. 3D imaging is a suitable tool to accurately locate and characterize in 3D the main contributors to the RCS. However, this is a non-invertible Fourier synthesis problem because the number of unknowns is larger than the number of data. Conventional methods such as the Polar Format Algorithm, which consists of data reformatting including zero-padding followed by a fast inverse Fourier transform, provide results of limited quality.

In this work, we propose a new high resolution method, named SPRITE (for SParse Radar Imaging TEchnique), which considerably increases the quality of the estimated RCS maps. It is based on a regularization scheme that accounts for information of sparsity and support. The solution is then defined as the minimizer of a penalized and constrained criterion. Optimization is ensured by an appropriate adaptation of the ADMM (Alternating Direction Method of Multiplier) algorithm that is able to quickly perform calculations using fast Fourier transforms.

Finally, the method is evaluated on both simulated and real data. Compared to the conventional method, the resolution is significantly increased and the images can support a better RCS analysis and control.

Keywords

SPRITE, HR 3D Imaging, RCS, Inverse Problem, Sparse Regularization, ADMM

TABLE DES MATIÈRES

Table des matières **5**

Introduction **13**

1 L'imagerie radar pour l'analyse et le contrôle de SER **17**

1.1 Le besoin de caractériser la SER 18

1.2 L'outil principal : l'imagerie radar 31

1.3 Des solutions au problème d'imagerie 50

2 Les approches parcimonieuses en inversion **61**

2.1 La notion de parcimonie 62

2.2 Des algorithmes de minimisation de critères non différentiables contraints . . 65

3 SPRITE : une méthode haute résolution pour l'imagerie radar 3D **87**

3.1 Les connaissances *a priori* sur l'objet de rétro-diffusion 88

3.2 La construction d'un critère pénalisé et contraint 96

3.3 La minimisation du critère à l'aide de l'ADMM 101

3.4 Des compléments sur la méthode SPRITE 107

4 L'évaluation de la méthode SPRITE **111**

4.1 Une étude préliminaire de la méthode SPRITE 112

4.2 L'analyse de SER 125

4.3 Le contrôle 3D de SER 136

Conclusions générales et perspectives **141**

A Processus de calibration des mesures de SER **145**

B Inférence bayésienne **149**

C Figures additionnelles **153**

D Démonstrations des formules proposées dans le document **159**

D.1 Preuves Chapitre 1 159

D.2 Preuves Chapitre 2 165

D.3 Preuves Chapitre 3 171

E Quelques définitions & théorèmes d'analyse convexe **183**

Références **189**

LISTE DES SYMBOLES

s Scalaire s (lettre minuscule)

$\boldsymbol{v}$ Vecteur $\boldsymbol{v}$ (lettre en gras)

$\|\boldsymbol{v}\|_1$ Norme ℓ_1 de $\boldsymbol{v}$, $\|\boldsymbol{v}\|_1 = \sum_n |v_n|$

$\|\boldsymbol{v}\|_2$ Norme euclidienne de $\boldsymbol{v}$, $\|\boldsymbol{v}\|_2 = \sqrt{\sum_n |v_n|^2}$

$\boldsymbol{M}$ Matrice $\boldsymbol{M}$ (lettre capitale en gras)

$\boldsymbol{M}^t$ Transposée de $\boldsymbol{M}$

$\boldsymbol{M}^*$ Conjuguée de $\boldsymbol{M}$

$\boldsymbol{M}^\dagger$ Transposée-conjuguée de $\boldsymbol{M}$, $\boldsymbol{M}^\dagger = \boldsymbol{M}^{t*}$

$\boldsymbol{M}^+$ Pseudo-Inverse de $\boldsymbol{M}$, $\boldsymbol{M}^+ = (\boldsymbol{M}^\dagger \boldsymbol{M})^{-1} \boldsymbol{M}^\dagger$

$\boldsymbol{W}_{M,N}$ Matrice de Transformée de Fourier Discrète rectangulaire de taille $M \times N$

$\boldsymbol{F}_N$ Matrice de Transformée de Fourier Discrète carrée de taille $N \times N$ normalisée, $\boldsymbol{F}_N = \boldsymbol{W}_{N,N}/\sqrt{N}$ et $\boldsymbol{F}^\dagger = \boldsymbol{F}^{-1}$

$\boldsymbol{F}_{3D}$ Matrice de Transformée de Fourier Discrète 3D carrée de taille $N_x N_y N_z \times N_x N_y N_z$ normalisée, $\boldsymbol{F}_{\text{3D}} = \boldsymbol{F}_{N_x} \otimes \boldsymbol{F}_{N_y} \otimes \boldsymbol{F}_{N_z}$ et $\boldsymbol{F}_{\text{3D}}^\dagger = \boldsymbol{F}_{\text{3D}}^{-1}$

$\mathrm{Im}\,\{\boldsymbol{x}\}$ Partie imaginaire de $\boldsymbol{x}$

$\otimes$ Produit de Kronecker

$\preccurlyeq$ Inégalité élément par élément

$\triangleq$ Égalité par définition

$\mathrm{dom}\ f$ Domaine effectif de la fonction f

∇f Gradient de la fonction f

∂f Sous-différentiel de la fonction f

$\mathrm{Conv}(\mathbb{E})$ Ensemble des fonctions définies sur $\mathbb{E}$ à valeurs dans $\mathbb{R} \cup \{-\infty, +\infty\}$ convexes et propres

$\mathrm{C\overline{onv}}(\mathbb{E})$ Partie de $\mathrm{Conv}(\mathbb{E})$ formée des fonctions fermées

prox_f Opérateur proximal de la fonction f

$\mathcal{P}_{\mathcal{C}}$ Opérateur de projection euclidienne sur l'ensemble $\mathcal{C}$

S_{α} Opérateur de seuillage doux de paramètre α

$\mathcal{I}_{\mathcal{C}}$ Fonction indicatrice de l'ensemble $\mathcal{C}$

N_x(resp. N_y, N_z) Nombre de points de la grille cartésienne du domaine spatial selon x(resp. y, z)

M_x(resp. M_y, M_z) Nombre de points de la grille cartésienne du k-space selon k_x(resp. k_y, k_z)

c Vitesse de la lumière dans le vide (en $\mathrm{m} \cdot \mathrm{s}^{-1}$), $c \approx 3 \cdot 10^8$ $\mathrm{m} \cdot \mathrm{s}^{-1}$

λ Longueur d'onde $\lambda = c/f$ (en m)

f Fréquence (en Hz)

$\boldsymbol{k}$ Vecteur d'onde

$\|\boldsymbol{k}\|$ Nombre d'onde, $\|\boldsymbol{k}\| = 2\pi f/c$

φ Angle de roulis

θ Angle de gisement

$\boldsymbol{\sigma}$ Vecteur des données

$\breve{\boldsymbol{\sigma}}$ Vecteur des données reformatées

$\boldsymbol{a}_{\mathrm{3D}}$ Carte de rétro-diffusion 3D de dimension $N_x \times N_y \times N_z$

$\boldsymbol{a}$ Carte 3D vectorisée de dimension $N \times 1$, $\boldsymbol{a} = \mathrm{vec}\, \boldsymbol{a}_{\mathrm{3D}}$

$\boldsymbol{a}^v$ Carte synthétique « vraie »

$\hat{\boldsymbol{a}}_{\mathcal{M}}$ Carte estimée par la méthode $\mathcal{M}$

LISTE DES ACRONYMES

ADL Ascension Duale Lagrangienne.
ADMM Alternating Direction Method of Mutipliers.
ARV Analyseur de Réseau Vectoriel.

CAO Conception Assistée par Ordinateur.
CEA Commissariat à l'Énergie Atomique et aux Énergies Alternatives.
CEP Conducteur Électrique Parfait.
CESTA Centre d'Études Scientifiques et Techniques d'Aquitaine.

EM Électromagnétique.
ESPRIT Estimation of Signal Parameters *via* Rotational Invariance Technique.

FBS Forward Backward Splitting.
FEP Fonction d'Étalement du Point.
FGP Fast Gradient Projection.
FISTA Fast Iterative Soft Thresholding Algorithm.

HR Haute Résolution.

ISAE Institut Supérieur de l'Aéronautique et de l'Espace.
ISTA Iterative Soft Thresholding Algorithm.

MFISTA Monotone Fast Iterative Soft Thresholding Algorithm.
ML Multiplicateurs de Lagrange.
MUSIC Multiple Signal Classifier.

ONERA Office National d'Études et de Recherches Aérospatiales.

PFA Polar Format Algorithm.

RADAR RAdio Detection And Ranging.
RAM Radar Absorbing Materials *i.e.* matériaux absorbant les ondes Électromagnétiques.
REQM Racine carrée de l'Erreur Quadratique Moyenne.
RPF Rétro-Projection Filtrée.

RSB Rapport Signal sur Bruit.

SAR Synthetic Aperture Radar (Radar à Synthèse d'Ouverture).

SASB Small Angle Small Bandwidth (Excursions fréquentielle et angulaires limitées).

SER Surface Équivalente Radar.

SPRITE SParse Radar Imaging TEchnique.

TFD Transformée de Fourier Discrète.

TFDI Transformée de Fourier Discrète Inverse.

TFR Transformée de Fourier Rapide.

TFRI Transformée de Fourier Rapide Inverse.

TOS Taux d'Ondes Stationnaires.

REMERCIEMENTS

En premier lieu, je tiens à remercier vivement mes directeurs de thèse, Jean – François GIOVANNELLI et Pierre MINVIELLE qui m'ont épaulé durant ces trois années. Vos compétences, votre enthousiasme et votre bonne humeur ont largement contribué à la réussite de cette thèse. Merci Jean-François pour tes précieux conseils, tant scientifiques qu'humains, qui m'ont permis de comprendre que bien souvent le cheminement est plus important que l'objectif. Merci également de m'avoir donné l'opportunité d'enseigner. Pierre, je te remercie pour ta disponibilité et ton infaillible soutien durant et après ma thèse. Je pense avoir rarement rencontré des personnes aussi bienveillantes que toi. Si nous sommes aujourd'hui collègues je sais que je te le dois en grande partie et je ne l'oublierai jamais ! J'espère avoir rapidement l'occasion de travailler à nouveau avec vous !

Je remercie les membres du jury pour l'intérêt qu'ils ont porté à mes travaux de thèse : Florence TUPIN et Amélie LITMAN qui ont accepté de rapporter mon manuscrit, Nicolas MALLEJAC qui a évalué mon travail et Hervé CARFANTAN qui a présidé le jury et m'a apporté des remarques très constructives et pertinentes.

J'adresse également mes remerciements à mes anciens chefs du DSGA et du SSPP du CEA-CESTA ainsi qu'à mes collègues du LOR et de l'IMS. Ce fut un honneur et un plaisir d'évoluer parmi des personnes aussi compétentes et sympathiques que vous. Un énorme merci à Anne-Pascale et Martine pour leur travail quotidien qui facilite le nôtre !

Je remercie aussi mes nouveaux chefs du DAIA et du SAFA ainsi que mes collègues du LCF pour leurs conseils avisés en fin de thèse et la confiance qu'ils m'ont témoignée en m'accueillant chaleureusement à leur côté !

J'ai également une pensée pour tous les post-doctorants, doctorants, stagiaires et élèves que j'ai croisés au cours de ma thèse à l'IMS et au CEA. J'adresse une mention spéciale à ma co-bureau Alexia qui m'a supporté, ou plutôt que j'ai supportée, pendant presque trois ans (merci aussi pour cette phrase tirée *in extenso* de tes remerciements de thèse !). Je te remercie également, ainsi que Greg, pour votre précieuse aide dans la confection de mon pot de thèse !

Je remercie mes amis (en particulier mon fidèle Coco !) qui, ne comprenant pas un traitre mot de ce que je fais me trouvent très intelligent... si seulement ils savaient !

En bon corse, je termine ces remerciements avec ma famille ! Merci pour le soutien indéfectible et l'amour que vous me témoignez quotidiennement. Merci à ma belle sœur Colyne pour son aide et son dynamisme, en particulier lors de ma soutenance. Merci à mes grands mères pour leur tendresse et leur dévouement de chaque jour. Merci à mon père et à ma mère pour qui je n'ai pas de mots assez forts pour exprimer toute ma gratitude et mon amour. Vous êtes mes exemples et mes piliers. Merci de m'avoir encouragé et soutenu depuis toujours et

de m'avoir donné des racines et des ailes. Merci à mon petit Lu'. Tu marches dans mes pas en préparant ton doctorat en médecine. Je t'aime et je suis infiniment fier de t'avoir comme petit frère. Enfin, je conclurai en ayant une pensée pour mes grands pères. Ils ont contribué à faire de moi la personne que je suis. Je sais à quel point ils étaient fiers de mon parcours. Réussir a toujours été pour moi un moyen de leur exprimer en retour toute l'admiration que j'avais pour eux. Je leur dédie cette thèse.

A voi Babbó e Dade, vi tengu cari per sempre... Cusì sìa...

INTRODUCTION

Contexte

Une cible illuminée par un radar est soumise à un champ EM (Électromagnétique). Elle ré-émet une partie du rayonnement intercepté dans toutes les directions. Le pouvoir rétro-diffuseur de la cible est caractérisé par une grandeur scalaire appelée SER (Surface Équivalente Radar). Son analyse et son contrôle sont primordiaux dans de nombreux domaines tels que la caractérisation, la détection ou encore l'identification d'objets.

L'imagerie 3D de SER est un outil particulièrement adapté pour l'analyse et le contrôle de SER. Il s'agit d'estimer des cartes 3D de rétro-diffusion dans le but d'identifier les zones qui contribuent à la réflectivité globale d'une cible d'intérêt. Pour cela, le CEA-CESTA a développé un instrument de mesure appelé Arche 3D, capable d'acquérir rapidement des données sous différents angles et fréquences. L'Arche 3D mesure des coefficients de rétro-diffusion sur une grille irrégulière et incomplète du k-space (*i.e.* le domaine des fréquences spatiales) de l'objet d'intérêt. Ces informations sont toutefois entachées d'un bruit de mesure. L'imagerie de SER consiste donc en un problème de synthèse de Fourier qui entre dans la classe des *problèmes inverses* [Idier, 2001; Tarantola, 2005]. Néanmoins, il présente la particularité d'être non inversible car les données dans le k-space sont incomplètes. Le problème est sous-déterminé : le nombre de voxels de la carte à estimer (*i.e. le nombre d'inconnues*) est supérieur au nombre de données.

Les méthodes conventionnelles d'imagerie radar telles que le Polar Format Algorithm 3D ou la rétro-projection filtrée 3D permettent de former des cartes 3D. Elles reposent sur un reformatage des données avec bourrage de zéro pour les fréquences spatiales non observées, suivi d'une TFRI (Transformée de Fourier Rapide Inverse) 3D. Elles produisent néanmoins des images dont la résolution est limitée.

Afin d'améliorer la qualité des cartes, il est possible d'envisager l'imagerie sous l'*approche problème inverse*. A ce titre, une large partie du travail présenté dans ce document concerne l'étude de méthodes HR (Haute Résolution). Elles reposent sur une *régularisation*, c'est-à-dire la prise en compte d'informations *a priori*. Cette régularisation peut s'exprimer à travers l'expression d'un critère. L'optimisation est ensuite assurée par des algorithmes de minimisation.

Dans cet esprit, des méthodes HR reposant sur une régularisation avec un *a priori* favorisant la parcimonie ont été proposées ces dernières années pour l'imagerie radar 2D [Cetin, 2001; Zhu, 2013] et 3D [Austin *et al.*, 2011; Xiao-Chun, 2010]. Elles permettent d'améliorer sensiblement la résolution des cartes de rétro-diffusion. Cependant, les images produites sont constituées d'un faible nombre de points brillants. Ceci n'est pas complètement cohérent des interactions onde-cible.

Contributions

La contribution majeure de cette thèse concerne le développement d'une nouvelle méthode HR pour l'imagerie 3D de SER. Dénommée SPRITE (SParse Radar Imaging TEchnique), elle permet d'accroitre considérablement la qualité des cartes 3D de rétro-diffusion estimées.

L'approche repose sur une régularisation du problème de synthèse de Fourier consistant en la prise en compte de plusieurs informations *a priori* cohérentes des interactions ondecible :

- la projection de la carte sur la direction du vecteur d'onde est parcimonieuse.
- la carte peut être décomposée en un faible nombre de facettes brillantes connexes.
- le coefficient de rétro-diffusion complexe est constant sur chaque facette.
- l'énergie de la carte est relativement faible.
- l'extension spatiale de la carte est limitée.

Ces *a priori* sont intrinsèquement liés à la notion de parcimonie et d'une information sur l'extension EM de la cible. La solution est alors définie comme le minimiseur d'un critère composite contraint. Ce critère est composé d'un terme d'attache aux données et de plusieurs pénalités à norme ℓ_1 et ℓ_2 qui sont cohérentes des *a priori* considérés. L'optimisation est assurée par une adaptation appropriée de l'algorithme primal-dual ADMM (Alternating Direction Method of Mutipliers) pour mener efficacement les calculs à l'aide de transformées de Fourier rapides.

Une autre contribution consiste en l'évaluation de cette méthode sur des données synthétiques et réelles. Comparativement à la méthode conventionnelle, nous montrons que la résolution est drastiquement accrue. Ceci produit une amélioration notable de l'analyse et du contrôle de SER.

Organisation du document

Le premier chapitre est consacré à la définition des notions d'analyse et de contrôle de SER. Le problème d'imagerie de SER et les méthodes conventionnelles de formation de cartes 3D sont exposés. La notion d'approche problème inverse est introduite et quelques méthodes HR reposant sur des *a priori* parcimonieux sont présentées.

Le second chapitre porte sur la notion de régularisation parcimonieuse non différentiable. Différents algorithmes d'optimisation de critères non différentiables sont recensés.

Le troisième chapitre détaille la contribution majeure de cette thèse : le développement d'une nouvelle méthode HR pour l'imagerie 3D de SER appelée SPRITE.

Le quatrième chapitre concerne l'évaluation de la méthode SPRITE sur des données synthétiques et réelles.

Publications et présentations

- **Articles en préparation**
  BENOUDIBA--CAMPANINI T., MINVIELLE P., GIOVANNELLI J.-F. ET MASSALOUX P. (2018), *SPRITE : A New Regularization Method for 3D High Resolution Radar Cross Section Imaging*, en préparation, IEEE AES
- **Conférences internationales avec actes et comité de lecture**
  BENOUDIBA--CAMPANINI T., MINVIELLE P., GIOVANNELLI J.-F. ET MASSALOUX P. (2017), *A New Regularization Method for Radar Cross Section Imaging* [1], European Conference on Antennas and Propagation 2017
- **Conférences nationales avec actes et comité de lecture**
  BENOUDIBA--CAMPANINI T., MINVIELLE P. ET GIOVANNELLI J.-F. (2017), *Régularisation et parcimonie pour l'imagerie radar haute résolution 1D* [2], 20-ièmes Journées Nationales Micro-Ondes 2017
- **Journées et séminaires nationaux sans actes et sans comité de lecture**
  BENOUDIBA--CAMPANINI T.,GIOVANNELLI J.-F., MINVIELLE P. (2015), *Méthode Semi-Quadratique et Multiplicateurs de Lagrange pour l'Imagerie RADAR*, Journées d'Imagerie Optique Non Conventionnelle 2015
  BENOUDIBA--CAMPANINI T., GIOVANNELLI J.-F., MINVIELLE P. (2015), *Régularisation par Pénalisation et Contrainte Convexe pour la Restauration d'Images*, Séminaire Groupe Signal-Image du Laboratoire de l'Intégration du Matériau au Système de Bordeaux

1. sélectionné pour les Best Paper Awards.
2. sélectionné dans les meilleurs papiers de conférence.

CHAPITRE 1

L'IMAGERIE RADAR POUR L'ANALYSE ET LE CONTRÔLE DE SER

Dans la première section, nous présentons la notion de SER. Nous montrons à travers divers exemples pourquoi il est important d'analyser et de contrôler la SER. Dans la seconde section, nous détaillons le processus de mesure de l'hologramme de SER, nous introduisons la notion d'objet et de carte de rétro-diffusion et nous formulons le modèle d'observation liant les données à l'objet d'intérêt. Nous présentons également le problème d'imagerie de SER. Enfin, nous présentons les méthodes conventionnelles et nous soulignons leurs limites en terme de résolution. Nous expliquons alors que l'approche problème inverse fournit un cadre rigoureux pour développer des méthodes HR. Nous listons enfin quelques méthodes reposant sur la notion de parcimonie, qui permettent d'accroître la résolution des cartes estimées.

Sommaire

1.1 Le besoin de caractériser la SER 18
1.1.1 La notion de SER 18
1.1.2 La phénoménologie des interactions onde-objet 21
1.1.3 L'analyse et le contrôle de SER 25
1.2 L'outil principal : l'imagerie radar 31
1.2.1 La mesure de l'hologramme de SER 31
1.2.2 L'objet de rétro-diffusion 35
1.2.3 Le modèle direct 36
1.2.4 L'imagerie de SER 48
1.3 Des solutions au problème d'imagerie 50
1.3.1 Les méthodes classiques et leur limites 50
1.3.2 L'approche problème inverse 53
1.3.3 La parcimonie en imagerie radar 57

1.1 Le besoin de caractériser la SER

1.1.1 La notion de SER

Une cible éclairée par un radar est soumise à un champ EM. Elle ré-émet une partie du rayonnement intercepté dans toutes les directions de l'espace. La SER, désignée par la lettre s, est une grandeur scalaire homogène à une surface, qui caractérise le pouvoir rétro-diffuseur de la cible. Elle correspond à la surface qu'aurait la cible si elle rétro-diffusait de façon isotrope les ondes EM pour renvoyer un écho de même puissance que celui effectivement reçu par le récepteur.

Le dictionnaire IEEE [Barton *et al.*, 1998] définit formellement la SER comme 4π fois le rapport entre la puissance rétro-diffusée par unité d'angle solide et la densité de puissance incidente. En exprimant la puissance et la densité de puissance en fonction du champ électrique, [Knott *et al.*, 2004] montre qu'en champ lointain, la SER peut s'exprimer comme suit :

$$s = 4\pi R^2 \frac{|E_r|^2}{|E_i|^2} \tag{1.1}$$

où E_i est l'amplitude du champ électrique incident au niveau de la cible, E_r l'amplitude du champ rétro-diffusé mesuré au niveau de l'antenne de réception et R la distance entre la cible et le radar.

L'unité de mesure de la SER est le mètre carré (m^2) mais, compte tenu de sa grande dynamique, les radaristes l'expriment généralement en décibel (dB) :

$$s_{\mathrm{dBm^2}} = 10 \log_{10}(s) \tag{1.2}$$

La SER d'une cible passive dépend de nombreux paramètres :

- la configuration du radar (monostatique ou bistatique) et la position de l'émetteur et du récepteur relativement à la cible,
- les paramètres de fonctionnement du radar : la fréquence de l'onde émise (plus précisément le rapport entre la taille de la cible et la longueur d'onde), la polarisation de l'émetteur et du récepteur,
- les paramètres de la cible : volume, géométrie, matériaux constitutifs,
- l'orientation angulaire de la cible (gisement, site, roulis) par rapport à l'émetteur et au récepteur, aussi appelée *attitude*.

Influence de la fréquence

La SER dépend de la fréquence de l'onde incidente. Son évolution peut-être séparée en trois domaines principaux. Ces derniers sont définis par le rapport entre la dimension principale D de la cible et la longueur d'onde λ de l'onde incidente.

- La *région de Rayleigh* ($D \ll \lambda$) : dans cette région la SER varie comme $1/\lambda^4$ et la cible est dite *électriquement petite*.

- La *région de résonance* ou *région de Mie* ($D \approx \lambda$) : dans cette région la SER oscille selon la longueur d'onde.
- La *région optique* ou *région des hautes fréquences* ($D \gg \lambda$) : dans cette région la SER dépend peu de λ et la cible est dite *électriquement grande*.

La Figure 1.1 présente la SER d'une sphère de rayon r et illustre ces trois domaines fréquentiels.

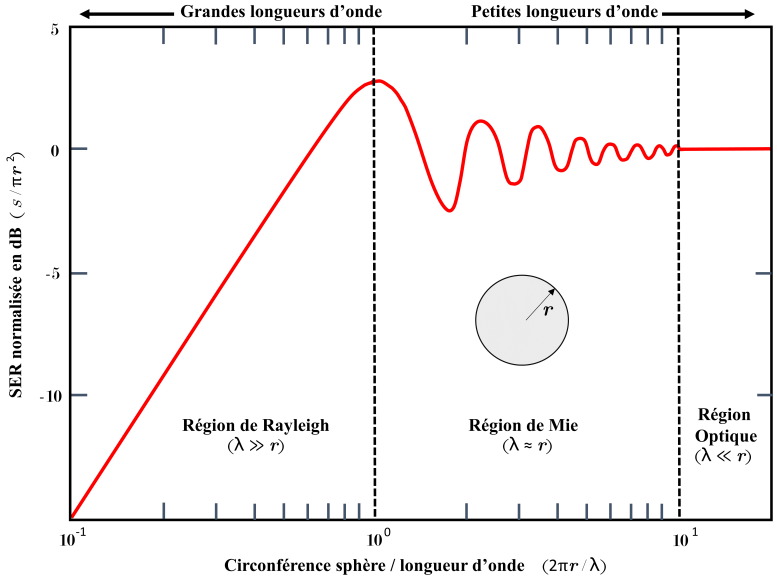


FIGURE 1.1 – Évolution de la SER d'une sphère de rayon r [Knott *et al.*, 2004].

Zones de rayonnement

L'expression (1.1) est valide pour des ondes planes en champ lointain. Nous précisons ici ce que cela signifie. Les *surfaces équiphases* (aussi appelées *front d'onde*) d'une onde se propageant dans un milieu isotrope et homogène sont des sphères. On parle alors d'*onde sphérique*. En *champ lointain*, lorsque la distance entre la source et la cible est suffisamment grande, le front d'onde peut-être considéré comme localement plan et l'onde est assimilée à une *onde plane*. Dans ce cas, les champs électrique $\boldsymbol{E}$ et magnétique $\boldsymbol{H}$ sont en phase, orthogonaux entre eux et à la direction de propagation.

[Balanis, 2016] montre que les champs entourant une cible rayonnante peuvent être partitionnés en trois zones distinctes, illustrées par la Figure 1.2.

- La *zone de champ proche réactif* ou *zone d'induction* ($R < R_1$) est l'espace entourant immédiatement la cible. Dans cette région, les composantes non rayonnantes d'induction prédominent.
- La *zone de champ proche radiatif* ou *zone de Fresnel* ($R_1 < R < R_2$)
- La *zone de champ lointain* ou *zone de Fraunhofer* ($R_2 < R$) est une zone suffisamment éloignée de la cible pour que les ondes soient considérées planes.

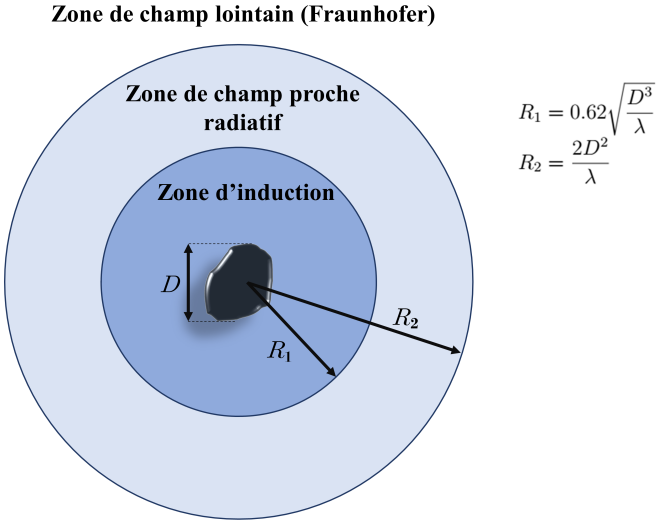


FIGURE 1.2 – Zones de rayonnement.

Influence de polarisation

Une onde EM plane est également caractérisée par sa polarisation. Cette dernière définit l'évolution de la direction des champs électrique et magnétique au cours d'une période de l'onde. En particulier, une onde est dite à *polarisation rectiligne* si ses champs restent parallèles à deux directions fixes perpendiculaires. Toute onde EM peut se décomposer comme la somme de deux ondes à polarisation rectiligne. La plupart des radars sont conçus de façon à transmettre des hyperfréquences avec une polarisation horizontale (H) ou verticale (V) comme illustré Figure 1.3. De même, le récepteur reçoit l'onde rétro-diffusée avec une polarisation H ou V. Cependant, certains radars peuvent recevoir les deux. Quatre combinaisons de polarisations en transmission et en réception sont alors possibles :

- HH : polarisation horizontale en transmission et en réception,
- VV : polarisation verticale en transmission et en réception,
- HV : polarisation horizontale en transmission et verticale en réception,
- VH : polarisation verticale en transmission et horizontale en réception.

Les combinaisons HH et VV sont appelées *polarisations parallèles* et les combinaisons HV et VH *polarisations croisées*.

Une description plus complète de l'interaction entre une onde incidente et la cible est donnée par la matrice complexe $\boldsymbol{S}$, appelée *matrice de Sinclair* en radar ou matrice de *Jones* en optique :

$$\boldsymbol{E}_r = \boldsymbol{S}\boldsymbol{E}_i \tag{1.3}$$

Comme le champ $\boldsymbol{E}$ n'a pas de composante selon le vecteur de propagation de l'onde $\boldsymbol{k}$, il se décompose en deux directions indépendantes, ou polarisations :

$$\begin{bmatrix} \boldsymbol{E}_r^H \\ \boldsymbol{E}_r^V \end{bmatrix} = \begin{bmatrix} \sigma^{HH} & \sigma^{HV} \\ \sigma^{VH} & \sigma^{VV} \end{bmatrix} \begin{bmatrix} \boldsymbol{E}_i^H \\ \boldsymbol{E}_i^V \end{bmatrix} \tag{1.4}$$

où chaque σ est le coefficient de rétro-diffusion complexe ou SER complexe (avec une information d'amplitude et de phase) [Fuhs, 1983; Knott *et al.*, 2004] qui s'écrit, pour une polarisation en émission et en réception donnée :

$$\sigma = 2\sqrt{\pi}R\frac{E_r}{E_i} \tag{1.5}$$

Notons que la SER définie par l'expression (1.1) est le module carré de σ :

$$s = |\sigma|^2 \tag{1.6}$$

Dans la suite du document, σ sera appelé SER quand il n'y a pas d'ambiguïté avec s.

REMARQUE

Dans une configuration monostatique, $\sigma^{HV} = \sigma^{VH}$.

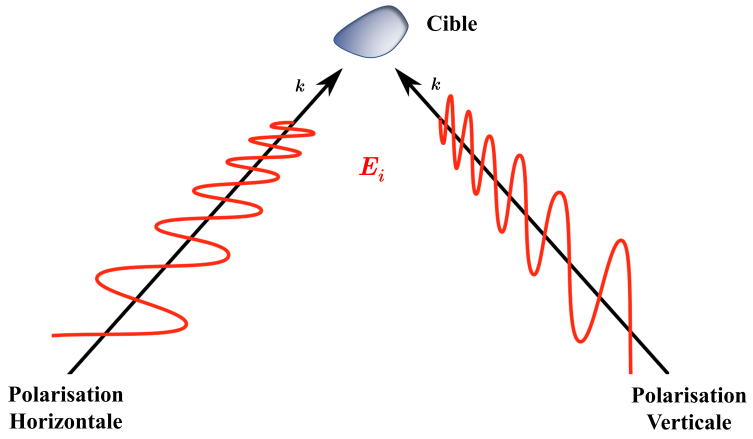


FIGURE 1.3 – Illustration polarisation H et V pour le champ électrique incident

1.1.2 La phénoménologie des interactions onde-objet

Dans cette section, nous répertorions, dans un contexte hautes fréquences, les différents mécanismes de rétro-diffusion qui peuvent intervenir lorsqu'une cible est éclairée par une onde EM. Ces derniers sont classés dans l'ordre décroissant de leur contribution à la SER. Cette section s'inspire de l'ouvrage de [Knott *et al.*, 2004, chapitres 6 et 14].

Les spéculaires

Une rétro-diffusion spéculaire est analogue à une réflexion sur un miroir ou sur une surface polie en optique. Dans un contexte monostatique et en champ lointain, la *direction spéculaire* est opposée au vecteur d'onde [Knott *et al.*, 2004] et s'écrit :

$$\hat{\boldsymbol{d}} = -\hat{\boldsymbol{k}} \tag{1.7}$$

où $\boldsymbol{k}$ est le *vecteur d'onde* de module (appelé *nombre d'onde*) $\|\boldsymbol{k}\| = 2\pi f/c$ [1] et $\hat{\boldsymbol{k}} = \boldsymbol{k}/\|\boldsymbol{k}\|$ indique la direction de propagation de l'onde incidente. Les *facettes spéculaires* sont alors

1. c étant la vitesse de la lumière dans le vide

définies comme les surfaces de la cible dont la normale extérieure $\hat{\boldsymbol{n}}$ est colinéaire au vecteur d'onde [Nouvel *et al.*, 2004].

Ces facettes spéculaires incluent des surfaces planes orthogonales au vecteur d'onde ainsi que des surfaces simplement ou doublement courbées au sens de la courbure de Gauss. Ainsi, comme illustré sur la Figure 1.4, si la surface est doublement courbée, la facette spéculaire peut s'appréhender comme un point brillant. De plus, l'amplitude de l'onde rétro-diffusée et sa persistance selon l'angle de vue dépendent également de la courbure de Gauss de la surface [Mendenhall, 1992]. La Figure 1.5 illustre cela pour une cible (une maquette de bateau métallique) constituée de surfaces planes et de surfaces courbées. L'énergie rétro-diffusée par les surfaces planes (1) dépend de l'angle de vue : elle est maximale, et donc les surfaces planes contribuent fortement à la SER, quand la direction spéculaire est confondue avec la normale à la surface. Cependant, l'écho s'atténue rapidement quand l'angle d'observation s'en écarte. La facette spéculaire apparait alors comme une *facette brillante* (1). Pour les surfaces simplement courbées, la facette spéculaire se résume à un *segment brillant* le long de la génératrice de la surface comme l'illustre la cheminée cylindrique du bateau (2). Enfin, la facette spéculaire se réduit à un *point brillant* pour une surface doublement courbée comme illustré par le radôme sphérique du bateau (3). Dans une configuration monostatique, l'énergie rétro-diffusée varie très peu selon l'angle de vue pour des sphéroïdes. D'une manière générale et sous incidence normale, une facette spéculaire participe plus à la SER de la cible qu'un segment ou un point brillant.

Les contributeurs spéculaires sont ceux qui participent le plus à la SER d'une cible. Cependant, de nombreux autres mécanismes d'interactions entre l'onde et l'objet peuvent intervenir. Certains, parmi les plus importants, sont évoqués ci-après et illustrés sur la Figure 1.6.

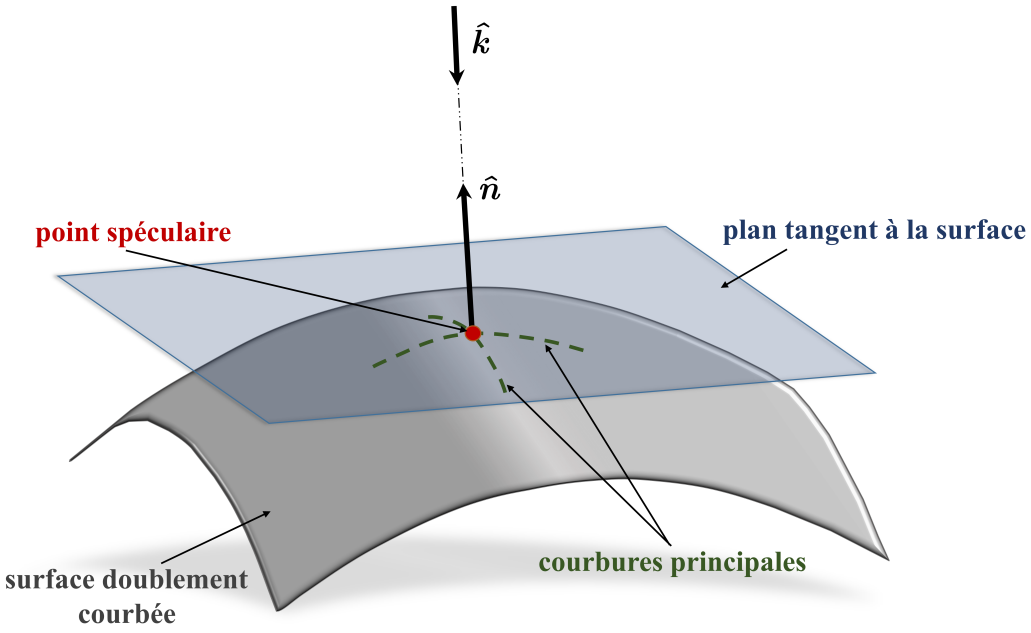


FIGURE 1.4 – Illustration d'un point spéculaire

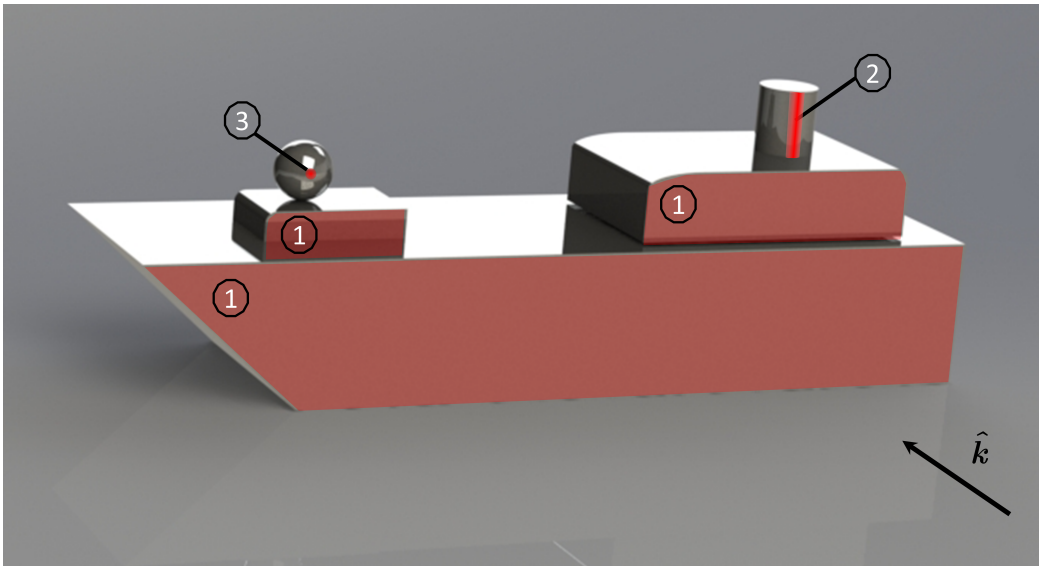


FIGURE 1.5 – Exemple de spéculaires pour une cible présentant des surfaces planes et courbées

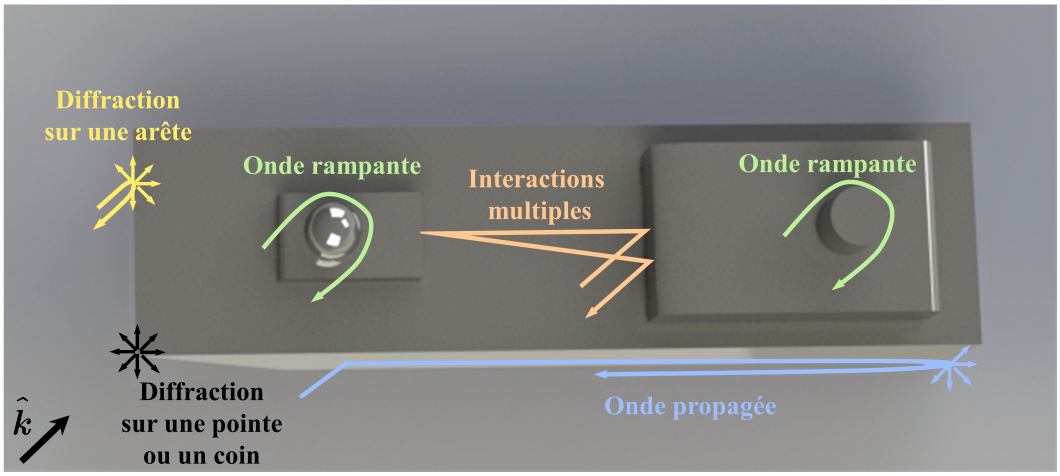


FIGURE 1.6 – Bateau (vue de dessus) : illustration de différents mécanismes de rétro-diffusion

Les ondes propagées

Une onde EM incidente peut engendrer des *ondes de courant* ou *ondes propagées*, qui vont se propager à la surface d'une cible CEP (Conducteur Électrique Parfait) le long d'un chemin géodésique, jusqu'à atteindre une discontinuité, comme une arête ou un sommet. L'énergie est alors rayonnée dans toutes les directions de l'espace dont une partie potentiellement non négligeable vers le radar. Une onde propagée en sens inverse peut également apparaitre. A l'origine, ces ondes étaient supposées être caractéristiques de longues surfaces planes éclairées à de faibles angles d'incidence mais des études [Knott *et al.*, 2004] ont montré que le phénomène existe également pour des objets considérablement plus courts et pour des angles d'incidence importants (jusqu'à 25°). Les ondes propagées ne sont néanmoins significatives que lorsqu'une composante de $\boldsymbol{E}_i$ est parallèle à la surface.

Les ondes rampantes

Une *onde rampante* est une onde propagée évoluant dans les zones d'ombre (*i.e.* les parties de la cible non éclairées par l'onde incidente) des cibles simplement ou doublement courbées. Tout comme l'onde propagée, l'onde rampante suit les géodésiques de surfaces et peut être rétro-diffusée par des discontinuités, même si ces dernières se situent dans une zone d'ombre. Si elle ne rencontre aucune discontinuité le long du chemin géodésique, l'onde rampante « tourne » autour de l'objet et finit par émerger de la zone d'ombre. Une partie de l'énergie est alors rétro-diffusée vers le radar. Néanmoins, à l'inverse de l'onde propagée, dont l'amplitude augmente généralement avec la distance parcourue, l'intensité de l'onde rampante décroît quand la distance parcourue augmente. Ceci s'explique par le fait que l'onde propagée évolue en zone éclairée et est « alimentée » par l'énergie fournie par l'onde incidente, alors que l'onde rampante évolue en zone d'ombre et rétro-diffuse de l'énergie à chaque fois qu'elle traverse une transition entre la zone d'ombre et la zone éclairée [Fuhs, 1983].

Les effets des discontinuités de surface ou de propriétés matériaux

Comme mentionné pour les ondes propagées, toute discontinuité (cavités, nervures, jointures...) sur une surface lisse et continue constitue une source d'écho. Il en est de même pour les discontinuités de courbure de surface, et cela même si la jonction entre les surfaces est lisse. Cependant, ces dernières ne sont généralement pas traitées car les autres mécanismes d'interaction sus-cités lui sont d'un ordre supérieur.

De la même manière, une anisotropie, une inhomogénéité ou une rupture des propriétés EM d'un matériau sont autant d'éléments qui constituent des sources d'écho. Ainsi, si une cible hétérogène est constituée de matériaux dont les milieux ont des indices de réfractions [2] différents (et en particulier quand cette différence est importante), les interfaces entre ces matériaux sont similaires à des discontinuités [Knott *et al.*, 2004]. La Figure 1.7 illustre ce phénomène pour un matériau inhomogène composé de plexiglas entourant un métal. Des spéculaires peuvent alors se produire à l'interface entre l'air et le plexiglas mais aussi à l'interface entre le plexiglas et le métal.

2. $n = \sqrt{\varepsilon_r \mu_r}$ où ε_r et μ_r sont respectivement la permittivité relative et la perméabilité relative du milieu. Dans le cas d'un milieu non magnétique $n = \sqrt{\varepsilon_r}$.

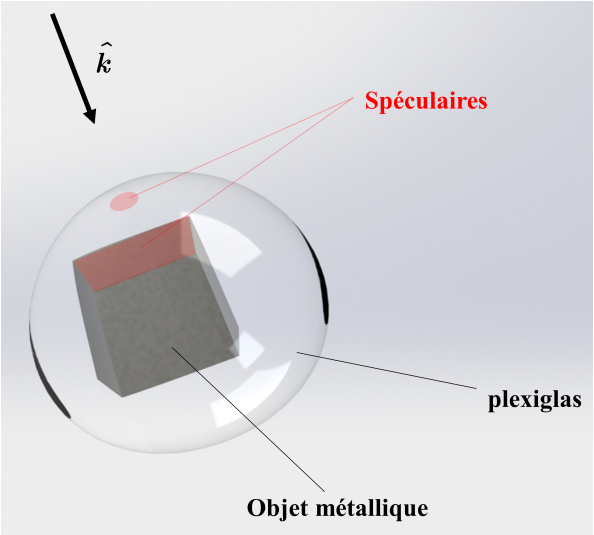


FIGURE 1.7 – Illustration d'une rétro-diffusion induite par une discontinuité de propriétés matériaux dans le cas d'une cible inhomogène

Les interactions multiples

Différente parties d'une cible peuvent potentiellement interagir entres elles. Généralement, ces interactions ne sont significatives que lorsque les parties se trouvent à quelques longueurs d'onde les unes des autres ou lorsqu'il y a une concentration de l'énergie sur une partie de la cible [Knott *et al.*, 2004]. Les interactions peuvent aussi être multiples en présence de cavités ou de guides d'onde.

Les diffractions

Les bords, les arêtes, les pointes, les coins et les sommets sont autant de discontinuités qui rétro-diffusent l'onde incidente dans toutes les directions de l'espace. La distribution angulaire de l'énergie diffusée dépend alors de l'angle d'arrivée de l'onde incidente et de l'angle d'observation ainsi que de la nature de la discontinuité. Ceci reste vrai que l'onde incidente provienne directement du radar ou d'une onde propagée ou rampante. Cependant, la signature radar de ces contributeurs reste faible car l'énergie est diffusée dans toutes les directions et seule une faible partie est interceptée par le radar.

1.1.3 L'analyse et le contrôle de SER

Dans la première section, nous avons expliqué que la SER est une grandeur scalaire caractérisant globalement le pouvoir de rétro-diffusion d'une cible. Si elle renseigne sur les capacités que doit avoir un radar pour détecter une cible, cette grandeur seule ne suffit pas à décrire et à caractériser en détail les interactions entre l'onde et l'objet. Dans la seconde section nous avons montré sommairement qu'il est possible de séparer les interactions onde-cible selon plusieurs phénomènes. Ceci nous permet seulement d'expliquer macroscopiquement certains mécanismes de rétro-diffusion. Néanmoins, lorsqu'une cible complexe est soumise à un champ EM, ces mécanismes sont souvent nombreux et variés. Pour mieux les appré-

hender, le radariste décompose généralement les cibles en un ensemble d'éléments simples dont l'étude est plus aisée. Si une telle décomposition est un procédé courant et raisonnable en hautes fréquences [Knott *et al.*, 2004], elle ne permet pas de comprendre et de maîtriser totalement toutes les interactions.

Dans cette section, nous définissons les notions d'analyse et de contrôle de SER et nous soulignons leur importance à travers divers exemples.

DÉFINITION : **Analyse de SER**

> L'*analyse de SER* est une démarche visant à expliquer et décrire plus finement les mécanismes de rétro-diffusion qui interviennent lorsqu'une cible est soumise à un champ EM. Elle consiste essentiellement à identifier, localiser et caractériser de manière précise les principaux contributeurs qui concourent à la SER totale de la cible. Elle repose généralement sur des méthodes d'imagerie radar comme nous l'expliquons dans la section suivante. *In fine*, l'analyse de SER conduit à une meilleure compréhension des phénomènes de rétro-diffusion complexes.

DÉFINITION : **Contrôle de SER**

> Le *contrôle de SER* est une démarche connexe, qui s'appuie souvent sur l'analyse de SER. Il s'agit de vérifier ou certifier que la SER d'un objet est conforme à un certain niveau de spécification en la comparant par exemple à celle d'objets nominaux ou à des standards.

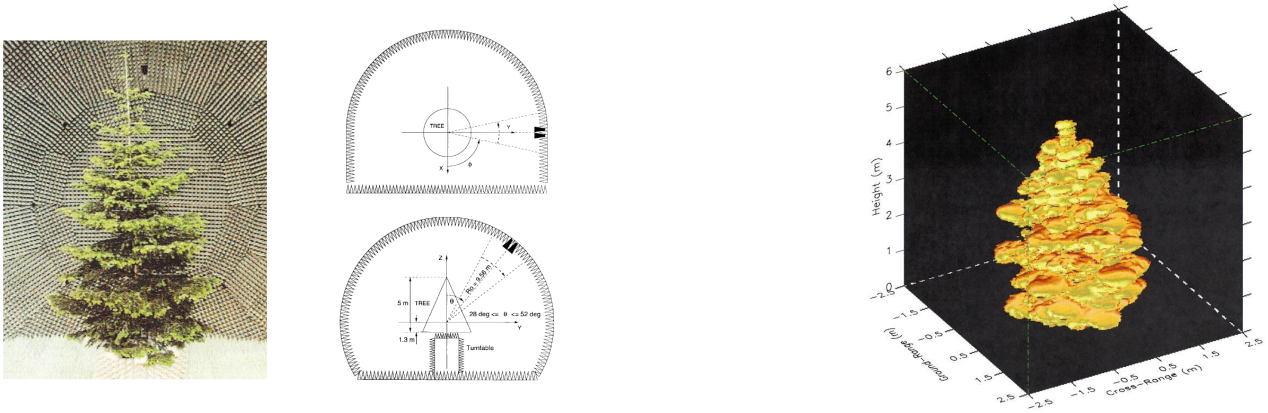
Nous faisons maintenant une liste illustrée et non exhaustive de quelques domaines où l'analyse et le contrôle de SER sont primordiaux.

Caractérisation de cibles ou de scènes complexes

La caractérisation précise de cibles ou de scènes complexes est d'une importance capitale dans de nombreuses applications. C'est par exemple le cas en imagerie SAR (Synthetic Aperture Radar ou Radar à Synthèse d'Ouverture) dont le champ d'utilisation très vaste inclut entre autres : les sciences de la Terre (géographie, hydrologie, océanographie, géologie, écologie, surveillance des forêts...), ou encore les applications militaires (surveillance de scènes urbaines, détection de véhicules, navires ou avions...). Les scènes imagées peuvent alors être particulièrement compliquées à analyser.

Par exemple, les mesures SAR de canopée en conditions naturelles sont complexes à analyser car elles dépendent de nombreux paramètres tels que : les conditions météorologiques, les interactions sol-arbres, la présence de signaux parasites, l'effet d'ombrage provoqué par les autres arbres ou le relief... Les données acquises dans ces conditions sont donc difficilement analysables en l'état. Dans ce domaine, [Fortuny, 2001] s'intéresse à l'analyse 3D et à la caractérisation de la SER d'arbres afin de mieux comprendre les images de canopée fournies par les SAR. Afin de calibrer et de contrôler pleinement les paramètres de mesure, ses expériences sont réalisées en laboratoire, dans une chambre anéchoïque Figure 1.8a.

L'analyse de la SER d'un arbre isolé n'exprime pas à elle seule toute la complexité de la scène imagée par un SAR. Néanmoins, [Fortuny, 2001] montre qu'elle permet d'expli-



(a) (b)

FIGURE 1.8 – (a) Photographie et schéma du dispositif de mesure [Fortuny et Sieber, 1999] (b) Carte 3D de rétro-diffusion d'un sapin en polarisation HH [Fortuny et Sieber, 1999]

quer une partie des interactions locales. Ainsi, les cartes 3D de rétro-diffusion produites Figure 1.8b, permettent, pour chaque polarisation, de préciser les interactions EM et les principaux contributeurs à la SER de l'arbre. Par ailleurs, une analyse plus fine met en évidence une forte dépendance de la SER à la structure de l'arbre (taille, forme, orientation du tronc, branchage et feuillage...) ainsi qu'à la polarisation choisie (essentiellement expliquée par le caractère fortement anisotrope des arbres).

Toujours en imagerie SAR mais dans un tout autre registre, [Tupin *et al.*, 2002] et [Soergel, 2010] s'intéressent à l'étude de la signature des scènes urbaines. Dans cet esprit, [Kempf *et al.*, 2007] et [Peichl *et al.*, 2008] ont mené des études visant à caractériser la SER de différentes structures présentes dans les villes, telles que des murs constitués de différents matériaux. Les résultats obtenus participent ensuite à la création de bases de données regroupant la SER de différentes cibles en zone urbaine. Ils permettent finalement d'enrichir les méthodes de détection et de reconnaissance non supervisées qui s'affranchissent de l'intervention d'un opérateur humain [Bhattacharyya et Sengupta, 1991].

Détection et identification de cibles

Pour détecter et identifier convenablement une cible à l'aide de dispositifs radar, il peut être nécessaire d'analyser précisément sa SER.

Par exemple, dans le domaine automobile, avec l'essor des véhicules autonomes, il est impératif de détecter et identifier des piétons afin d'éviter les collisions. Dans ce but, comme illustré Figure 1.9, [Chen et Chen, 2014] étudie la SER d'humains pour des fréquences variant entre 76 et 77 GHz. Cette recherche a pour objectif d'augmenter le taux de détection des radars anti-collision de piétons, embarqués dans la plupart des nouvelles automobiles. La détermination de la SER d'un individu est particulièrement difficile à mener car elle dépend de sa morphologie, de sa posture ou encore de ses vêtements. Dans l'article, les auteurs présentent également une analyse détaillée de la contribution des différentes parties du corps à la SER globale.

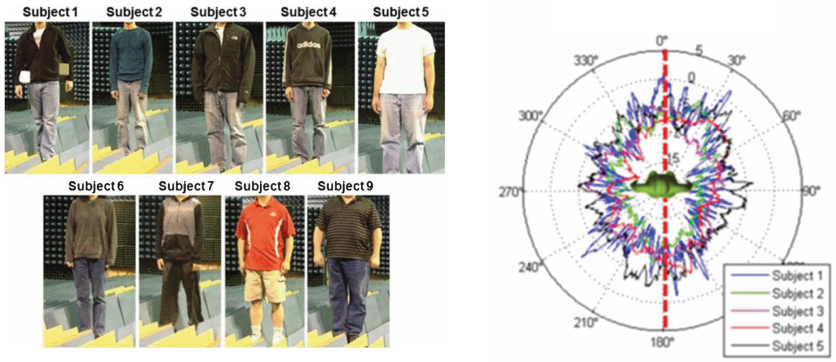
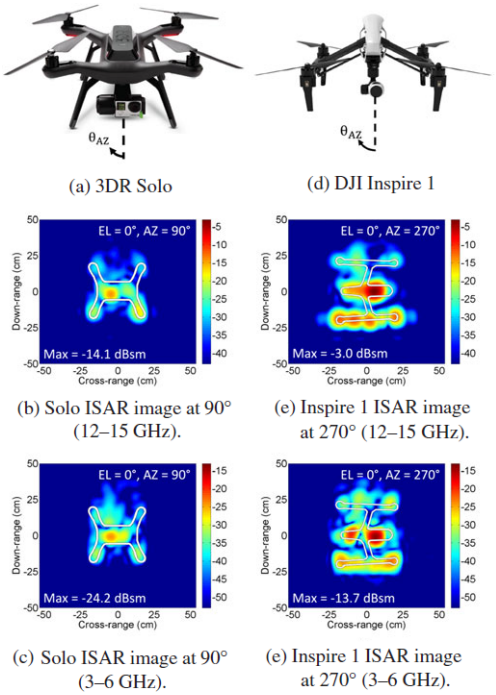


FIGURE 1.9 – Comparaison de la SER de 9 individus adultes [Chen et Chen, 2014]

Face à la multiplication des drones et à la menace potentielle qu'ils peuvent représenter, de nombreuses études ont été menées pour améliorer leur détection, leur identification et leur classification. Nous pouvons, par exemple, citer les articles de [Birch *et al.*, 2015; Herschfelt *et al.*, 2017]. L'analyse de la SER est rendue difficile par les faibles niveaux de SER de ces cibles. Un examen approfondi est donc nécessaire. Par exemple, [Ritchie *et al.*, 2015] ont étudié plus spécifiquement l'influence sur la SER des pales des drones alors que [Guay, 2016] et [Li et Ling, 2017] étudient l'influence de la fréquence d'analyse ainsi que de la forme des drones, comme illustré sur la Figure 1.10.



(a) 3DR Solo
(d) DJI Inspire 1
(b) Solo ISAR image at 90° (12–15 GHz).
(e) Inspire 1 ISAR image at 270° (12–15 GHz).
(c) Solo ISAR image at 90° (3–6 GHz).
(e) Inspire 1 ISAR image at 270° (3–6 GHz).

FIGURE 1.10 – Influence de la forme du drone et de la fréquence d'analyse sur la carte 2D de rétro-diffusion [Li et Ling, 2017]

Augmentation ou réduction de la SER

La maîtrise du niveau de SER est certainement l'aspect le plus critique dans le développement d'objets rétro-diffusants. Selon les applications, l'objectif peut être de minimiser la SER afin que les cibles ne soient pas détectables par les radar ou au contraire de maximiser leur signature radar.

Dans un contexte de réduction de SER, [Ton, 2013], [Li *et al.*, 2015a] ou [Vega *et al.*, 2016] s'intéressent à l'étude de la SER des éoliennes. Ces dernières peuvent être installées à proximité de systèmes d'émission et de réception micro-ondes (radio, télévision, système GPS, système de communication, radar de contrôle aérien, radar météorologique ou militaire...) qui requièrent souvent la détection et le traitement de signaux très faibles. La forte SER induite par les mâts et les pales, peut masquer ou altérer les signaux utiles, voire saturer les récepteurs. Pour assurer un bon fonctionnement des systèmes radar situés à proximité des parcs éoliens, il est donc capital, dans un premier temps, d'analyser la SER des éoliennes pour comprendre les processus de rétro-diffusion, puis, dans un second temps, de contrôler et réduire leur SER afin de limiter l'incidence sur les systèmes environnants. Par exemple, [Rashid et Brown, 2011] propose d'appliquer des matériaux absorbant les ondes EM (RAM) sur les pales. Après avoir divisé la pale en quatre sections, comme illustré Figure 1.11, ils appliquent alternativement des RAM sur chaque partie et mènent une série de mesures pour chaque configuration. Ceci permet de mieux appréhender la contribution de chaque section à la SER global de la pale. Finalement, la Figure 1.12 compare les SER d'une pale non traitée à celle d'une pale dont les quatre sections ont été recouvertes par des RAM.

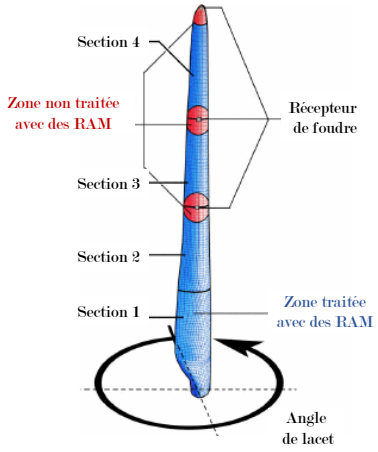


FIGURE 1.11 – Pale d'éolienne divisée en 4 sections [Rashid et Brown, 2011]

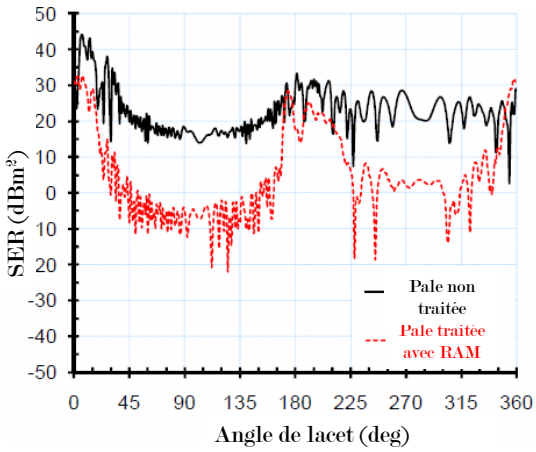
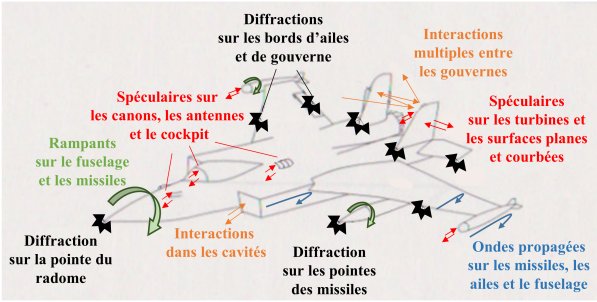


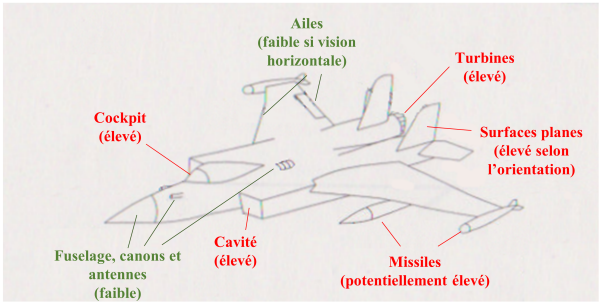
FIGURE 1.12 – Comparaison SER pale traitée et non traitée avec RAM [Rashid et Brown, 2011]

Dans un contexte militaire, les cibles (avions, chars, bateaux, missiles...) doivent généralement respecter certaines spécifications. Dans cet objectif, il est nécessaire d'identifier précisément les principaux contributeurs à la SER. Dès l'avènement des premiers systèmes radar dans la première moitié du XX^{e} siècle, de très nombreuses recherches, [Crispin, 1968; Fritsch, 1965; Knott *et al.*, 2004; Ruck *et al.*, 1970; Skolnik, 1990; Stone, 1989; White, 1998] sont ainsi consacrées à l'étude de la minimisation de la signature radar des cibles militaires. La Figure 1.13 illustre la démarche de [Fuhs, 1983] et [Bhattacharyya et Sengupta, 1991]

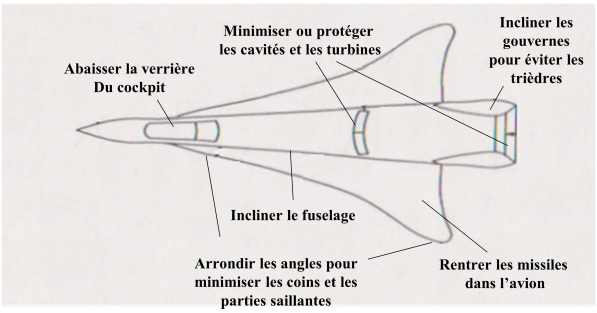
pour étudier la SER du MIG-31 (un intercepteur soviétique). Elle consiste à identifier les principaux contributeurs, à définir leur importance dans la SER totale de l'avion et finalement à proposer des solutions permettant de la réduire, notamment par des modifications de design, tout en maintenant des performances aérodynamiques acceptables.



1ère étape: Identification des principaux contributeurs



2ième étape: Identification du niveau de contribution à la SER totale



3ième étape: Amélioration du design pour réduire la SER

FIGURE 1.13 – Analyse approfondie de la SER du MIG-31
[Fuhs, 1983] et [Bhattacharyya et Sengupta, 1991]

A contrario, dans certaines applications, il peut être nécessaire d'augmenter la SER des cibles. C'est par exemple le cas dans le domaine des transports (aviation civile, transport maritime,...) où il s'agit de détecter et de suivre les véhicules. Dans cet objectif, [Shang et Shen, 2017; Shang *et al.*, 2015; Wood, 1998] ont par exemple proposé des techniques consistant en des cavités, des dièdres ou d'autres structures de fortes SER.

Contrôle non destructif

Le *contrôle non destructif* consiste en un ensemble de méthodes permettant de caractériser l'état de structures ou de matériaux, sans les détruire. Dans un contexte radar nous

pouvons citer les *radar à pénétration de sol* (ou *Ground Penetrating Radar* en anglais) qui permettent d'étudier et de caractériser la composition et la structure des sols [Langley *et al.*, 2007] ou de mettre en évidence la présence d'objets enfouis [Griffiths et Baker, 2007]. Ici encore, l'analyse précise de la SER est indispensable pour cartographier la scène et en tirer des informations.

Évaluation et validation de codes de calcul de SER

Les outils numériques de calcul de SER sont des moyens puissants quand il s'agit de prédire la SER d'une cible. Néanmoins, ils doivent être confrontés aux mesures de SER afin de mettre en évidence leurs éventuelles limites et de proposer des améliorations. L'étude de la diffraction sur la base d'un cône illustre parfaitement cela car elle reste partiellement décrite par les différentes méthodes de calcul. En effet, même si des avancées successives avec [Keller, 1961] puis [Bechtel, 1969] et enfin [Burnside et Peters, 1972] ont permis de mieux appréhender et prédire les mécanismes de diffraction à la base d'un cône (notamment pour une incidence axiale) ; [Knott et Senior, 1973] ont montré, à l'aide d'une série de mesures, qu'il est nécessaire de considérer des termes d'ordres supérieurs dans les méthodes de calcul pour bien rendre compte du phénomène.

1.2 L'outil principal : l'imagerie radar

Dans les sections précédentes, nous avons défini la notion de SER, répertorié les principaux mécanismes de rétro-diffusion qui interviennent dans un contexte hautes fréquences et souligné l'importance de l'analyse et du contrôle de SER à travers divers exemples.

Dans les sections qui suivent nous détaillons le processus de mesure de la SER et nous introduisons la notion d'*objet de rétro-diffusion*. Nous explicitons alors le modèle d'observation liant les données à l'objet à estimer. Nous présentons ensuite le problème d'imagerie de SER. Enfin, nous présentons quelques solutions classiques à ce problème.

1.2.1 La mesure de l'hologramme de SER

1.2.1.1 La notion d'hologramme

En optique, on appelle *holographie* le procédé consistant à enregistrer la phase et l'amplitude de l'onde diffractée par un objet. L'*hologramme* est alors le produit de l'holographie. Un exemple d'hologramme est donné Figure 1.14.

DÉFINITION : **Hologramme de SER**

> En radar, par analogie avec l'optique et par abus de langage, nous désignons par *hologramme de SER* l'ensemble des mesures de la SER complexe σ d'une cible pour différentes fréquences f et différents angles d'observation[a] θ et φ [Lonnqvist *et al.*, 2006].

a. Ces angles sont précisés ensuite sur la Figure 1.16.

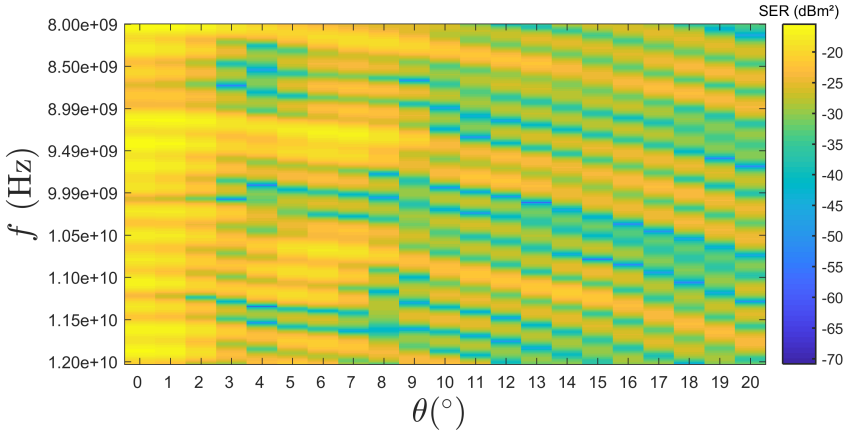


FIGURE 1.14 – Amplitude de l'hologramme (en dBm^2) d'un cône avec pastille (cf. **Section 4.2.2.1** p. 130) pour $\varphi = 0^\circ$

1.2.1.2 La mesure de SER

Principe général

L'objectif d'une instrumentation de mesure de SER est d'obtenir un hologramme pour une gamme de fréquences et des secteurs angulaires d'intérêt. Il s'agit pour cela d'acquérir, à polarisation donnée et avec une très grande précision, une séquence de M SER complexes :

$$\sigma^{[m]} = \sigma(f_m, \theta_m, \varphi_m), \quad m = 0, \cdots, M-1 \tag{1.8}$$

Nous rappelons que σ dépend en effet de l'attitude de la cible (*i.e.* son orientation angulaire par rapport au radar) et des paramètres propres au radar (fréquence et polarisation).

Finalement, l'hologramme mesuré est constitué d'un ensemble de M échantillons que nous pouvons écrire dans un vecteur :

$$\boldsymbol{\sigma} = \left[\sigma^{[0]}, \sigma^{[1]}, \cdots, \sigma^{[M-1]}\right]^{\mathrm{t}} \tag{1.9}$$

La mesure de la SER consiste à illuminer une cible immobile avec une onde EM monochromatique plane, dont la fréquence varie sur un domaine d'intérêt. La cible est généralement placée sur un positionneur qui permet de faire varier l'angle de présentation. À la réception, l'ARV (Analyseur de Réseau Vectoriel) mesure l'amplitude et la phase du champ EM rétro-diffusé. La SER complexe (1.5) est alors proportionnelle au ratio entre l'amplitude complexe du champ mesuré et celle du champ émis.

Une cible illuminée par une onde EM rétro-diffuse l'énergie dans toutes les directions spatiales. Seule une partie de cette énergie, proportionnelle à la SER, est captée par le radar. Cependant, dans un cadre expérimental en laboratoire, des réflexions sur les parois du local peuvent perturber la mesure. En général, la mesure de SER s'effectue donc dans une chambre anéchoïque où les parois sont tapissées de RAM – usuellement des pyramides de mousse de polyuréthane contenant des particules de carbone – afin de réduire les échos. Le lecteur est invité à se référer à [Knott *et al.*, 2004, chapitres 8, 9 et 11] pour de plus amples informations sur la mesure de SER et les RAM. De la même manière, le système de positionnement,

permettant d'obtenir avec précision les attitudes d'intérêt de la cible, doit être le plus discret possible afin de ne pas parasiter la mesure.

En pratique, et malgré ces précautions, l'information utile provenant de la cible est presque toujours masquée par la signature de l'environnement. Ceci est d'autant plus vrai si le niveau de SER de la cible est faible. En effet, les murs du local, bien que recouverts de RAM, n'absorbent pas la totalité des ondes EM. L'ARV mesure donc également la signature propre de la chambre en plus de celle de la cible. Par ailleurs, dans le cas d'un radar monostatique, la désadaptation de l'antenne entraîne une réflexion au niveau de l'élément rayonnant. Ce phénomène d'ondes stationnaires, caractérisé par un TOS (Taux d'Ondes Stationnaires) de l'antenne, engendre une propagation de la puissance réfléchie dans le dispositif de détection, ce qui masque l'information utile.

Un processus de calibration est donc préalable à toute mesure de SER. Il consiste généralement en une soustraction vectorielle et un étalonnage. Nous le décrivons de manière plus détaillée dans l'**Annexe A**.

L'Arche 3D : le dispositif de mesure de SER du CEA

Le CEA-CESTA a développé un moyen de mesure 3D de SER, dénommé « Arche 3D » [Massaloux *et al.*, 2014]. Représenté sur la Figure 1.15, il consiste en une antenne positionnée au dessus de la cible d'intérêt. La rotation de l'arche motorisée peut être commandée, afin d'obtenir un angle de gisement θ. Par ailleurs, la cible est positionnée sur un mât de support lui-même posé sur un plateau rotatif afin de faire varier l'angle de roulis φ autour de l'axe vertical. Afin de limiter sa signature, le mât de support doit être relativement « transparent » aux ondes EM mais aussi très rigide pour positionner parfaitement l'objet sous test. Pour cela il est généralement choisi en plexiglas ou en polystyrène expansé dont la permittivité est proche de celle de l'air.

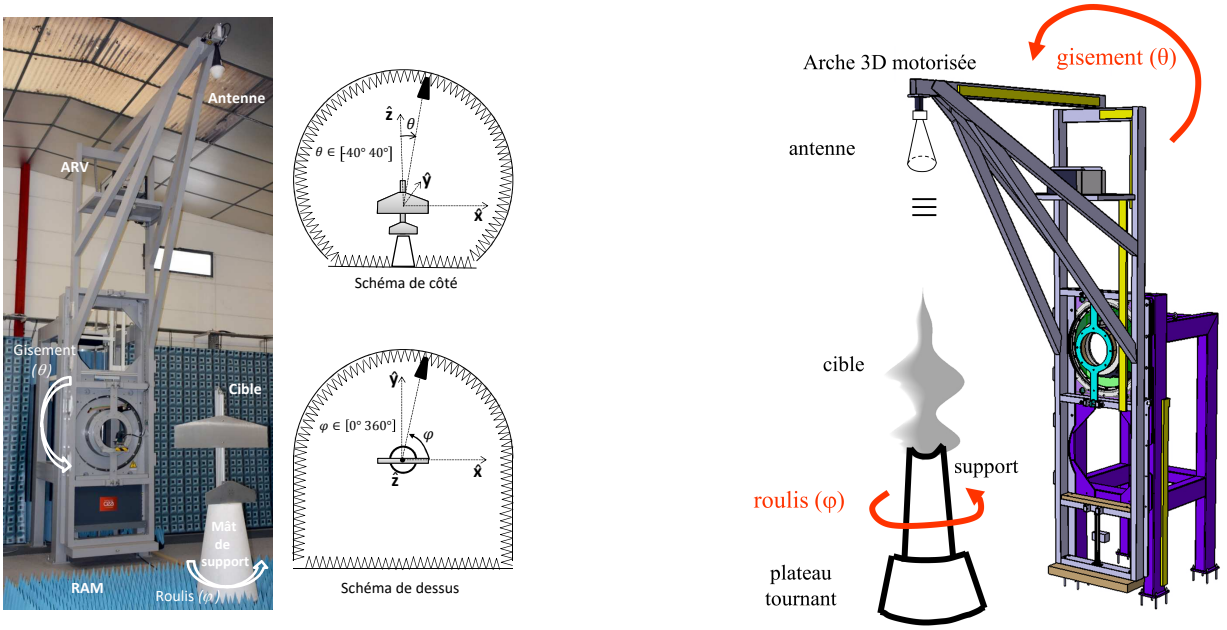


FIGURE 1.15 – Photographie et schéma de l'Arche 3D

L'Arche 3D permet des acquisitions rapides et précises de la SER pour différents angles d'aspect de la cible et différentes fréquences. L'acquisition est schématiquement représentée sur la Figure 1.16. Les directions successives d'illumination résultent de la conjugaison de la rotation de l'arche avec un gisement θ et de la rotation de la cible avec un roulis φ. Le référentiel de la cible est défini par $(O, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ où O est l'origine du repère correspondant au centroïde de la cible.

Selon la polarisation linéaire de l'antenne, les trois modes d'acquisitions sont possibles.

- **HH** : émission horizontale (champ $\boldsymbol{E}_i$ selon $\hat{\boldsymbol{x}}$ à $\theta = 0°$) – réception horizontale.
- **VV** : émission verticale (champ $\boldsymbol{E}_i$ selon $\hat{\boldsymbol{y}}$ à $\theta = 0°$) – réception verticale.
- **HV** : émission horizontale – réception verticale.

Ces derniers peuvent également être menés conjointement au cours de l'acquisition *via* une antenne bipolarisation.

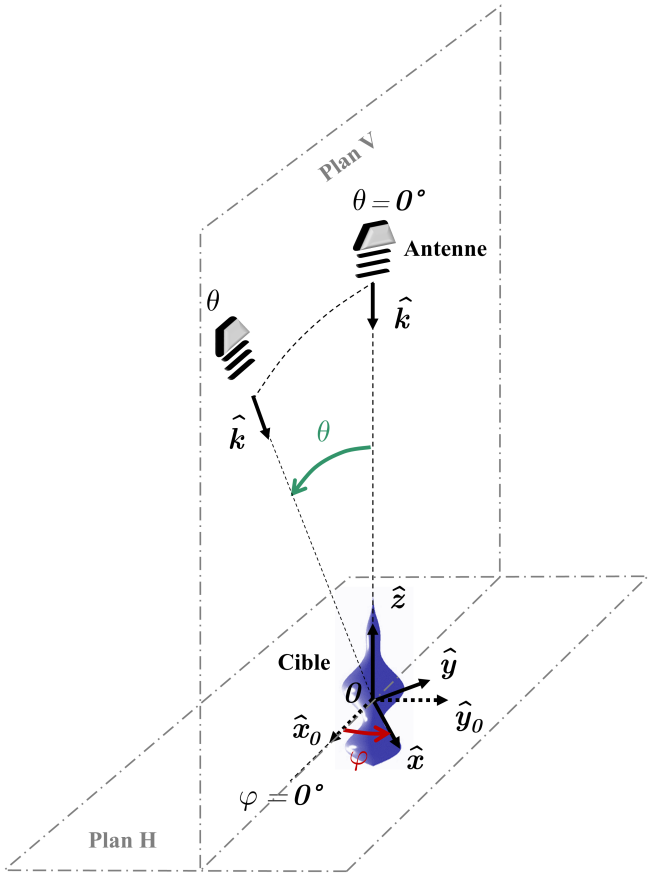


FIGURE 1.16 – Schéma des acquisitions

Dans le référentiel de la cible, le vecteur d'onde s'écrit :

$$\boldsymbol{k} = \|\boldsymbol{k}\| \begin{bmatrix} -\sin(\theta)\cos(\varphi) \\ -\sin(\theta)\sin(\varphi) \\ -\cos(\theta) \end{bmatrix} = \begin{bmatrix} -k_x \\ -k_y \\ -k_z \end{bmatrix} \tag{1.10}$$

où nous rappelons que $\|\boldsymbol{k}\| = 2\pi f/c$ est le nombre d'onde.

DÉFINITION : **Espace des vecteurs d'onde ou k-space**

L'*espace des vecteurs d'onde* ou *k-space* est l'espace des fréquences spatiales $[k_x, k_y, k_z]^{\mathrm{t}}$ dans lequel s'exprime le vecteur d'onde $\boldsymbol{k}$.

Pour des questions de praticité et de compacité des notations lors de l'écriture du modèle direct, le k-space est généralement préféré au domaine (f, θ, φ). La SER $\sigma(f, \theta, \varphi)$ s'exprime alors en fonction de $\boldsymbol{k}$ et est notée $\sigma(\boldsymbol{k})$. Le vecteur $\boldsymbol{\sigma}$ (1.9) est donc constitué d'un ensemble de SER mesurées pour une succession de M vecteurs d'onde $\boldsymbol{k}$. La Figure 1.17 illustre la position des mesures dans le domaine (f, θ, φ) et dans le k-space.

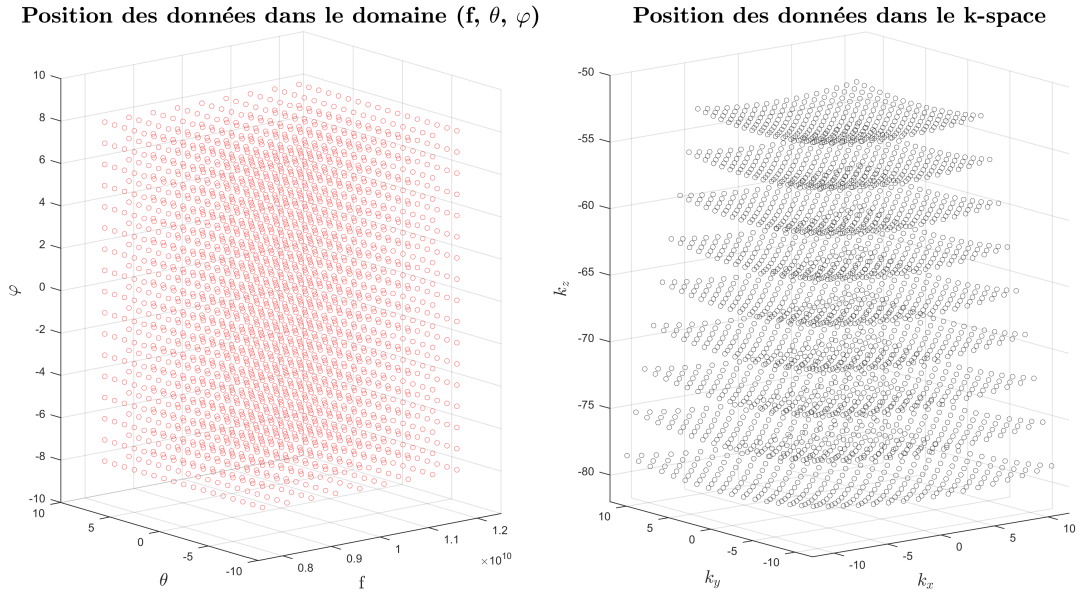


FIGURE 1.17 – Positions des données dans le domaine (f, θ, φ) et dans le k-space

Finalement, pour résumer, l'instrument mesure des coefficients de l'objet d'intérêt sur une grille irrégulière et incomplète du k-space.

REMARQUE

Parmi les dispositifs de mesures développés dans d'autres laboratoires nous pouvons par exemple citer [Fortuny, 2001] et [Eyraud *et al.*, 2008] qui sont entourés de RAM et permettent des acquisitions bi-statiques.

1.2.2 L'objet de rétro-diffusion

Maintenant que le processus de mesure a été décrit, nous définissons les notions d'*objet de rétro-diffusion* et de *carte de rétro-diffusion*.

DÉFINITION : **Objet et carte de rétro-diffusion**

Nous appelons *objet de rétro-diffusion* (ou simplement *objet*) la densité spatiale de réflectivité EM. La *carte de rétro-diffusion* (ou simplement *carte*) est la discrétisation de l'objet de rétro-diffusion sur une grille cartésienne régulière.

Nous notons $a(\boldsymbol{r}, \boldsymbol{k})$ l'objet de rétro-diffusion, où $\boldsymbol{r} = [x, y, z]^{\mathrm{t}}$ est le vecteur position dans le repère cartésien $(O, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ illustré sur la Figure 1.16. L'objet $a(\boldsymbol{r}, \boldsymbol{k})$ est anisotrope (*i.e.* il dépend des angles de vue) et coloré (*i.e* il dépend de la fréquence), il dépend donc du vecteur d'onde d'observation $\boldsymbol{k}$.

Considérons le contexte SASB (Small Angle Small Bandwidth) suivant :

Contexte SASB

Les excursions fréquentielles et angulaires des vecteurs d'onde $\boldsymbol{k}$ sont relativement limitées autour d'un vecteur d'onde d'intérêt $\boldsymbol{k}_c$.

Dans le contexte particulier SASB, nous pouvons supposer que $a(\boldsymbol{r}, \boldsymbol{k})$ ne dépend plus de $\boldsymbol{k}$ [Minvielle *et al.*, 2017; Özdemir, 2012]. Ainsi $a(\boldsymbol{r}, \boldsymbol{k}) \approx a(\boldsymbol{r}, \boldsymbol{k}_c)$ pour des $\boldsymbol{k}$ qui varient peu autour du vecteur d'onde d'intérêt $\boldsymbol{k}_c$. Par usage, le vecteur d'onde d'intérêt $\boldsymbol{k}_c$ est appelé *vecteur d'onde central*. Afin de simplifier les notations qui suivent, nous notons dans la suite $a(\boldsymbol{r}) \triangleq a(\boldsymbol{r}, \boldsymbol{k}_c)$.

Dans la suite du document nous nous intéressons donc aux méthodes d'estimation de l'objet $a(\boldsymbol{r})$ à partir des mesures $\boldsymbol{\sigma}$. Pour cela, nous explicitons d'abord le modèle direct liant chaque σ mesuré à l'objet a.

1.2.3 Le modèle direct

Dans cette section, nous présentons le modèle direct c'est-à-dire le modèle liant la SER complexe σ à l'objet $a(\boldsymbol{r}, \boldsymbol{k})$. Le développement des équations menant à ce modèle ne fait pas partie du champ de ce travail et n'est donc pas détaillé.

Dans un premier temps, nous explicitons le contexte dans lequel le modèle est valide.

Contexte du modèle

- Configuration et paramètres de fonctionnement du radar :
  - ▶ configuration monostatique
  - ▶ illumination en champ lointain avec des ondes électromagnétiques supposées quasi-planes sur la cible,
  - ▶ contexte hautes fréquences où l'optique physique s'applique,
- Nature des cibles :
  - ▶ relativement convexes avec une surface fermée,
  - ▶ surface lisse relativement aux longueurs d'ondes considérées.

En plus de ce contexte, nous faisons un choix sur l'orientation du repère spatial qui simplifie les expressions : la direction longitudinale $\hat{\boldsymbol{z}}$ est choisie colinéaire au vecteur d'onde central $\hat{\boldsymbol{k}}_c$.

1.2.3.1 Formulation continue

Dans le contexte présenté ci-dessus, [Borden, 1999] (par exemple) montre en s'appuyant sur les équations de Maxwell qu'une onde incidente engendre des courants électriques à la surface de la cible. Ces courants surfaciques produisent alors une onde rétro-diffusée dont l'amplitude et la phase dépendent entre autres de la surface de la cible, de son attitude et du vecteur d'onde $\boldsymbol{k}$. Des considérations d'optique physique en hautes fréquences pour des cibles lisses mènent alors au modèle direct continu (1.11) dans un contexte tri-dimensionnel général.

$$\sigma(\boldsymbol{k}) = \iiint_{\mathcal{D}} a(\boldsymbol{r}, \boldsymbol{k}) \exp\left\{2\jmath\pi \boldsymbol{k}\cdot\boldsymbol{r}\right\} d\boldsymbol{r} \tag{1.11}$$

où $\mathcal{D}$ est le support spatial de $a(\boldsymbol{r}, \boldsymbol{k})$, aussi appelé *champ de vue* et $\boldsymbol{k}\cdot\boldsymbol{r}$ est le produit scalaire entre le vecteur d'onde et le vecteur position. Notons que si le champ de vue est l'ensemble de l'espace (*i.e.* $\mathcal{D} = \mathbb{R}^3$), la SER correspond à la transformée de Fourier inverse de l'objet de rétro-diffusion. A ce titre, en tant que domaine dual du domaine spatial, le k-space est aussi appelé *domaine de Fourier*.

REMARQUE

> L'expression (1.11) peut aussi dériver de l'approximation de Born qui est plus adaptée pour des problèmes volumiques impliquant des cibles semi-transparentes et de grandes longueurs d'onde [Hodge et Schultz, 1965; Uberall, 1964].

Comme expliqué dans la **Section 1.2.4**, dans le contexte SASB, l'objet $a(\boldsymbol{r}, \boldsymbol{k})$ ne dépend plus de $\boldsymbol{k}$. Pour des faibles excursions angulaires et fréquentielles, (1.11) s'écrit :

$$\begin{aligned}\sigma(\boldsymbol{k}) &= \iiint_{\mathcal{D}} a(\boldsymbol{r}) \exp\left\{2\jmath\pi \boldsymbol{k}\cdot\boldsymbol{r}\right\} d\boldsymbol{r} \\ &= \iiint_{\mathcal{D}} a(x,y,z) \exp\left\{-4\jmath\pi f[k_x x + k_y y + k_z z]/c\right\} \mathrm{d}x\mathrm{d}y\mathrm{d}z\end{aligned} \tag{1.12}$$

Par ailleurs, dans le contexte des mesures de l'Arche 3D présenté dans la **Section 1.2.1.2**, (1.12) s'écrit plus précisément :

$$\sigma(\boldsymbol{k}) = \iiint_{\mathcal{D}} a(x,y,z) \exp\left\{-4\jmath\pi f\left(x\sin\theta\cos\varphi + y\sin\theta\sin\varphi + z\cos\theta\right)/c\right\} \mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{1.13}$$

Dans un contexte mono-dimensionnel, si $\hat{\boldsymbol{k}}$ est colinéaire à $\hat{\boldsymbol{z}}$, *i.e.* pour $\theta = 0°$, (1.12) s'écrit :

$$\sigma(\boldsymbol{k}) = \int_{\mathcal{D}_z} a(z) \exp\{-4\jmath\pi f z/c\} \, \mathrm{d}z \tag{1.14}$$

REMARQUE

Dans un contexte radar, [Borden, 1999; Chassay, 1983; Cheney et Borden, 2009; Le Chevalier et Fugier-Garrel, 1979; Le Chevalier et Pouit, 1983] ont montré que l'optique physique et l'approximation de la phase stationnaire mènent à très hautes fréquences à l'optique géométrique. Ils ont alors proposé un modèle communément appelé *modèle de points brillants* qui dérive du modèle (1.11). Dans ce dernier, la cible est modélisée comme une collection d'un faible nombre M de contributeurs isotropes. Chaque contributeur individuel est alors défini par sa position $\boldsymbol{r}_m = [x_m, y_m, z_m]^{\mathrm{t}}$ et par un coefficient de rétro-diffusion complexe a_m. Finalement, le modèle point brillants indique que l'énergie rétro-diffusée vers le radar résulte de la superposition cohérente des échos provenant de chaque contributeur et il s'écrit :

$$\sigma(\boldsymbol{k}) = \sum_{m=0}^{M-1} a_m \exp\{2\jmath \boldsymbol{k} \cdot \boldsymbol{r}_m\} \tag{1.15}$$

Les modèles (1.11), et *a fortiori* (1.12) et (1.15), présentent l'avantage d'être linéaires ce qui permet de simplifier le problème inverse d'imagerie que nous présentons ensuite. Il s'agit cependant d'une approximation qui ne prend pas en compte de nombreux phénomènes physiques tels que les interactions multiples ou les ondes rampantes. Lors du processus d'inversion, ces derniers apparaissent donc comme des *artefacts d'imagerie* que les radaristes savent analyser [Borden, 2002; Cheney et Borden, 2009]. Notons que des modèles plus complexes ont également été proposés pour considérer, par exemple, les interactions multiples [Chassay, 1982] et pallier ainsi certains des artefacts.

1.2.3.2 Formulation discrète

L'objet $a(\boldsymbol{r})$ n'est généralement pas envisagé comme une fonction de la variable spatiale continue. En effet, des considérations pratiques telles que l'utilisation de dispositifs de visualisation ou encore de méthodes rapides de traitement d'image, incitent à diviser le champ de vue en un certain nombre de voxels. Dans la suite, nous considérerons que le domaine spatial est discrétisé en une grille cartésienne régulière centrée sur l'origine du repère $(O, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$, avec N_x (resp. N_y et N_z) points selon l'axe $\hat{\boldsymbol{x}}$ (resp. $\hat{\boldsymbol{y}}$ et $\hat{\boldsymbol{z}}$), comme illustré sur la Figure 1.18. Notons qu'il s'agit d'un choix simple et direct de discrétisation communément adopté en imagerie radar. En effet, ce choix reste cohérent de la notion d'image que nous décrivons par la suite. Il permet en outre de « matricialiser » le modèle direct comme nous le montrons ci-après.

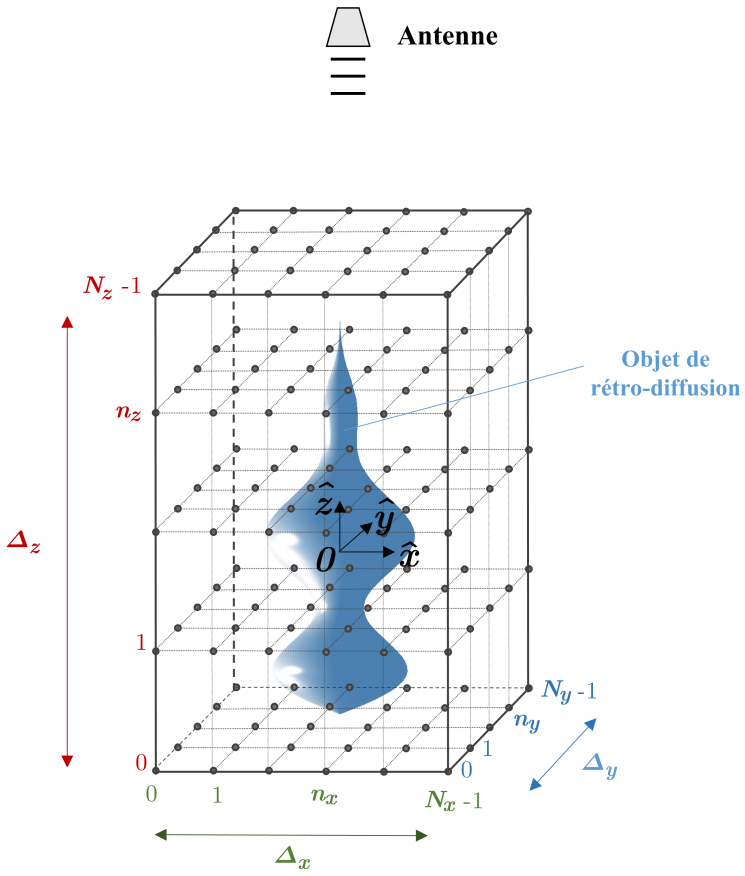


FIGURE 1.18 – Discrétisation du domaine spatial

REMARQUE

> Un autre choix aurait pu être de décomposer $a(\boldsymbol{r})$ sur une famille de fonctions ou *dictionnaire*. Par exemple, dans un contexte d'imagerie SAR, [Tria, 2005] effectue une décomposition en ondelettes. Néanmoins, il indique que les temps de calcul pour appliquer l'analyse en ondelettes peuvent être longs.

Nous notons $a(n_x, n_y, n_z)$ les éléments de l'objet $a(\boldsymbol{r})$ évaluées aux points $\left[x^{[n_x]}, y^{[n_y]}, z^{[n_z]}\right]$. Les indices n_x, n_y et n_z prennent les valeurs suivantes :

$$\begin{cases} n_x = 0, \cdots, N_x - 1 \\ n_y = 0, \cdots, N_y - 1 \\ n_z = 0, \cdots, N_z - 1 \end{cases}.$$

L'expression des points $x^{[n_x]}$, $y^{[n_y]}$ et $z^{[n_z]}$ est donnée par :

$$\begin{cases} x^{[n_x]} = x^{[0]} + n_x \delta_x \\ y^{[n_y]} = y^{[0]} + n_y \delta_y \\ z^{[n_z]} = z^{[0]} + n_z \delta_z \end{cases}.$$

où les pas d'échantillonnage spatiaux δ_x, δ_y et δ_z s'écrivent :

$$\begin{cases} \delta_x = \Delta x / (N_x - 1) \\ \delta_y = \Delta y / (N_y - 1) \\ \delta_z = \Delta z / (N_z - 1) \end{cases}$$

et Δx, Δy et Δz sont les extensions du champ de vue selon les directions spatiales.

Le vecteur position associé au point $\left[x^{[n_x]}, y^{[n_y]}, z^{[n_z]}\right]$ dans le repère $(O, \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ est noté :

$$\boldsymbol{r}_{n_x n_y n_z} = x^{[n_x]}\hat{\boldsymbol{x}} + y^{[n_y]}\hat{\boldsymbol{y}} + z^{[n_z]}\hat{\boldsymbol{z}}.$$

Le modèle direct discret résulte alors de la discrétisation de l'équation (1.12) :

$$\sigma^{[m]} = \sum_{n_x=0}^{N_x-1} \sum_{n_y=0}^{N_y-1} \sum_{n_z=0}^{N_z-1} a(n_x, n_y, n_z) \exp\left\{-2\jmath\pi \boldsymbol{k}^{[m]} \cdot \boldsymbol{r}_{n_x n_y n_z}\right\} \tag{1.16}$$

Par ailleurs chaque mesure est entachée d'un bruit. Nous pouvons alors considérer un vecteur de bruit défini par :

$$\boldsymbol{b} = \left[b^{[0]}, b^{[1]}, \cdots, b^{[M-1]}\right]^{\mathrm{t}}. \tag{1.17}$$

En plus du bruit de mesure, nous supposons que le vecteur $\boldsymbol{b}$ exprime également les erreurs de modélisation. Nous supposons également que ces erreurs sont additives.

Le *modèle d'observation*, dont un schéma est donné par la Figure 1.19, s'écrit alors :

$$\sigma^{[m]} = \sum_{n_x=0}^{N_x-1} \sum_{n_y=0}^{N_y-1} \sum_{n_z=0}^{N_z-1} a(n_x, n_y, n_z) \exp\left\{-2\jmath\pi \boldsymbol{k}^{[m]} \cdot \boldsymbol{r}_{n_x n_y n_z}\right\} + b^{[m]} \tag{1.18}$$

Le modèle d'observation (1.18) peut également s'exprimer sous la forme matricielle suivante :

$$\boldsymbol{\sigma} = \boldsymbol{H}\boldsymbol{a} + \boldsymbol{b} \tag{1.19}$$

où $\boldsymbol{a}$ est un vecteur colonne regroupant les éléments de $a(n_x, n_y, n_z)$ et $\boldsymbol{H}$ la matrice qui modélise le système, de terme générique $\exp\left\{-2\jmath\pi \boldsymbol{k}^{[m]} \cdot \boldsymbol{r}_{n_x n_y n_z}\right\}$ où $m = 0, \cdots, M-1$. Chaque ligne de la matrice $\boldsymbol{H}$ correspond à une observation dans le k-space et chaque colonne est associée à un point du domaine spatial.

$$\boldsymbol{H} = \begin{pmatrix} \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[0]} \cdot \boldsymbol{r}_{000}} & \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[0]} \cdot \boldsymbol{r}_{100}} & \cdots & \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[0]} \cdot \boldsymbol{r}_{(N_x-1)(N_y-1)(N_z-1)}} \\ \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[1]} \cdot \boldsymbol{r}_{000}} & \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[1]} \cdot \boldsymbol{r}_{100}} & \cdots & \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[1]} \cdot \boldsymbol{r}_{(N_x-1)(N_y-1)(N_z-1)}} \\ \vdots & \vdots & & \vdots \\ \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[m]} \cdot \boldsymbol{r}_{000}} & \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[m]} \cdot \boldsymbol{r}_{100}} & \cdots & \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[m]} \cdot \boldsymbol{r}_{(N_x-1)(N_y-1)(N_z-1)}} \\ \vdots & \vdots & & \vdots \\ \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[M-1]} \cdot \boldsymbol{r}_{000}} & \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[M-1]} \cdot \boldsymbol{r}_{100}} & \cdots & \mathrm{e}^{-2\jmath\pi \boldsymbol{k}^{[M-1]} \cdot \boldsymbol{r}_{(N_x-1)(N_y-1)(N_z-1)}} \end{pmatrix} \tag{1.20}$$

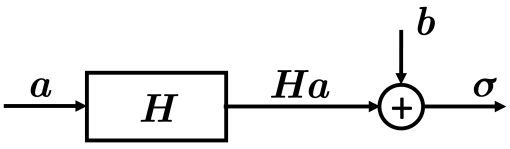


FIGURE 1.19 – Schéma du modèle d'observation

1.2.3.3 Reformatage des données

Comme les données ne sont pas sur une grille cartésienne régulière dans le k-space, la matrice dense $\boldsymbol{H}$ (1.20) n'a pas de structure de Tœplitz ou circulante. En grande dimension, ceci peut être un frein pour mener rapidement les calculs des méthodes d'inversion présentées par la suite. Une technique largement utilisée en traitement du signal, et en particulier en traitement des données radar, consiste à « reformater » les données dans le k-space sur une grille cartésienne régulière, en effectuant une transformation sphérique / rectangulaire. Cette démarche, dont un schéma de principe est illustré sur la Figure 1.20, est aussi connue sous le nom de *regridding*, *resampling* ou encore *mapping* en anglais [Doerry, 2012].

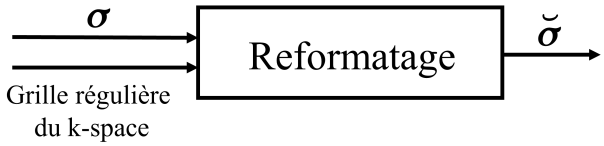


FIGURE 1.20 – Schéma de principe du reformatage des données.
Notations : $\boldsymbol{\sigma}$ données non regriddées et $\breve{\boldsymbol{\sigma}}$ données regriddées sur une grille cartésienne du k-space.

Pour mettre en œuvre le reformatage des données, plusieurs options sont possibles et différentes méthodes de regridding 3D ont été étudiées dans la littérature. L'objet de ce document n'est pas de les explorer ou de les comparer en détail. Aussi, le lecteur est invité à se référer à [Özdemir, 2012] qui présente plus spécifiquement deux méthodes usuellement employées en imagerie radar : l'*interpolation trilinéaire* et l'*interpolation plus proche voisin*.

Ces dernières consistent à parcourir l'ensemble des points de la grille cartésienne régulière du k-space et à leur affecter une valeur résultant de l'interpolation des mesures voisines. L'inconvénient de ces deux méthodes est qu'elles complètent toute la grille du k-space y compris les fréquences spatiales non observées éloignés de la zone effectivement mesurée. Des données « artificielles » sont donc ajoutées au k-space.

Dans notre cas, pour éviter d'ajouter des données « factices » et mieux respecter les données réellement mesurées, nous utilisons une autre méthode que nous appelons : *plus proche voisin lacunaire*. Contrairement à l'interpolation plus proche voisin présentée dans [Özdemir, 2012], notre approche consiste à parcourir l'ensemble des acquisitions (et non plus les points de la grille régulière) et à « déplacer » chaque mesure pour l'affecter uniquement au point le plus proche de la grille cartésienne régulière du k-space. Cette approche possède plusieurs caractéristiques :

- elle laisse des « trous » dans le k-space pour les points de la grille éloignés des mesures,
- plusieurs mesures peuvent être attribuées à un même point de la grille. Si tel est le cas,

ces dernières peuvent être moyennées [3],

- elle conserve le nombre de données et au maximum M points sont affectés sur la grille après l'étape de regridding.

[Sitko, 2013] compare les méthodes illustrées sur la Figure 1.21. Il montre que l'image formée à partir des données reformatées par la méthode plus proche voisin lacunaire est de meilleure qualité que celle résultant des deux autres méthodes. Ceci s'explique par le fait que les données ne sont que peu déplacées et le domaine de Fourier est donc moins « déformé ». Il montre également que le regridding plus proche voisin lacunaire est plus rapide que les deux autres et ne nuit donc pas aux temps de calcul. En effet, à l'inverse des méthodes classiques de regridding qui nécessitent de balayer tous les points de la grille du k-space, ici seules les acquisitions sont parcourues. Une implémentation matricielle sous Matlab (plutôt que l'usage de boucles) permet notamment d'améliorer considérablement les temps de calcul.

REMARQUE

Quelque soit la méthode choisie, une approximation est introduite car les données ne sont pas considérées à leur position initiale dans le k-space. Cependant, cette erreur diminue quand la grille régulière est raffinée.

Nous attirons l'attention du lecteur sur le fait qu'une large partie de ce qui suit est valide que les mesures soient positionnées ou non sur une grille régulière. Cependant, lorsqu'elles le sont, la matrice $\boldsymbol{H}$ a une structure particulière comme détaillée dans la **Section 1.2.3.4**. Ceci permet de mener les calculs numériques à l'aide de TFR (Transformée de Fourier Rapide) et donc de réduire drastiquement le coût calculatoire des méthodes que nous présentons ensuite.

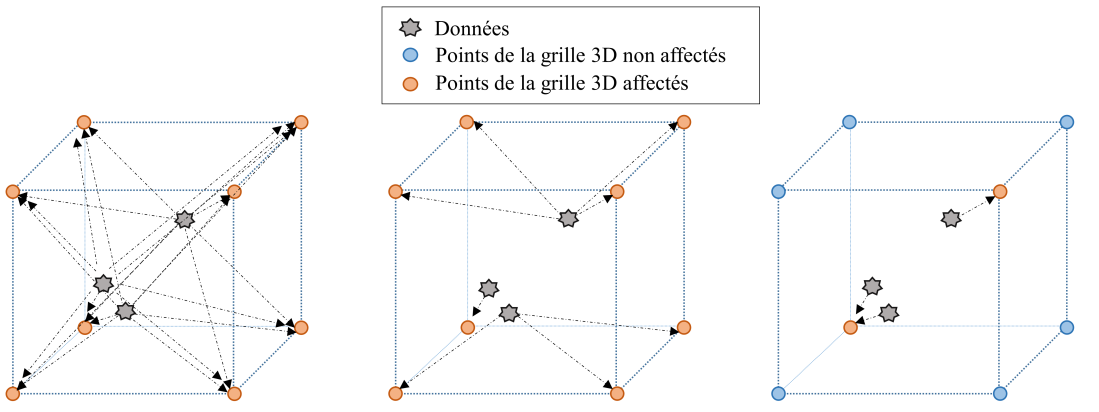


FIGURE 1.21 – Présentation des différentes méthodes de regridding

La Figure 1.22 illustre la position des données regriddées après une interpolation trilinéaire qui remplit la grille régulière du k-space et après un regridding plus proche voisin lacunaire qui laisse des « trous » dans la grille.

3. Nous justifions ce moyennage des données dans l'**Annexe D** (p. 162). Nous ajoutons qu'il s'agit d'une approximation permettant de mener les calculs par TFD. En effet, rigoureusement, il faudrait pondérer par le nombre de points affectés pour tenir compte de la variance du bruit plus faible pour les données reformatées.

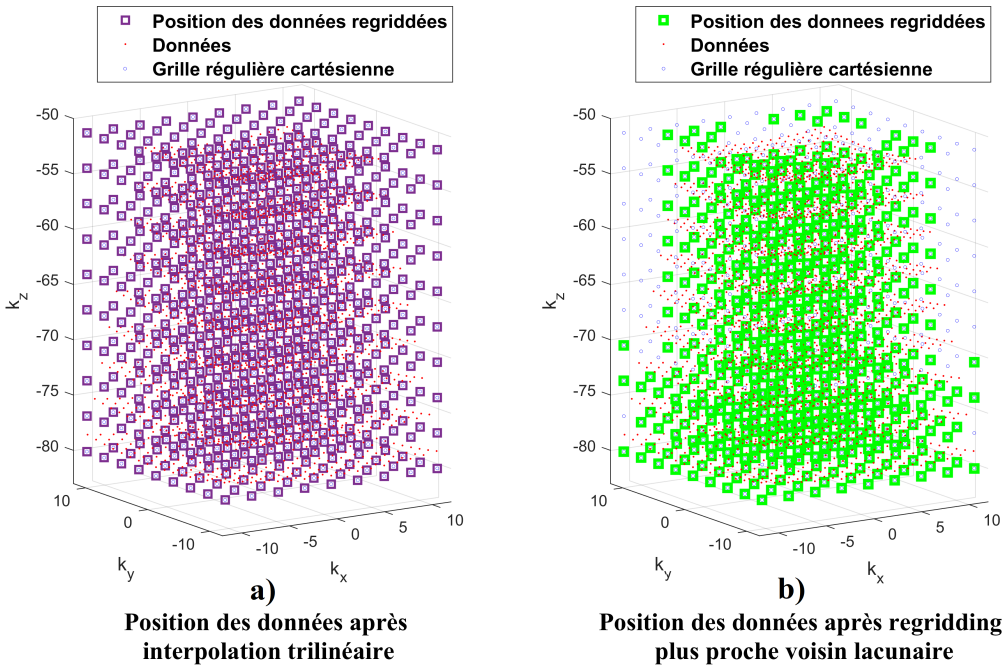


FIGURE 1.22 – Comparaison de la position des données regriddées après : a) interpolation trilinéaire et b) regridding plus proche voisin lacunaire.

1.2.3.4 Réécriture de la matrice $\boldsymbol{H}$

A l'issue de l'étape de regridding, les acquisitions sont sur une grille cartésienne régulière dans le k-space avec M_x points (resp. M_y et M_z) selon l'axe $\hat{\boldsymbol{k}}_x$ (resp. $\hat{\boldsymbol{k}}_y$ et $\hat{\boldsymbol{k}}_z$). Un point de la grille du k-space est défini par $\left[k_x^{[m_x]}, k_y^{[m_y]}, k_z^{[m_z]}\right]$.

Les indices m_x, m_y et m_z prennent les valeurs suivantes :

$$\begin{cases} m_x = 0, \cdots, M_x - 1 \\ m_y = 0, \cdots, M_y - 1. \\ m_z = 0, \cdots, M_z - 1 \end{cases}$$

L'expression des points $k_x^{[m_x]}$, $k_y^{[m_y]}$ et $k_z^{[m_z]}$ est donnée par :

$$\begin{cases} k_x^{[m_x]} = k_x^{[0]} + m_x \delta k_x \\ k_y^{[m_y]} = k_y^{[0]} + m_y \delta k_y \\ k_z^{[m_z]} = k_z^{[0]} + m_z \delta k_z \end{cases}$$

où les pas d'échantillonnage du k-space δk_x, δk_y et δk_z s'écrivent :

$$\begin{cases} \delta k_x = \Delta k_x / M_x - 1 \\ \delta k_y = \Delta k_y / M_y - 1 \\ \delta k_z = \Delta k_z / M_z - 1 \end{cases}$$

et Δk_x, Δk_y et Δk_z sont les extensions de la grille du k-space selon les directions des fréquences spatiales.

Dans la suite, nous choisissons (cf. p.160) :

$$\begin{cases} \delta k_x \delta_x = 1/N_x \\ \delta k_y \delta_y = 1/N_y \\ \delta k_z \delta_z = 1/N_z \end{cases} \tag{1.21}$$

Dans ce cas, la matrice $\boldsymbol{H}$ s'écrit [4] :

$$\boldsymbol{H} = \alpha\sqrt{N}\, \boldsymbol{\Delta}_\Psi \boldsymbol{S} \boldsymbol{F}_{3D} \boldsymbol{\Delta}_\Phi \tag{1.22}$$

- $\alpha = \exp\left\{-2\jmath\pi(k_x^{[0]}x^{[0]} + k_y^{[0]}y^{[0]} + k_z^{[0]}z^{[0]})\right\}$ est un coefficient complexe exprimant le déphasage à l'origine des grilles du k-space et spatiale.
- $N = N_x N_y N_z$ est le nombre d'éléments du vecteur $\boldsymbol{a}$ c'est-à-dire le nombre de voxels de la carte.
- $\boldsymbol{S}$ est une matrice de sélection binaire qui prélève uniquement les coefficients associés aux points de la grille du k-space affectés lors de l'étape de regridding.
- $\boldsymbol{F}_{3D} = \boldsymbol{F}_{N_x} \otimes \boldsymbol{F}_{N_y} \otimes \boldsymbol{F}_{N_z}$ est la matrice de TFD 3D [5] où

$$\boldsymbol{F}_\aleph = \frac{1}{\sqrt{\aleph}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{\aleph-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(\aleph-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{\aleph-1} & \omega^{2(\aleph-1)} & \cdots & \omega^{(\aleph-1)(\aleph-1)} \end{pmatrix}$$

  est la matrice de TFD (Transformée de Fourier Discrète) normalisée et $\omega = \mathrm{e}^{-2\jmath\pi/\aleph}$ est la racine $\aleph - $ ième de l'unité.
- $\boldsymbol{\Delta}_\Psi$ est une matrice diagonale de déphasage dans le k-space. Le $(m_x m_y m_z) -$ ième élément de sa diagonale est $\Psi = \exp\left\{-2\jmath\pi(x^{[0]}\delta k_x m_x + y^{[0]}\delta k_y m_y + z^{[0]}\delta k_z m_z)\right\}$.
- $\boldsymbol{\Delta}_\Phi = \boldsymbol{\Delta}_{\Phi_y} \otimes \boldsymbol{\Delta}_{\Phi_x} \otimes \boldsymbol{\Delta}_{\Phi_z}$ est une matrice diagonale de déphasage dans le domaine spatial où :
  - $\boldsymbol{\Delta}_{\Phi_x} = \mathrm{diag}\left\{ \exp\left\{-2\jmath\pi k_x^{[0]}\delta_x n_x\right\},\ n_x = 0, \cdots, N_x - 1\right\}$
  - $\boldsymbol{\Delta}_{\Phi_y} = \mathrm{diag}\left\{ \exp\left\{-2\jmath\pi k_y^{[0]}\delta_y n_y\right\},\ n_y = 0, \cdots, N_y - 1\right\}$
  - $\boldsymbol{\Delta}_{\Phi_z} = \mathrm{diag}\left\{ \exp\left\{-2\jmath\pi k_z^{[0]}\delta_z n_z\right\},\ n_z = 0, \cdots, N_z - 1\right\}$

Comme nous le montrerons dans la suite du document, la présence de la matrice de TFD $\boldsymbol{F}_{3D}$ dans l'expression de $\boldsymbol{H}$ nous permet d'utiliser des TFR.

4. Les calculs menant à cette expression sont donnés dans l'**Annexe D** (p. 159).
5. $\otimes$ est le produit de Kronecker

1.2.3.5 Fonction d'étalement du point

La FEP (Fonction d'Étalement du Point) , aussi appelée *réponse impulsionnelle spatiale*, est la réponse de l'instrument à une source ponctuelle.

Comme nous l'avons montré précédemment, si l'ensemble du k-space était observé, la SER serait la TFD de la carte de rétro-diffusion. Cependant, seule une partie du domaine de Fourier est observée. En effet, avant regridding, les données sont sur une portion de boule du k-space comme illustré sur la Figure 1.17. Après reformatage sur une grille régulière cartésienne, elles occupent une certaine portion du k-space comme illustré sur la Figure 1.22. Comme expliqué ensuite, le fait que le k-space ne soit pas totalement observé induit une convolution dans le domaine spatial entre chaque élément de la carte « parfaite »[6] et un certain noyau de convolution.

Pour illustrer cela nous pouvons réécrire (1.19) sous une forme spatial-spatial. Après reformatage des données, en multipliant (1.19) à gauche par $\boldsymbol{H}^\dagger = \alpha^* \sqrt{N} \boldsymbol{\Delta}_\Phi^* \boldsymbol{F}_{3D}^\dagger \boldsymbol{S}^t \boldsymbol{\Delta}_\Psi^*$, nous avons :

$$\boldsymbol{H}^\dagger \breve{\boldsymbol{\sigma}} = \boldsymbol{H}^\dagger \boldsymbol{H} \boldsymbol{a} + \boldsymbol{H}^\dagger \boldsymbol{b} \,. \tag{1.23}$$

- Nous posons $\bar{\boldsymbol{x}} = \boldsymbol{H}^\dagger \breve{\boldsymbol{\sigma}} = \alpha^* \sqrt{N} \boldsymbol{\Delta}_\Phi^* \boldsymbol{F}_{3D}^\dagger \boldsymbol{S}^t \boldsymbol{\Delta}_\Psi^* \breve{\boldsymbol{\sigma}}$. On l'obtient à partir des données regriddées, en les déphasant dans le domaine fréquentiel, en les complétant par des zéros, en prenant leur TFDI (Transformée de Fourier Discrète Inverse) puis en déphasant dans le domaine spatial.
- Nous posons également $\boldsymbol{A} = \boldsymbol{H}^\dagger \boldsymbol{H}$. Nous montrons dans l'**Annexe D** (p. 173) que $\boldsymbol{A}$ est une matrice de convolution circulante. Elle s'obtient en constituant un domaine de Fourier à partir des seules données observées, comme si tous les coefficients de Fourier observés étaient unitaires, puis en complétant par des zéros et en prenant une TFDI.

  Par exemple, si les données sont sur un parallélépipède rectangle du k-space – comme c'est le cas après une interpolation sur une grille cartésienne régulière du k-space Figure 1.22 a) – la FEP est un noyau de Dirichlet 3D[7] en trois dimensions comme illustré sur la Figure 1.23 a). [Özdemir, 2012], par exemple, montre que la largeur de la FEP est égale à $\pi/\Delta k_x$ selon la direction $\hat{k}_x$ (resp. $\pi/\Delta k_y$ et $\pi/\Delta k_z$ selon les directions $\hat{k}_y$ et $\hat{k}_z$). Le premier lobe secondaire est environ -13 dB en dessous du lobe principal. Ceci est illustré en 1D sur la Figure 1.24.

  Si les données ne sont pas sur un parallélépipède rectangle – comme c'est le cas après le regridding plus proche voisin lacunaire présenté sur la Figure 1.22 b) – la FEP n'est plus un noyau de Dirichlet 3D mais présente toujours un large lobe principal et de nombreux lobes secondaires élevés comme illustré sur la Figure 1.23 b).

  REMARQUE

  > Si tout le domaine de Fourier est observé, la FEP est une impulsion de Kronecker et le système est « parfait ».

6. *i.e.* correspondant à une observation de tout le domaine de Fourier
7. L'équivalent d'un sinus cardinal 3D discret.

- Pour terminer, nous posons $\bar{\boldsymbol{b}} = \boldsymbol{H}^{\dagger}\boldsymbol{b}$. Notons que si $\boldsymbol{b}$ est gaussien et centré alors $\bar{\boldsymbol{b}}$ l'est aussi. En revanche, la structure de corrélation est modifiée : par exemple, si $\boldsymbol{b}$ est blanc, $\bar{\boldsymbol{b}}$ ne l'est plus et sa structure de corrélation est même non-inversible (car il y a des relations déterministes entre ses composantes).

Finalement, la relation (1.19) prend alors la forme :

$$\bar{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{a} + \bar{\boldsymbol{b}}\,, \tag{1.24}$$

ce qui constitue bien une forme spatial-spatial. Ce modèle est donc convolutif et son inversion pose un problème de déconvolution. Cet aspect convolutif est illustré en 1D sur la Figure 1.24. Dans le domaine de l'astro-interférométrie, $\bar{\boldsymbol{x}}$ est appelé la « carte sale » ou « *dirty map* » et $\boldsymbol{A}$ est appelé le « lobe sale » ou « *dirty beam* ».

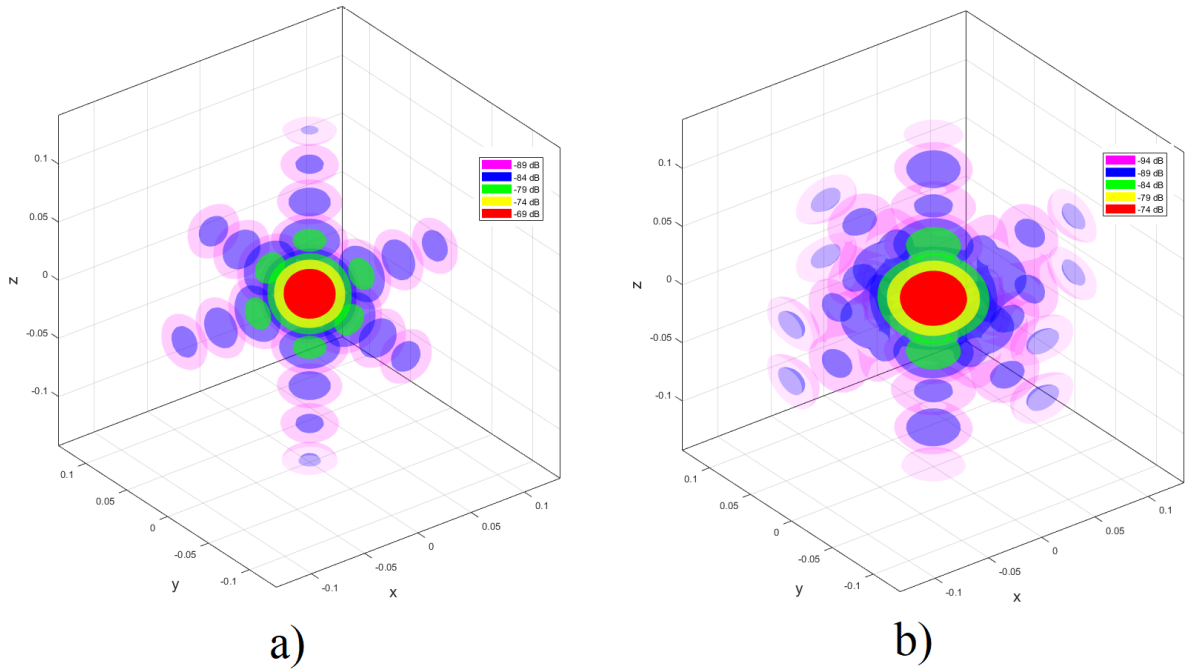


FIGURE 1.23 – Comparaison des fonctions d'étalement du point 3D (iso-surfaces en dB).
a) Noyau de Dirichlet 3D. b) Exemple de FEP après regridding plus proche voisin lacunaire.

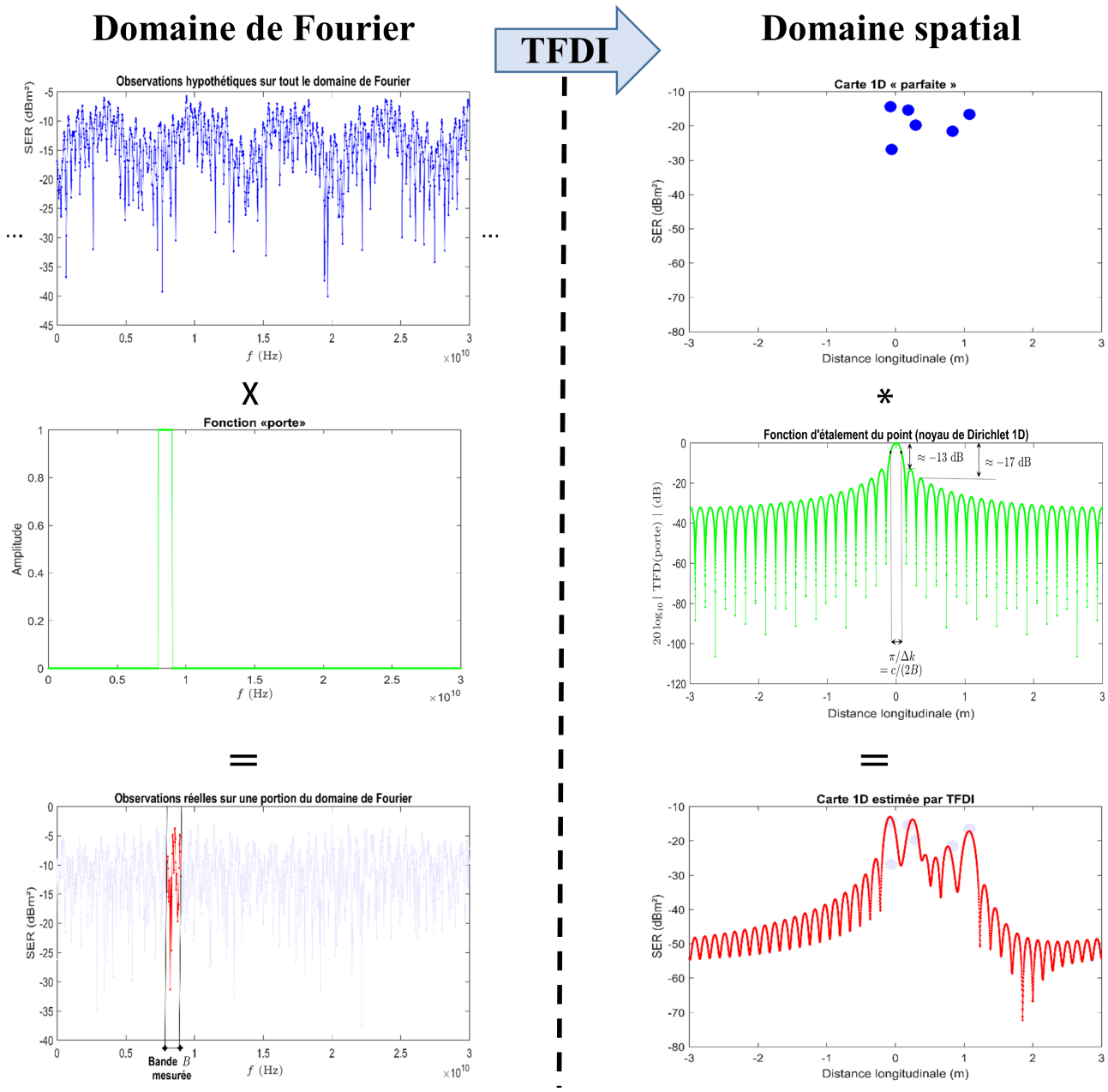


FIGURE 1.24 – Illustration 1D de la convolution spatiale avec un noyau de Dirichlet.

1.2.4 L'imagerie de SER

Après avoir présenté le processus de mesure de SER, la notion d'objet et de carte de rétro-diffusion puis le modèle direct liant les mesures à l'objet, nous formulons dans cette section le problème d'imagerie de SER.

Le dictionnaire IEEE [Jay, 1984] définit une *image* comme « la distribution spatiale d'une propriété physique telle qu'un rayonnement, une charge électrique, une conductivité ou encore une réflectivité, cartographiée à partir de la distribution de cette même propriété ou d'une autre propriété physique ». Communément, on appelle *image radar* la distribution spatiale de la réflectivité EM d'une cible [Mensa, 1981; Sidiropoulos *et al.*, 2014]. Il s'agit donc d'un outil essentiel pour la compréhension et la description précise des mécanismes de rétro-diffusion et finalement pour l'analyse et le contrôle de la SER.

REMARQUE

> L'imagerie radar ne se résume pas à la seule imagerie de SER. Ainsi, dans un contexte d'étude de propriétés des matériaux, l'objectif est par exemple d'estimer des cartes de permittivité ou de la conductivité de la cible [Eyraud *et al.*, 2015; Geffrin *et al.*, 2015; Pastorino, 2015; Yu *et al.*, 2009].

Dans ce travail, nous nous intéressons exclusivement à l'imagerie de SER dans un contexte monostatique. Il s'agit d'estimer une carte de rétro-diffusion dans le domaine spatial à partir de la mesure de l'hologramme de SER d'une cible c'est-à-dire de quelques coefficients de Fourier de l'objet. En cela, l'imagerie de SER constitue un problème de *synthèse de Fourier* qui rentre dans la classe des *problèmes inverses*.

Il est intéressant de noter que d'autres domaines éloignés de l'imagerie radar et reposant sur des principes physiques différents, posent le même type de problème.

- En astro-interférométrie, l'exploitation d'un réseau d'antennes ou de télescopes permet de simuler numériquement un système de plus grande taille (synthèse d'ouverture) afin de gagner en résolution. Par un principe de type « goniométrie », l'interaction entre deux de ces capteurs fournit un coefficient de Fourier de l'objet d'intérêt. L'exploitation de tous les couples de capteurs et éventuellement de la rotation de la terre permet de remplir en partie le plan de Fourier.
- En Imagerie par Résonance Magnétique (IRM), un champ magnétique intense génère un phénomène oscillatoire dont la fréquence est fonction de l'intensité du champ. Un gradient de ce champ (légères variations avec l'espace) est alors produit ce qui génère des oscillations dont la fréquence varie spatialement. Les signaux mesurés regroupent naturellement toutes les fréquences et leur analyse par transformée de Fourier restitue la répartition spatiale.
- Dans le domaine médical, la plupart des systèmes de scanner reposent sur une physique de projection de Radon et sur l'emploi du théorème coupe centrale. Les données peuvent alors être reliées à l'objet par une transformée de Fourier.
- L'analyse spectrale recouvre plusieurs techniques de description des signaux dans le domaine fréquentiel et permet en particulier d'obtenir les caractéristiques de la réponse d'un système linéaire à partir d'une fonction de transfert.

Ainsi donc, dans ces exemples, les données acquises sont des échantillons de la transformée de Fourier de l'objet d'intérêt et la reconstruction des cartes pose également un problème de synthèse de Fourier.

Dans la suite, et en cohérence avec les méthodes inverses que nous détaillons après, nous appelons indifféremment *observations* ou *données* les SER complexes $\sigma(\boldsymbol{k})$ mesurées. Afin de reconstruire une carte 3D il faut disposer de plusieurs observations $\sigma(\boldsymbol{k})$ correspondant à différents vecteurs d'onde, c'est-à-dire à différentes fréquences f et à différents angles de vue θ et φ. Ceci permet de lever les ambiguïtés longitudinales et transverses [Chen et Martorella, 2014; Garat *et al.*, 1993; Knott *et al.*, 2004]. Cependant, pour que l'objet $a(\boldsymbol{r}, \boldsymbol{k})$ ne dépende plus de $\boldsymbol{k}$, nous nous plaçons dans le contexte SASB (p. 36) où les vecteurs d'onde $\boldsymbol{k}$ varient peu autour d'un vecteur d'onde central $\boldsymbol{k}_c$ donné. La carte $\boldsymbol{a}$ à estimer est alors associée à $\boldsymbol{k}_c$.

REMARQUE

> Si les mesures s'effectuent en large bande fréquentielle, nous sortons du contexte SASB et l'objet de rétro-diffusion est dit *non stationnaire* (*i.e.* il dépend de $\boldsymbol{k}$). Il faut alors tenir compte de sa dépendance à la fréquence. Sans rentrer dans les détails qui sont en dehors du champ de ce travail, [Ros, 1994] et [Vignaud, 1996], par exemple, ont montré que le problème d'imagerie se ramène alors à une analyse temps-fréquence.

En résumé, l'imagerie de SER consiste à estimer $a(\boldsymbol{r})$ à partir des observations $\sigma(\boldsymbol{k})$ pour des excursions fréquentielles et angulaires restreintes de $\boldsymbol{k}$ autour de $\boldsymbol{k}_c$ et de la connaissance du système $\boldsymbol{H}$. Il s'agit d'un problème inverse qui présente plusieurs difficultés :

1. les données dans l'espace de Fourier sont incomplètes et bruitées,
2. le problème est posé en 3D et en grande dimension.

Le système d'observation se modélise néanmoins par un système linéaire (1.19). Après regridding des données, il repose sur des transformations simples (1.22) : déphasages, transformation de Fourier et troncature.

Cependant ce problème est sous-déterminé et il est impossible d'inverser directement le modèle pour trouver une solution. En effet, le nombre d'inconnues *i.e.* de voxels des cartes à estimer, est supérieur au nombre de données. Ainsi, la matrice $\boldsymbol{H}$, définie par (1.20), n'est pas carrée et elle est donc non inversible.

A titre d'exemple, le Tableau 1.1 présente le nombre d'inconnues pour quelques dimensions typiques de cartes HR. En pratique, dans notre cas, le nombre de données M est quant à lui de l'ordre de 100 000. Ainsi, le nombre d'inconnues peut être de 10 à plus de 10000 fois supérieur au nombre de données.

$N_x \times N_y \times N_z$	$128 \times 128 \times 128$	$256 \times 256 \times 256$	$512 \times 512 \times 512$	$1024 \times 1024 \times 1024$
Nombre d'inconnues	> 2 millions	> 16 millions	> 130 millions	> 1 milliard
Taille mémoire	> 30 Mo	> 270 Mo	> 2 Go	> 17 Go

TABLEAU 1.1 – Exemples du nombre d'inconnues et de la taille mémoire occupée pour quelques tailles typiques de cartes HR

1.3 Des solutions au problème d'imagerie

Dans cette partie nous présentons l'état de l'art sur les solutions apportées au problème d'imagerie. Nous commençons par les méthodes conventionnelles de résolution limitée puis nous listons quelques méthodes HR.

1.3.1 Les méthodes classiques et leur limites

Depuis l'avènement du radar au début du XX^{e} siècle la plupart des travaux ont essentiellement concerné le dispositif de mesure. Portée par de nombreuses recherches en physique, électronique, électromagnétisme ou encore dans le domaine de l'ingénierie radar, la partie matériel ou « hardware » a ainsi connu de nombreuses avancées.

A contrario, la partie « traitement des données » a réellement connu un regain d'intérêt dans les années 1960 avec l'essor de l'informatique plus performante et la formulation de l'algorithme de TFR [Cooley et Tukey, 1965]. Par ailleurs, c'est seulement dans les années 1980 que le traitement des données radar a réellement été appréhendé sous l'angle du traitement du signal [Munson *et al.*, 1983; Munson et Visentin, 1989].

Nous présentons ci-après quelques méthodes d'imagerie classiquement utilisées dans le domaine du radar pour apporter une solution au problème d'imagerie de SER.

1.3.1.1 Méthodes d'imagerie conventionnelles

Les méthodes d'imagerie conventionnelles reposent sur une TFRI des données. Parmi elles, deux sont usuellement employées :

- le PFA (Polar Format Algorithm),
- la RPF (Rétro-Projection Filtrée) ou FBP (Filtered Backprojection) en anglais.

1.3.1.1.1 Le Polar Format Algorithm

Le *Polar Format Algorithm* (PFA) [Mensa, 1981; Özdemir, 2012] fonctionne de la manière suivante :

- les données $\boldsymbol{\sigma}$ sont dans un premier temps reformatées sur une grille cartésienne du k-space avec l'une des méthodes présentées dans la **Section 1.2.3.3**. Ceci permet de former le vecteur $\breve{\boldsymbol{\sigma}}$ des données reformatées.
- les fréquences spatiales non observées sont mises à 0 (on parle de *bourrage de zéro*).
- une TFRI 3D des données reformatées $\breve{\boldsymbol{\sigma}}$ est ensuite réalisée pour estimer $\boldsymbol{a}$.

1.3.1.1.2 Rétro-projection filtrée

La *rétro-projection* est une méthode initialement développée en tomographie médicale et géophysique. Elle permet de reconstruire une structure à partir des projections spatiales

autour de l'objet. Dans un contexte radar, elle repose généralement sur une formulation tomographique du problème d'imagerie [Cetin et Karl, 1997; Munson *et al.*, 1983]. Le principe de la rétro-projection est d'*épandre* (*i.e.* de rétro-projeter) les projections en exploitant le théorème de la coupe centrale [Kak et Slaney, 2001]. Ce dernier indique que la transformée de Fourier d'une projection est une ligne passant par l'origine dans le domaine des fréquences spatiales. Le k-space peut donc se construire à partir de l'ensemble des projections autour de l'objet.

Tout comme le PFA, cette méthode permet de reconstruire l'objet. Pour cela, la transformée de Fourier de chaque projection spatiale est calculée puis le k-space est complété. Les données peuvent alors également être reformatées sur une grille cartésienne régulière pour tirer profit des TFR. Le nombre de projections étant limité en pratique, le remplissage du domaine fréquentiel n'est pas complet, notamment pour les hautes fréquences. Les zones vides du k-space sont alors complétées par interpolation à partir des points les plus proches ou par bourrage de 0. Si les données sont sur une grille cartésienne régulière du k-space, une TFRI 3D est réalisée pour estimer la carte $\boldsymbol{a}$. Notons que pour pouvoir effectuer une reconstruction correcte, les projections doivent balayer au moins 180°[Dubois, 1998]. Plus le nombre de projections est grand, plus le k-space est complété et plus la reconstruction est précise.

La RPF est une méthode qui dérive de la rétro-projection. Elle consiste à réduire les basses fréquences et à augmenter les hautes fréquences en filtrant les projections. En pratique, ce filtrage est réalisé dans le domaine fréquentiel.

1.3.1.1.3 Comparaison entre le PFA et la RPF

Le PFA est généralement employé quand les mesures sont effectuées pour des bandes fréquentielles et angulaires limitées. La RPF, en raison de son pré-traitement des données, est quant à elle couramment utilisée quand le dispositif de mesure tourne tout autour de l'objet. Avant reformatage, les portions de k-space mesurées sont donc généralement différentes. Pour la RPF, il s'agit d'une boule complète en raison de la rotation de l'instrument autour de la cible. Pour le PFA, et comme illustré sur la Figure 1.17, il s'agit d'une portion de boule déterminée par les excursions angulaires et fréquentielles des acquisitions. A ce titre, le PFA peut s'inscrire dans un contexte SASB si les bandes fréquentielles et angulaires sont petites.

En résumé, bien que des différences existent entre le PFA et la RPF, essentiellement en raison de leurs domaines d'application, ces méthodes reposent toutes deux sur un reformatage des données dans le k-space avec bourrage de 0 pour les fréquences spatiales non observées puis une TFRI 3D. Les cartes produites par ces deux méthodes sont d'ailleurs similaires [Cetin, 2001]. Dans la suite, nous appelons *méthode conventionnelle* le PFA.

1.3.1.1.4 Limites des méthodes conventionnelles

Comme expliqué dans la **Section 1.2.3.5**, le fait que seule une portion du k-space soit observée conduit à une convolution spatiale de la carte avec un noyau de convolution. La largeur du lobe principal de cette FEP dépend des bandes fréquentielles et angulaires des mesures ainsi que du type de reformatage employé [Özdemir, 2012]. Deux contributeurs proches

peuvent donc ne pas être séparés comme illustré en 1D sur la Figure 1.24. Par ailleurs, le niveau des lobes secondaires est élevé. La résolution des images produites avec ces méthodes est donc limitée. Ceci est illustré sur la Figure 1.25 qui présente la carte de rétro-diffusion d'un cône avec pastille (cf. cf. **Section 4.2.2.1** p. 130) obtenue avec la méthode PFA. Les interactions avec les principaux contributeurs (pointe tronquée, pastilles, base du cône) ont une résolution limitée à cause de la FEP.

REMARQUE

> Avant de réaliser la TFRI 3D, une pondération spectrale appelée *apodisation* peut être effectuée. Ce fenêtrage des données permet de réduire le niveau des lobes secondaires. Généralement, en radar, on choisit une fenêtre de Hanning, de Hamming, de Blackman ou de Kaiser [Kolawole, 2002; Özdemir, 2012]. Des apodisations non linéaires ont également été étudiées [Stankwitz *et al.*, 1995]. Cependant, comme illustré sur la Figure 1.26 pour un fenêtrage de Blackman, la largeur des lobes principaux est augmentée dans le même temps, ce qui dégrade la résolution et ne permet pas de séparer des contributeurs proches.

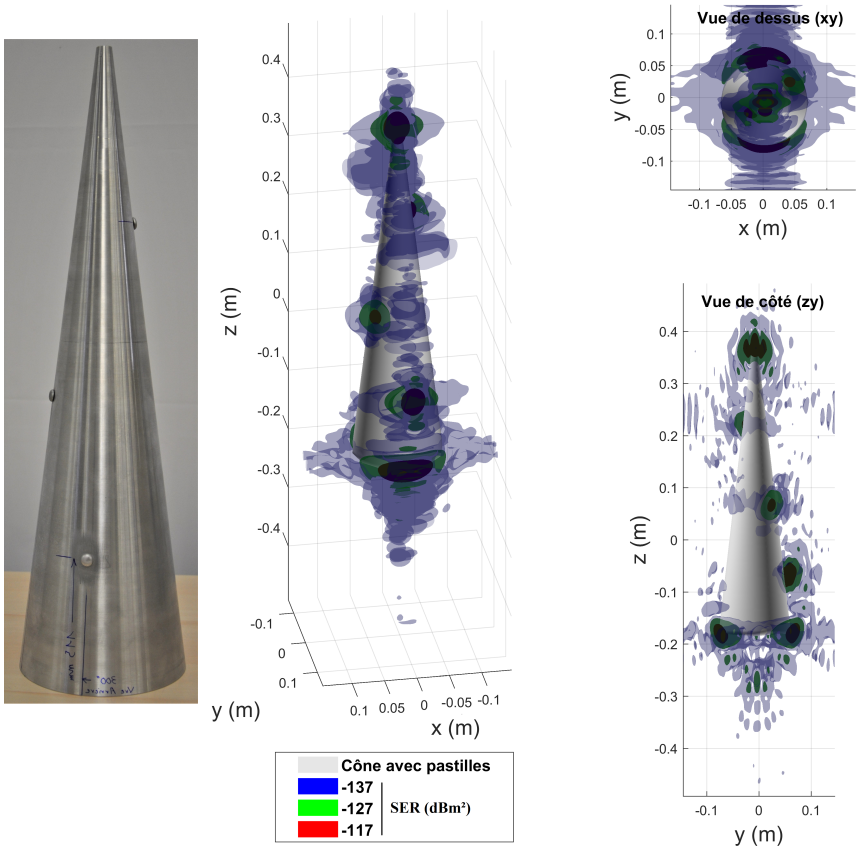


FIGURE 1.25 – Carte de rétro-diffusion d'un cône avec pastilles estimée avec la méthode PFA f =[8 GHz : 3,9 MHz : 12 GHz], φ = [-20 ° : 1° : 20 °], θ = [-20 ° : 4° : 20 °], polarisation HH sans fenêtrage

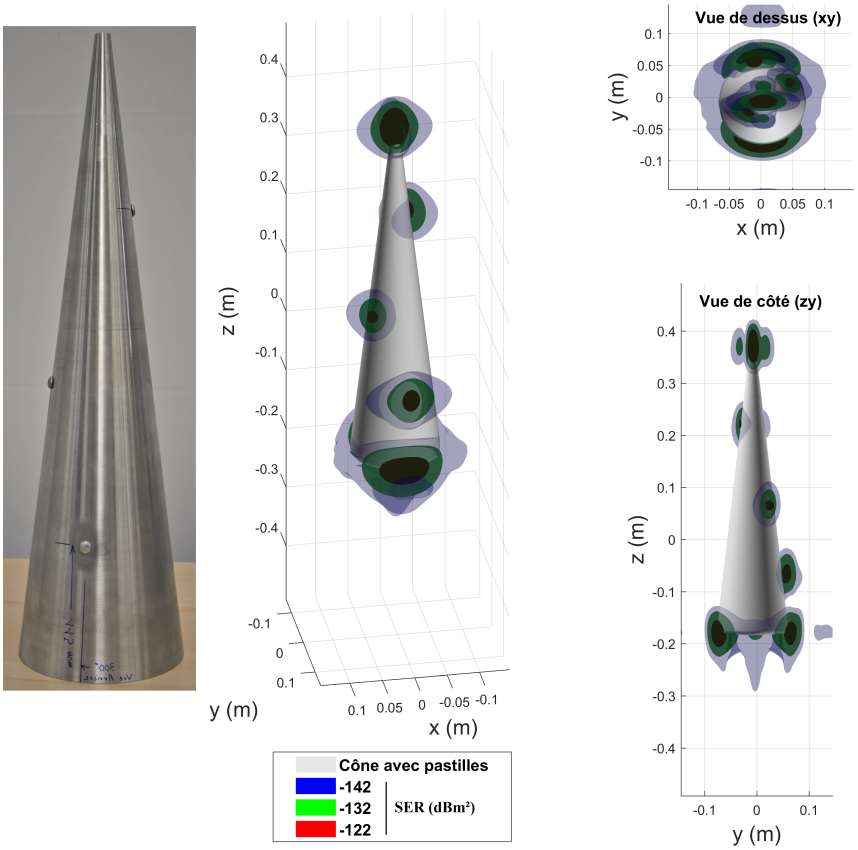


FIGURE 1.26 – Carte de rétro-diffusion d'un cône avec pastilles estimée avec la méthode PFA f =[8 GHz : 3,9 MHz : 12 GHz], φ = [-20 ° : 1° : 20 °], θ = [-20 ° : 4° : 20 °], polarisation HH fenêtrage de Blackman

1.3.2 L'approche problème inverse

Dans cette section, nous montrons que l'*approche problème inverse* fournit un cadre rigoureux pour définir des solutions au problème d'imagerie de SER. Dans ce formalisme il est possible de développer des méthodes permettant d'estimer les cartes avec une plus grande résolution.

1.3.2.1 Généralités sur l'inversion

Le *modèle d'observation* s'exprime sous la forme :

$$\boldsymbol{\sigma} = \boldsymbol{H}\boldsymbol{a} + \boldsymbol{b} \tag{1.25}$$

où la matrice $\boldsymbol{H}$ est appelée *matrice d'observation* et modélise le lien physique entre l'inconnue $\boldsymbol{a}$ et les données $\boldsymbol{\sigma}$. Cette formulation est commune à de nombreux domaines. [Giovannelli et Idier, 2013] présentent ainsi des exemples en imagerie médicale, en astrophysique ou encore en imagerie radar. $\boldsymbol{H}$ peut être par exemple une matrice de convolution (c'est par exemple le cas dans la relation spatial-spatial (1.23)). Connaissant le modèle direct (1.25),

l'objectif est alors de déterminer une estimée $\hat{\boldsymbol{a}}$ de l'objet d'intérêt $\boldsymbol{a}$ à partir des observations $\boldsymbol{\sigma}$: il s'agit dès lors d'un *problème inverse*.

REMARQUE

> Le modèle d'observation que nous avons présenté dans la **Section 1.2.3** rentre dans ce cadre. Dans ce modèle, $\boldsymbol{\sigma}$ est le vecteur des SER complexes, $\boldsymbol{a}$ l'objet de rétro-diffusion que l'on cherche à estimer et $\boldsymbol{H}$ la matrice d'observation définie par (1.20).

Si la matrice $\boldsymbol{H}$ est carrée et inversible et que le bruit est négligeable, une solution $\hat{\boldsymbol{a}}$ existe et est unique :

$$\hat{\boldsymbol{a}} = \boldsymbol{H}^{-1}\boldsymbol{\sigma} \tag{1.26}$$

Cependant, en pratique, dans notre cas, le problème est bruité et sous-déterminé : le nombre d'inconnues est supérieur au nombre de données. Dans ce cas, la matrice $\boldsymbol{H}$ n'est pas inversible et la solution (1.26) n'existe pas.

Solution des moindres carrés

Une première approche repose sur la minimisation de l'écart entre les mesures $\boldsymbol{\sigma}$ et la sortie du modèle $\boldsymbol{H}\boldsymbol{a}$. La métrique communément employée pour quantifier cet écart est la norme ℓ_2 et elle conduit au critère des *moindres carrés* :

$$\mathcal{J}_{\text{MC}}(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 \tag{1.27}$$

Une solution qui minimise le critère $\mathcal{J}_{\text{MC}}$, *i.e.* qui correspond le mieux aux données, s'écrit :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{J}_{\text{MC}}(\boldsymbol{a}) \tag{1.28}$$

Le minimiseur $\hat{\boldsymbol{a}}$ peut être calculée en annulant le gradient de $\mathcal{J}_{\text{MC}}(\boldsymbol{a})$. Il faut alors résoudre le système linéaire suivant :

$$\boldsymbol{H}^{\dagger}\boldsymbol{\sigma} = \boldsymbol{H}^{\dagger}\boldsymbol{H}\hat{\boldsymbol{a}} \tag{1.29}$$

Néanmoins, dans notre cas, ce système ne peut pas être résolu car $\boldsymbol{H}^{\dagger}\boldsymbol{H}$ est de rang déficient : le nombre d'inconnues est plus grand que le nombre de données. En conséquence, une infinité de cartes $\hat{\boldsymbol{a}}$ minimisent et annulent le critère des moindres carrés $\mathcal{J}_{\text{MC}}(\boldsymbol{a})$. En d'autres termes, une infinité de cartes est cohérente des données.

Pseudo-inverse

Dans des cas sous-déterminés comme le notre, parmi l'infinité de solutions minimisant le critère des moindres carrés, une approche consiste à chercher celle qui est de norme minimale qui s'écrit sous la forme :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmin}} \left\{ \begin{array}{l} \|\boldsymbol{a}\|_2^2 \\ \text{s.c. } \boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a} = 0 \end{array} \right. \tag{1.30}$$

REMARQUE

Cette solution est appelée *solution des moindres carrés à norme minimale* car si la contrainte $\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a} = 0$ est respectée le critère des moindres carrés est nul.

On peut montrer (par exemple avec une approche reposant sur les multiplicateurs de Lagrange [Llibre, 2008; Minvielle *et al.*, 2017]) que l'estimée $\hat{\boldsymbol{a}}$ s'écrit :

$$\hat{\boldsymbol{a}} = \boldsymbol{H}^{+}\boldsymbol{\sigma} \tag{1.31}$$

où $\boldsymbol{H}^{+} = \boldsymbol{H}^{\dagger}(\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}$ est connue sous le nom d'*inverse généralisée* ou encore *pseudo-inverse de Moore-Penrose* [Penrose et Todd, 1955]. Cette solution peut être considérée comme la « solution la moins énergique cohérente des données ».

Dans le cas de l'imagerie de SER, nous montrons dans l'**Annexe D** (p. 163) qu'après reformatage des données dans le k-space, en exploitant la structure (1.22) de la matrice $\boldsymbol{H}$, la solution des moindres carrés à norme minimale s'écrit :

$$\hat{\boldsymbol{a}} = \alpha^{*}\frac{1}{\sqrt{N}}\boldsymbol{\Delta}_{\Phi}^{*}\boldsymbol{F}_{3D}^{\dagger}\boldsymbol{S}^{t}\boldsymbol{\Delta}_{\Psi}^{*}\breve{\boldsymbol{\sigma}} \tag{1.32}$$

La solution $\hat{\boldsymbol{a}}$ consiste donc en un déphasage dans le domaine spatial ($\boldsymbol{\Delta}_{\Phi}^{*}$) de la TFRI 3D ($\boldsymbol{F}_{3D}^{\dagger}$) des données complétées avec des zéros ($\boldsymbol{S}^{t}$) et déphasées dans le k-space ($\boldsymbol{\Delta}_{\Psi}^{*}$). Finalement, si les données sont sur une grille cartésienne régulière du k-space (après une étape de regridding par exemple), la solution au sens des moindres carrés à norme minimale est la solution obtenue avec le PFA. D'une certaine manière, c'est la solution la plus « compromettante » puisqu'elle met à zéro les coefficients de la solution associée aux vecteurs propres de l'opérateur d'observation qui sont liés à des directions non observées. Elle est robuste dans le sens où elle réduit l'impact possible du bruit en empêchant sa propagation à travers des composantes non observés.

Comme expliqué dans la **Section 1.3.1.1.4**, cette méthode présente l'avantage de fournir rapidement une solution grâce à l'emploi de la TFRI 3D. Cependant, la complétion de la grille du k-space par des zéros (les fréquences spatiales non observées sont mises à zéro), limite la résolution des cartes estimées. Ces dernières présentent de larges lobes principaux et des lobes secondaires élevés qui nuisent à l'interprétation fine des mécanismes de rétro-diffusion.

1.3.2.2 Régularisation

Pour trouver une meilleure solution il est possible de *régulariser* le problème, c'est-à-dire de prendre en compte des informations *a priori* sur l'objet inconnu. Cette information permettra de compenser, en partie, les informations manquantes dans les données. La suite des travaux est consacrée aux approches régularisées par pénalisation et/ou contrainte.

Une méthode de régularisation communément employée consiste à considérer un *terme d'attache aux données* auquel est ajouté une pénalité dont la réduction favorise l'information *a priori* que nous avons sur la solution recherchée. La solution régularisée s'obtient par minimisation d'un critère composite :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \Upsilon(\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}) + \mu\mathcal{R}(\boldsymbol{a}) \tag{1.33}$$

Nous détaillons ci-après les éléments intervenant dans ce critère :

- $\Upsilon(\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a})$ est le terme d'*attache aux données*. Le choix de la fonction Υ dépend de la statistique du bruit additif $\boldsymbol{b}$. [Carlavan *et al.*, 2010] donnent quelques exemples de fonctions généralement utilisées :
  - $\Upsilon(\cdot) = \|\cdot\|_2^2$ pour un bruit blanc gaussien [Tikhonov et Arsenin, 1977]
  - $\Upsilon(\cdot) = \|\cdot\|_1$ pour un bruit blanc laplacien [Ben Hamza et Krim, 2001] ou pour un bruit impulsionnel [Nikolova, 2004]
  - $\Upsilon(\cdot) = \|\cdot\|_\infty$ pour un bruit uniforme ou de quantification [Weiss *et al.*, 2006]
- $\mathcal{R}(\boldsymbol{a})$ est la pénalité. Sa réduction favorise l'information *a priori* ajoutée sur la solution recherchée. Une large partie du travail qui suit concerne le choix et l'étude des pénalités.
- $\mu > 0$ est un hyperparamètre, appelé *paramètre de régularisation*. Il permet de contrôler le compromis entre le terme d'attache aux données et la pénalité. Plus μ est petit et plus la solution obtenue satisfait les données ; plus μ est grand et plus la solution est régularisée et tient compte de l'*a priori*.

REMARQUE

> Une approche bayésienne équivalente existe le plus souvent pour la solution (1.33). En effet, comme expliqué dans l'**Annexe B**, $\hat{\boldsymbol{a}}$ peut s'interpréter comme l'estimateur du *Maximum A Posteriori* (MAP) où $\Upsilon(\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a})$ correspond à l'anti-log-vraisemblance et traduit l'adéquation aux données et $\mathcal{R}(\boldsymbol{a})$ est l'anti-logarithme de la densité *a priori*. L'approche bayésienne offre notamment des perspectives pour régler l'hyperparamètre μ [Idier, 2001].

Dans la suite nous nous intéressons plus particulièrement au cas déterministe des moindres carrés pénalisés. Les critères s'écrivent :

$$\mathcal{J}(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu\mathcal{R}(\boldsymbol{a}) \tag{1.34}$$

Régularisation de Tikhonov

La *régularisation de Tikhonov*, étudiée par [Tikhonov et Arsenin, 1977] dans un contexte continu et par [Hunt, 1975] dans un contexte discret, est certainement la méthode de régularisation la plus employée pour la résolution des problèmes inverses. Le *critère de Tikhonov* s'écrit :

$$\mathcal{J}(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu \|\boldsymbol{M}\boldsymbol{a}\|_2^2 \tag{1.35}$$

où la matrice $\boldsymbol{M}$ est choisie en fonction de l'information *a priori* considérée. La *solution de Tikhonov* s'obtient en annulant le gradient du critère et s'écrit :

$$\begin{aligned} \hat{\boldsymbol{a}} &= \underset{\boldsymbol{a}}{\operatorname{argmin}} \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu \|\boldsymbol{M}\boldsymbol{a}\|_2^2 \\ &= \left(\boldsymbol{H}^\dagger\boldsymbol{H} + \mu\boldsymbol{M}^\dagger\boldsymbol{M}\right)^{-1}\boldsymbol{H}^\dagger\boldsymbol{\sigma} \end{aligned} \tag{1.36}$$

L'effet de la régularisation dépend du choix de la matrice $\boldsymbol{M}$.

- Lorsque $\mu = 0$ ou $\boldsymbol{M} = \boldsymbol{0}$, nous retrouvons la solution non régularisée des moindres carrés (1.29), pourvu que $\left(\boldsymbol{H}^\dagger \boldsymbol{H}\right)^{-1}$ soit inversible (ce qui n'est pas le cas si le système est sous déterminé comme expliqué avant).
- Lorsque $\boldsymbol{M}$ est la matrice identité $\boldsymbol{I}$, les solutions de faible norme sont favorisées.
- Lorsque la matrice $\boldsymbol{M}$ est un opérateur de différences finies (typiquement un gradient discret), des solutions « lisses » sont favorisées.

REMARQUE

L'expression (1.36) peut être calculée par TFD. Nous parlons alors également de *filtrage de Wiener* comme expliqué par exemple par [**?**].

La régularisation de Tikhonov n'est pas particulièrement adaptée au problème d'imagerie de SER. En effet, comme expliqué dans la **Section 1.1.2**, l'objet d'intérêt est généralement constitué d'un nombre relativement limité de spéculaires. Les méthodes favorisant un faible nombre de contributeurs à la SER sont donc privilégiées. Ces méthodes conduisent à des solutions parcimonieuses, *i.e.* les cartes estimées sont constituées d'un faible nombre d'éléments non nuls. Ceci permet généralement d'accroître la résolution des cartes de rétro-diffusion.

REMARQUE

Nous précisons dans le **Chapitre 2** la notion de parcimonie et les méthodes de régularisation pour la favoriser (notamment l'emploi de la norme ℓ_1).

1.3.3 La parcimonie en imagerie radar

Dans cette section nous présentons quelques méthodes utilisées dans le domaine du radar et reposant sur l'idée de parcimonie.

L'idée de parcimonie dans le domaine du radar n'est pas nouvelle. Par exemple, [Tsao et Steinberg, 1988] en 1D puis par la suite [Stuff *et al.*, 2004] en 3D, ont proposé l'usage d'algorithmes gloutons tels que CLEAN [Högbom, 1974] pour estimer un faible nombre de contributeurs à la SER. Cet algorithme – ainsi que ses nombreuses variantes (HOGBOM, CLARK, MX, SDI, MRC, MULTI) – suppose que l'objet d'intérêt est constitué d'un certain nombre d'éléments ponctuels. L'algorithme permet alors d'estimer la position et l'amplitude des principaux contributeurs, selon un processus itératif décrit par exemple dans [Chen et Martorella, 2014][chapitre 6].

Dès la fin du XVIII^{e} siècle, en ingénierie hydraulique, [De Prony, 1795] a proposé d'utiliser des techniques de régression linéaire, considérées comme les premières approches HR [Sahnoun, 2012], pour estimer les fréquences d'une somme d'exponentielles. De nombreuses améliorations ont été apportées par la suite [Kumaresan et Tufts, 1982; Pisarenko, 1973; Rahman et Kai-Bor Yu, 1987]. Dans cette veine, des méthodes HR reposant sur la décomposition de l'espace de données en un sous-espace signal et un sous-espace bruit ont été développées à la fin du XX^{e} siècle. Parmi elles, les plus connues sont : les algorithmes MUSIC (Multiple Signal Classifier) [Schmidt, 1986] et ESPRIT (Estimation of Signal Parameters *via* Rotational

Invariance Technique) [Roy *et al.*, 1986] et leurs variantes respectives (root-MUSIC [Barabell, 1983] et TLS-ESPRIT [Roy *et al.*, 1989]). L'étude de ces méthodes est hors du champ de ce travail, le lecteur est invité à se reporter aux articles mentionnés et à [Sahnoun, 2012] pour obtenir de plus amples détails.

[Berthoumieu, 1996; Burrows, 2004; Poulalion, 1997] se sont intéressés aux méthodes d'estimation spectrale HR MUSIC et ESPRIT dans un contexte radar 2D. Dans cette voie, [Moch *et al.*, 2017] ont récemment proposé une nouvelle méthode d'imagerie radar 3D : une carte est obtenue par la méthode PFA avec une fenêtre d'apodisation de Bartlett puis les contributeurs principaux sont extraits en appliquant un masque déterminé par une extension 3D de MUSIC.

REMARQUE

> Les techniques sus-mentionnées ne sont pas à proprement parler des méthodes d'imagerie dans le sens où il ne s'agit pas de déterminer une carte mais la position et l'amplitude d'un ensemble de contributeurs ponctuels comme illustré sur la Figure 1.27.

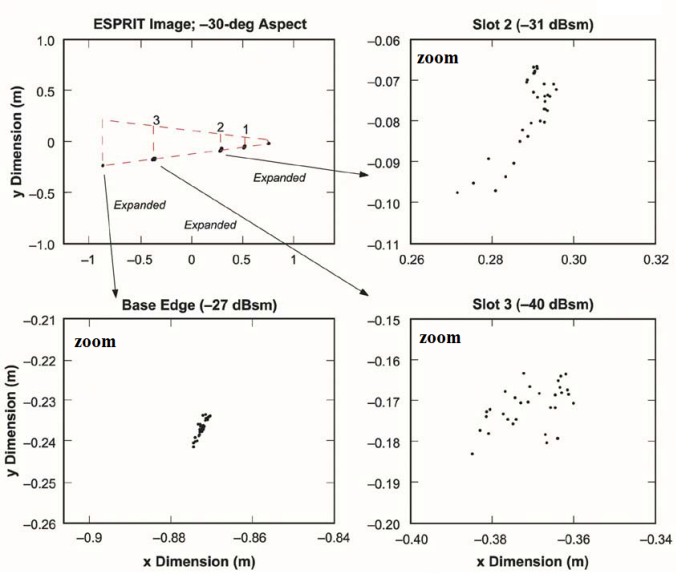


FIGURE 1.27 – Principaux contributeurs à la rétro-diffusion d'un cône avec rainures estimés par la méthode ESPRIT et zooms au niveau des rainures [Burrows, 2004].

L'idée d'aborder l'imagerie radar sous l'approche problème inverse en y intégrant des *a priori* favorisant la parcimonie est quant à elle apparue plus tardivement. Les articles de [Austin et Moses, 2008], [Herman et Strohmer, 2009], [Potter *et al.*, 2010] puis [Austin *et al.*, 2011; Xiao-Chun, 2010] en 3D, sont, à notre connaissance, les premiers à avoir véritablement présenté le problème de l'imagerie radar sous cet angle. Ils proposent une méthode reposant sur la construction puis l'optimisation d'un critère avec pénalité ℓ_1 non contraint. La solution recherchée est définie par :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmin}} \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu \|\boldsymbol{a}\|_1 \tag{1.37}$$

La Figure 1.28 illustre une solution obtenue avec cette méthode. La carte est constituée d'un ensemble de points brillants similairement aux résultats obtenus avec les méthodes d'estimation spectrale.

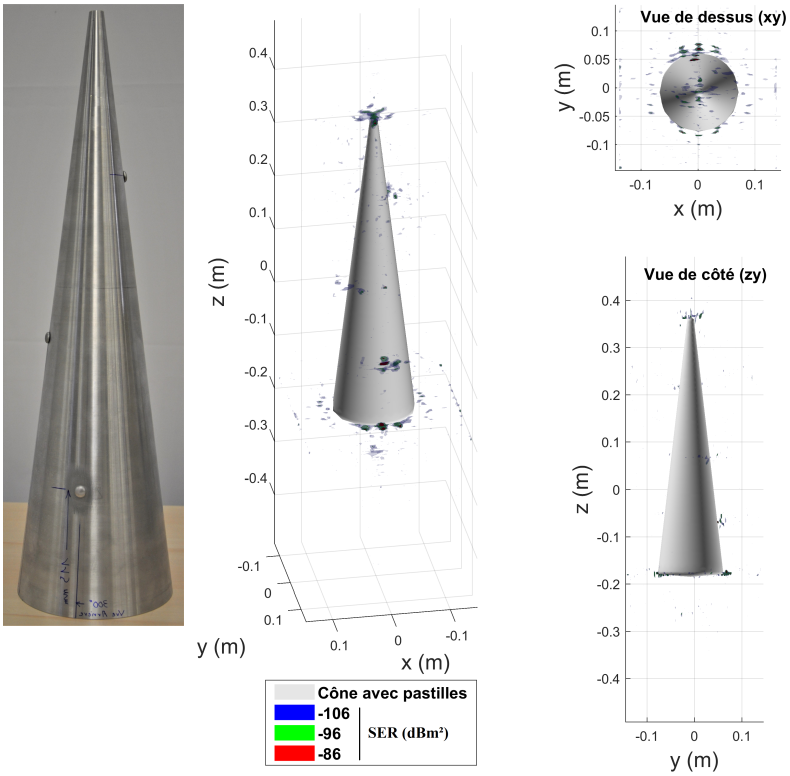


FIGURE 1.28 – Carte de rétro-diffusion d'un cône avec pastilles estimée avec la méthode ℓ_1 f =[8 GHz : 3,9 MHz : 12 GHz], φ = [-20 ° : 1° : 20 °], θ = [-20 ° : 4° : 20 °], polarisation HH, μ =0,01

Récemment, [Lasserre *et al.*, 2016] ont proposé une approche semblable, pour l'imagerie radar doppler 2D, en adoptant une stratégie bayésienne avec un *a priori* laplacien. Ce formalisme offre des perspectives quant au réglage des hyperparamètres.

REMARQUE

> A notre sens, l'emploi des méthodes spectrales et des méthodes de régularisation parcimonieuse sus-citées se justifie en 1D pour estimer les contributeurs du modèle points brillants (1.15). Néanmoins, cela est moins évident dans un contexte multi-dimensionnel. En effet, comme expliqué dans la **Section 1.1.2**, la notion de spéculaires est plutôt naturellement liée à celle de *facettes brillantes* et non pas seulement à celle de *points brillants*.

Dans le sens de cette remarque, [Ciuciu, 2000] (dans le contexte de l'imagerie radar doppler 2D) et [Cetin, 2001; Ren *et al.*, 2016] (dans un contexte d'imagerie SAR 2D), ont proposé des approches bi-modèles parcimonieuses permettant de reconstruire des zones homogènes et des points brillants. Pour contourner les difficultés d'optimisation liées à la non différentiabilité de la norme ℓ_1, ils ont considéré une approximation lisse des termes non différentiables. Le critère est alors minimisé à l'aide de méthodes semi-quadratiques (comme expliqué **Section 2.2.1**). [Zhu, 2013] a proposé une approche similaire dans un formalisme bayésien.

Conclusion du Chapitre 1

Dans un premier temps, nous avons présenté la notion de SER. Il s'agit d'une grandeur qui permet de caractériser le pouvoir rétro-diffuseur d'une cible soumise à un champ EM. Nous avons répertorié les différents mécanismes de rétro-diffusion et les interactions onde-objet qui interviennent en hautes fréquences. Enfin nous avons montré à travers divers exemples qu'il est important d'analyser et de contrôler la SER dans de nombreuses applications.

Dans un second temps, nous avons détaillé le processus de mesure de l'hologramme de SER, introduit la notion d'objet et de carte de rétro-diffusion et formulé le modèle d'observation liant les données à l'objet d'intérêt. Nous avons également présenté le problème d'imagerie de SER en expliquant qu'il s'agit d'un problème de synthèse de Fourier non inversible.

Finalement, nous avons présenté quelques méthodes de l'état de l'art. Tout d'abord les méthodes conventionnelles reposant sur un regridding avec bourrage de zéros et TFRI 3D ; nous avons souligné leurs limites en terme de résolution. Nous avons ensuite expliqué que l'approche problème inverse fournit un cadre rigoureux pour développer des méthodes HR. Enfin, nous avons listé quelques méthodes reposant sur la notion de parcimonie et qui permettent d'obtenir des cartes de rétro-diffusion plus résolues. Ces méthodes reposent sur l'idée de *points brillants* et ne sont pas totalement cohérentes des interactions onde/cible. En effet, les spéculaires s'expriment également comme des *facettes brillantes*.

Dans le **Chapitre 3** nous proposons une méthode permettant d'estimer des cartes plus résolues que celles produites avec les méthodes actuelles et permettant d'améliorer leur adéquation à l'objet à estimer *i.e.* respectant plus la physique des interactions onde-cible. Cette méthode repose également sur une régularisation parcimonieuse. Ainsi, dans le **Chapitre 2**, nous nous intéressons plus spécifiquement aux méthodes de régularisations parcimonieuses et nous présentons quelques algorithmes d'optimisation.

CHAPITRE 2

LES APPROCHES PARCIMONIEUSES EN INVERSION

Dans la première section, nous faisons un bref historique sur la notion de parcimonie. Nous présentons ensuite les pénalités, telle que la norme ℓ_1, qui permettent de favoriser la parcimonie dans les approches d'inversions pénalisées mais conduisent à des critères non différentiables. Dans la seconde section, nous répertorions quelques algorithmes et méthodes d'optimisation des critères non différentiables et contraints. Ces derniers sont classés selon quatre grandes familles : les méthodes de lissage des non différentiabilités, les méthodes de sous-gradients et de sous-gradients projetés, les méthodes proximales et les méthodes primales-duales. Une attention particulière est portée à l'algorithme ADMM (Alternating Direction Method of Multipliers).

Sommaire

2.1 La notion de parcimonie 62
2.1.1 Un bref historique 62
2.1.2 Pénalités favorisant la parcimonie 62
2.2 Des algorithmes de minimisation de critères non différentiables contraints 65
2.2.1 Méthode de lissage de la norme ℓ_1 66
2.2.2 Méthode de sous-gradient et de sous-gradient projeté 68
2.2.3 Méthodes de gradient proximal 70
2.2.4 Méthodes primales-duales 73

REMARQUE

Dans ce chapitre, les définitions et les algorithmes sont présentés dans le cas de variables réelles afin de simplifier les écritures. La plupart des notions introduites ont une extension quasi immédiate dans le cas de variables complexes, l'important étant que le critère soit à valeurs réelles pour conserver notamment les relations d'ordre. Notons qu'il est également possible de se ramener au cas réel en exprimant $\boldsymbol{a} \in \mathbb{C}^N$ comme $(\boldsymbol{a}_{\text{re}}^t, \boldsymbol{a}_{\text{im}}^t)^t \in \mathbb{R}^{2N}$ c'est à dire en formant un vecteur deux fois plus grand où sont concaténées les parties réelles et imaginaires de $\boldsymbol{a}$.

2.1 La notion de parcimonie

2.1.1 Un bref historique

En traitement du signal et de l'image, la *parcimonie* traduit le fait qu'un objet est constitué d'un faible nombre d'éléments. On dit qu'un signal est *parcimonieux au sens strict* si la plupart de ses coefficients sont nuls et qu'il est *parcimonieux au sens large* si la plupart de ses coefficients a une valeur faible [Sbai, 2013] .

Dès les années 1970, dans le domaine des géosciences, [Claerbout et Muir, 1973] ont proposé une méthode de reconstruction parcimonieuse appliquée à la sismographie. Elle repose sur une déconvolution avec une pénalité à norme ℓ_1 (nous détaillons ensuite l'intérêt de cette norme pour favoriser une parcimonie au sens large). Des améliorations ont été proposées ensuite pour mieux gérer le bruit d'observation [Oldenburg *et al.*, 1983; Taylor *et al.*, 1979], et il a été confirmé empiriquement que la norme ℓ_1 permet de favoriser la parcimonie.

Des résultats plus rigoureux ont été apportés dans la fin des années 1980 avec [Donoho et Logan, 1992; Donoho et Stark, 1989]

Au milieu des années 1990, la recherche de solutions parcimonieuses *via* la minimisation de critère à pénalité ℓ_1, s'est élargie à de nombreux domaines. De nombreuses méthodes ont été développées dans ce sens dont : le *LASSO* [Tibshirani, 1996], le *Basis Pursuit* [Chen *et al.*, 1998] et la *variation totale* [Rudin *et al.*, 1992].

Depuis les années 2000, les applications et les recherches au sujet de la minimisation de critères favorisant la parcimonie se sont poursuivies. [Donoho et Huo, 2001; Elad et Bruckstein, 2002; Gribonval et Nielsen, 2003] ont par exemple fourni une analyse plus rigoureuse du Basis Pursuit.

Plus récemment, de nombreuses études en compression de données [Candes *et al.*, 2006; Tsaig et Donoho, 2006], en imagerie médicale [Graff et Sidky, 2015; Lustig *et al.*, 2007] ou encore en radar [Ender, 2010] se sont portées sur le domaine émergent de l'*acquisition comprimée* (ou *compressive sensing* en anglais) [Donoho, 2006]. Il s'agit d'estimer un signal à partir d'un faible nombre d'acquisitions.

Si elle n'est donc pas nouvelle, la notion de *parcimonie* a connu ces dernières décennies un véritable regain d'intérêt dans la communauté du traitement du signal et de l'image.

Nous présentons maintenant plus en détail, et en particulier dans un formalisme déterministe, comment il est possible de favoriser la parcimonie.

2.1.2 Pénalités favorisant la parcimonie

Une norme $\|\cdot\|_p$, $0 \leqslant p \leqslant 1$, permet naturellement de favoriser la parcimonie. Nous illustrons cela dans un contexte bi-dimensionnel sur la Figure 2.1. Sur la figure, le point bleu représente la solution des moindres carrés. Elle n'est pas parcimonieuse car le vecteur $\boldsymbol{a}$ a deux composantes a_1 et a_2 non nulles. Les traits en pointillés représentent quant à eux les iso-contours de $\|\boldsymbol{a}\|_p$ pour $p = 2$, $p = 1$ et $0 < p < 1$. Le point rouge représente le

minimiseur de la norme $\|\boldsymbol{a}\|_p$ sous contrainte que l'attache aux données soit inférieure à un certain seuil ε :

$$\hat{\boldsymbol{a}} = \underset{\|\boldsymbol{\sigma}-\boldsymbol{H}\boldsymbol{a}\|_2^2<\varepsilon}{\operatorname{argmin}} \|\boldsymbol{a}\|_p \tag{2.1}$$

Nous constatons que si $0 \leqslant p \leqslant 1$, le vecteur $\boldsymbol{a}$ a une seule composante non nulle, il est donc plus parcimonieux que lorsque $p = 2$. Comme nous le montrons ensuite, le problème contraint (2.1) peut également s'exprimer sous une forme pénalisée, dont la pénalité est $\|\boldsymbol{a}\|_p$, en introduisant un hyperparamètre μ qui a un rôle analogue à celui du seuil ε.

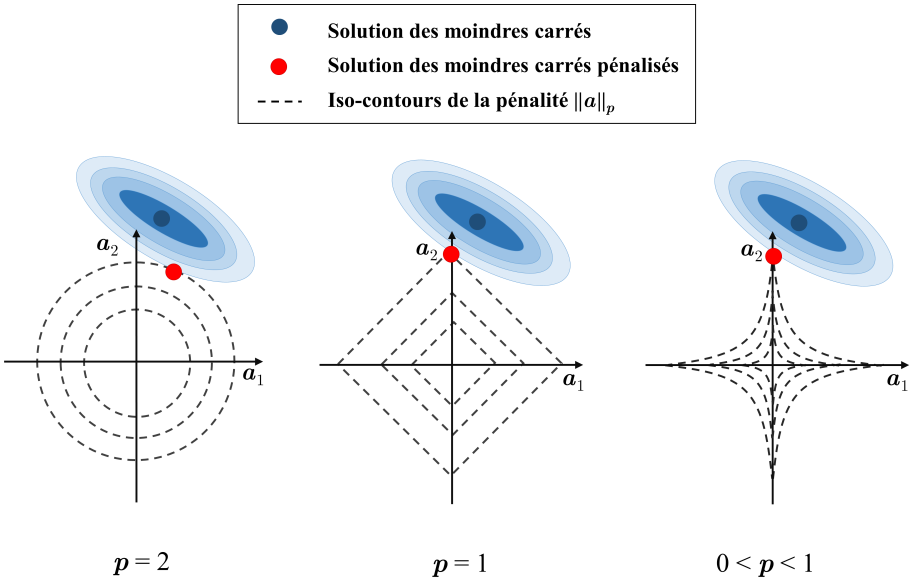


FIGURE 2.1 – Interprétation géométrique de la promotion de la parcimonie pour $0 \leqslant p \leqslant 1$

La pseudo-norme ℓ_0 est définie par :

$$\|\boldsymbol{a}\|_0 = \text{Card}\{i \mid a_i \neq 0\} \tag{2.2}$$

où Card est le cardinal de l'ensemble. Elle « compte » le nombre d'éléments non nuls de $\boldsymbol{a}$.

La minimisation du critère (2.3) suivant permet de trouver une solution parcimonieuse.

$$\mathcal{J}_0(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu \|\boldsymbol{a}\|_0 \tag{2.3}$$

où $\mu > 0$ est l'hyperparamètre contrôlant le compromis entre le terme d'attache aux données et la pénalité favorisant la parcimonie. Cependant, la pseudo-norme ℓ_0 n'est pas convexe et le critère peut être multi-modal. Par ailleurs, la minimisation de (2.3) constitue un problème combinatoire NP-complet [Natarajan, 1995]. De nombreux algorithmes gloutons, que nous ne détaillons pas ici, tels que l'*Orthogonal Least Squares* (OLS) [Chen *et al.*, 1989], le *Matching Pursuit* (MP) [Mallat et Zhifeng Zhang, 1993], l'*Orthogonal Matching Pursuit* (OMP) [Davis *et al.*, 1994] ou encore le *Single Best Replacement* (SBR) [Soussen *et al.*, 2011], permettent de minimiser ce critère de manière approchée. Cependant, et malgré des avancées récentes telles que [Bourguignon *et al.*, 2016] ou [Soubies *et al.*, 2017], la minimisation exacte et rapide du critère (2.3) en grande dimension demeure une question ouverte.

Pour cela, en pratique, quand il s'agit de traiter des problèmes en grande dimension, la pseudo-norme ℓ_0 est souvent relaxée par la norme ℓ_1 [Candes *et al.*, 2006; Fuchs, 2004] et le critère (2.3) s'écrit :

$$\mathcal{J}_1(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu \|\boldsymbol{a}\|_1 . \tag{2.4}$$

Le problème de minimisation s'écrit alors :

$$\min_{\boldsymbol{a}} \mathcal{J}_1(\boldsymbol{a}). \tag{$\mathcal{P}_1$}$$

et une solution du problème ($\mathcal{P}_1$) s'écrit :

$$\hat{\boldsymbol{a}} = \operatorname*{argmin}_{\boldsymbol{a}} \mathcal{J}_1(\boldsymbol{a}) \tag{2.5}$$

Dans la littérature, le problème d'optimisation ($\mathcal{P}_1$) est connu sous le nom de *Basis Pursuit Denoising* (BPDN) [Chen *et al.*, 1998] ou *Least Absolute Shrinkage and Selection Operator* (LASSO) en statistiques [Hastie *et al.*, 2015; Tibshirani, 2011].

Par ailleurs, dans certains cas, l'objet d'intérêt $\boldsymbol{a}$ n'est pas directement parcimonieux mais il existe un opérateur $\boldsymbol{M}$ tel que $\boldsymbol{M}\boldsymbol{a}$ est parcimonieux. Dans ce cas, $\boldsymbol{a}$ peut-être estimé en minimisant le critère :

$$\mathcal{J}_1(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu \|\boldsymbol{M}\boldsymbol{a}\|_1 . \tag{2.6}$$

La matrice $\boldsymbol{M}$ est choisie en fonction de l'information *a priori* considérée. Par exemple, une méthode de régularisation connue sous le nom de *variation totale* [Rudin *et al.*, 1992], consiste à choisir une matrice $\boldsymbol{M}$ égale à un opérateur de différences finies tel qu'un gradient discret. Ceci permet de privilégier des solutions « constantes par morceaux » donnant un aspect « cartoon » aux images estimées en faisant ressortir les contours [Hawe *et al.*, 2012]. En statistiques, cette méthode est aussi connue sous le nom de *fused lasso* [Tibshirani *et al.*, 2005].

REMARQUE

Si $\boldsymbol{M} = \boldsymbol{I}$ (la matrice identité), le critère (2.6) est exactement le critère (2.4).

REMARQUE

Comme indiqué dans l'**Annexe B**, l'idée générale présentée ci-avant, de régulariser un problème en minimisant un critère composite peut également s'interpréter dans le cadre de l'inférence bayésienne. Dans ce contexte, la norme ℓ_1 peut être associée à un *a priori* laplacien [Donoho et Tsaig, 2008]. D'autres lois à queues lourdes telles que certaines lois alpha stables, la loi gaussienne généralisée avec un facteur de forme inférieur à 1 ou la loi de Student, permettent également de favoriser la parcimonie en se rapprochant plus du comportement de la pseudo-norme ℓ_0.
De la même manière, en parallèle du développement des méthodes déterministes mentionnées ci-dessus, des techniques équivalentes ont été proposées dans un formalisme bayésien. Par exemple, en restauration d'images de contours nous pouvons citer les travaux de [Geman et Geman, 1984; Urago *et al.*, 1992] et les études sur les *a priori* Bernoulli-gaussien [Condat, 2013; Kormylo et Mendel, 1983] et Bernoulli-laplacien [Chaari *et al.*, 2013; Costa *et al.*, 2015].

Enfin, si la physique du problème étudié le justifie, il peut être intéressant de contraindre la solution recherchée à appartenir à un certain ensemble. Cette démarche permet généralement d'accroitre sensiblement la résolution. Par exemple, [Henrot, 2013] en imagerie médicale ou [Giovannelli et Coulais, 2005] en astronomie, ont considéré l'usage de contrainte de positivité et de support. Le problème de minimisation général du critère pénalisé et contraint s'écrit alors :

$$\min_{\boldsymbol{a}} \begin{cases} \mathcal{J}_1(\boldsymbol{a}) \\ \text{s.c. } \boldsymbol{a} \in \mathcal{C} \end{cases} \tag{$\mathcal{P}\text{c}_1$}$$

où $\mathcal{C}$ est l'ensemble décrivant les contraintes et une solution du problème ($\mathcal{P}\text{c}_1$) s'écrit :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}\in\mathcal{C}}{\operatorname{argmin}}\, \mathcal{J}_1(\boldsymbol{a}) \tag{2.7}$$

Généralement, la minimisation du critère (2.6) (et *a fortiori* sa version contrainte) présente plusieurs difficultés.

- Ces critères sont non-différentiables à cause des pénalités à norme ℓ_1. Ceci empêche l'utilisation immédiate de méthodes reposant sur un calcul du gradient du critère.
- La minimisation du critère contraint implique que la solution recherchée appartient à un certain ensemble. La détermination de la solution (2.7) peut donc être plus complexe que celle de la solution (2.5).

De plus, une autre difficulté liée plus généralement aux approches d'inversions pénalisées est que le problème de minimisation est souvent posé en grande dimension. Ceci est particulièrement vrai pour des applications réalistes en imagerie 3D comme expliqué dans la **Section 1.2.4** (p. 48).

2.2 Des algorithmes de minimisation de critères non différentiables contraints

Dans la section précédente, nous avons présenté une façon de promouvoir la parcimonie dans un cadre déterministe. Celle-ci repose sur la minimisation d'un critère composite (2.6) éventuellement contraint. Il consiste en un terme d'attache aux données et une pénalité de norme ℓ_1 favorisant la parcimonie. Comme expliqué dans le chapitre suivant, cette idée générale sera exploitée pour construire un critère dont la minimisation permet de résoudre le problème d'imagerie 3D HR de SER.

Dans cette section, nous listons quelques algorithmes, bien connus dans la littérature afférente, qui permettent de calculer les solutions (2.5) et (2.7).

Parmi la grande variété d'algorithmes existants, nous nous focalisons sur les méthodes de premier ordre reposant sur des calculs itératifs de gradients ou de sous-gradients. Ces dernières ont généralement un coût calculatoire par itération raisonnable qui permet de traiter des problèmes en grande dimension. Les techniques d'ordre supérieur telles que le *Second*

Order Cone Programming [Boyd et Vandenberghe, 2004], qui dérivent généralement de la méthode de Newton et reposent sur le calcul de l'inverse du Hessien, ne sont pas présentées dans ce document.

Dans un premier temps, nous présentons une méthode consistant à lisser la norme ℓ_1 du critère (2.6) afin de le rendre différentiable. Nous présentons ensuite des méthodes de sous-gradient puis de gradient proximal. Enfin, afin de tenir compte des contraintes dans le problème ($\mathcal{P}c_1$), nous détaillons une classe de méthodes appelées *primales-duales*.

Cette étude ne se veut pas exhaustive et seules les principales caractéristiques des méthodes sont soulignées. Nous incitons donc le lecteur à se référer aux articles cités afin d'obtenir de plus amples précisions, notamment sur les preuves de convergence. A noter que l'ouvrage très récent de [Beck, 2017] détaille plusieurs des méthodes décrites ci-après.

2.2.1 Méthode de lissage de la norme ℓ_1

Nous rappelons que l'objectif est de résoudre le problème ($\mathcal{P}_1$) et sa version contrainte ($\mathcal{P}c_1$). Nous cherchons donc à minimiser le critère (2.6) rappelé ci-dessous :

$$\mathcal{J}_1(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu \|\boldsymbol{M}\boldsymbol{a}\|_1 .$$

Comme expliqué précédemment, une des principales difficultés est la non-différentiabilité de ce critère.

Une première approche, proposée notamment par [Bertsekas, 1975], consiste à modifier le critère non lisse initial $\mathcal{J}_1$ en lissant ses parties non-différentiables. Nous nous intéressons alors à un nouveau critère $\mathcal{J}_\epsilon$ qui doit être construit de façon à ce que pour tout $\boldsymbol{a}$:

$$\begin{cases} \lim\limits_{\epsilon\to 0} \mathcal{J}_\epsilon(\boldsymbol{a}) = \mathcal{J}_1(\boldsymbol{a}) \\ \text{et } \lim\limits_{\epsilon\to 0} \mathcal{J}_\epsilon(\hat{\boldsymbol{a}}_\epsilon) \geqslant \mathcal{J}_1(\hat{\boldsymbol{a}}) \end{cases} \tag{2.8}$$

avec $\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}\in\mathcal{C}}{\operatorname{argmin}}\, \mathcal{J}_1(\boldsymbol{a})$, $\hat{\boldsymbol{a}}_\epsilon = \underset{\boldsymbol{a}\in\mathcal{C}}{\operatorname{argmin}}\, \mathcal{J}_\epsilon(\boldsymbol{a})$ et $\epsilon > 0$. En effet, [Glowinski *et al.*, 1976] ont montré qu'alors :

$$\lim_{\epsilon\to 0} \hat{\boldsymbol{a}}_\epsilon = \hat{\boldsymbol{a}}. \tag{2.9}$$

La valeur absolue dans le critère $\mathcal{J}_1$ est généralement approchée par une fonction différentiable convexe qui vérifie (2.8). Classiquement, on choisit une des deux fonctions suivantes :

- la fonction hyperbolique (utilisée par exemple par [Ciuciu, 2000]) :

$$\varphi_{\text{hyp}_\epsilon}(a) = \sqrt{|a|^2 + \epsilon^2}.$$

- le potentiel de Huber [Huber, 1996] (utilisé par exemple par [Henrot, 2013]). Il est défini par morceaux :

$$\varphi_{\text{hub}_\epsilon}(a) = \begin{cases} a^2 & \text{si } |a| \leqslant \epsilon \\ 2|a|\epsilon - \epsilon^2 & \text{sinon} \end{cases}$$

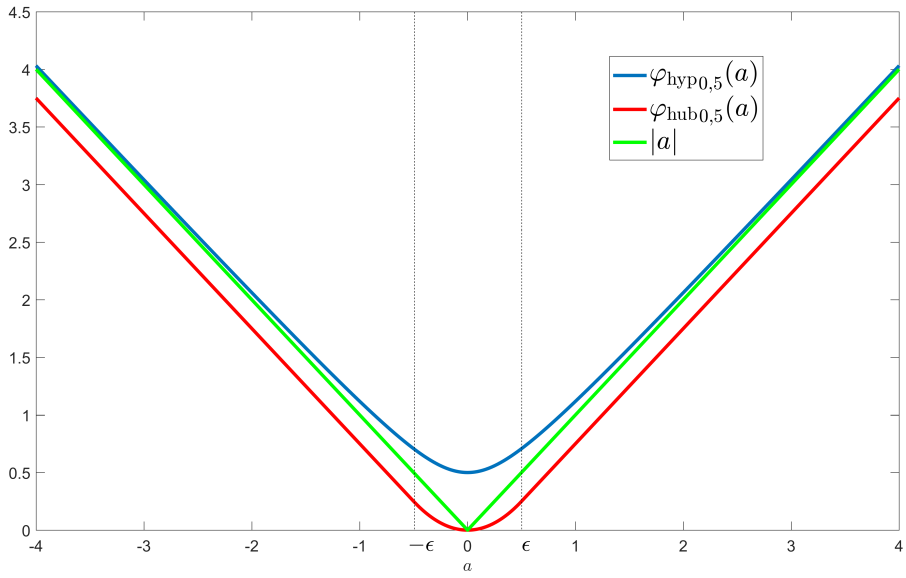


FIGURE 2.2 – Évolution de $\varphi_{\mathrm{hyp}_\epsilon}(a)$ et $\varphi_{\mathrm{hub}_\epsilon}(a)$ pour $\epsilon = 0{,}5$

Notons que ces deux fonctions ont un comportement quadratique autour de l'origine et linéaire quand $|a|$ est grand devant ϵ, comme illustré sur la Figure 2.2.

La minimisation du critère $\mathcal{J}_\epsilon$ peut alors être menée par des techniques rapides développées pour des fonctions lisses : généralement une descente de gradient. Néanmoins, des études menées notamment par [Vogel et Oman, 1996], ont montré que cette méthode n'est pas très performante quand il s'agit de minimiser des critères non-quadratiques tel que $\mathcal{J}_\epsilon$. Il est alors plus efficace de construire un *critère augmenté* $\mathcal{K}_\epsilon(\boldsymbol{a}, \boldsymbol{\beta})$ en introduisant des variables auxiliaires $\boldsymbol{\beta}$. La construction de tels critères doit cependant respecter les conditions suivantes :

- la minimisation de $\mathcal{K}_\epsilon$ par rapport à $\boldsymbol{a}$ et $\boldsymbol{\beta}$ doit être équivalente à celle de $\mathcal{J}_\epsilon$, *i.e.* le minimiseur $\tilde{\boldsymbol{a}}$ de $\mathcal{K}_\epsilon$ doit être le même que celui qui minimise $\mathcal{J}_\epsilon$,
- la minimisation de $\mathcal{K}_\epsilon$ par rapport à $\boldsymbol{a}$ doit être plus rapide que celle de $\mathcal{J}_\epsilon$,
- la minimisation de $\mathcal{K}_\epsilon$ par rapport à $\boldsymbol{\beta}$ doit être simple et rapide à calculer.

De nombreuses publications telles que [Blake et Zisserman, 1987], [Geman et Reynolds, 1992], [Geman et Yang, 1995] ou encore [Idier, 2001] ont montré que des critères *semi-quadratiques* (*i.e.* qui sont quadratiques vis-à-vis de $\boldsymbol{a}$ et dont la minimisation est séparable vis-à-vis de $\boldsymbol{\beta}$), répondent aux conditions énoncées ci-dessus. Le lecteur est invité à consulter ces articles pour obtenir de plus amples informations.

Enfin, dans le même esprit, [Nesterov, 2005] puis [Weiss *et al.*, 2009] ont proposé plus récemment d'autres approximations lisses du critère non différentiable initial. La minimisation est ensuite assurée par une méthode de gradient projeté accélérée.

Pour conclure, cette approche consistant à lisser les parties non-différentiables du critère initial présente l'avantage de se ramener à des fonctions lisses pour lesquelles la littérature fournit de très nombreux algorithmes de minimisation. La faiblesse principale de cette démarche est que le nouveau critère à minimiser est une approximation du critère (2.6) initial.

Cette approximation dépend d'un paramètre ϵ qui peut être difficile à régler. Ce dernier doit toutefois être suffisament petit pour continuer à favoriser des solutions parcimonieuses. Or, quand ϵ tend vers 0, les algorithmes itératifs de minimisation du critère lisse $\mathcal{J}_\epsilon$ ou du critère augmenté $\mathcal{K}_\epsilon$ peuvent converger lentement vers une solution.

2.2.2 Méthode de sous-gradient et de sous-gradient projeté

Les méthodes de sous-gradient et de sous-gradient projeté [Poliyak, 1987] permettent de résoudre des problèmes d'optimisation convexe non différentiable sans approximation du critère initial. Elles s'inspirent des méthodes de descente de gradient et de gradient projeté développées pour des critères différentiables.

Sous-gradient

De manière générale, la méthode de sous-gradient permet de chercher une solution de la forme :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{J}(\boldsymbol{a}) \tag{2.10}$$

où $\mathcal{J}$ est une fonction L-lipschitzienne, convexe et fermée [1] qui admet un ensemble de minimiseurs $\widetilde{\mathcal{A}} \neq \emptyset$. Nous supposons pour l'instant que le critère (2.6) est une telle fonction. L'algorithme de sous-gradient est donné par l'**Algorithme 1**.

Algorithme 1 : Algorithme de sous-gradient

Entrées : $k = 0$: initialisation du compteur ;
$\boldsymbol{a}^{(0)} \in \mathcal{A}$: le point de départ
Sorties : $\hat{\boldsymbol{a}}$: un élément de $\widetilde{\mathcal{A}}$

Répéter jusqu'à convergence
- Calculer $\boldsymbol{s}^{(k)} \in \partial\mathcal{J}(\boldsymbol{a}^{(k)})$
- Chercher un pas $\tau^{(k)} > 0$ tel que $\mathcal{J}\left(\boldsymbol{a}^{(k)} - \tau^{(k)} \frac{\boldsymbol{s}^{(k)}}{\|\boldsymbol{s}^{(k)}\|_2}\right) < \mathcal{J}(\boldsymbol{a}^{(k)})$
- $\boldsymbol{a}^{(k+1)} = \boldsymbol{a}^{(k)} - \tau^{(k)} \frac{\boldsymbol{s}^{(k)}}{\|\boldsymbol{s}^{(k)}\|_2}$
- $k = k + 1$

Contrairement à une descente de gradient classique, le sous-gradient est ici normalisé. En effet, dans le cas différentiable, le gradient tend vers 0 lorsque nous nous rapprochons du minimiseur alors que ce n'est plus le cas pour une fonction non différentiable. De plus, dans le cas non-différentiable, une direction opposée à un sous-gradient n'est pas forcément une direction de descente. Même si l'**Algorithme 1** n'est pas un algorithme de descente, un choix adéquat des $\tau^{(k)}$ permet de rendre la méthode convergente. Ainsi, si la suite des $\tau^{(k)}$ vérifie :

1. Les définitions de ces propriétés, du sous-gradient et du sous-différentiel $\partial\mathcal{J}$ sont rappelées dans l'**Annexe E** (p. 183).

$$\sum_{k=0}^{+\infty} \tau^{(k)} = +\infty \quad \text{et} \quad \sum_{k=0}^{+\infty} \left(\tau^{(k)}\right)^2 < +\infty$$

alors la suite des $\boldsymbol{a}^{(k)}$ converge vers un minimiseur de $\mathcal{J}$ [Weiss, 2008]. Un choix simple qui satisfait les deux conditions est $\tau^{(k)} = 1/\sqrt{k}$ mais la convergence peut-être lente car le pas tend vers 0.

Sous-gradient projeté

De manière générale, la méthode de sous-gradient projeté permet de gérer les contraintes et de chercher une solution de la forme :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a} \in \mathcal{C}}{\operatorname{argmin}}\, \mathcal{J}(\boldsymbol{a}) \tag{2.11}$$

où $\mathcal{C}$ est un ensemble convexe décrivant les contraintes (aussi appelé *domaine admissible*) et $\mathcal{J}$ est toujours une fonction L-lipschitzienne, convexe et fermée. L'algorithme de sous-gradient projeté est donné par l'**Algorithme 2**.

Algorithme 2 : Algorithme de sous-gradient projeté

Entrées : $k = 0$: initialisation du compteur ;
$\boldsymbol{a}^{(0)} \in \mathcal{A}$: le point de départ
Sorties : $\hat{\boldsymbol{a}}$: un élément de $\widetilde{\mathcal{A}}$

Répéter jusqu'à convergence
- Calculer $\boldsymbol{s}^{(k)} \in \partial\mathcal{J}(\boldsymbol{a}^{(k)})$
- Chercher un pas $\tau^{(k)} > 0$ tel que $\mathcal{J}\left(\boldsymbol{a}^{(k)} - \tau^{(k)} \dfrac{\boldsymbol{s}^{(k)}}{\|\boldsymbol{s}^{(k)}\|_2}\right) < \mathcal{J}(\boldsymbol{a}^{(k)})$
- $\boldsymbol{a}^{(k+1)} = \mathcal{P}_{\mathcal{C}}\left(\boldsymbol{a}^{(k)} - \tau^{(k)} \dfrac{\boldsymbol{s}^{(k)}}{\|\boldsymbol{s}^{(k)}\|_2}\right)$
- $k = k + 1$

$\mathcal{P}_{\mathcal{C}}$ est l'opérateur de projection euclidienne sur le domaine admissible $\mathcal{C}$. Cette projection est facile à calculer dans le cas de contraintes simples telles qu'une contrainte de support de type « boîte »[2], mais elle peut être plus difficile à calculer dans d'autres cas. L'application de cet algorithme à des contraintes plus générales ainsi qu'une étude sur le pas a été réalisée par [Weiss, 2008].

Pour conclure, [Carlavan *et al.*, 2010] indique que les méthodes de sous-gradient sont essentiellement intéressantes pour des applications où le temps de calcul est déterminant car une dizaine d'itérations avec des pas pré-calculés permet généralement d'obtenir une approximation grossière de la solution. Cependant, les auteurs précisent que ces méthodes sont à éviter pour des applications où la précision est primordiale.

2. définie par (E.2) (p. 186).

2.2.3 Méthodes de gradient proximal

Les méthodes de gradient proximal, étudiées notamment par [Lions et Mercier, 1979], permettent de déterminer une solution de la forme :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{J}(\boldsymbol{a}) \tag{2.12}$$

où $\mathcal{J}$ peut être décomposé de la manière suivante : $\mathcal{J}(\boldsymbol{a}) = f(\boldsymbol{a}) + g(\boldsymbol{a})$ avec :

- f est convexe, fermée, différentiable et de gradient ∇f L-Lipschitz,
- g est convexe, fermée, possiblement non différentiable, telle que $\operatorname{dom}(g) \neq \emptyset$[3].

Nous notons $\widetilde{\mathcal{A}}$ l'ensemble des minimiseurs de $\mathcal{J}$, que nous supposons non vide, et $\widetilde{\boldsymbol{a}}$ un élément de cet ensemble. Nous pouvons nous ramener au critère (2.6) en posant :

$$f(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 \;\text{ et }\; g(\boldsymbol{a}) = \mu \|\boldsymbol{M}\boldsymbol{a}\|_1.$$

Forward-Backward splitting

[Combettes et Wajs, 2005] ont montré que pour tout pas τ positif non nul, les conditions d'optimalité de $\mathcal{J}$ s'écrivent :

$$\begin{aligned} 0 \in \partial g(\widetilde{\boldsymbol{a}}) + \nabla f(\widetilde{\boldsymbol{a}}) &\Leftrightarrow -\nabla f(\widetilde{\boldsymbol{a}}) \in \partial g(\widetilde{\boldsymbol{a}}) \\ &\Leftrightarrow \widetilde{\boldsymbol{a}} - \tau \nabla f(\widetilde{\boldsymbol{a}}) - \widetilde{\boldsymbol{a}} \in \tau \partial g(\widetilde{\boldsymbol{a}}) \\ &\Leftrightarrow \widetilde{\boldsymbol{a}} = \operatorname{prox}_{\tau g}\left(\widetilde{\boldsymbol{a}} - \tau \nabla f(\widetilde{\boldsymbol{a}})\right) \end{aligned}$$

Le passage de l'avant dernière ligne à la dernière s'explique par la définition de l'opérateur proximal prox qui est rappelée dans l'**Annexe E** (p. 183). L'algorithme FBS (Forward Backward Splitting) est alors donné par l'**Algorithme 3**.

Algorithme 3 : Algorithme FBS

Entrées : $k = 0$: initialisation du compteur ;
$\boldsymbol{a}^{(0)} \in \operatorname{dom}(g)$: le point de départ

Sorties : $\hat{\boldsymbol{a}}$: un élément de $\widetilde{\mathcal{A}}$

Répéter jusqu'à convergence

$$\boldsymbol{a}^{(k+1)} = \underbrace{\operatorname{prox}_{\tau g}}_{\text{étape backward}} \Big(\underbrace{\boldsymbol{a}^{(k)} - \tau \nabla f(\boldsymbol{a}^{(k)})}_{\text{étape forward}}\Big)$$

$$k = k + 1$$

L'algorithme itère deux étapes :

- l'étape explicite *forward* consiste en une simple descente de gradient,
- l'étape implicite *backward* est une étape proximale.

3. La définition du domaine effectif dom est rappelée dans l'**Annexe E** (p. 183).

REMARQUE

- Si g est la fonction nulle, l'**Algorithme 3** se résume à un algorithme de descente de gradient.
- Si g est une norme ℓ_1, l'unique étape de l'**Algorithme 3** se réduit à une descente de gradient à pas fixe suivie d'un seuillage doux [a]. Dans ce cas, l'algorithme FBS est appelé ISTA (Iterative Soft Thresholding Algorithm).
- Si f est la fonction nulle, l'**Algorithme 3** se résume à un algorithme de point proximal.

a. Le seuillage doux est défini dans l'**Annexe E** (p. 183)

Si cette méthode présente l'avantage d'avoir un faible coût calculatoire par itération elle présente également plusieurs faiblesses. Premièrement, elle ne s'applique qu'à une classe restreinte de problème car il faut que f soit différentiable et il faut être en mesure de calculer facilement $\mathrm{prox}_{\tau g}$. Enfin, [Beck et Teboulle, 2009b] ont montré que pour tout pas τ strictement positif et inférieur à $2/L$ (où L est la constante de Lipschitz de ∇f), l'algorithme converge faiblement vers un minimiseur $\widetilde{\boldsymbol{a}}$ avec une vitesse de convergence sous-linéaire en $\mathcal{O}\big(1/k\big)$. Ceci peut être un frein à de nombreuses applications concrètes en grande dimension.

Fast Iterative Soft Thresholding Algorithm

La méthode FISTA (Fast Iterative Soft Thresholding Algorithm) proposée par [Beck et Teboulle, 2009b] est une version inertielle du FBS qui repose sur les travaux menés par [Nesterov, 1983] pour des fonctions lisses. Par souci de simplicité, nous utilisons la même notation que [Beck et Teboulle, 2009b] :

$$\begin{aligned} \mathrm{p}_L(\boldsymbol{y}) &\triangleq \mathrm{prox}_{(1/L)g}\left(\boldsymbol{y} - \frac{1}{L}\nabla f(\boldsymbol{y})\right) \\ &= \underset{\boldsymbol{a}}{\mathrm{argmin}} \left(g(\boldsymbol{a}) + \frac{L}{2}\left\|\boldsymbol{a} - \left(\boldsymbol{y} - \frac{1}{L}\nabla f(\boldsymbol{y})\right)\right\|_2^2 \right) \end{aligned}$$

L'algorithme FISTA est donné par l'**Algorithme 4**. Contrairement à FBS où l'opérateur proximal est appliqué à $\boldsymbol{a}^{(k)}$, dans FISTA, l'opérateur proximal est appliqué à $\boldsymbol{y}^{(k)}$ qui résulte d'une combinaison linéaire des deux itérations précédentes $\boldsymbol{a}^{(k)}$ et $\boldsymbol{a}^{(k-1)}$.

[Beck et Teboulle, 2009b, Lemme 4.1] ont montré que FISTA peut atteindre une vitesse de convergence en $\mathcal{O}\big(1/k^2\big)$ [4] pour toutes les sur-relaxations du FBS telles que la suite des $\tau^{(k)}$ satisfait :

$$\tau^{(1)} = 1 \text{ et } \left(\tau^{(k+1)}\right)^2 - \tau^{(k+1)} \leqslant \left(\tau^{(k)}\right)^2 \text{ pour } k \geqslant 1. \tag{2.13}$$

Notons que le choix de [Beck et Teboulle, 2009b] pour l'**Algorithme 4** consiste à prendre un $t^{(k+1)}$ vérifiant l'égalité dans l'inégalité large (2.13).

Plus récemment, [Chambolle et Dossal, 2015, Théorème 3] ont apporté la preuve que la suite des $\boldsymbol{a}^{(k)}$ donnée par FISTA converge faiblement vers un minimiseur $\widetilde{\boldsymbol{a}}$ de $\mathcal{J}$ si pour tout $\vartheta > 2$:

$$\tau^{(k)} = \left(k + \vartheta - 1\right)/\vartheta. \tag{2.14}$$

4. Ce qui est optimal pour une méthode d'ordre 1 d'après [Nesterov, 1983].

Algorithme 4 : Algorithme FISTA

Entrées : L : un majorant de la constante de Lipschitz de ∇f ;
$k = 0$: initialisation du compteur ;
$\boldsymbol{y}^{(1)} = \boldsymbol{a}^{(0)} \in \mathrm{dom}(g)$: les points de départ ;
$\tau^{(1)} = 1$: le pas initial
Sorties : $\hat{\boldsymbol{a}}$: un élément de $\widetilde{\mathcal{A}}$

Répéter jusqu'à convergence

$$\boldsymbol{a}^{(k)} = \mathrm{p}_L(\boldsymbol{y}^{(k)})$$
$$\tau^{(k+1)} = \left(1 + \sqrt{1 + 4\left(\tau^{(k)}\right)^2}\right)/2$$
$$\boldsymbol{y}^{(k+1)} = \boldsymbol{a}^{(k)} + \frac{\tau^{(k)} - 1}{\tau^{(k+1)}}\left(\boldsymbol{a}^{(k)} - \boldsymbol{a}^{(k-1)}\right)$$
$$k = k + 1$$

REMARQUE

Si $g = \mathcal{I}_{\mathcal{C}}(\boldsymbol{a})$, l'étape $\boldsymbol{a}^{(k)} = \mathrm{p}_L(\boldsymbol{y}^{(k)})$ dans l'**Algorithme 4** s'écrit désormais $\boldsymbol{a}^{(k)} = \mathcal{P}_{\mathcal{C}}\left(\boldsymbol{y}^{(k)} - \frac{1}{L}\nabla f(\boldsymbol{y}^{(k)})\right)$ [a]. L'algorithme FISTA correspond alors exactement à la méthode FGP (Fast Gradient Projection) développée par [Nesterov, 1983] pour des problèmes de minimisation sous contrainte impliquant des critères $\mathcal{J}$ lisses et convexes.

a. $\mathcal{I}_{\mathcal{C}}$ est la fonction indicatrice de l'ensemble convexe $\mathcal{C}$ et $\mathcal{P}_{\mathcal{C}}$ l'opérateur de projection sur $\mathcal{C}$ qui sont définis par (E.1) (p. 185).

Tout comme l'algorithme FBS dont il dérive, l'algorithme FISTA présente quelques faiblesses. Premièrement, il ne s'applique qu'à des problèmes où au moins une partie du critère est différentiable. Cette hypothèse est restrictive dans le cas où l'on souhaite résoudre, par exemple, un problème où le terme d'attache aux données et la pénalité sont tous les deux non différentiables (typiquement une norme ℓ_1).

De plus, si g est une fonction non lisse, nous constatons que les étapes de l'**Algorithme 4** ne permettent pas de prendre en compte des contraintes. Ainsi, l'algorithme FISTA ne peut pas naturellement résoudre des problèmes contraints [Beck et Teboulle, 2009a]. Pour résoudre ce genre de problèmes, [Beck et Teboulle, 2009a] ont proposé une méthode qui repose sur une approche duale lagrangienne, initialement étudiée par [Chambolle, 2004]. Cette méthode consiste à construire le dual du problème contraint initial afin de se ramener à un problème d'optimisation convexe quadratique (maximisation d'une fonction quadratique concave sous contraintes linéaires) qui peut être résolu rapidement à l'aide de l'algorithme FGP.

Pour conclure, nous pouvons noter que d'autres versions plus sophistiquées de FISTA ont été développées par la suite pour pallier certains problèmes. Parmi les plus utilisées dans la littérature nous pouvons citer les suivantes.

- MFISTA (Monotone Fast Iterative Soft Thresholding Algorithm) [Beck et Teboulle, 2009a] qui est une version monotone de FISTA où une étape d'évaluation du critère $\mathcal{J}$ est ajoutée. En effet, FISTA n'est pas un algorithme de descente contrairement à FBS, et en pratique la suite des $\mathcal{J}(\boldsymbol{a}^{(k)})$ est rarement décroissante.
- FISTA avec retour sur trace (ou *backtracking* en anglais) [Beck et Teboulle, 2009b], quand la constante de Lipschitz L n'est pas connue ou difficilement calculable ce qui est notamment le cas des problèmes en grande dimension impliquant la norme ℓ_1.

2.2.4 Méthodes primales-duales

Dans les sections précédentes nous avons présenté quelques méthodes permettant de minimiser des critères non différentiables. Ces dernières ne sont pas totalement adaptées au problème contraint ($\mathcal{P}\mathrm{c}_1$) posé en grande dimension et donc à la minimisation du critère que nous proposons dans le prochain chapitre pour résoudre le problème d'imagerie 3D HR de SER.

En effet, la première méthode présentée consiste en une approximation du critère initial par lissage de la norme ℓ_1. Cependant, nous souhaitons justement conserver ces singularités afin de continuer à favoriser des solutions parcimonieuses.

La méthode de sous-gradient projeté est, quant à elle, à privilégier si l'objectif est l'obtention d'une approximation grossière de la solution. Cependant, en grande dimension, elle ne permet généralement pas d'atteindre une précision satisfaisante dans un temps raisonnable.

Les méthodes proximales telles que l'algorithme FISTA nécessitent l'évaluation d'un gradient à chaque itération. La version initiale de FISTA ne permet pas de traiter les problèmes contraints ($\mathcal{P}\mathrm{c}_1$). Pour cela, et comme expliqué avant, la littérature préconise de « mixer » FISTA avec des méthodes duales.

Dans cette partie, nous présentons une autre classe de méthodes appelées *méthodes primales-duales*. Dans un premier temps, nous introduisons quelques résultats généraux mais fondamentaux afin de préciser le fonctionnement de ces méthodes.

Dans un second temps, nous nous focalisons sur l'ADMM qui est un algorithme appartenant à cette classe de méthodes. Nous montrons qu'il est particulièrement adapté à la résolution du problème contraint ($\mathcal{P}\mathrm{c}_1$) et donc à la minimisation du critère que nous proposons dans le chapitre suivant.

Dans la suite, et par souci de simplicité pour ne pas alourdir les notations, nous présentons les résultats dans un contexte de variables réelles. Ils ont bien entendu été étendus aux variables complexes (en particulier, les transposées qui apparaissent avec la norme euclidienne deviennent des transposées conjuguées pour une norme hermitienne). Des preuves rigoureuses de validité de ces algorithmes pour des variables complexes peuvent être trouvées par exemple dans [Li *et al.*, 2015b].

Pour rédiger cette partie nous nous sommes appuyés, entre autres, sur [Boyd et Vandenberghe, 2004] qui est un ouvrage de référence sur l'ADMM.

2.2.4.1 Notions de dualité et notations

Problème primal

Nous considérons le problème général de minimisation sous contrainte d'égalité suivant :

$$\min_{\boldsymbol{a}} \left\{ \begin{array}{l} \mathcal{J}(\boldsymbol{a}) \\ \text{s.c. } \boldsymbol{A}\boldsymbol{a} = \boldsymbol{c} \end{array} \right. \tag{$\mathrm{P_p}$}$$

où $\boldsymbol{A}$ est une matrice et $\boldsymbol{c}$ le vecteur des contraintes d'égalité. Le critère $\mathcal{J}$ est convexe. ($\mathrm{P_p}$) est appelé *problème primal* car c'est celui que nous souhaitons résoudre initialement.

Lagrangien et fonction duale

Le lagrangien $\mathcal{L}$ pour le problème ($\mathrm{P_p}$) est la fonction définie par :

$$\mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) \triangleq \mathcal{J}(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c}) \tag{2.15}$$

où $\boldsymbol{\gamma}$ est la *variable duale* appelée *multiplicateur de Lagrange*. Notons que pour tout $\boldsymbol{a}$, $\mathcal{L}$ est une fonction affine de $\boldsymbol{\gamma}$.

La fonction duale Θ, définie sur $\Gamma = \{\boldsymbol{\gamma} | \inf_{\boldsymbol{a}} \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) > -\infty\}$, s'écrit :

$$\begin{aligned} \Theta(\boldsymbol{\gamma}) &\triangleq \inf_{\boldsymbol{a}} \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) \\ &= -\mathcal{J}^*(-\boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma}) - \boldsymbol{c}^{\mathrm{t}}\boldsymbol{\gamma} \end{aligned} \tag{2.16}$$

où $\mathcal{J}^*$ est la conjuguée convexe de $\mathcal{J}$ rappelée dans l'**Annexe E** (p. 185). Par ailleurs, en tant qu'enveloppe inférieure d'une famille de fonctions affines, $\Theta(\boldsymbol{\gamma})$ est concave sur l'ensemble Γ. Une preuve de l'égalité dans (2.16) et de la concavité de la fonction duale est donnée dans l'**Annexe D** (p. 165).

Optimum primal et dual

Nous notons $\widetilde{p}$ la valeur du primal optimal (*i.e.* la valeur que prend $\mathcal{J}(\boldsymbol{a})$ quand le problème primal est résolu) :

$$\begin{aligned} \widetilde{p} &\triangleq \inf_{\boldsymbol{a}} \sup_{\boldsymbol{\gamma} \succcurlyeq 0} \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) \\ &= \inf_{\boldsymbol{a}} \{\mathcal{J}(\boldsymbol{a}) | \boldsymbol{A}\boldsymbol{a} \preccurlyeq \boldsymbol{c}\} \end{aligned} \tag{2.17}$$

Et nous notons $\widetilde{d}$ la valeur du dual optimal (*i.e.* le maximum de la fonction Θ) :

$$\begin{aligned} \widetilde{d} &\triangleq \sup_{\boldsymbol{\gamma}} \inf_{\boldsymbol{a}} \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) \\ &= \sup_{\boldsymbol{\gamma}} \Theta(\boldsymbol{\gamma}) \end{aligned} \tag{2.18}$$

Dualité

Les résultats présentés sur la dualité reposent essentiellement sur la notion de point selle et sur le théorème Min-Max Inégalité qui sont présentés dans l'**Annexe E** (p. 183). En partant de l'expression (2.16), nous pouvons écrire pour tout $\boldsymbol{\gamma}$ dans l'ensemble Γ et pour tout $\boldsymbol{a}$:

$$\Theta(\boldsymbol{\gamma}) \leqslant \mathcal{J}(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c})$$

et *a fortiori*, quand la contrainte est respectée (*i.e.* $\boldsymbol{A}\boldsymbol{a} = \boldsymbol{c}$), nous avons :

$$\Theta(\boldsymbol{\gamma}) \leqslant \mathcal{J}(\boldsymbol{a}) \tag{2.19}$$

Nous déduisons alors de l'inégalité (2.19) que :

$$\sup_{\boldsymbol{\gamma}} \Theta(\boldsymbol{\gamma}) \leqslant \inf_{\boldsymbol{a}} \{\mathcal{J}(\boldsymbol{a}) \mid \boldsymbol{A}\boldsymbol{a} = \boldsymbol{c}\} \tag{2.20}$$

Par ailleurs, en exploitant (2.15) nous avons :

$$\sup_{\boldsymbol{\gamma} \succcurlyeq \mathbf{0}} \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) = \sup_{\boldsymbol{\gamma} \succcurlyeq \mathbf{0}} \{\mathcal{J}(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c})\} = \begin{cases} \mathcal{J}(\boldsymbol{a}) & \text{si } (\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c}) \preccurlyeq \mathbf{0} \text{ ou si } \boldsymbol{\gamma} = \mathbf{0} \\ +\infty & \text{sinon} \end{cases} \tag{2.21}$$

Finalement, nous déduisons des expressions (2.16) – (2.21) que :

$$\begin{gathered} \sup_{\boldsymbol{\gamma}} \inf_{\boldsymbol{a}} \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) \leqslant \inf_{\boldsymbol{a}} \sup_{\boldsymbol{\gamma} \succcurlyeq \mathbf{0}} \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) \\ \widetilde{d} \leqslant \widetilde{p} \end{gathered} \tag{2.22}$$

Cette inégalité large est toujours vérifiée. Elle est connue sous le nom de *dualité faible*. La différence $\widetilde{d} - \widetilde{p}$ est appelée *saut de dualité*. Si le saut de dualité est nul (*i.e.* $\widetilde{d} = \widetilde{p}$) on dit que la *dualité est forte* et il existe un couple $(\widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{\gamma}})$ qui est un point selle de $\mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma})$. La Figure 2.3 est une illustration de l'existence d'un point selle dans le cas d'une dualité forte.

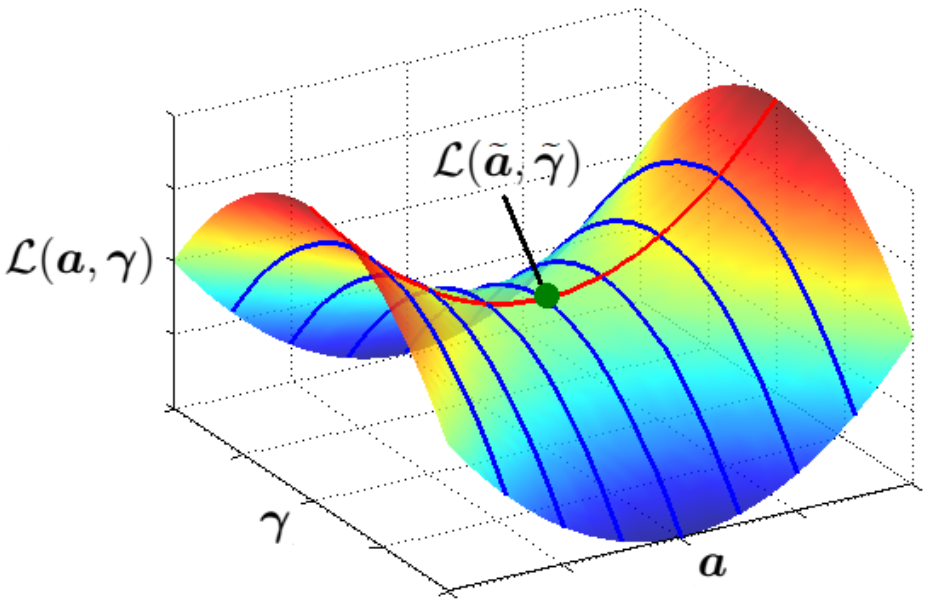


FIGURE 2.3 – Illustration point selle

En optimisation, la *qualification de contraintes* est l'étude des conditions d'optimalité. Sans rentrer dans les détails, nous indiquons au lecteur intéressé que les conditions de Karush,

Kuhn et Tucker (KKT) [Kuhn et Tucker, 1951] ainsi que le théorème de Slater [Slater, 1950] fournissent des conditions nécessaires et suffisantes sur $\mathcal{J}$ et sur les contraintes pour que la dualité soit forte. L'exemple classique dans lequel le théorème de Slater est vérifié et les conditions KKT sont respectées, ce qui assure une dualité forte, est le cas où le critère $\mathcal{J}$ est convexe et les contraintes d'égalité sont affines.

Problème dual

Le *problème dual* s'écrit :

$$\hat{\boldsymbol{\gamma}} = \underset{\boldsymbol{\gamma}}{\operatorname{argmax}}\, \Theta(\boldsymbol{\gamma}) \tag{P_d}$$

- Si la **dualité est forte** : comme $\Theta(\boldsymbol{\gamma})$ est concave, résoudre (P_d) (ce qui signifie chercher $\widetilde{d} = \sup_{\boldsymbol{\gamma}} \Theta(\boldsymbol{\gamma})$) permet également de chercher $\widetilde{p} = \inf_{\boldsymbol{a}} \{\mathcal{J}(\boldsymbol{a}) | \boldsymbol{A}\boldsymbol{a} = \boldsymbol{c}\}$ et donc de résoudre le problème primal (P_p).
- Si la **dualité est faible** : chercher $\widetilde{d}$ permet de trouver la meilleure (plus grande) borne inférieure de $\widetilde{p}$ et donc d'approcher la résolution de (P_p).

Par ailleurs, nous pouvons noter que si :

- $\widetilde{d} = +\infty$, alors (P_p) n'admet pas de solution.
- $\widetilde{p} = -\infty$, alors (P_d) n'admet pas de solution.

2.2.4.2 Algorithme d'ascension duale lagrangienne

L'algorithme ADL (Ascension Duale Lagrangienne) permet de résoudre le problème (P_p) sous la condition que $\mathcal{J}$ soit strictement convexe. Il est présenté ci-après, **Algorithme 5**.

Algorithme 5 : Algorithme Ascension duale lagrangienne

Entrées : $k = 0$: initialisation du compteur ;
$\boldsymbol{a}^{(0)}$ et $\boldsymbol{\gamma}^{(0)}$: les points de départ ;
$\alpha^{(0)} > 0$: un pas initial ;
Sorties : $\hat{\boldsymbol{a}}$: une approximation de $\widetilde{\boldsymbol{a}}$
$\hat{\boldsymbol{\gamma}}$: une approximation de $\widetilde{\boldsymbol{\gamma}}$

Répéter jusqu'à convergence
$\boldsymbol{a}^{(k+1)} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}^{(k)})$
$\boldsymbol{\gamma}^{(k+1)} = \boldsymbol{\gamma}^{(k)} + \alpha^{(k)}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})$
$k = k + 1$

À chaque itération, la mise à jour de $\boldsymbol{a}$ consiste à trouver le minimiseur du lagrangien et la mise à jour de $\boldsymbol{\gamma}$ repose sur un algorithme de montée dans la direction du gradient de la

fonction duale $\Theta(\boldsymbol{\gamma})$ où $\alpha^{(k)} > 0$ est le pas à l'itération k. Ce gradient, appelé *résidu pour la contrainte d'égalité* dans le cas d'une contrainte d'égalité, vaut :

$$\nabla\Theta(\boldsymbol{\gamma}^{(k)}) = \boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c} \tag{2.23}$$

Une preuve est donnée dans l'**Annexe D** (p. 166). Ceci présuppose évidemment que Θ est différentiable. Cette assertion est vérifiée si $\mathcal{J}$ est strictement convexe, [Rockafellar, 1970][Théorèmes 26.1 et 26.3]. La méthode d'ascension duale lagrangienne peut toutefois être utilisée, sous certaines conditions, même si Θ n'est pas différentiable. Dans ce cas, le résidu $\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c}$ n'est pas le gradient mais un sous-gradient de Θ. Nous parlons alors de *méthode duale de sous-gradient* et la convergence n'est plus monotone.

L'inconvénient de l'**Algorithme 5** est qu'il n'est assuré de converger que si $\mathcal{J}$ est une fonction fermée strictement convexe. Cette hypothèse peut être trop restrictive pour des applications où le critère est convexe au sens large.

Malgré ses faiblesses, il n'en demeure pas moins que cette approche est la pierre angulaire des méthodes duales plus sophistiquées que nous détaillons ensuite. Il nous paraît judicieux d'illustrer la notion de dualité et le principe de fonctionnement de cet algorithme dans un cas simple où $\mathcal{J}$ est une fonction quadratique strictement convexe.

Cas d'un objectif quadratique strictement convexe

Pour illustrer le propos, nous considérons le problème de minimisation sous contrainte d'égalité avec un objectif $\mathcal{J}$ quadratique et strictement convexe :

$$\min_{\boldsymbol{a}} \begin{cases} \mathcal{J}(\boldsymbol{a}) = \boldsymbol{a}^{\mathrm{t}}\boldsymbol{P}\boldsymbol{a} + \boldsymbol{q}^{\mathrm{t}}\boldsymbol{a} + r \\ \text{s.c. } \boldsymbol{A}\boldsymbol{a} = \boldsymbol{c} \end{cases} \tag{2.24}$$

avec $\boldsymbol{P}$ une matrice réelle, symétrique, définie positive ; $\boldsymbol{q}$ un vecteur réel et r un scalaire réel.

L'étape de mise à jour de $\boldsymbol{a}$ dans l'**Algorithme 5** s'écrit :

$$\boldsymbol{a}^{(k+1)} \triangleq \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}^{(k)}) = -\frac{1}{2}\boldsymbol{P}^{-1}(\boldsymbol{q} + \boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma}^{(k)}) \tag{2.25}$$

La preuve est donnée dans l'**Annexe D** (p. 167).

Ce problème quadratique étant simple, nous pouvons calculer explicitement les minimiseurs théoriques respectivement du problème primal et du problème dual :

$$\begin{aligned} \widetilde{\boldsymbol{a}} &= \frac{1}{2}\boldsymbol{P}^{-1}\left[-\boldsymbol{q} + 2\boldsymbol{A}^{\mathrm{t}}\left((\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}})^{-1}(\frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{q} + \boldsymbol{c})\right)\right] \\ \text{et } \widetilde{\boldsymbol{\gamma}} &= -2\left(\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\right)^{-1}\left(\frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{q} + \boldsymbol{c}\right) \end{aligned} \tag{2.26}$$

La preuve est également donnée dans l'**Annexe D** (p. 168).

A titre d'exemple, nous prenons les valeurs numériques suivantes dans le problème (2.24),

$$\boldsymbol{P} = \begin{pmatrix} 5 & 0.2 \\ 0.2 & 5 \end{pmatrix}; \boldsymbol{q} = \begin{pmatrix} 1 \\ -5 \end{pmatrix}; r = 5; \boldsymbol{A} = \begin{pmatrix} -5 & -1 \end{pmatrix}; \boldsymbol{c} = 10$$

et les entrées de l'**Algorithme 5** suivantes :

$$\boldsymbol{a}^{(0)} = \begin{pmatrix} -4.5 \\ 5 \end{pmatrix}; \ \boldsymbol{\gamma}^{(0)} = -9; \ \alpha = 0.01 \text{ (un pas fixe).}$$

A convergence nous obtenons les résultats suivants :

$$\hat{\boldsymbol{a}} \approx \begin{pmatrix} -2.04 \\ 0.195 \end{pmatrix} \text{ et } \hat{\boldsymbol{\gamma}} \approx -3.86.$$

La Figure 2.4 représente les iso-contours de $\mathcal{J}(\boldsymbol{a})$ et l'évolution de la suite des $\boldsymbol{a}^{(k)}$ ($\circ$) qui convergent vers $\hat{\boldsymbol{a}}$ ($\star$). La Figure 2.5 illustre la résolution du problème dual et plus spécifiquement l'évolution de $\Theta(\boldsymbol{\gamma})$ en fonction de $\boldsymbol{\gamma}$ et l'évolution de la suite des $\boldsymbol{\gamma}^{(k)}$ ($\circ$) qui convergent vers $\hat{\boldsymbol{\gamma}}$ ($\star$). Sur ces deux figures, nous pouvons vérifier que les valeurs estimées $(\hat{\boldsymbol{a}}, \hat{\boldsymbol{\gamma}})$ correspondent aux valeurs théoriques $(\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{\gamma}})$ ($\bullet$) calculées avec (2.26).

$$\widetilde{\boldsymbol{a}} = \begin{pmatrix} -2.04 \\ 0.195 \end{pmatrix} \text{ et } \widetilde{\boldsymbol{\gamma}} = -3.86.$$

La Figure 2.6 représente :

- L'évolution de la suite des $\Theta(\boldsymbol{\gamma}^{(k)})$ (—) qui converge vers un maximum $\Theta(\hat{\boldsymbol{\gamma}}) = \hat{d}$, à mesure que le problème dual est résolu.
- L'évolution de la suite des $\mathcal{J}(\boldsymbol{a}^{(k)})$ (—) qui converge vers $\mathcal{J}(\hat{\boldsymbol{a}}) = \hat{p}$, à mesure que le problème primal est résolu.
- L'évolution du terme ${\boldsymbol{\gamma}^{(k)}}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}^{(k)} - \boldsymbol{c})$ (—) lié à la contrainte d'égalité dans le lagrangien, qui converge vers 0 à mesure que la contrainte est respectée.

REMARQUE

Nous constatons que $\Theta(\boldsymbol{\gamma}^{(k)})$ et $f(\boldsymbol{\gamma}^{(k)})$ convergent vers la même valeur :

$$\widetilde{p} = \mathcal{J}(\widetilde{\boldsymbol{a}}) = \Theta(\widetilde{\boldsymbol{\gamma}}) = \widetilde{d}.$$

En cohérence avec le théorème de Slater et les conditions KKT dans le cas de contraintes d'égalité affines et d'une fonction objectif strictement convexe, nous remarquons qu'il n'y a pas de saut de dualité et que la dualité est forte pour ce problème.

La Figure 2.7 représente quant à elle l'évolution du résidu pour la contrainte d'égalité $\nabla\Theta(\boldsymbol{\gamma}) = \boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c}$ qui converge vers 0 à mesure que le problème dual est résolu.

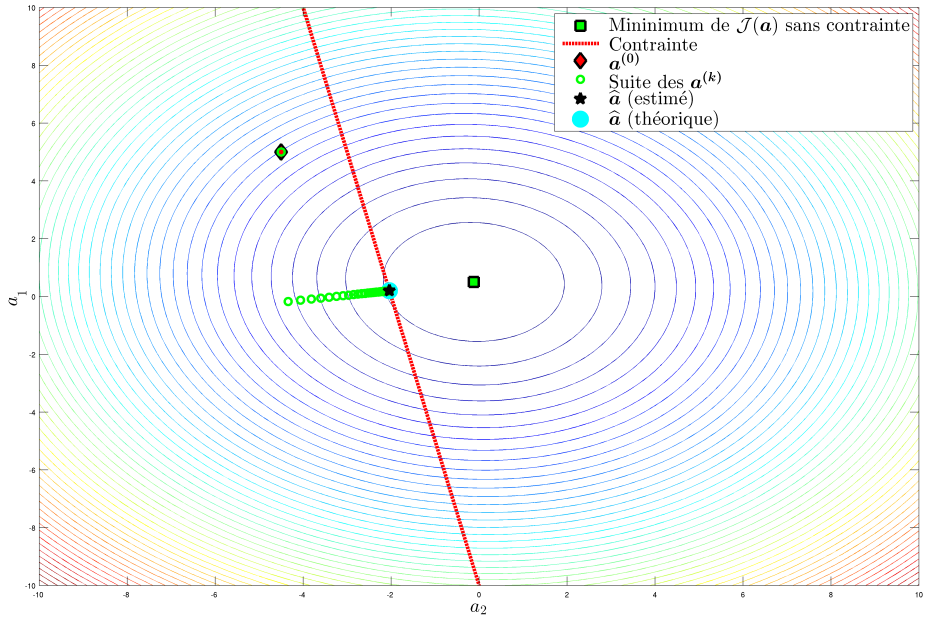


FIGURE 2.4 – Résolution du problème primal : illustration 2D

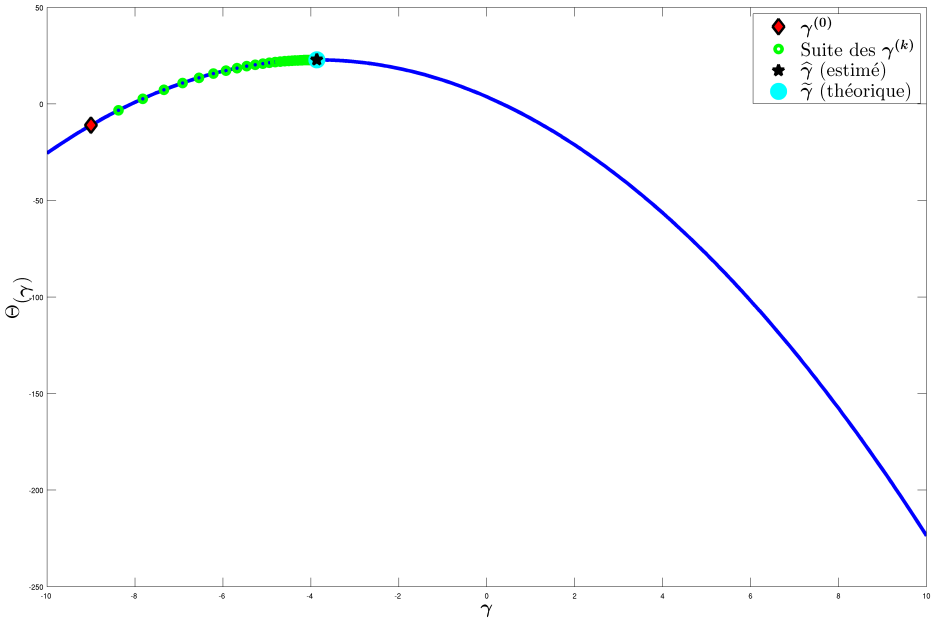


FIGURE 2.5 – Résolution du problème dual par la méthode ADL

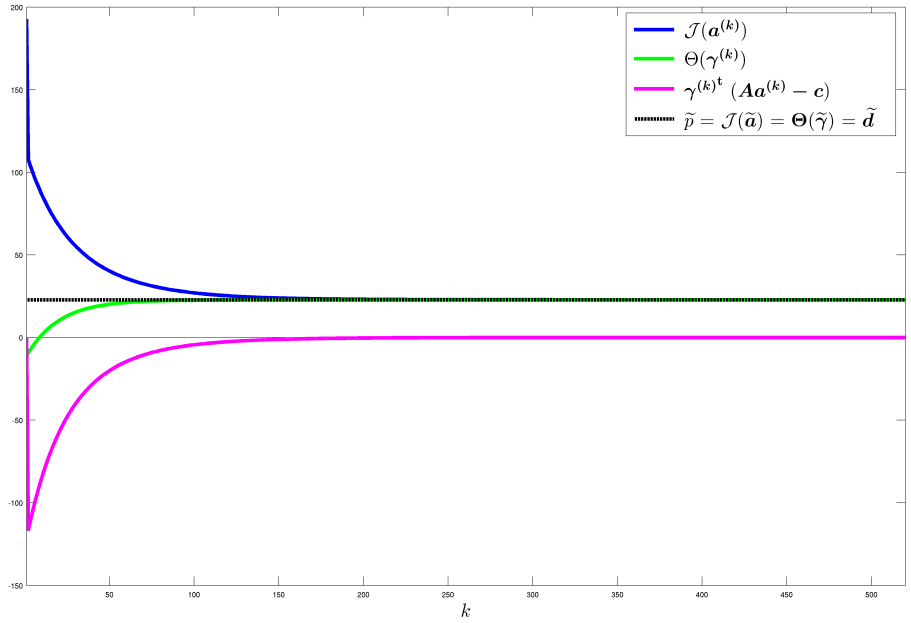


FIGURE 2.6 – Évolution de $\mathcal{J}(\boldsymbol{a}^{(k)})$, $\Theta(\boldsymbol{\gamma}^{(k)})$ et $\boldsymbol{\gamma}^{(k)^{\mathrm{t}}}(\boldsymbol{A}\boldsymbol{a}^{(k)}-\boldsymbol{c})$

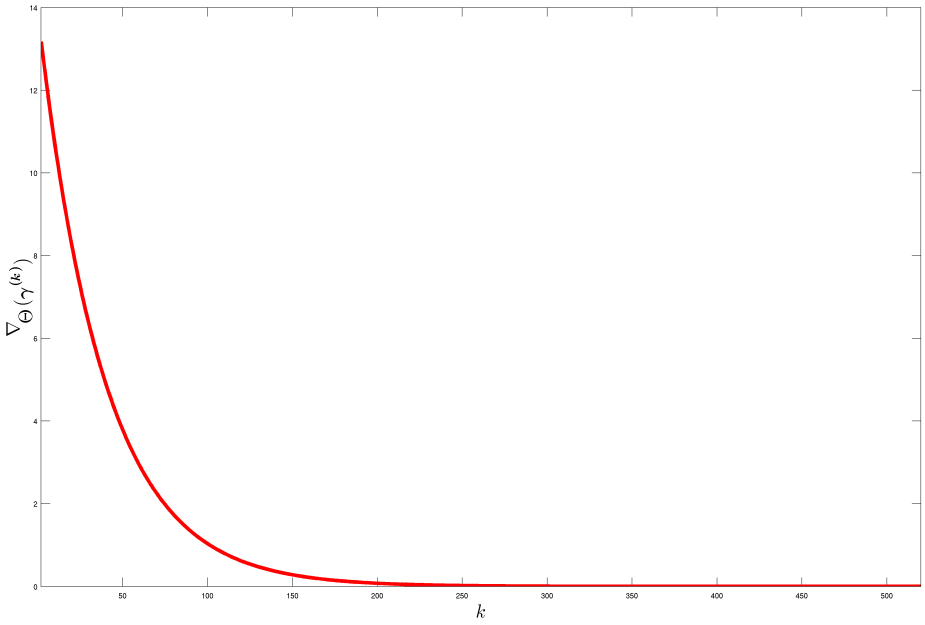


FIGURE 2.7 – Évolution de $\nabla\Theta\left(\boldsymbol{\gamma}^{(k)}\right)$

Décomposition duale

Outre son apparente simplicité, l'ADL présente un autre avantage : il est possible d'effectuer une *décomposition duale* sous certaines conditions. Pour cela il faut que :

- $\mathcal{J}$ soit séparable et puisse s'écrire :

$$\mathcal{J}(\boldsymbol{a}) = \sum_{p=1}^{P} \mathcal{J}_p(\boldsymbol{a}_p)$$

où $\boldsymbol{a} = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_P)^{\mathrm{t}}$ et où chaque $\boldsymbol{a}_p$ est un sous-vecteur de $\boldsymbol{a}$.

- la matrice $\boldsymbol{A}$ puisse être partitionnée en P blocs de sorte que :

$$\boldsymbol{A} = [\boldsymbol{A}_1, \cdots, \boldsymbol{A}_P] \Rightarrow \boldsymbol{A}\boldsymbol{a} = \sum_{p=1}^{P} \boldsymbol{A}_p \boldsymbol{a}_p.$$

Le lagrangien $\mathcal{L}$ est alors lui aussi séparable en $\boldsymbol{a}$ et il s'écrit :

$$\begin{aligned} \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}) &\triangleq \mathcal{J}(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}} \boldsymbol{A}\boldsymbol{a} - \boldsymbol{\gamma}^{\mathrm{t}} \boldsymbol{c} \\ &= \sum_{p=1}^{P} \mathcal{L}_p(\boldsymbol{a}_p, \boldsymbol{\gamma}) = \sum_{p=1}^{P} \left(\mathcal{J}_p(\boldsymbol{a}_p) + \boldsymbol{\gamma}^{\mathrm{t}} \boldsymbol{A}_p \boldsymbol{a}_p - \frac{1}{P} \boldsymbol{\gamma}^{\mathrm{t}} \boldsymbol{c} \right) \end{aligned} \tag{2.27}$$

La mise à jour de $\boldsymbol{a}$ conduit alors à P problèmes séparés qui peuvent être traités en parallèle. A chaque itération, il suffit donc d'effectuer une *partition* de $\boldsymbol{A}$ et $\boldsymbol{a}$ pour calculer les P étapes de mise à jour des ${\boldsymbol{a}_p}^{(k+1)}$ (en parallèle) ; puis une *fusion* des contributions résiduelles $\boldsymbol{A}_p {\boldsymbol{a}_p}^{(k+1)}$ afin de calculer le résidu global $\boldsymbol{A}\boldsymbol{a}^{(k+1)}$, puis enfin, le multiplicateur lagrangien global $\boldsymbol{\gamma}^{(k+1)}$. Plus précisément, pour une fonction objectif séparable, l'**Algorithme 5** prend la forme de l'**Algorithme 6** suivant.

Algorithme 6 : ADL pour une fonction objectif séparable

Entrées : $k = 0$: initialisation du compteur ;
$\boldsymbol{a}^{(0)}$ et $\boldsymbol{\gamma}^{(0)}$: les points de départ ;
$\alpha^{(0)} > 0$: un pas initial ;
Sorties : $\hat{\boldsymbol{a}}$: une approximation de $\widetilde{\boldsymbol{a}}$
$\hat{\boldsymbol{\gamma}}$: une approximation de $\widetilde{\boldsymbol{\gamma}}$

Répéter jusqu'à convergence
  pour p allant de 0 à P **faire**
    Partitionner $\boldsymbol{a}$ en P sous-vecteurs $\boldsymbol{a}_p$ et $\boldsymbol{A}$ en P blocs $\boldsymbol{A}_p$
    ${\boldsymbol{a}_p}^{(k+1)} = \underset{\boldsymbol{a}_p}{\operatorname{argmin}}\, \mathcal{L}(\boldsymbol{a}_p, \boldsymbol{\gamma}^{(k)})$
  Fusionner les $\boldsymbol{A}_p {\boldsymbol{a}_p}^{(k+1)}$
  $\boldsymbol{\gamma}^{(k+1)} = \boldsymbol{\gamma}^{(k)} + \alpha^{(k)}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})$
  $k = k + 1$

2.2.4.3 Lagrangien augmenté & méthode des multiplicateurs de Lagrange

Lagrangien augmenté

Les méthodes reposant sur un lagrangien augmenté ont été introduites afin de remédier aux faiblesses de la méthode ADL qui impose que le critère $\mathcal{J}$ soit strictement convexe. Ainsi, la convergence est désormais vérifiée même si $\mathcal{J}$ est convexe au sens large.

L'idée de considérer un lagrangien augmenté repose sur un constat simple, le problème primal (P_p) :

$$\min_{\boldsymbol{a}} \left\{ \begin{array}{l} \mathcal{J}(\boldsymbol{a}) \\ \text{s.c. } \boldsymbol{A}\boldsymbol{a} = \boldsymbol{c} \end{array} \right. \tag{P_p}$$

peut se réécrire sous la forme équivalente :

$$\min_{\boldsymbol{a}} \left\{ \begin{array}{l} \mathcal{J}(\boldsymbol{a}) + \dfrac{\rho}{2} \left\| \boldsymbol{A}\boldsymbol{a} - \boldsymbol{c} \right\|_2^2 \\ \text{s.c. } \boldsymbol{A}\boldsymbol{a} = \boldsymbol{c} \end{array} \right. \tag{2.28}$$

qui fait intervenir un terme quadratique additif $\rho \left\| \boldsymbol{A}\boldsymbol{a} - \boldsymbol{c} \right\|_2^2 / 2$.

Les deux problèmes sont strictement équivalents car pour chaque $\boldsymbol{a}$ respectant la contrainte (*i.e.* $\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c} = 0$), le terme quadratique ajouté s'annule.

Le lagrangien augmenté associé au problème (2.28) est alors défini par :

$$\mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{\gamma}) \triangleq \mathcal{J}(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c}) + \frac{\rho}{2} \| \boldsymbol{A}\boldsymbol{a} - \boldsymbol{c} \|^2 \tag{2.29}$$

où $\rho > 0$ est appelé *paramètre de pénalité*.

La fonction duale fait intervenir le lagrangien augmenté (2.29) :

$$\Theta_\rho(\boldsymbol{\gamma}) \triangleq \inf_{\boldsymbol{a}} \mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{\gamma}) \tag{2.30}$$

REMARQUE

Si $\rho = 0$, $\mathcal{L}_0$ est le lagrangien classique non augmenté $\mathcal{L}$ défini par (2.15) et Θ_0 la fonction duale classique définie par (2.16).

Méthode des multiplicateurs de Lagrange

La méthode des ML (Multiplicateurs de Lagrange) correspond à l'extension de l'ADL pour un lagrangien augmenté. Cependant, contrairement à cette dernière, la méthode des ML requiert des hypothèses plus souples sur le critère $\mathcal{J}$ pour garantir la convergence. Ainsi, l'**Algorithme 7** suivant est garanti de converger si $\mathcal{J}$ est convexe (au sens large), fermée et propre. Notons également que désormais le choix d'un pas fixe $\rho > 0$ assure la convergence de l'**Algorithme 7** ce qui n'était pas forcément le cas avec l'ADL. Une preuve de convergence de la méthode des ML est donnée dans l'**Annexe D** (p. 169).

L'inconvénient principal de la méthode des ML est que, même si $\boldsymbol{A}$ peut-être convenablement partitionnée et que $\mathcal{J}$ est séparable, $\mathcal{L}_\rho$ ne l'est pas. En effet, l'ajout du terme $\left\| \boldsymbol{A}\boldsymbol{a} - \boldsymbol{c} \right\|_2^2$ fait apparaître des termes croisés ($-2\boldsymbol{c}^t\boldsymbol{A}\boldsymbol{a}$) dans $\mathcal{L}_\rho$. L'étape de mise à jour de $\boldsymbol{a}$ ne peut donc pas être menée en parallèle comme c'est le cas pour l'ADL.

Algorithme 7 : Méthode des multiplicateurs de Lagrange

Entrées : $k = 0$: initialisation du compteur ;
$\boldsymbol{a}^{(0)}$ et $\boldsymbol{\gamma}^{(0)}$: les points de départ ;
$\rho > 0$: un pas fixe ;
Sorties : $\hat{\boldsymbol{a}}$: une approximation de $\widetilde{\boldsymbol{a}}$
$\hat{\boldsymbol{\gamma}}$: une approximation de $\widetilde{\boldsymbol{\gamma}}$

Répéter jusqu'à convergence
$\boldsymbol{a}^{(k+1)} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{\gamma}^{(k)})$
$\boldsymbol{\gamma}^{(k+1)} = \boldsymbol{\gamma}^{(k)} + \rho(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})$
$k = k + 1$

2.2.4.4 Alternating Direction Method of Multipliers

L'ADMM n'est pas un algorithme récent dans le sens où ses prémices remontent aux années 1970 avec [Glowinski et Marroco, 1975] et [Gabay et Mercier, 1976]. Des études ont ensuite montré qu'il était équivalent ou tout du moins très proche de nombreuses autres méthodes, dont certaines plus anciennes, telles que le *Douglas-Rachford splitting* [Douglas et Rachford, 1956], la méthode des *projections alternées de Dykstra* [Boyle et Dykstra, 1986], l'algorithme itératif de Bregman pour des problèmes ℓ_1 [Yin *et al.*, 2008], les méthodes proximales et bien d'autres méthodes encore.

L'ADMM repose sur la méthode des multiplicateurs de Lagrange tout en continuant à tirer profit de la décomposition duale. Il vise à résoudre des problèmes de la forme :

$$\min_{\boldsymbol{a}} \begin{cases} f(\boldsymbol{a}) + g(\boldsymbol{v}) \\ \text{s.c. } \boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{v} = \boldsymbol{c} \end{cases} \tag{$\mathrm{P}_{\mathrm{ADMM}}$}$$

où $\boldsymbol{A}$ et $\boldsymbol{B}$ sont deux matrices, $\boldsymbol{c}$ un vecteur et f et g deux fonctions convexes.

REMARQUE

Ce problème se ramène au problème non contraint ($\mathcal{P}_1$) en posant :

$$f(\boldsymbol{a}) = \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2,\ g(\boldsymbol{v}) = \mu\,\|\boldsymbol{v}\|_1,\ \boldsymbol{A} = \boldsymbol{M}, \boldsymbol{B} = \boldsymbol{I} \text{ et } \boldsymbol{c} = \boldsymbol{0}.$$

Le lagrangien augmenté $\mathcal{L}_\rho$ pour le problème ($\mathrm{P}_{\mathrm{ADMM}}$) s'écrit classiquement :

$$\mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{v}, \boldsymbol{\gamma}) \triangleq f(\boldsymbol{a}) + g(\boldsymbol{v}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{v} - \boldsymbol{c}) + \frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{v} - \boldsymbol{c}\|_2^2 \tag{2.31}$$

Comme le préconise [Boyd, 2010] et afin de simplifier l'écriture et l'implémentation ensuite, nous pouvons réécrire $\mathcal{L}_\rho$ sous une forme « normalisée » plus compacte complètement

équivalente. Plus précisément, nous notons $\boldsymbol{r} \triangleq \boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{v} - \boldsymbol{c}$ le *résidu pour la contrainte d'égalité*, puis, par simple complétion du carré, en combinant les termes linéaires et quadratiques de (2.31) il vient :

$$\begin{aligned}\mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{v}, \boldsymbol{\gamma}) &\triangleq f(\boldsymbol{a}) + g(\boldsymbol{v}) + \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{r} + \frac{\rho}{2}\|\boldsymbol{r}\|_2^2 \\ &= f(\boldsymbol{a}) + g(\boldsymbol{v}) + \frac{\rho}{2}\|\boldsymbol{r} + \frac{1}{\rho}\boldsymbol{\gamma}\|_2^2 - \frac{1}{2\rho}\|\boldsymbol{\gamma}\|_2^2\end{aligned} \tag{2.32}$$

En notant $\boldsymbol{u} = (1/\rho)\,\boldsymbol{\gamma}$ la *variable duale normalisée* et en remplaçant dans (2.32) nous obtenons :

$$\begin{aligned}\mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{v}, \boldsymbol{u}) &= f(\boldsymbol{a}) + g(\boldsymbol{v}) + \frac{\rho}{2}\|\boldsymbol{r} + \boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2 \\ &= f(\boldsymbol{a}) + g(\boldsymbol{v}) + \frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{v} - \boldsymbol{c} + \boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2\end{aligned} \tag{2.33}$$

Cette forme normalisée permet de simplifier les écritures qui suivent. La fonction duale s'écrit quant à elle toujours :

$$\Theta_\rho(\boldsymbol{u}) \triangleq \inf_{\boldsymbol{a}} \mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{v}, \boldsymbol{u}) \tag{2.34}$$

L'algorithme ADMM s'écrit alors comme suit.

Algorithme 8 : ADMM

Entrées : $k = 0$: initialisation du compteur ;
$\boldsymbol{a}^{(0)}$, $\boldsymbol{v}^{(0)}$ et $\boldsymbol{u}^{(0)}$: les points de départ ;
$\rho > 0$: le pas fixe ;
Sorties : $\hat{\boldsymbol{a}}$: une approximation de $\widetilde{\boldsymbol{a}}$;
$\hat{\boldsymbol{u}}$: une approximation de $\widetilde{\boldsymbol{u}}$

Répéter jusqu'à convergence

$$\boldsymbol{a}^{(k+1)} = \operatorname*{argmin}_{\boldsymbol{a}} \mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{v}^{(k)}, \boldsymbol{u}^{(k)}) = \operatorname*{argmin}_{\boldsymbol{a}} \left(f(\boldsymbol{a}) + \frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{v}^{(k)} - \boldsymbol{c} + \boldsymbol{u}^{(k)}\|_2^2 \right)$$

$$\boldsymbol{v}^{(k+1)} = \operatorname*{argmin}_{\boldsymbol{v}} \mathcal{L}_\rho(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}, \boldsymbol{u}^{(k)}) = \operatorname*{argmin}_{\boldsymbol{v}} \left(g(\boldsymbol{v}) + \frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{a}^{(k+1)} + \boldsymbol{B}\boldsymbol{v} - \boldsymbol{c} + \boldsymbol{u}^{(k)}\|_2^2 \right)$$

$$\boldsymbol{u}^{(k+1)} = \boldsymbol{u}^{(k)} + \rho(\boldsymbol{A}\boldsymbol{a}^{(k+1)} + \boldsymbol{B}\boldsymbol{v}^{(k+1)} - \boldsymbol{c})$$

$$k = k + 1$$

Cet algorithme tire son nom du fait que $\boldsymbol{a}$ et $\boldsymbol{v}$ sont mis à jour de façon alternée. A chaque itération, l'étape de mise-à-jour de $\boldsymbol{a}$ et $\boldsymbol{v}$ est similaire à un passage de Gauss-Seidel et s'écrit :

$$(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}^{(k+1)}) = \operatorname*{argmin}_{\boldsymbol{a},\boldsymbol{v}} \mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{v}, \boldsymbol{u}^{(k)}).$$

La mise a jour de $\boldsymbol{u}$ repose quant à elle, encore une fois, sur une méthode de montée dans la direction du gradient de $\Theta_\rho(\boldsymbol{u})$.

De nombreuses preuves de convergence de l'ADMM ont été données dans la littérature. Celle présentée par [Boyd, 2010][pp 16-17 et Annexe A] (qui repose sur la preuve de convergence de la méthode des ML), présente l'avantage d'être relativement concise et générale. Les conditions générales de convergence, qui s'appliquent à de très nombreux problèmes, posées par [Boyd, 2010] sont les suivantes.

Condition 1 f et g sont convexes, fermées et propres.
Condition 2 le lagrangien non augmenté $\mathcal{L}_0$ associé au problème, admet au moins un point selle (*i.e.* la dualité est forte). Autrement dit, il existe au moins un couple $(\widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{v}})$ tel que $\boldsymbol{A}\widetilde{\boldsymbol{a}} + \boldsymbol{B}\widetilde{\boldsymbol{v}} = \boldsymbol{c}$ et $f(\widetilde{\boldsymbol{a}}) < \infty$, $g(\widetilde{\boldsymbol{v}}) < \infty$.

Si ces deux conditions sont respectées, alors les itérées de l'ADMM satisfont les affirmations suivantes.

- *Convergence du résidu pour la contrainte d'égalité* : $\boldsymbol{r}^{(k)} \to \boldsymbol{0}$, pour $k \to \infty$, où $\boldsymbol{r}^{(k)} = \boldsymbol{A}\boldsymbol{a}^{(k)} + \boldsymbol{B}\boldsymbol{v}^{(k)} - \boldsymbol{c}$ est le résidu primal pour la contrainte d'égalité à l'itération k. Ceci signifie que la contrainte est asymptotiquement respectée.
- *Convergence du primal* : $f(\boldsymbol{a}^{(k)}) + g(\boldsymbol{v}^{(k)}) \to \widetilde{p}$, pour $k \to \infty$,
  où $\widetilde{p} = \inf_{\boldsymbol{a},\boldsymbol{v}} \{f(\boldsymbol{a}) + g(\boldsymbol{v}) | \boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{v} = \boldsymbol{c}\}$ est la valeur optimale du primal pour le problème (P_{ADMM}). Ceci signifie que la fonction objectif des itérées tend asymptotiquement vers le primal optimal (*i.e.* le problème primal est résolu).
- *Convergence du dual* : $\boldsymbol{u}^{(k)} \to \widetilde{\boldsymbol{u}}$, pour $k \to \infty$,
  où $\widetilde{\boldsymbol{u}}$ est la valeur optimale du dual. Ceci signifie que le problème dual est asymptotiquement résolu.

Entre autres, un des points forts de l'ADMM est que, sous certaines conditions de sous-optimalité détaillées dans [Eckstein et Bertsekas, 1992], l'algorithme converge même si la minimisation de $\boldsymbol{a}$ et $\boldsymbol{v}$ est inexacte. Ceci est particulièrement intéressant si la mise-à-jour de $\boldsymbol{a}$ et/ou $\boldsymbol{v}$ est obtenue par une méthode itérative car seules quelques sous-itérations pour minimiser ces variables sont nécessaires à chaque itération de l'ADMM.

REMARQUE

> Sur le même principe et sous les mêmes hypothèses, [Li *et al.*, 2015b], par exemple, a étendu la validité de l'ADMM et la preuve de convergence de [Boyd, 2010] à des fonctions réelles à variables complexes en partant de la forme normalisée du Lagrangien augmenté.

Sans rentrer dans les détails qui sont expliqués rigoureusement dans [Boyd, 2010][p. 19], nous signalons que l'étude des conditions d'optimalité de l'ADMM pour le problème (P_{ADMM}) permet de considérer le critère d'arrêt raisonnable suivant :

$$\left\|\boldsymbol{r}^{(k)}\right\|_2 \leqslant \varepsilon_{\text{pri}} \text{ et } \left\|\boldsymbol{s}^{(k)}\right\|_2 \leqslant \varepsilon_{\text{dual}} \tag{2.35}$$

où $\boldsymbol{s}^{(k)} = \rho \boldsymbol{A}^{\text{t}} \boldsymbol{B}(\boldsymbol{v}^{(k)} - \boldsymbol{v}^{(k-1)})$ est le résidu dual à l'itération k et $\varepsilon_{\text{pri}} > 0$ et $\varepsilon_{\text{dual}} > 0$ sont des tolérances sur le primal et le dual.

REMARQUE

De nombreuses méthodes dérivées de l'ADMM, ont été développées récemment, souvent pour résoudre un problème spécifique. Certaines, comme par exemple *C-SALSA* [Afonso *et al.*, 2011] ont des vitesses de convergence supérieures. Dans la même idée, et toujours pour accélérer la convergence de l'ADMM, il est possible d'envisager des variantes où le paramètre de pénalité ρ varie au cours des itérations [Xu *et al.*, 2017] (mais la convergence n'est plus garantie) ainsi que des méthodes pour déterminer un ρ optimal [Lin *et al.*, 2017].

Conclusion du Chapitre 2

Dans ce chapitre, nous avons brièvement présenté un historique de la notion de parcimonie. Nous avons montré que la minimisation de critères pénalisés avec une norme ℓ_1 permet de favoriser la parcimonie. Enfin, nous avons comparé quelques algorithmes de minimisation de critères non différentiables et contraints. Nous nous sommes en particulier focalisés sur l'algorithme primal-dual ADMM.

Pour conclure cette revue des différents algorithmes permettant de minimiser des critères non différentiables et possiblement contraints, nous insistons sur le fait que l'ADMM est un algorithme simple à implémenter et qui est plus général que d'autres méthodes (telle que FISTA par exemple) car les hypothèses de convergence sont plus souples. A l'inverse d'autres méthodes, il permet également de minimiser des critères non différentiables et contraints.

Enfin, de nombreuses études, dont notamment [Nithya et Leela Lakshmi, 2014; Tao *et al.*, 2015], ont été menées pour comparer la vitesse de convergence de l'ADMM à celle d'autres algorithmes. Il a été montré empiriquement que si la mise à jour de l'objet d'intérêt peut-être menée rapidement – et nous montrerons dans le prochain chapitre que c'est le cas pour notre problème d'imagerie de SER – alors l'ADMM converge aussi sinon plus rapidement que d'autres méthodes concurrentes.

Dans le prochain chapitre, nous présentons la contribution majeure de cette thèse : le développement d'une nouvelle méthode d'imagerie 3D HR de SER. Elle s'appuie sur la construction d'un critère avec pénalité ℓ_1 et contraint favorisant la parcimonie qui est minimisé à l'aide d'une adaptation de l'ADMM.

CHAPITRE 3

SPRITE : UNE MÉTHODE HAUTE RÉSOLUTION POUR L'IMAGERIE RADAR 3D

Ce chapitre s'articule autour de la contribution centrale de cette thèse : le développement d'une nouvelle méthode HR appelée SPRITE (Sparse Radar Imaging TEchnique) pour l'imagerie 3D de SER. Elle repose sur une régularisation du problème consistant en la prise en compte d'informations a priori de parcimonie et d'une information de support. La solution est alors définie comme le minimiseur d'un critère pénalisé et contraint. L'optimisation est assurée par l'algorithme ADMM dont une adaptation aux spécificités du problème permet de mener efficacement les calculs à l'aide de transformées de Fourier rapides.

Dans la première section nous détaillons les informations a priori dont nous disposons sur l'objet de rétro-diffusion ; elles expriment la réalité physique des interactions onde-cible. Dans la seconde section nous définissons la solution comme le minimiseur d'un critère pénalisé et contraint qui intègre les a priori énoncés. La troisième section concerne la minimisation du critère avec l'ADMM et la prise en compte des spécificités du problème pour amener efficacement les calculs. Enfin, la dernière section présente des compléments sur la méthode SPRITE .

Sommaire

3.1 Les connaissances *a priori* sur l'objet de rétro-diffusion 88
3.1.1 Un nombre réduit de plans spéculaires 88
3.1.2 Un nombre réduit de facettes spéculaires connexes 91
3.1.3 Le coefficient de rétro-diffusion constant sur chaque facette 91
3.1.4 L'énergie globale relativement faible 92
3.1.5 L'extension spatiale limitée 93
3.1.6 La transposition des *a priori* à la carte discrète 94
3.2 La construction d'un critère pénalisé et contraint 96
3.2.1 Les pénalités et la contrainte de support 96
3.2.2 L'écriture du critère et de la solution 101
3.3 La minimisation du critère à l'aide de l'ADMM 101
3.3.1 La réécriture de la solution sous forme ADMM équivalente 101
3.3.2 L'écriture de l'algorithme et mise à jour des variables 102
3.3.3 La gestion de la mise à jour de la variable $\boldsymbol{a}$ 105
3.4 Des compléments sur la méthode SPRITE 107
3.4.1 La version mono-dimensionnelle 107
3.4.2 L'initialisation des variables 108
3.4.3 Le choix des paramètres 108
3.4.4 Le critère d'arrêt 109

3.1 Les connaissances *a priori* sur l'objet de rétro-diffusion

Nous rappelons que, dans le contexte SASB explicité **Section 1.2.2** (p. 36), l'objet de rétro-diffusion $a(\boldsymbol{r}, \boldsymbol{k})$ varie peu avec le vecteur d'onde $\boldsymbol{k}$. Dans la suite, nous nous intéressons à $a(\boldsymbol{r}, \boldsymbol{k}_c)$ l'objet de rétro-diffusion correspondant au vecteur d'onde central $\boldsymbol{k}_c$, pour lequel nous souhaitons former l'image. Pour simplifier les notations, nous posons :

$$a(\boldsymbol{r}) \triangleq a(\boldsymbol{r}, \boldsymbol{k}_c). \tag{3.1}$$

Dans les sections qui suivent, nous expliquons comment la phénoménologie des interactions onde-cible, présentée dans la **Section 1.1.2** (p. 21), renseigne sur les différents mécanismes de rétro-diffusion qui interviennent et comment ces derniers peuvent se traduire sous forme de connaissances *a priori* sur l'objet de rétro-diffusion $a(\boldsymbol{r})$.

Les arguments énoncés dans les sections suivantes sont valides dans le contexte du modèle que nous avons explicité dans la **Section 1.2.3** (p. 36) et que nous rappelons ci-dessous.

Contexte du modèle

- Configuration et paramètres de fonctionnement du radar :
  - ▶ configuration monostatique,
  - ▶ illumination en champ lointain avec des ondes électromagnétiques supposées quasi-planes sur la cible,
  - ▶ contexte hautes fréquences où l'optique physique s'applique.
- Nature des cibles :
  - ▶ relativement convexes avec une surface fermée,
  - ▶ surface lisse relativement aux longueurs d'ondes considérées.

En plus de ce contexte, nous faisons un choix sur l'orientation du repère spatial : la direction longitudinale $\hat{\boldsymbol{z}}$ est choisie colinéaire au vecteur d'onde central $\hat{\boldsymbol{k}}_c$.

Nous formulons maintenant les différentes informations *a priori* sur $a(\boldsymbol{r})$ que nous intègrerons à la méthode d'inversion dans le but d'augmenter la fidélité des cartes à la réalité physique et d'accroître ainsi leur résolution.

3.1.1 Un nombre réduit de plans spéculaires

Comme illustré sur la Figure 3.1, pour une direction de vecteur d'onde $\hat{\boldsymbol{k}}_c$ donnée, il y a un nombre fini et faible de plans $\mathcal{P}$ tangents à la surface de la cible et de normale extérieure $\hat{\boldsymbol{n}}$ colinéaire à $\hat{\boldsymbol{k}}_c$. Dans le [**Contexte du modèle**], nous avons expliqué **Section 1.1.2** que seuls ces plans sont susceptibles de contenir des facettes spéculaires.

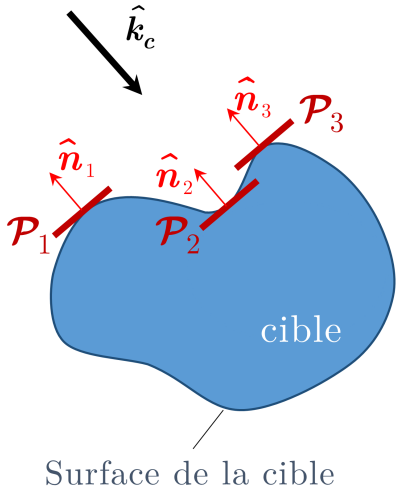


FIGURE 3.1 – Plans tangents à la surface et de normale extérieure colinéaire à $\hat{\boldsymbol{k}}_c$

Nous notons $\mathcal{P}_z$ la fonction de $z \in \mathbb{R}$ suivante qui correspond à une projection de $a(\boldsymbol{r})$ sur l'axe $\hat{\boldsymbol{z}}$:

$$\mathcal{P}_z(z) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} a(x,y,z)\mathrm{d}x\mathrm{d}y. \tag{3.2}$$

Comme nous l'avons expliqué ci-dessus, le nombre de plans susceptibles de contenir des facettes spéculaires est fini et faible. Or, par choix, $\hat{\boldsymbol{z}}$ est colinéaire à $\hat{\boldsymbol{k}}_c$. Nous en déduisons que le nombre de points n'annulant pas $\mathcal{P}_z$ est réduit. Nous formulons alors l'*a priori* **(AP1)** suivant.

(AP1)

La projection de $a(\boldsymbol{r})$ sur l'axe $\hat{\boldsymbol{z}}$ est nulle sauf pour un nombre limité de points.

Cet *a priori* est illustré sur la Figure 3.2 où peu de plans orthogonaux à $\hat{\boldsymbol{z}}$ (colinéaire à $\hat{\boldsymbol{k}}_c$) contiennent des facettes spéculaires. Comme expliqué dans l'étude phénoménologique **Section 1.1.2**, nous rappelons que ces facettes spéculaires se réduisent à des segments et des points brillants selon la courbure de la surface de la cible.

REMARQUE

L'*a priori* (AP1) est cohérent du modèle point brillant mono-dimensionnel (1.15) qui stipule qu'une cible est constituée d'un faible nombre de réflecteurs selon la direction longitudinale. Ceci est illustré sur la Figure 3.3 : les points brillants se situent aux coordonnées longitudinales des plans contenant des facettes spéculaires c'est-à-dire pour les points z qui n'annulent pas $\mathcal{P}_z$.

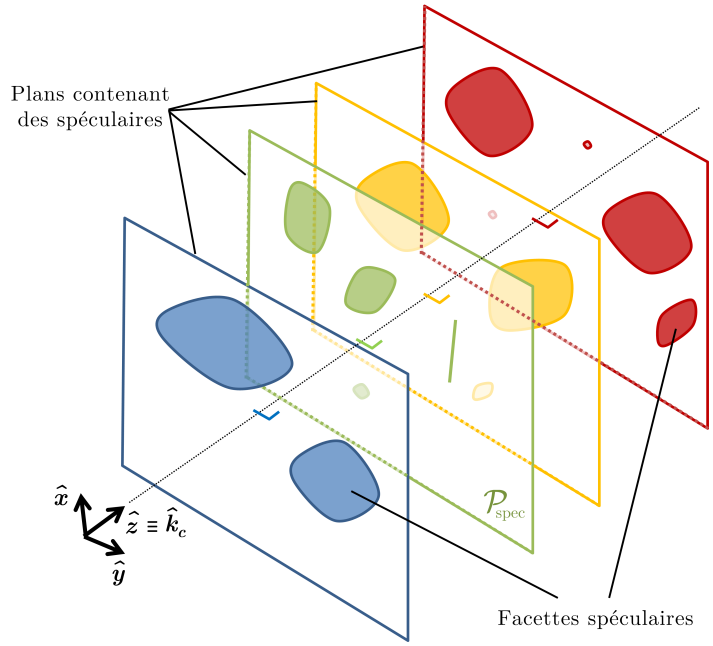


FIGURE 3.2 – Peu de plans orthogonaux à $\hat{z}$ contiennent des facettes spéculaires.

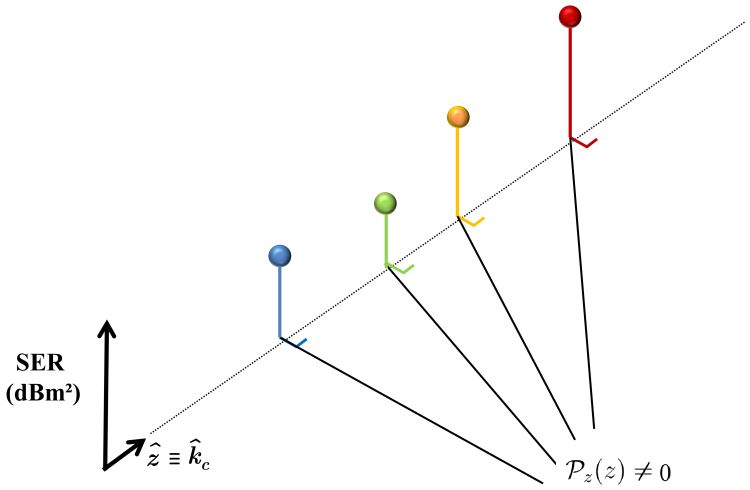


FIGURE 3.3 – Projection sur l'axe $\hat{z}$ en cohérence avec le modèle point brillant 1D.

3.1.2 Un nombre réduit de facettes spéculaires connexes

Nous rappelons que les facettes spéculaires se situent à la surface d'une cible homogène CEP ainsi qu'aux interfaces de propriétés matériaux pour des cibles non homogènes.

D'après le [**Contexte du modèle**], la surface de la cible est lisse relativement aux longueurs d'onde, nous en déduisons que sur chaque plan spéculaire, les facettes spéculaires sont connexes. Ceci signifie qu'une facette spéculaire est d'un « seul tenant » comme illustré Figure 3.4.

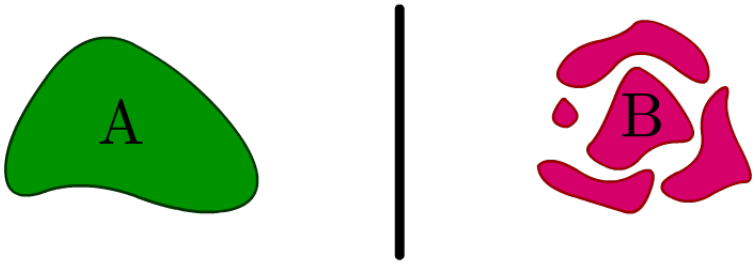


FIGURE 3.4 – La facette verte A est connexe contrairement à la facette rouge B

Nous formulons alors l'*a priori* **(AP2)** suivant.

(AP2)

L'objet $a(\boldsymbol{r})$ peut-être décomposé en un nombre restreint de facettes spéculaires connexes.

3.1.3 Le coefficient de rétro-diffusion constant sur chaque facette

Par définition, tout élément d'une facette spéculaire est à la même distance selon $\hat{\boldsymbol{z}}$ du radar. En champ lointain, dans le cas d'ondes planes, ceci a pour conséquence que la différence de phase induite par la distance longitudinale est la même sur toute la facette spéculaire.

Par ailleurs, une facette spéculaire est associée à une partie homogène de la surface lisse de la cible ou de l'interface de propriétés matériaux. Sur cette portion de surface les propriétés matériaux sont donc invariantes. Il en résulte que l'amplitude propre et la phase propre (*i.e.* l'atténuation d'amplitude et le déphasage induits par le matériau) ne varient pas sur la facette spéculaire [Knott *et al.*, 2004]. Ceci signifie que le coefficient de rétro-diffusion est constant sur chaque facette.

Nous formulons l'*a priori* **(AP3)** qui se déduit des deux arguments précédents.

(AP3)

$a(\boldsymbol{r})$ est constant sur chaque facette spéculaire.

REMARQUE

Nous signalons que cet *a priori* n'est pas nécessairement valide si la surface de l'objet n'est pas lisse relativement aux longueurs d'onde. En effet, dans des contextes SAR aéroporté ou satellitaire – qui diffèrent de nos objectifs d'analyse et de contrôle de SER – les scènes complexes imagées peuvent être considérées comme rugueuses relativement aux longueurs d'onde [Bryant *et al.*, 2007]. De nombreux contributeurs différents sont alors potentiellement présents au sein d'une même cellule de résolution. La sommation cohérente de leurs phases propres produit alors des interférences constructives ou destructives, provoquant ainsi une forte fluctuation d'une cellule de résolution à l'autre. Ce phénomène est connu sous le nom de *bruit de chatoiement* (ou *speckle* en anglais) [Moreira *et al.*, 2013].

Enfin, nous insistons sur le fait que l'*a priori* **(AP3)** stipule que le coefficient de rétro-diffusion est constant sur chaque facette spéculaire et non pas forcément sur chaque plan spéculaire $\mathcal{P}_{\text{spec}}$. Ainsi, et comme illustré Figure 3.5, dans le cas d'un objet non homogène, si deux facettes spéculaires provenant de deux matériaux différents sont sur un même plan $\mathcal{P}_{\text{spec}}$, alors le coefficient de rétro-diffusion est constant sur chaque facette mais n'est pas forcément égal d'une facette à l'autre car il dépend des propriétés du matériau.

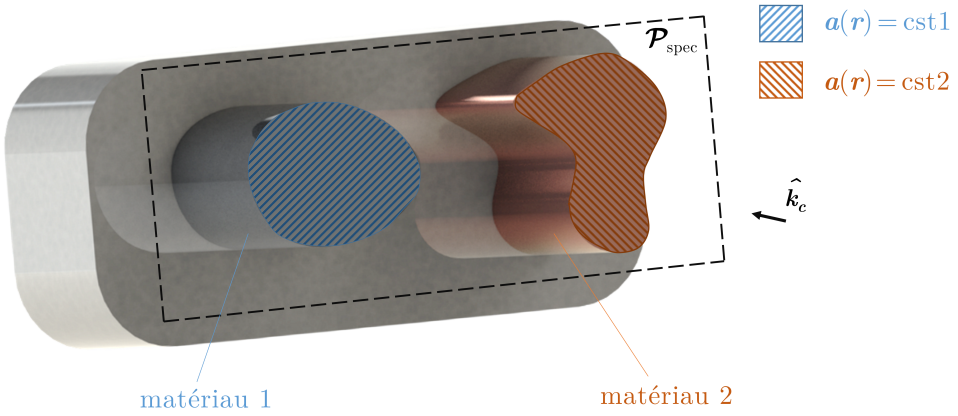


FIGURE 3.5 – Coefficient de rétro-diffusion constant sur chaque facette spéculaire

3.1.4 L'énergie globale relativement faible

L'énergie de l'objet de rétro-diffusion est définie par :

$$E = \iiint_{-\infty}^{+\infty} |\boldsymbol{a}(\boldsymbol{r})|^2 \mathrm{d}\boldsymbol{r} \tag{3.3}$$

Dans notre contexte d'analyse et de contrôle de SER, nous pouvons formuler l'*a priori* **(AP4)** suivant.

(AP4)

L'énergie de l'objet de rétro-diffusion $a(\boldsymbol{r})$ est relativement faible.

3.1.5 L'extension spatiale limitée

Le principe d'incertitude d'Heisenberg-Gabor [Gabor, 1946] stipule qu'un signal ne peut être à la fois limité spatialement et limité dans le k-space (domaine des fréquences spatiales). En effet, un signal spatialement limité a une bande fréquentielle infinie dans le domaine dual et vice-versa.

Le nombre d'observations M étant fini, la bande des fréquences spatiales des mesures $\boldsymbol{\sigma}(\boldsymbol{k})$ est nécessairement limitée dans le k-space. Par suite, et d'après le principe d'incertitude d'Heisenberg-Gabor sus-cité, l'extension de l'objet $a(\boldsymbol{r})$ que l'on cherche à reconstruire à partir de ces mesures est théoriquement illimitée dans le domaine spatial.

Cependant, les radaristes supposent généralement que l'objet de rétro-diffusion $a(\boldsymbol{r})$ est limité simultanément dans le domaine spatial et dans le domaine des fréquences spatiales. [Mahafza, 2013] définit plus rigoureusement le cadre dans lequel cette approximation est valide en introduisant la notion d'objet de rétro-diffusion *essentiellement limité*. Ainsi, $a(\boldsymbol{r})$ est dit essentiellement limité dans le domaine spatial relativement à un certain niveau ϵ faible, si son énergie est suffisamment faible en dehors d'une certaine extension spatiale $\boldsymbol{\Delta}_s$. Plus précisément, nous avons la relation suivante :

$$\iiint_{\boldsymbol{\Delta}_s} |\boldsymbol{a}(\boldsymbol{r})|^2 \mathrm{d}\boldsymbol{r} \geqslant (1-\epsilon) \iiint_{-\infty}^{+\infty} |\boldsymbol{a}(\boldsymbol{r})|^2 \mathrm{d}\boldsymbol{r} \tag{3.4}$$

Il est alors possible de considérer une certaine extension spatiale $\boldsymbol{\Delta}_s$ en dehors de laquelle $a(\boldsymbol{r})$ est quasiment nul :

$$a(\boldsymbol{r}) \approx 0 \text{ si } \boldsymbol{r} \notin \boldsymbol{\Delta}_s \tag{3.5}$$

Nous formulons alors l'*a priori* **(AP5)** suivant.

(AP5)

L'extension spatiale de l'objet de rétro-diffusion $a(\boldsymbol{r})$ que l'on cherche à reconstruire est limitée.

REMARQUE

Le support spatial ne se limite pas forcément à la seule géométrie de la cible imagée. En effet, comme expliqué dans la **Section 1.2.3**, de nombreux mécanismes de rétro-diffusion tels que les rampants ou les interactions multiples, peuvent apparaitre sous forme d'« artefacts » en dehors des limites de l'objet physique. Or, bien que de niveaux généralement inférieurs aux spéculaires, ces derniers contribuent également à la SER globale de la cible.
Cela est illustré sur la Figure 3.6 qui présente l'imagerie 1D (obtenue par TFRI) de la SER d'une sphère CEP de diamètre 1 m, pour une excursion fréquentielle de 200 à 1800 MHz, par pas de 10 MHz. Nous constatons que le support spatial hachuré en rouge et restreint aux limites de la sphère ne prend pas en compte le rampant contrairement au support hachuré en vert. Or, nous pouvons observer que la contribution du rampant à la SER globale est non négligeable.
En conclusion, l'analyse d'un expert en SER est indispensable pour déterminer correctement l'extension spatiale de l'objet de rétro-diffusion $a(\boldsymbol{r})$.

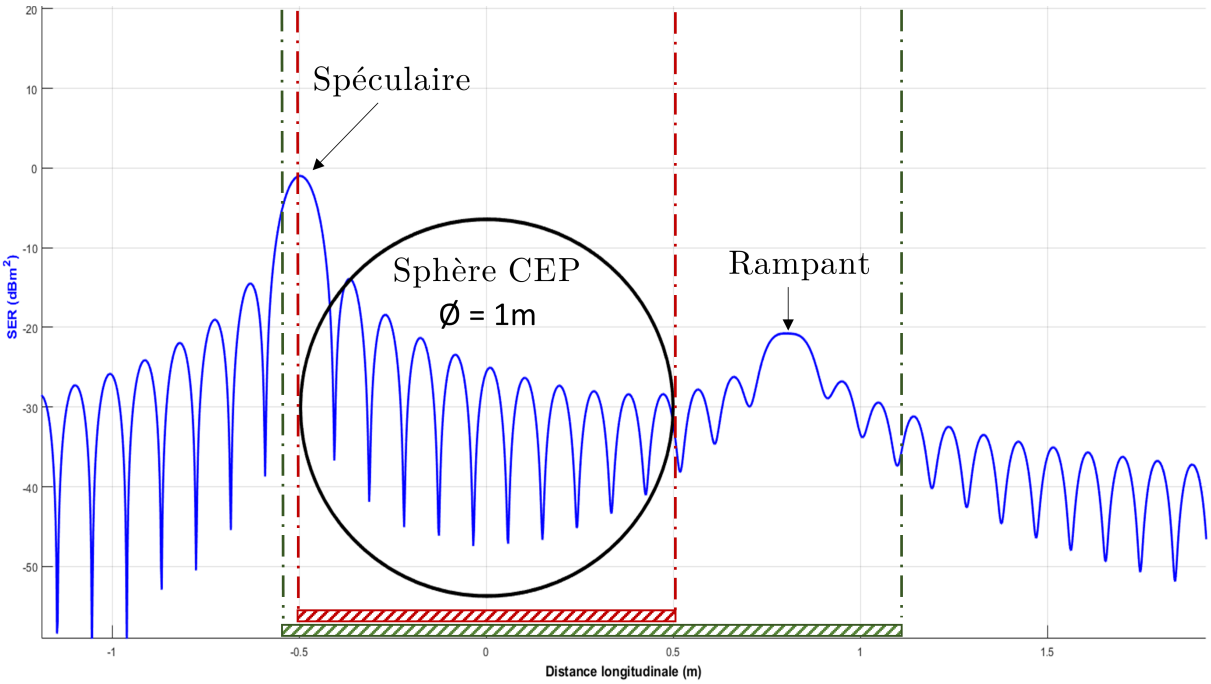


FIGURE 3.6 – Imagerie 1D avec la méthode PFA d'une sphère.

3.1.6 La transposition des *a priori* à la carte discrète

Les *a priori* mentionnés ci-avant ont été formulés pour un objet à variable continu $a(\boldsymbol{r})$. Or, comme expliqué dans la **Section 1.2.3.2** (p. 38), $a(\boldsymbol{r})$ est discrétisé sur une grille cartésienne régulière du domaine spatial.

Tous les *a priori* formulés dans un contexte continu peuvent être transposés au cas discret pour s'appliquer à la carte de rétro-diffusion $\boldsymbol{a}_{3\mathrm{D}}$ de dimension $N_x \times N_y \times N_z$.

Plus spécifiquement, la projection de $a(\boldsymbol{r})$ sur l'axe $\hat{\boldsymbol{z}}$ définie dans un contexte continu par (3.2) s'écrit de la façon suivante pour la carte de rétro-diffusion $\boldsymbol{a}_{\text{3D}}$ discrète :

$$\mathcal{P}_z(n_z) = \sum_{n_x=0}^{N_x-1} \sum_{n_y=0}^{N_y-1} \boldsymbol{a}_{\text{3D}}(n_x, n_y, n_z), \text{ pour } n_z \in [\![0, N_z - 1]\!] \tag{3.6}$$

et d'après **(AP1)**, $\mathcal{P}_z(n_z) \neq 0$ pour un nombre limité d'indices n_z.

Dans le cas discret, **(AP1)** s'écrit alors :

(AP1 discret)

La projection de la carte $\boldsymbol{a}_{\text{3D}}$ sur l'axe $\hat{\boldsymbol{z}}$ est nulle sauf pour un nombre relativement limité d'indices n_z.

De la même manière, les *a priori* **(AP2)** et **(AP3)** s'appliquent à la carte $\boldsymbol{a}_{\text{3D}}$ et s'écrivent :

(AP2 discret)

La carte $\boldsymbol{a}_{\text{3D}}$ peut-être décomposée en un nombre relativement restreint de facettes spéculaires connexes.

Notons qu'en « discret » la notion de connexité est implicitement liée à celle de voisinage.

(AP3 discret)

La carte de rétro-diffusion est constante par morceaux.

Les « morceaux » correspondent aux facettes spéculaires.

L'énergie de la carte de rétro-diffusion $\boldsymbol{a}_{\text{3D}}$ est définie par :

$$E = \sum_{n_x=0}^{N_x-1} \sum_{n_y=0}^{N_y-1} \sum_{n_z=0}^{N_z-1} |\boldsymbol{a}_{\text{3D}}(n_x, n_y, n_z)|^2 \tag{3.7}$$

et dans le cas discret, **(AP4)** s'écrit :

(AP4 discret)

L'énergie de la carte $\boldsymbol{a}_{\text{3D}}$ est relativement faible.

Enfin, l'expression de **(AP5)** dans le cas discret est analogue au cas continu :

(AP5 discret)

L'extension spatiale de la carte $\boldsymbol{a}_{\text{3D}}$ est limitée.

3.2 La construction d'un critère pénalisé et contraint

3.2.1 Les pénalités et la contrainte de support

L'objectif est de construire une carte de rétro-diffusion à partir des données $\boldsymbol{\sigma}$. De plus, il s'agit de tenir compte des *a priori* mentionnés ci-avant afin d'améliorer la résolution de la carte comparativement à celle obtenue avec la méthode conventionnelle. Pour cela nous proposons une approche consistant à minimiser un critère pénalisé et contraint. Dans la suite nous détaillons les différents termes qui interviennent dans ce critère.

Par souci de simplicité, dans la suite nous notons $\boldsymbol{a}$ le vecteur colonne de dimension $N \times 1$ (où $N = N_x N_y N_z$ est le nombre de voxels) obtenu par vectorisation de la carte $\boldsymbol{a}_{\text{3D}}$:

$$\boldsymbol{a} = \text{vec}(\boldsymbol{a}_{\text{3D}}). \tag{3.8}$$

Terme d'attache aux données $\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2$

L'attache aux données $\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2$ quantifie l'écart entre les données $\boldsymbol{\sigma}$ et la sortie modèle $\boldsymbol{H}\boldsymbol{a}$. Dans un sens, la réduction de ce terme permet de s'assurer que la solution $\hat{\boldsymbol{a}}$ est bien cohérente des données $\boldsymbol{\sigma}$. Notons que nous considérons une norme ℓ_2, qui est cohérente d'un bruit additif gaussien, car c'est le choix communément adopté pour décrire le bruit de mesure radar [Borden, 1992; Mensa, 1981].

Pénalité $\|\boldsymbol{P}\boldsymbol{a}\|_1$

Nous définissons $\boldsymbol{P}$ comme la matrice de projection sur l'axe $\hat{\boldsymbol{z}}$ colinéaire au vecteur d'onde central $\hat{\boldsymbol{k}}_c$. $\boldsymbol{P}$ est de dimension $N_z \times N$ et s'écrit :

$$\boldsymbol{P} = \left(\begin{array}{c|c|c|c} \underbrace{\begin{matrix}1 & & 1\\ 0 & & 0\\ & \vdots & \\ 0 & & 0\\ 0 & & 0\end{matrix}}_{N_x N_y} & \begin{matrix}0 & & 0\\ 1 & & 1\\ & \vdots & \\ 0 & & 0\\ 0 & & 0\end{matrix} & \cdots & \begin{matrix}0 & & 0\\ 0 & & 0\\ & \vdots & \\ 1 & & 1\\ 0 & & 0\end{matrix} \quad \begin{matrix}0 & & 0\\ 0 & & 0\\ & \vdots & \\ 0 & & 0\\ 1 & & 1\end{matrix} \end{array}\right) \Bigg\} N_z$$

La pénalité $\|\boldsymbol{P}\boldsymbol{a}\|_1$ permet de favoriser des cartes dont la projection sur $\hat{\boldsymbol{z}}$ est parcimonieuse, en cohérence avec l'*a priori* **(AP1 discret)**.

Pénalités $\|\boldsymbol{D}_x\boldsymbol{a}\|_1$ et $\|\boldsymbol{D}_y\boldsymbol{a}\|_1$

Nous définissons $\boldsymbol{D}_x$ comme la matrice multi-niveaux par blocs permettant de réaliser des différences entre pixels voisins selon l'axe $\hat{\boldsymbol{x}}$ (différence en ligne) pour tous les plans $n_z \in [\![0, N_z - 1]\!]$. Pour construire $\boldsymbol{D}_x$ nous définissons d'abord le bloc élémentaire $\boldsymbol{D}_x^{[0]}$ suivant, de dimension $N_x \times N_x$.

$$\boldsymbol{D}_x^{[0]} = \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 1 & & & -1 \end{pmatrix}$$

REMARQUE

L'ajout de la dernière ligne permet de rendre $\boldsymbol{D}_x^{[0]}$ circulante afin de simplifier les calculs numériques ensuite. Cette approximation circulante est licite si le champ de vue est choisi suffisamment grand et que l'objet d'intérêt est au centre de la grille spatiale. En effet, dans ce cas, le coefficient de rétro-diffusion est supposé nul sur les bords.

Nous définissons ensuite le bloc intermédiaire $\boldsymbol{D}_x^{[1]}$ suivant de dimension $N_xN_y \times N_xN_y$.

$$\boldsymbol{D}_x^{[1]} = \boldsymbol{D}_x^{[0]} \otimes \boldsymbol{I}_{N_y} = \begin{pmatrix} \boldsymbol{D}_x^{[0]} & \mathbf{O}_{N_y} & \cdots & \mathbf{O}_{N_y} \\ \mathbf{O}_{N_y} & \boldsymbol{D}_x^{[0]} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{O}_{N_y} \\ \mathbf{O}_{N_y} & \cdots & \mathbf{O}_{N_y} & \boldsymbol{D}_x^{[0]} \end{pmatrix}$$

Ce bloc intermédiaire permet de réaliser des différences entre pixels voisins selon l'axe $\hat{\boldsymbol{x}}$ (différence en ligne), pour un plan n_z donné.

La matrice $\boldsymbol{D}_x$, de dimension $N \times N$, est alors définie comme suit.

$$\boldsymbol{D}_x = \boldsymbol{D}_x^{[1]} \otimes \boldsymbol{I}_{N_z} = \boldsymbol{D}_x^{[0]} \otimes \boldsymbol{I}_{N_y N_z} = \begin{pmatrix} \boxed{\boldsymbol{D}_x^{[0]}} & \mathbf{O}_{N_y} & \cdots & \mathbf{O}_{N_y} & & & \\ \mathbf{O}_{N_y} & \boxed{\boldsymbol{D}_x^{[0]}} & \ddots & \vdots & & \mathbf{O}_{N_x N_y} & \\ \vdots & \ddots & \ddots & \mathbf{O}_{N_y} & & & \\ \mathbf{O}_{N_y} & \cdots & \mathbf{O}_{N_y} & \boxed{\boldsymbol{D}_x^{[0]}} & & & \\ & & & & \ddots & & \\ & \mathbf{O}_{N_x N_y} & & & \boxed{\boldsymbol{D}_x^{[0]}} & \mathbf{O}_{N_y} & \cdots & \mathbf{O}_{N_y} \\ & & & & \mathbf{O}_{N_y} & \boxed{\boldsymbol{D}_x^{[0]}} & \ddots & \vdots \\ & & & & \vdots & \ddots & \ddots & \mathbf{O}_{N_y} \\ & & & & \mathbf{O}_{N_y} & \cdots & \mathbf{O}_{N_y} & \boxed{\boldsymbol{D}_x^{[0]}} \end{pmatrix}$$

REMARQUE

> La matrice $\boldsymbol{D}_x$ est 3-niveaux circulante (car elle est bloc-diagonale et chaque sous-bloc encadré est circulant).

Nous construisons de manière similaire la matrice $\boldsymbol{D}_y$ qui est également multi-niveaux par blocs et réalise des différences entre pixels voisins selon l'axe $\hat{\boldsymbol{y}}$ (différence en colonne) pour tous les plans n_z.

Dans un premier temps, nous définissons le bloc élémentaire $\boldsymbol{D}_y^{[0]}$, de dimension $N_y \times N_y$ suivant.

$$\boldsymbol{D}_y^{[0]} = \begin{pmatrix} -1 & & & \boxed{1} \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \boxed{-1} \end{pmatrix}$$

Encore une fois, l'ajout de la dernière colonne permet de rendre $\boldsymbol{D}_y^{[0]}$ circulant. Cette approximation circulante reste valide pour les mêmes raisons que précédemment.

Nous définissons ensuite le bloc intermédiaire $\boldsymbol{D}_y^{[1]}$ de dimension $N_x N_y \times N_x N_y$ comme suit.

$$\boldsymbol{D}_y^{[1]} = \boldsymbol{I}_{N_x} \otimes \boldsymbol{D}_y^{[0]} = \begin{pmatrix} -\boldsymbol{I}_{N_x} & & & \boldsymbol{I}_{N_x} \\ \boldsymbol{I}_{N_x} & \ddots & & \\ & \ddots & \ddots & \\ & & \boldsymbol{I}_{N_x} & -\boldsymbol{I}_{N_x} \end{pmatrix}$$

Ce dernier permet de réaliser des différences entre pixels voisins selon l'axe $\hat{\boldsymbol{y}}$ (différence en colonne), pour un plan n_z donné.

La matrice $\boldsymbol{D}_y$, de dimension $N \times N$, est alors définie par :

$$\boldsymbol{D}_y = \boldsymbol{D}_y^{[1]} \otimes \boldsymbol{I}_{N_z} = \begin{pmatrix} \boxed{\begin{matrix} -\boldsymbol{I}_{N_x} & & \boldsymbol{I}_{N_x} \\ \boldsymbol{I}_{N_x} & \ddots & \\ & \ddots & \ddots \\ & & \boldsymbol{I}_{N_x} & -\boldsymbol{I}_{N_x} \end{matrix}} & & \boldsymbol{O}_{N_x N_y} \\ & \ddots & \\ \boldsymbol{O}_{N_x N_y} & & \boxed{\begin{matrix} -\boldsymbol{I}_{N_x} & & \boldsymbol{I}_{N_x} \\ \boldsymbol{I}_{N_x} & \ddots & \\ & \ddots & \ddots \\ & & \boldsymbol{I}_{N_x} & -\boldsymbol{I}_{N_x} \end{matrix}} \end{pmatrix}$$

La matrice $\boldsymbol{D}_y$ est elle aussi 3-niveaux circulante (car elle est bloc-diagonale et chaque sous-bloc encadré est circulant).

La réduction conjointe des termes $\|\boldsymbol{D}_x \boldsymbol{a}\|_1$ et $\|\boldsymbol{D}_y \boldsymbol{a}\|_1$ permet de favoriser, sur chaque plan n_z contenant des facettes spéculaires, des facettes connexes sur lesquelles l'objet de rétro-diffusion est constant, en cohérence avec les *a priori* **(AP2 discret)** et **(AP3 discret)**.

Pénalité $\|\boldsymbol{a}\|_2^2$

La pénalité $\|\boldsymbol{a}\|_2^2$ quantifie l'énergie de $\boldsymbol{a}$. Sa réduction favorise des solutions de faible norme, c'est pourquoi cette pénalité est souvent appelée *terme de rappel à zéro*. Cette pénalité est en cohérence avec l'*a priori* **(AP4 discret)**.

Remarque sur les pénalités $\|\boldsymbol{P}\boldsymbol{a}\|_1$, $\|\boldsymbol{D}_x \boldsymbol{a}\|_1$ et $\|\boldsymbol{D}_y \boldsymbol{a}\|_1$

Nous insistons sur les deux points suivants.

1. Les opérateurs $\boldsymbol{P}$, $\boldsymbol{D}_x$ et $\boldsymbol{D}_y$ s'appliquent au vecteur complexe $\boldsymbol{a}$ et non pas à son module $|\boldsymbol{a}|$. Comme expliqué en fin de **Section 3.1.3** (p. 91), cela diffère de certains domaines comme par exemple l'imagerie SAR aéroportée ou satellitaire où la scène

imagée est considérée comme « rugueuse » et où une cellule de résolution peut contenir de nombreux contributeurs provenant de sources différentes. Ceci conduit à une différence de phase d'une cellule de résolution à l'autre et donc à une inconstance du coefficient de rétro-diffusion complexe qui encode à la fois l'information de phase et d'amplitude. Il s'ensuit que dans ces applications, il est préférable de pénaliser le module $|\boldsymbol{a}|$ (car l'amplitude reste quant à elle assez constante d'un élément à l'autre [Cetin, 2001]), plutôt que le complexe $\boldsymbol{a}$. Bien évidemment, la pénalisation du module $|\boldsymbol{a}|$ est plus compliquée que celle du complexe $\boldsymbol{a}$ car les pénalités ne sont dès lors plus des normes d'une transformée linéaire de $\boldsymbol{a}$. Nous indiquons au lecteur intéressé – et sans rentrer dans plus de détails car nous n'y sommes pas confrontés ensuite – que [Onhon et Cetin, 2012] et [Sanders *et al.*, 2017] proposent une méthode appelée *correction de phase* pour contourner cette difficulté.

Dans un contexte d'analyse et de contrôle de SER, aux longueurs d'onde où les radars opèrent, nous pouvons considérer que la surface de la cible est « lisse ». De plus, et contrairement à l'imagerie SAR précédemment décrite, les contributeurs présents dans une cellule de résolution proviennent d'une même zone homogène de la cible et ont donc la même phase propre. Il en résulte qu'il est préférable de pénaliser directement le complexe $\boldsymbol{a}$ et de favoriser ainsi des facettes spéculaires sur lesquelles le coefficient de rétro-diffusion complexe est constant (*i.e.* la phase et l'amplitude sont constants).

2. Les opérateurs $\boldsymbol{P}$, $\boldsymbol{D}_x$ et $\boldsymbol{D}_y$ s'appliquent au vecteur $\boldsymbol{a}$ et non pas à la carte $\boldsymbol{a}_{3D}$. Le vecteur estimé $\hat{\boldsymbol{a}}$ est ensuite ré-écrit sous forme d'une carte 3D par opération inverse de l'opérateur vec :

$$\hat{\boldsymbol{a}}_{3D} = \mathrm{vec}^{-1}(\hat{\boldsymbol{a}}) \tag{3.9}$$

Contrainte de support

L'ensemble $\mathcal{C}$ décrivant les contraintes de support est cohérent de l'*a priori* **(AP5 discret)**. Le choix du support spatial repose sur la connaissance *a priori* de l'extension EM de l'objet d'intérêt.

Comme expliqué en fin de **Section 3.1.5**, la définition du support spatial, et donc de l'ensemble $\mathcal{C}$, peut bien entendu s'appuyer sur la connaissance de la géométrie de la cible imagée et sur l'expertise d'un analyste de SER. Cependant, il doit être judicieusement choisi afin de ne pas favoriser une solution erronée qui ne tiendrait pas compte de certains phénomènes comme les rampants ou les interactions multiples qui apparaissent comme des artefacts en dehors des limites de l'objet physique et qui participent également à la SER globale. On ne choisira donc pas de contraintes trop restrictives si nous disposons de peu d'information sur l'extension EM de l'objet. A cet égard, nous signalons qu'un choix de support coïncidant avec l'ensemble du domaine imagé revient de fait à ne pas contraindre le problème.

A contrario, nous montrons dans la **Section 4.2.2.2** que des connaissances *a priori* sur l'extension spatiale de l'objet de rétro-diffusion, menant à un choix justifié de contraintes de support spatial, permettent d'accroitre sensiblement la résolution, notamment transverse, de la carte de rétro-diffusion estimée.

3.2.2 L'écriture du critère et de la solution

Après avoir détaillé les différents termes permettant de construire une carte tenant compte des différents *a priori*, nous proposons le *critère pénalisé* suivant. En plus de la formulation des différents *a priori*, il s'agit de la contribution centrale de cette thèse.

$$\mathcal{J}(\boldsymbol{a}) = \frac{1}{2}\left\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\right\|_2^2 + \mu\left\|\boldsymbol{P}\boldsymbol{a}\right\|_1 + \lambda\left\|\boldsymbol{D}_x\boldsymbol{a}\right\|_1 + \lambda\left\|\boldsymbol{D}_y\boldsymbol{a}\right\|_1 + \frac{\nu}{2}\left\|\boldsymbol{a}\right\|_2^2 \tag{3.10}$$

Il inclut un terme d'attache aux données ainsi que plusieurs pénalités. La minimisation de ce critère permet de trouver une solution qui est à la fois fidèle aux données et qui satisfait les *a priori*. Les paramètres de pénalité λ, μ et ν sont positifs et permettent d'assurer un compromis entre l'attache aux données et les pénalités. La solution est alors définie comme :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}\in\mathbb{C}^N}{\operatorname{argmin}} \begin{cases} \dfrac{1}{2}\left\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\right\|_2^2 + \mu\left\|\boldsymbol{P}\boldsymbol{a}\right\|_1 + \lambda\left\|\boldsymbol{D}_x\boldsymbol{a}\right\|_1 + \lambda\left\|\boldsymbol{D}_y\boldsymbol{a}\right\|_1 + \dfrac{\nu}{2}\left\|\boldsymbol{a}\right\|_2^2 \\ \\ \text{s.c. } \boldsymbol{a} \in \mathcal{C} \end{cases} \tag{3.11}$$

où $\mathcal{C}$ est un ensemble fermé et convexe décrivant les contraintes de support.

Outre le fait que le problème est posé en grande dimension, une des difficultés de la minimisation est la non-différentiabilité du critère (3.10).

3.3 La minimisation du critère à l'aide de l'ADMM

L'étude et la revue des différents algorithmes permettant de minimiser un critère non différentiable et contraint, menées dans le chapitre précédent, nous encourage à utiliser l'ADMM pour les raisons décrites dans la section **Section 2.2.4.4** (p. 86). En effet, et comme nous allons le montrer ensuite, il s'avère que l'ADMM est particulièrement efficace pour minimiser le critère (3.11) car l'étape de mise à jour de la variable $\boldsymbol{a}$ est explicite et peut se calculer rapidement en tirant profit des spécificités de notre problème.

3.3.1 La réécriture de la solution sous forme ADMM équivalente

Dans un premier temps, pour gérer plus aisément les différentes pénalités et la contrainte, nous ré-écrivons la solution (3.11) sous une forme ADMM équivalente en introduisant les variables auxiliaires $\boldsymbol{v}_P \in \mathbb{C}^{N_z}$, $\boldsymbol{v}_x \in \mathbb{C}^N$, $\boldsymbol{v}_y \in \mathbb{C}^N$ et $\boldsymbol{v}_{\mathcal{C}} \in \mathbb{C}^N$.

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmin}} \begin{cases} \dfrac{1}{2}\left\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\right\|_2^2 + \mu\left\|\boldsymbol{v}_P\right\|_1 + \lambda\left\|\boldsymbol{v}_x\right\|_1 + \lambda\left\|\boldsymbol{v}_y\right\|_1 + \dfrac{\nu}{2}\left\|\boldsymbol{a}\right\|_2^2 + \mathcal{I}_{\mathcal{C}}(\boldsymbol{v}_{\mathcal{C}}) \\ \\ \text{s.c.} \begin{cases} \boldsymbol{v}_P = \boldsymbol{P}\boldsymbol{a} \\ \boldsymbol{v}_x = \boldsymbol{D}_x\boldsymbol{a} \\ \boldsymbol{v}_y = \boldsymbol{D}_y\boldsymbol{a} \\ \boldsymbol{v}_{\mathcal{C}} = \boldsymbol{a} \end{cases} \end{cases} \tag{3.12}$$

où $\mathcal{I}_{\mathcal{C}}$ est la fonction indicatrice de l'ensemble $\mathcal{C}$ qui s'écrit $\mathcal{I}_{\mathcal{C}}(\boldsymbol{v}_{\mathcal{C}}) = \begin{cases} 0 & \text{si } \boldsymbol{v}_{\mathcal{C}} \in \mathcal{C} \\ +\infty & \text{sinon} \end{cases}$.

Le lagrangien augmenté (sous sa forme normalisée) associé au problème (3.12) s'écrit comme suit.

$$\begin{aligned}\mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a},\boldsymbol{v},\boldsymbol{u}) = \frac{1}{2}\left\|\boldsymbol{\sigma}-\boldsymbol{H}\boldsymbol{a}\right\|_2^2 + \mu\left\|\boldsymbol{v}_P\right\|_1 + \lambda\left\|\boldsymbol{v}_x\right\|_1 + \lambda\left\|\boldsymbol{v}_y\right\|_1 + \frac{\nu}{2}\left\|\boldsymbol{a}\right\|_2^2 + \mathcal{I}_{\mathcal{C}}(\boldsymbol{v}_{\mathcal{C}}) \\ + \frac{\rho_P}{2}\left\|\boldsymbol{P}\boldsymbol{a}-\boldsymbol{v}_P+\boldsymbol{u}_P\right\|_2^2 - \frac{\rho_P}{2}\left\|\boldsymbol{u}_P\right\|_2^2 + \frac{\rho_D}{2}\left\|\boldsymbol{D}_x\boldsymbol{a}-\boldsymbol{v}_x+\boldsymbol{u}_x\right\|_2^2 - \frac{\rho_D}{2}\left\|\boldsymbol{u}_x\right\|_2^2 \\ + \frac{\rho_D}{2}\left\|\boldsymbol{D}_y\boldsymbol{a}-\boldsymbol{v}_y+\boldsymbol{u}_y\right\|_2^2 - \frac{\rho_D}{2}\left\|\boldsymbol{u}_y\right\|_2^2 + \frac{\rho_{\mathcal{C}}}{2}\left\|\boldsymbol{a}-\boldsymbol{v}_{\mathcal{C}}+\boldsymbol{u}_{\mathcal{C}}\right\|_2^2 - \frac{\rho_{\mathcal{C}}}{2}\left\|\boldsymbol{u}_{\mathcal{C}}\right\|_2^2\end{aligned} \tag{3.13}$$

- $\boldsymbol{\rho} = [\rho_P, \rho_D, \rho_{\mathcal{C}}]$ est le vecteur des paramètres de pénalité strictement positifs.
- $\boldsymbol{v} = [\boldsymbol{v}_P, \boldsymbol{v}_x, \boldsymbol{v}_y, \boldsymbol{v}_{\mathcal{C}}]$ est le vecteur des variables auxiliaires.
- $\boldsymbol{u} = [\boldsymbol{u}_P, \boldsymbol{u}_x, \boldsymbol{u}_y, \boldsymbol{u}_{\mathcal{C}}]$ est le vecteur des multiplicateurs de Lagrange.

3.3.2 L'écriture de l'algorithme et mise à jour des variables

L'algorithme ADMM permettant de résoudre le problème (3.12) s'écrit comme suit.

Algorithme 9 : Algorithme ADMM pour le problème (3.4)

Entrées : $k = 0$: initialisation du compteur ;
$\boldsymbol{v}_P^{(0)}, \boldsymbol{v}_x{}^{(0)}, \boldsymbol{v}_y{}^{(0)}, \boldsymbol{v}_{\mathcal{C}}^{(0)}$: les variables auxiliaires initiales ;
$\boldsymbol{u}_P^{(0)}, \boldsymbol{u}_x{}^{(0)}, \boldsymbol{u}_y{}^{(0)}, \boldsymbol{u}_{\mathcal{C}}^{(0)}$: les multiplicateurs de Lagrange initiaux ;
$(\rho_P, \rho_D, \rho_{\mathcal{C}}) > 0$: les paramètres de pénalité fixes ;
$(\mu, \lambda, \nu) \geqslant 0$: les hyperparamètres ;
Sorties : $\hat{\boldsymbol{a}}$: la carte estimée.

Répéter jusqu'à convergence

$$\begin{aligned}
\boldsymbol{a}^{(k+1)} &= \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}, \boldsymbol{v}^{(k)}, \boldsymbol{u}^{(k)}) \\
\boldsymbol{v}_P^{(k+1)} &= \underset{\boldsymbol{v}_P}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P, \boldsymbol{v}_x^{(k)}, \boldsymbol{v}_y^{(k)}, \boldsymbol{v}_{\mathcal{C}}^{(k)}, \boldsymbol{u}^{(k)}) \\
\boldsymbol{v}_x^{(k+1)} &= \underset{\boldsymbol{v}_x}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P^{(k+1)}, \boldsymbol{v}_x, \boldsymbol{v}_y^{(k)}, \boldsymbol{v}_{\mathcal{C}}^{(k)}, \boldsymbol{u}^{(k)}) \\
\boldsymbol{v}_y^{(k+1)} &= \underset{\boldsymbol{v}_y}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P^{(k+1)}, \boldsymbol{v}_x^{(k+1)}, \boldsymbol{v}_y, \boldsymbol{v}_{\mathcal{C}}^{(k)}, \boldsymbol{u}^{(k)}) \\
\boldsymbol{v}_{\mathcal{C}}^{(k+1)} &= \underset{\boldsymbol{v}_{\mathcal{C}}}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P^{(k+1)}, \boldsymbol{v}_x^{(k+1)}, \boldsymbol{v}_y^{(k+1)}, \boldsymbol{v}_{\mathcal{C}}, \boldsymbol{u}^{(k)}) \\
\boldsymbol{u}_P^{(k+1)} &= \boldsymbol{u}_P^{(k)} + \boldsymbol{P}\boldsymbol{a}^{(k+1)} - \boldsymbol{v}_P^{(k+1)} \\
\boldsymbol{u}_x^{(k+1)} &= \boldsymbol{u}_x^{(k)} + \boldsymbol{D}_x\boldsymbol{a}^{(k+1)} - \boldsymbol{v}_x^{(k+1)} \\
\boldsymbol{u}_y^{(k+1)} &= \boldsymbol{u}_y^{(k)} + \boldsymbol{D}_y\boldsymbol{a}^{(k+1)} - \boldsymbol{v}_y^{(k+1)} \\
\boldsymbol{u}_{\mathcal{C}}^{(k+1)} &= \boldsymbol{u}_{\mathcal{C}}^{(k)} + \boldsymbol{a}^{(k+1)} - \boldsymbol{v}_{\mathcal{C}}^{(k+1)} \\
k &= k+1
\end{aligned}$$

Nous détaillons maintenant la mise à jour de chaque variable. Les preuves sont données dans l'**Annexe D** (p. 171).

Mise à jour de $\boldsymbol{v}_P, \boldsymbol{v}_x, \boldsymbol{v}_y$ et $\boldsymbol{v}_\mathcal{C}$

$$\begin{aligned}\boldsymbol{v}_P^{(k+1)} &= \underset{\boldsymbol{v}_P}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P, \boldsymbol{v}_x^{(k)}, \boldsymbol{v}_y^{(k)}, \boldsymbol{v}_\mathcal{C}^{(k)}, \boldsymbol{u}_P^{(k)}, \boldsymbol{u}_x^{(k)}, \boldsymbol{u}_y^{(k)}, \boldsymbol{u}_\mathcal{C}^{(k)}) \\ &= \mathrm{S}_{\mu/\rho_P}\left(\boldsymbol{P}\boldsymbol{a}^{(k+1)} + \boldsymbol{u}_P^{(k)}\right)\end{aligned} \tag{3.14}$$

$$\begin{aligned}\boldsymbol{v}_x^{(k+1)} &= \underset{\boldsymbol{v}_x}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P^{(k+1)}, \boldsymbol{v}_x, \boldsymbol{v}_y^{(k)}, \boldsymbol{v}_\mathcal{C}^{(k)}, \boldsymbol{u}_P^{(k)}, \boldsymbol{u}_x^{(k)}, \boldsymbol{u}_y^{(k)}, \boldsymbol{u}_\mathcal{C}^{(k)}) \\ &= \mathrm{S}_{\lambda/\rho_D}\left(\boldsymbol{D}_x\boldsymbol{a}^{(k+1)} + \boldsymbol{u}_x^{(k)}\right)\end{aligned} \tag{3.15}$$

$$\begin{aligned}\boldsymbol{v}_y^{(k+1)} &= \underset{\boldsymbol{v}_y}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P^{(k+1)}, \boldsymbol{v}_x^{(k+1)}, \boldsymbol{v}_y, \boldsymbol{v}_\mathcal{C}^{(k)}, \boldsymbol{u}_P^{(k)}, \boldsymbol{u}_x^{(k)}, \boldsymbol{u}_y^{(k)}, \boldsymbol{u}_\mathcal{C}^{(k)}) \\ &= \mathrm{S}_{\lambda/\rho_D}\left(\boldsymbol{D}_y\boldsymbol{a}^{(k+1)} + \boldsymbol{u}_y^{(k)}\right)\end{aligned} \tag{3.16}$$

$$\begin{aligned}\boldsymbol{v}_\mathcal{C}^{(k+1)} &= \underset{\boldsymbol{v}_\mathcal{C}}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P^{(k+1)}, \boldsymbol{v}_x^{(k+1)}, \boldsymbol{v}_y^{(k+1)}, \boldsymbol{v}_\mathcal{C}, \boldsymbol{u}_P^{(k)}, \boldsymbol{u}_x^{(k)}, \boldsymbol{u}_y^{(k)}, \boldsymbol{u}_\mathcal{C}^{(k)}) \\ &= \mathcal{P}_\mathcal{C}(\boldsymbol{a}^{(k+1)} + \boldsymbol{u}_\mathcal{C}^{(k)})\end{aligned} \tag{3.17}$$

où $\mathcal{P}_\mathcal{C}$ est l'opérateur de projection euclidienne sur l'ensemble $\mathcal{C}$ et S est l'opérateur de seuillage doux pour une variable complexe qui sont définis dans l'**Annexe E** (p. 185).

Les mises-à-jour des variables auxiliaires $\boldsymbol{v}_P, \boldsymbol{v}_x, \boldsymbol{v}_y$ et $\boldsymbol{v}_\mathcal{C}$ sont très simples. Elles sont séparables et directes et peuvent être menées en parallèle.

- Les mises à jour de $\boldsymbol{v}_P, \boldsymbol{v}_x, \boldsymbol{v}_y$ consistent simplement à appliquer un seuillage doux S_{μ/ρ_P} ou $\mathrm{S}_{\lambda/\rho_D}$ élément par élément, c'est-à-dire à mettre à zéro les composantes dont le module est inférieur au seuil et à pondérer les autres comme illustré sur la Figure 3.7.

- La mise à jour de $\boldsymbol{v}_\mathcal{C}$ est également très simple. Elle consiste à mettre à zéro les composantes en dehors du support spatial et à conserver les autres.

Mise à jour de $\boldsymbol{u}_P, \boldsymbol{u}_x, \boldsymbol{u}_y$ et $\boldsymbol{u}_\mathcal{C}$

$$\begin{aligned}\boldsymbol{u}_P^{(k+1)} &= \boldsymbol{u}_P^{(k)} + \boldsymbol{P}\boldsymbol{a}^{(k+1)} - \boldsymbol{v}_P^{(k+1)} \\ \boldsymbol{u}_x^{(k+1)} &= \boldsymbol{u}_x^{(k)} + \boldsymbol{D}_x\boldsymbol{a}^{(k+1)} - \boldsymbol{v}_x^{(k+1)} \\ \boldsymbol{u}_y^{(k+1)} &= \boldsymbol{u}_y^{(k)} + \boldsymbol{D}_y\boldsymbol{a}^{(k+1)} - \boldsymbol{v}_y^{(k+1)} \\ \boldsymbol{u}_\mathcal{C}^{(k+1)} &= \boldsymbol{u}_\mathcal{C}^{(k)} + \boldsymbol{a}^{(k+1)} - \boldsymbol{v}_\mathcal{C}^{(k+1)}\end{aligned} \tag{3.18}$$

Les mises à jours des multiplicateurs de Lagrange $\boldsymbol{u}_P, \boldsymbol{u}_x, \boldsymbol{u}_y$ et $\boldsymbol{u}_\mathcal{C}$ correspondent à des montées dans la direction du gradient des fonctions duales respectives (*c.f.* **Section 2.2.4.4** (p. 83)). Elles sont également séparables et directes et peuvent être menées en parallèle. Ces mises à jours sont très simples car elles correspondent à des manipulations linéaires des variables auxiliaires $\boldsymbol{v}_P, \boldsymbol{v}_x, \boldsymbol{v}_y$ et $\boldsymbol{v}_\mathcal{C}$ calculées avant.

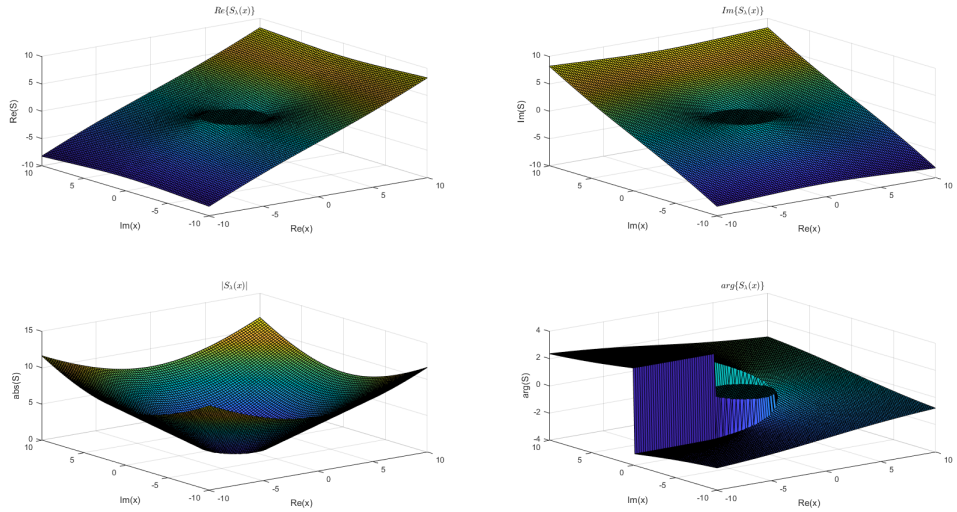


FIGURE 3.7 – Seuillage doux complexe pour un seuil $\lambda = 2{,}5$.

Mise à jour de $\boldsymbol{a}$

$$\boldsymbol{a}^{(k+1)} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{\rho}}(\boldsymbol{a}, \boldsymbol{v}^{(k)}, \boldsymbol{u}^{(k)}) = \boldsymbol{G}^{-1}\boldsymbol{d}^{(k)} \tag{3.19}$$

Nous montrons dans l'**Annexe D** (p. 172) que le lagrangien est quadratique et strictement convexe selon $\boldsymbol{a}$ et que nous avons :

$$\begin{aligned}
\boldsymbol{G} &= \boldsymbol{H}^\dagger\boldsymbol{H} + \rho_P\boldsymbol{P}^\dagger\boldsymbol{P} + \rho_D(\boldsymbol{D}_x^\dagger\boldsymbol{D}_x + \boldsymbol{D}_y^\dagger\boldsymbol{D}_y) + (\nu + \rho_{\mathcal{C}})\boldsymbol{I}_N \\
\boldsymbol{d}^{(k)} &= \boldsymbol{H}^\dagger\boldsymbol{\sigma} + \boldsymbol{t}^{(k)} \\
\boldsymbol{t}^{(k)} &= \rho_P\boldsymbol{P}^\dagger(\boldsymbol{v}_P^{(k)} - \boldsymbol{u}_P^{(k)}) + \rho_D\left[\boldsymbol{D}_x^\dagger(\boldsymbol{v}_x^{(k)} - \boldsymbol{u}_x^{(k)}) + \boldsymbol{D}_y^\dagger(\boldsymbol{v}_y^{(k)} - \boldsymbol{u}_y^{(k)})\right] + \rho_{\mathcal{C}}(\boldsymbol{v}_{\mathcal{C}}^{(k)} - \boldsymbol{u}_{\mathcal{C}}^{(k)})
\end{aligned} \tag{3.20}$$

REMARQUE

1. La matrice $\boldsymbol{G}$ est inversible. En effet, les matrices $\boldsymbol{H}^\dagger\boldsymbol{H}$, $\boldsymbol{P}^\dagger\boldsymbol{P}$, $\boldsymbol{D}_x^\dagger\boldsymbol{D}_x$ et $\boldsymbol{D}_y^\dagger\boldsymbol{D}_y$ sont semi-définies positives. Par somme, et comme $(\nu + \rho_{\mathcal{C}}) > 0$, la matrice $\boldsymbol{G}$ est définie positive. La preuve est donnée dans l'**Annexe D** (p. 172).
2. $\boldsymbol{d}$ est un vecteur colonne de taille $N \times 1$ (*i.e.* de la même taille que la carte vectorisée).

À hyperparamètre ν et à paramètres de pénalités ρ_P, ρ_D et $\rho_{\mathcal{C}}$ donnés, la matrice $\boldsymbol{G}^{-1}$ ne varie pas au cours des itérations. Il en est de même pour $\boldsymbol{H}^\dagger\boldsymbol{\sigma}$, aussi connu sous le nom de *carte sale* (ou *dirty map* en anglais). Néanmoins, en grande dimension, la taille des matrices impliquées dans (3.19) est telle qu'il est impossible de les former, de les mémoriser et encore moins de les inverser.

Le principal coût calculatoire repose donc sur la mise à jour de l'objet d'intérêt $\boldsymbol{a}$. Nous expliquons ensuite les stratégies envisageables pour mettre à jour rapidement $\boldsymbol{a}$ à chaque itération.

3.3.3 La gestion de la mise à jour de la variable $\boldsymbol{a}$

Calculs rapides dans le domaine fréquentiel

Cette partie s'inspire du travail de [Hunt, 1973] qui a proposé de circulariser les matrices Tœplitz par bloc Tœplitz (dans un contexte bi-dimensionnel) pour les diagonaliser dans le domaine de Fourier et mener ainsi rapidement les calculs à l'aide de TFR.

Nous avons montré **Section 1.2.3.4** (p. 43) que si les données sont sur une grille cartésienne régulière du k-space (après un reformatage par exemple) et si $\delta k_x \delta_x = 1/N_x$ (idem pour y et z), alors la matrice $\boldsymbol{H}$ s'écrit :

$$\boldsymbol{H} = \alpha\sqrt{N}\,\boldsymbol{\Delta}_\Psi \boldsymbol{S}\boldsymbol{F}_{3D}\boldsymbol{\Delta}_\Phi$$

Nous montrons alors dans l'**Annexe D** (p. 173) que $\boldsymbol{H}^\dagger\boldsymbol{H}$ est une matrice 3-niveaux circulante qui peut donc être diagonalisée dans le domaine de Fourier et nous avons :

$$\boldsymbol{H}^\dagger\boldsymbol{H} = \boldsymbol{F}_{3D}^\dagger \boldsymbol{\Lambda}_{\boldsymbol{H}^\dagger\boldsymbol{H}} \boldsymbol{F}_{3D} \tag{3.21}$$

où $\boldsymbol{\Lambda}_{\boldsymbol{H}^\dagger\boldsymbol{H}}$ est une matrice diagonale qui s'écrit comme suit.

$$\boldsymbol{\Lambda}_{\boldsymbol{H}^\dagger\boldsymbol{H}} = \boldsymbol{F}_{3D}\boldsymbol{H}^\dagger\boldsymbol{H}\boldsymbol{F}_{3D}^\dagger = (N_xN_yN_z)\boldsymbol{\Pi}_{3\text{D}}\boldsymbol{S}^t\boldsymbol{S}\boldsymbol{\Pi}_{3\text{D}}^\dagger \tag{3.22}$$

avec $\boldsymbol{\Pi}_{3\text{D}} = \boldsymbol{F}_{3D}\boldsymbol{\Delta}_\Phi^*\boldsymbol{F}_{3D}^\dagger$ une matrice de décalage circulaire si et seulement si $P = k^{[0]}/\delta k \in \mathbb{Z}$ (cf. preuve p. 178). En effet, sous cette hypothèse, les éléments de la matrice diagonale $\boldsymbol{S}^t\boldsymbol{S}$ sont décalés circulairement, sinon le décalage n'est pas entier.

De la même manière nous montrons dans l'**Annexe D** (p. 173) que $\boldsymbol{P}^\dagger\boldsymbol{P}$, $\boldsymbol{D}_x^\dagger\boldsymbol{D}_x$ et $\boldsymbol{D}_y^\dagger\boldsymbol{D}_y$ sont 3-niveaux circulantes et donc diagonalisables dans le domaine de Fourier de sorte que :

$$\begin{aligned}
\boldsymbol{D}_x^\dagger\boldsymbol{D}_x &= \boldsymbol{F}_{3D}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{D}_x{}^\dagger\boldsymbol{D}_x}\boldsymbol{F}_{3D}\\
\boldsymbol{D}_y^\dagger\boldsymbol{D}_y &= \boldsymbol{F}_{3D}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{D}_y{}^\dagger\boldsymbol{D}_y}\boldsymbol{F}_{3D}\\
\boldsymbol{P}^\dagger\boldsymbol{P} &= \boldsymbol{F}_{3D}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{P}^\dagger\boldsymbol{P}}\boldsymbol{F}_{3D}
\end{aligned} \tag{3.23}$$

où $\boldsymbol{\Lambda}_{\boldsymbol{D}_x{}^\dagger\boldsymbol{D}_x}$, $\boldsymbol{\Lambda}_{\boldsymbol{D}_y{}^\dagger\boldsymbol{D}_y}$ et $\boldsymbol{\Lambda}_{\boldsymbol{P}^\dagger\boldsymbol{P}}$ sont des matrices diagonales.

En tant que somme de matrices 3-niveaux circulantes (dont la taille des sous-blocs est similaire d'une matrice à l'autre), $\boldsymbol{G}$ est également une matrice 3-niveaux circulante. Par suite, nous montrons dans l'**Annexe D** (p. 173) que $\boldsymbol{G}$ est diagonalisable dans le domaine fréquentiel et peut s'écrire :

$$\boldsymbol{G} = \boldsymbol{F}_{3D}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{G}}\boldsymbol{F}_{3D} \tag{3.24}$$

où, $\boldsymbol{\Lambda}_{\boldsymbol{G}} = \boldsymbol{\Lambda}_{\boldsymbol{H}^\dagger\boldsymbol{H}} + \rho_P\boldsymbol{\Lambda}_{\boldsymbol{P}^\dagger\boldsymbol{P}} + \rho_D(\boldsymbol{\Lambda}_{\boldsymbol{D}_x{}^\dagger\boldsymbol{D}_x} + \boldsymbol{\Lambda}_{\boldsymbol{D}_y{}^\dagger\boldsymbol{D}_y}) + (\nu + \rho_{\mathcal{C}})\boldsymbol{I}_N$ est une matrice diagonale.

Finalement, en exploitant le résultat (3.24), la mise à jour (3.19) de $\boldsymbol{a}$ peut être effectuée dans le domaine fréquentiel (une preuve est donnée dans l'**Annexe D** (p. 181)) :

$$\boldsymbol{F}_{3D}\boldsymbol{a}^{(k+1)} = \alpha N\boldsymbol{\Lambda}_{\boldsymbol{G}}^{-1}\left(\boldsymbol{\Pi}_{3\text{D}}\boldsymbol{S}^t\boldsymbol{\Delta}_\Psi^*\breve{\boldsymbol{\sigma}} + \boldsymbol{l}\right) \tag{3.25}$$

où le vecteur $\boldsymbol{l}$ s'écrit :

$$\boldsymbol{l} = \boldsymbol{F}_{3D}\left[\rho_P\boldsymbol{P}^\dagger(\boldsymbol{v}_P^{(k)} - \boldsymbol{u}_P^{(k)}) + \rho_D\left(\boldsymbol{D}_x^\dagger(\boldsymbol{v}_x^{(k)} - \boldsymbol{u}_x^{(k)}) + \boldsymbol{D}_y^\dagger(\boldsymbol{v}_y^{(k)} - \boldsymbol{u}_y^{(k)})\right) + \rho_{\mathcal{C}}(\boldsymbol{v}_{\mathcal{C}}^{(k)} - \boldsymbol{u}_{\mathcal{C}}^{(k)})\right].$$

La mise à jour (3.25) de $\boldsymbol{a}$ dans le domaine fréquentiel est rapide car :

- $\boldsymbol{\Lambda_G}$ est une matrice diagonale dont l'inversion est directe et qui peut être calculée une seule fois en amont.
- $\boldsymbol{\Pi_{3D} S^t \Delta_\Psi^* \breve{\sigma}}$ est l'expression de la carte sale dans le domaine fréquentiel. Ce terme peut également être calculé une seule fois en amont rapidement et consiste simplement à déphaser les données $\boldsymbol{\sigma}$, à les compléter avec des 0 puis à les permuter. Notons que la carte sale est la carte obtenue avec la méthode PFA (1.32) (p. 55).
- Les produits matrice-vecteur $\boldsymbol{P^\dagger}(\boldsymbol{v_P^{(k)}} - \boldsymbol{u_P^{(k)}})$, $\boldsymbol{D_x^\dagger}(\boldsymbol{v_x^{(k)}} - \boldsymbol{u_x^{(k)}})$ et $\boldsymbol{D_y^\dagger}(\boldsymbol{v_y^{(k)}} - \boldsymbol{u_y^{(k)}})$ impliquent des matrices creuses et sont rapidement calculables.
- $\boldsymbol{l}$ est calculé avec une TFR 3D en pratique.

Enfin, en multipliant (3.25) à gauche par $\boldsymbol{F_{3D}^{-1}}$, ce qui est réalisé en pratique par une TFRI 3D, nous retrouvons la mise à jour de $\boldsymbol{a}$ dans le domaine spatial :

$$\boldsymbol{a}^{(k+1)} = \alpha N \boldsymbol{F}_{3D}^{-1} \boldsymbol{\Lambda}_{\boldsymbol{G}}^{-1} \left(\boldsymbol{\Pi}_{3D} \boldsymbol{S}^t \boldsymbol{\Delta}_\Psi^* \breve{\boldsymbol{\sigma}} + \boldsymbol{l} \right) \tag{3.26}$$

Pour conclure, si le reformatage des données sur une grille cartésienne régulière du k-space n'est pas indispensable, il permet néanmoins de mener rapidement les calculs de mise à jour de $\boldsymbol{a}$ à chaque itération de l'**Algorithme 9**. Dans la section suivante, nous expliquons succinctement qu'il est possible, au prix d'un coût calculatoire plus élevé, d'utiliser d'autres méthodes qui ne nécessitent pas de reformatage des données.

Minimisation d'une fonction quadratique et résolution d'un système de Tœplitz

Nous rappelons que le lagrangien augmenté (3.13) est une fonction quadratique strictement convexe de $\boldsymbol{a}$ comme démontré dans l'**Annexe D** (p. 171). La littérature au sujet de la minimisation de fonctions quadratiques convexes est vaste et de nombreux algorithmes itératifs existent. Parmi les plus connues nous pouvons citer les méthodes de descente : algorithme de descente du gradient, algorithme du gradient conjugué, méthode de Newton, méthode de quasi-Newton, méthode de Gauss-Newton ou encore la descente par bloc de coordonnées. Ces dernières reposent sur l'évaluation de la fonction et de son gradient (voire de son Hessien).

De prime abord, ces derniers pourraient s'avérer intéressants pour la mise à jour de $\boldsymbol{a}$ car la convergence lors d'une itération de l'ADMM n'est pas nécessaire comme expliqué dans la **Section 2.2.4.4** (p. 85). Nous montrons l'**Annexe D** (p. 171) que le gradient du lagrangien augmenté par rapport à $\boldsymbol{a}$ est donné par :

$$\nabla_{\boldsymbol{a}} \mathcal{L}_{\boldsymbol{\rho}} = \boldsymbol{G}\boldsymbol{a} - \boldsymbol{d} \tag{3.27}$$

A chaque itération, $\boldsymbol{a}^{(k+1)}$ annule le gradient (3.27) et nous devons dès lors résoudre le système linéaire suivant :

$$\boldsymbol{G}\boldsymbol{a}^{(k+1)} - \boldsymbol{d}^{(k)} = \boldsymbol{0} \tag{3.28}$$

Comme expliqué précédemment, en grande dimension, il n'est généralement pas intéressant (voire impossible) de former les matrices pour effectuer les calculs matrices vecteurs. Si les données ne sont pas sur une grille régulière du k-space, la matrice $\boldsymbol{H^\dagger H}$ générale est

3-niveaux Tœplitz (*i.e.* Tœplitz par bloc Tœplitz par bloc Tœplitz). De la même manière, les matrices $\boldsymbol{P}^\dagger\boldsymbol{P}$, $\boldsymbol{D}_x^\dagger\boldsymbol{D}_x$, $\boldsymbol{D}_y^\dagger\boldsymbol{D}_y$ (et bien entendu $\boldsymbol{I}_N$) sont 3-niveaux circulantes (*i.e.* circulante par bloc circulante par bloc circulante) comme démontré dans l'**Annexe D** (p. 173). Elles sont donc *a fortiori* 3-niveaux Tœplitz. En tant que somme de matrices 3-niveaux Tœplitz (avec des sous-blocs de même dimension d'une matrice à l'autre) la matrice $\boldsymbol{G}$ est également 3-niveaux Tœplitz. (3.28) est donc un système linéaire de Tœplitz.

Pour ce genre de systèmes, il existe des méthodes plus performantes que l'*élimination de Gauss-Jordan* dont la complexité est en $\mathcal{O}\big(N^3\big)$. Sans la détailler, nous pouvons citer la *méthode de Levinson-Durbin* – et dans le cas d'espèce son extension par bloc pour des matrices de Tœplitz multi-niveaux [Musicus, 1988] – dont la complexité est en $\mathcal{O}\big(N^2\big)$.

Par ailleurs, notons que dans (3.28), le produit $\boldsymbol{G}\boldsymbol{a}^{(k+1)}$ correspond à une convolution discrète tri-dimensionnelle car $\boldsymbol{G}$ est 3-niveaux Tœplitz. Il est donc possible de le calculer par TFR et de réduire la complexité de ce calcul à $\mathcal{O}\big(N \ln N\big)$.

En grande dimension, l'évaluation de $\boldsymbol{d}^{(k)}$ à chaque itération est un autre point de difficulté. En particulier, et même s'il doit être réalisé une seule fois en amont, le calcul de la carte sale $\boldsymbol{H}^\dagger\boldsymbol{\sigma}$ n'est pas direct car la matrice $\boldsymbol{H}$ (1.20) est pleine et n'a pas de structure de Tœplitz ou circulante.

3.4 Des compléments sur la méthode SPRITE

Dans cette partie nous apportons certains compléments sur la méthode SPRITE. En particulier nous présentons sa version mono-dimensionnelle, nous expliquons comment sont initialisées les variables, comment nous choisissons les paramètres et nous explicitons le critère d'arrêt de l'algorithme.

3.4.1 La version mono-dimensionnelle

Afin d'améliorer la résolution des cartes 1D (*i.e.* la SER selon la distance longitudinale), nous avons proposé dans [Benoudiba-Campanini *et al.*, 2017a] et [Benoudiba-Campanini *et al.*, 2017b] une méthode haute-résolution reposant sur la minimisation d'un critère pénalisé et contraint :

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}\in\mathbb{C}^{N_z}}{\operatorname{argmin}} \begin{cases} \dfrac{1}{2}\left\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\right\|_2^2 + \mu\left\|\boldsymbol{a}\right\|_1 + \dfrac{\nu}{2}\left\|\boldsymbol{a}\right\|_2^2 \\ \\ \text{s.c. } \boldsymbol{a} \in \mathcal{C} \end{cases} \tag{3.29}$$

Nous pouvons constater que le problème (3.29) est une version « dégénérée » du problème (3.11) tri-dimensionnel où :

- $N_x = N_y = 1$ car le problème est mono-dimensionnel.
- $\boldsymbol{P} = \boldsymbol{I}_{N_z}$ par définition de la matrice $\boldsymbol{P}$ dans le cas où $N_x = N_y = 1$. Ceci peut aussi se comprendre par le fait que $\boldsymbol{a}$ n'a que des composantes selon l'axe longitudinal $\hat{\boldsymbol{z}}$, ainsi $\boldsymbol{P}\boldsymbol{a} = \boldsymbol{a}$.

- $\boldsymbol{D}_x = \boldsymbol{D}_y = \mathbf{0}_{Nz \times N_z}$ car la notion de différence en ligne et en colonne n'existe pas en 1D où $\boldsymbol{a}$ n'a que des composantes selon $\hat{\boldsymbol{z}}$.

Considérations faites des éléments mentionnés ci-dessus, le problème (3.29) se résout d'une manière analogue à celle décrite par l'**Algorithme 9**.

Dans la suite nous appelons donc SPRITE 1D la méthode permettant d'estimer des cartes mono-dimensionnelles.

3.4.2 L'initialisation des variables

La première étape de l'**Algorithme 9** consiste en la mise à jour de $\boldsymbol{a}^{(1)}$ et ne nécessite pas d'initialiser $\boldsymbol{a}^{(0)}$. Les variables auxiliaires initiales $\boldsymbol{v}^{(0)}$ ainsi que les multiplicateurs de Lagrange initiaux $\boldsymbol{u}^{(0)}$ sont initialisés à $\mathbf{0}$.

Une alternative est de calculer préalablement $\hat{\boldsymbol{a}}_{\text{PFA}}$ avec la méthode PFA puis d'initialiser les variables auxiliaires avec (3.14) – (3.17) et les multiplicateurs de Lagrange comme dans l'**Algorithme 9**. Cependant, en pratique, il s'avère que cette démarche n'est pas plus avantageuse que la précédente.

3.4.3 Le choix des paramètres

Paramètres de pénalité

Comme expliqué **Section 2.2.4.4** (p. 83), les paramètres de pénalité ρ_P, ρ_D et $\rho_{\mathcal{C}}$ strictement positifs, permettent de contrôler le compromis entre la résolution du problème primal et du problème dual aux cours des itérations. Ce sont donc des paramètres algorithmiques importants mais qui n'influent que sur la vitesse de convergence.

Récemment, et comme mentionné **Section 2.2.4.4**, des stratégies adaptatives ont été développées où ces paramètres varient au cours des itérations. Cependant comme le précise [Boyd, 2010], la convergence n'est alors plus garantie. Par ailleurs, si ρ_P, ρ_D et $\rho_{\mathcal{C}}$ ne sont pas fixes, il faut recalculer la matrice $\boldsymbol{\Lambda}_{\boldsymbol{G}}$ apparaissant dans (3.24) à chaque itération.

Pour ces raisons, nous optons pour des paramètres de pénalités fixes qui sont choisis de façon à ce que la norme du résidu primal et la norme du résidu dual soient dans le même ordre de grandeur [Boyd, 2010]. Empiriquement, nous constatons que les choix de paramètres suivants sont satisfaisants.

$$\rho_P = N_z/2 \quad \text{et} \quad \rho_D = \rho_{\mathcal{C}} = N/2$$

Ceci s'explique par le fait qu'ils permettent d'établir un compromis entre la valeur propre moyenne des matrices $\boldsymbol{P}^\dagger\boldsymbol{P}$, $\boldsymbol{D}_x^\dagger\boldsymbol{D}_x$ et $\boldsymbol{D}_y^\dagger\boldsymbol{D}_y$ et $\boldsymbol{I}_N$ et la valeur propre moyenne de $\boldsymbol{H}^\dagger\boldsymbol{H}$.

Enfin, notons que ρ_D est le même pour les contraintes d'égalité impliquant $\boldsymbol{D}_x$ et $\boldsymbol{D}_y$ dans l'expression (3.13). Ce choix permet de ne pas favoriser une direction particulière entre $\hat{\boldsymbol{x}}$ et $\hat{\boldsymbol{y}}$ lorsque nous effectuons les différences entre pixels voisins.

Hyperparamètres

Le choix des hyperparamètres λ, μ et ν est crucial car il détermine la solution $\hat{\boldsymbol{a}}$. Cependant le réglage est beaucoup plus complexe que celui des paramètres de pénalité $\boldsymbol{\rho}$. En effet, à notre connaissance, pour des problèmes impliquant des normes ℓ_1 tels que (3.11), la question de la détermination de ces hyperparamètres est un problème encore ouvert. Il n'existe pas à ce jour de consensus sur une méthode formelle permettant de les déterminer en même temps que $\hat{\boldsymbol{a}}$.

Dans notre cas, les hyperparamètres sont réglés par calibration comme expliqué dans la **Section 4.1.3.3**.

Nous signalons également, et ceci sera étayé dans les perspectives de ce travail de thèse, que des stratégies d'inférence bayésienne ont été formulées pour ce genre de problèmes. Elles fournissent un cadre plus approprié à l'estimation conjointe des hyperparamètres.

3.4.4 Le critère d'arrêt

A chaque itération, les résidus primaux pour chaque contrainte d'égalité dans (3.12) s'écrivent comme suit.

$$\begin{aligned}
\boldsymbol{r}_{\boldsymbol{P}}^{(k)} &= \boldsymbol{P}\boldsymbol{a}^{(k)} - \boldsymbol{v}_P^{(k)} \\
\boldsymbol{r}_x^{(k)} &= \boldsymbol{D}_x\boldsymbol{a}^{(k)} - \boldsymbol{v}_x^{(k)} \\
\boldsymbol{r}_y^{(k)} &= \boldsymbol{D}_y\boldsymbol{a}^{(k)} - \boldsymbol{v}_y^{(k)} \\
\boldsymbol{r}_{\mathcal{C}}^{(k)} &= \boldsymbol{a}^{(k)} - \boldsymbol{v}_{\mathcal{C}}^{(k)}
\end{aligned} \tag{3.30}$$

De la même manière, à chaque itération, les résidus duaux s'écrivent :

$$\begin{aligned}
\boldsymbol{s}_{\boldsymbol{P}}^{(k)} &= \rho_P \boldsymbol{P}^\dagger(\boldsymbol{v}_P^{(k)} - \boldsymbol{v}_P^{(k-1)}) \\
\boldsymbol{s}_x^{(k)} &= \rho_D \boldsymbol{D}_x^\dagger(\boldsymbol{v}_x^{(k)} - \boldsymbol{v}_x^{(k-1)}) \\
\boldsymbol{s}_y^{(k)} &= \rho_D \boldsymbol{D}_y^\dagger(\boldsymbol{v}_y^{(k)} - \boldsymbol{v}_y^{(k-1)}) \\
\boldsymbol{s}_{\mathcal{C}}^{(k)} &= \rho_{\mathcal{C}}(\boldsymbol{v}_{\mathcal{C}}^{(k)} - \boldsymbol{v}_{\mathcal{C}}^{(k-1)})
\end{aligned} \tag{3.31}$$

Comme expliqué **Section 2.2.4.4** (p. 83), chacun des résidus primaux et duaux de (3.30) et (3.31) tend vers 0 quand k tend vers ∞.

En pratique, et comme l'indique [Bristow et Lucey, 2014], si la dualité est forte – ce qui est le cas ici car le critère et l'ensemble des contraintes $\mathcal{C}$ sont convexes – l'étude de la convergence des résidus primaux est suffisante. Ceci est particulièrement vrai si les paramètres de pénalités sont tels que la norme de chaque résidu primal et la norme du résidu dual associé sont proches. Nous posons alors le résidu primal global suivant :

$$\boldsymbol{r}^{(k)} = \boldsymbol{r}_{\boldsymbol{P}}^{(k)} + \boldsymbol{r}_x^{(k)} + \boldsymbol{r}_y^{(k)} + \boldsymbol{r}_{\mathcal{C}}^{(k)} \tag{3.32}$$

et nous avons $\boldsymbol{r}^{(k)} \to \boldsymbol{0}$ quand $k \to \infty$.

Comme le préconise [Boyd, 2010], nous choisissons la condition d'arrêt suivante pour l'**Algorithme 9** :

$$\left\|\boldsymbol{r}^{(k)}\right\|_2 \leqslant \varepsilon \tag{3.33}$$

où $\varepsilon > 0$ est une tolérance choisie suffisamment petite.

Un autre choix envisageable, pourrait être de contrôler la norme euclidienne de chaque résidu indépendamment.

Enfin, nous signalons qu'en pratique, nous fixons également un nombre maximum d'itération (typiquement une centaine d'itérations) en plus de de la condition d'arrêt (3.32).

Conclusion du Chapitre 3

Dans ce chapitre, nous avons présenté la contribution centrale de cette thèse. Nous avons premièrement explicité les *a priori* sur l'objet de rétro-diffusion. Nous avons ensuite développé d'une nouvelle méthode d'imagerie 3D HR de SER dénommée SPRITE. Elle s'appuie sur la construction d'un critère avec plusieurs pénalités ℓ_1 et contraint qui permet de tenir compte d'informations *a priori* sur la solution recherchée. Le critère est ensuite minimisé à l'aide d'une adaptation de l'ADMM au problème et qui permet de mener efficacement les calculs à l'aide de TFR. Enfin, des compléments ont été apportés sur la méthode SPRITE quant au choix des paramètres ou au critère d'arrêt de l'algorithme de minimisation.

Dans le chapitre suivant, nous évaluons la méthode SPRITE sur des données synthétiques et réelles et nous montrons son apport en terme d'analyse et de contrôle de SER.

CHAPITRE 4

L'ÉVALUATION DE LA MÉTHODE SPRITE

Dans la première section, la méthode SPRITE est étudiée et validée pour des données issues d'un objet de rétro-diffusion synthétique. Les résultats obtenus sont comparés à ceux de la méthode conventionnelle PFA. L'influence des hyperparamètres et la façon de les calibrer sont ensuite analysés. Dans la seconde section, nous montrons que cette méthode peut être employée pour analyser plus précisément la SER des cibles. Nous illustrons cela pour des résultats issus de données simulées par un code de calcul EM 3D et de données réelles mesurées avec l'Arche 3D. Enfin, dans la dernière section, nous montrons que la méthode SPRITE est un outil efficace pour contrôler la SER d'une cible présentant des défauts.

Sommaire

4.1 Une étude préliminaire de la méthode SPRITE 112
4.1.1 Une validation élémentaire 112
4.1.2 Robustesse au bruit 120
4.1.3 Influence des hyperparamètres et calibration 121
4.2 L'analyse de SER 125
4.2.1 Imagerie d'un drone 125
4.2.2 Imagerie d'un cône avec pastilles 129
4.3 Le contrôle 3D de SER 136

4.1 Une étude préliminaire de la méthode SPRITE

4.1.1 Une validation élémentaire

Dans un premier temps, nous illustrons l'apport de la méthode proposée dans un contexte d'imagerie radar 1D. Pour cela, nous présentons des résultats obtenus avec la méthode SPRITE 1D sur des données simulées pour un objet de rétro-diffusion synthétique 1D. Dans un second temps, la méthode SPRITE est validée dans un contexte d'imagerie radar 3D.

Objet synthétique 1D

Nous considérons trois contributeurs ponctuels, dont l'amplitude, la phase propre et la position sont détaillées dans le Tableau 4.1 suivant :

Contributeurs	①	②	③
Position longitudinale (m)	0	$0,20$	$0,21$
Amplitude propre (dBm^2)	-30	-20	-20
Phase propre (rad)	$\pi/4$	$\pi/5$	$\pi/5$

TABLEAU 4.1 – Description des contributeurs ponctuels

Les contributeurs ② et ③ sont volontairement choisis proches du point de vue de la distance longitudinale et de l'amplitude complexe (amplitude et phase) afin de comparer le pouvoir de résolution de la méthode SPRITE 1D et de la méthode conventionnelle.

Données simulées

Les données sont simulées avec le modèle points brillants mono-dimensionnel (1.15) (p. 38), pour une bande fréquentielle de 1 GHz à 6 GHz, par pas de 5 MHz. Ceci représente 1000 données. Un bruit blanc gaussien centré additif est ajouté aux données simulées. Le RSB (Rapport Signal sur Bruit) vrai en entrée est de 11,7 dB.

Paramètres de la méthode conventionnelle et de SPRITE 1D

Les caractéristiques de la grille spatiale sont indiquées dans le Tableau 4.2 et l'origine du repère est au milieu de la carte.

N_z	Δ_z
16384	$0,4$ m

TABLEAU 4.2 – Grille spatiale pour l'imagerie 1D

Le Tableau 4.3 détaille les paramètres de la méthode SPRITE 1D. Le support spatial est choisi égal à l'ensemble du domaine imagé Δ_z (le problème est donc non contraint).

μ	ρ_{P}	ε
10	8192	10^{-5}

TABLEAU 4.3 – Paramètres de la méthode SPRITE 1D pour l'objet synthétique 1D

Résultats

La Figure 4.1 présente la SER estimée selon la distance longitudinale, par la méthode PFA et par la méthode SPRITE 1D que nous proposons. Notons qu'en 1D, la méthode conventionnelle consiste en un déphasage des données dans le k-space, suivi d'une TFRI 1D sur N_z points avec bourrage de 0 et d'un déphasage dans le domaine spatial.

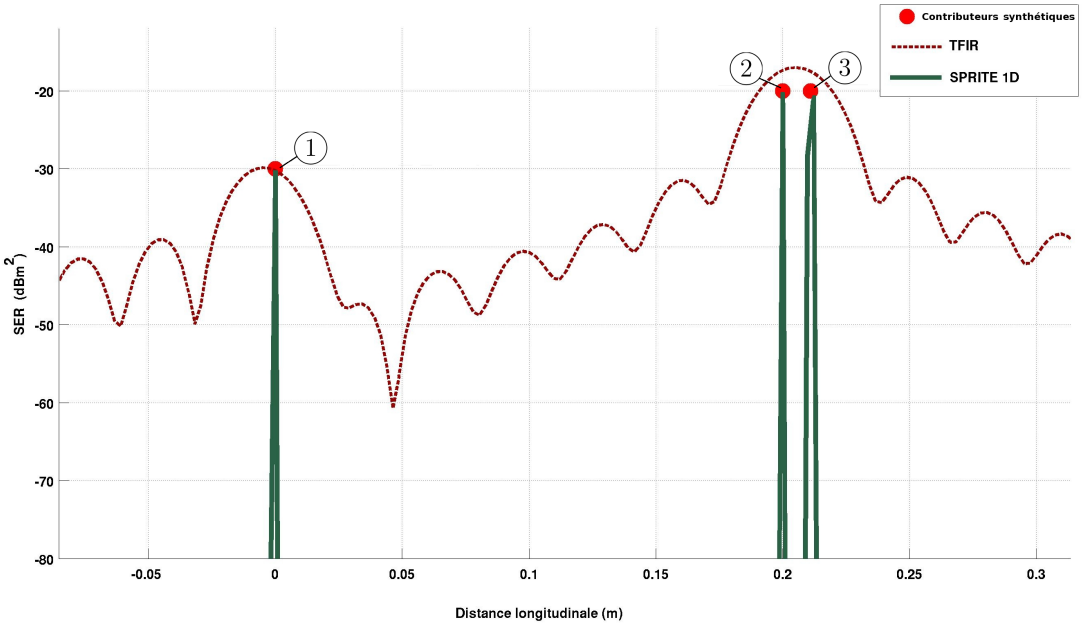


FIGURE 4.1 – Profil des SER estimées par TFRI et SPRITE 1D

Sur la Figure 4.1, nous constatons que la méthode SPRITE 1D (–) permet d'accroitre considérablement la résolution comparativement à la méthode conventionnelle (--). En effet, la largeur des lobes principaux et le niveau des lobes secondaires sont conjointement réduits. SPRITE 1D permet ainsi d'estimer beaucoup plus précisément la SER et la localisation des contributeurs synthétiques (1), (2) et (3) (●).

Par ailleurs, nous remarquons qu'à l'inverse de la méthode conventionnelle, SPRITE 1D permet de distinguer les contributeurs (2) et (3). Le niveau des contributeurs est également bien estimé pour les deux contributeurs avec SPRITE 1D. A l'inverse, avec la méthode conventionnelle, le lobe principal a un niveau surestimé car il résulte de la somme des lobes principaux et secondaires des deux contributeurs qui se superposent. Ces résultats illustrent donc l'intérêt de l'usage de méthodes parcimonieuses en imagerie radar et dans le cas d'espèce l'apport de la méthode SPRITE 1D dans un contexte mono-dimensionnel. Nous constatons cependant un doublement des pics quand le contributeur n'est pas localisé sur la grille comme c'est le cas pour le contributeur (3) sur la Figure 4.1.

Le temps de calcul pour produire la carte de SER estimée avec la méthode SPRITE 1D présentée sur la Figure 4.1 est de 5,1 secondes [1]. L'essentiel de la charge calculatoire repose sur le calcul des TFRI.

1. MATLAB © 2013b, Processeur Intel® Xeon® E5-2620 v3 2.4 GHz et 32 Go de mémoire RAM

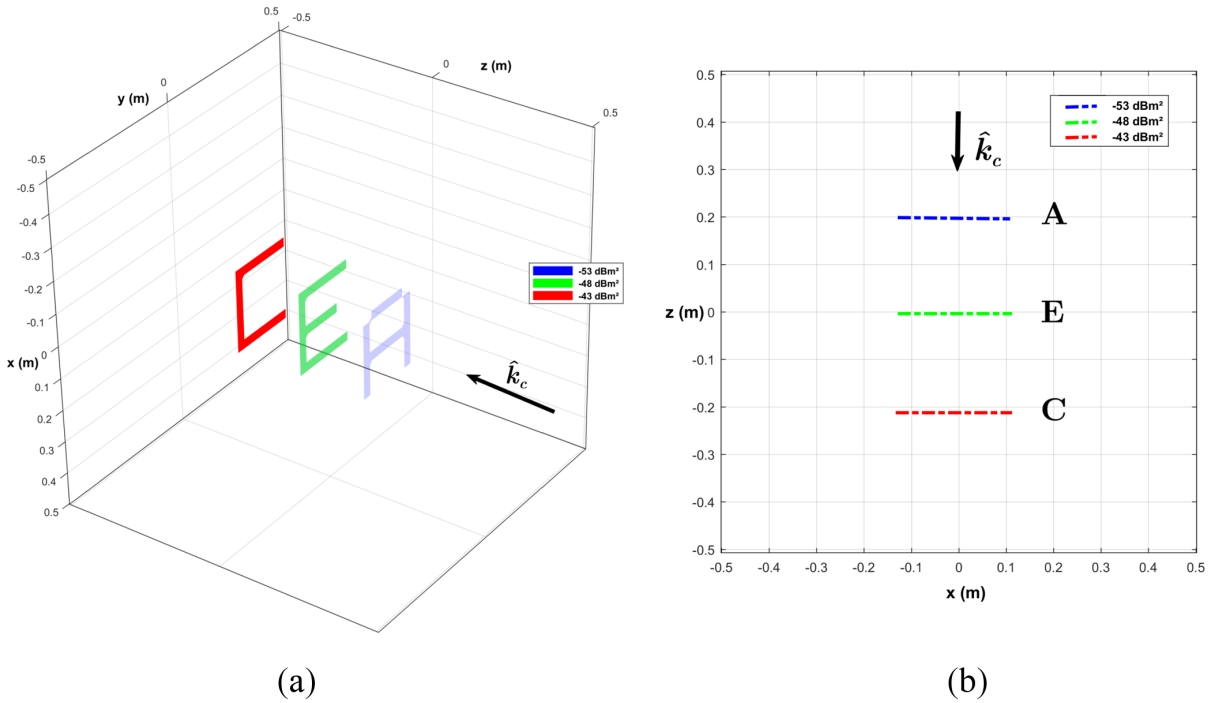
Carte synthétique 3D

Nous validons maintenant la méthode SPRITE dans un contexte d'imagerie radar 3D. Pour cela, nous considérons une carte synthétique 3D composée de trois facettes spéculaires dont la normale est colinéaire à la direction du vecteur d'onde central $\hat{\boldsymbol{k}}_c$. Ces dernières sont sur les plans $n_z = 30$, $n_z = 50$ et $n_z = 70$ localisés à trois positions longitudinales (selon $\hat{\boldsymbol{z}}$ colinéaire à $\hat{\boldsymbol{k}}_c$) différentes et forment les lettres « **C** », « **E** » et « **A** ». Un coefficient de rétro-diffusion complexe, constant sur toute la surface, est associé à chaque facette spéculaire. Chacune a donc une amplitude et une phase propre comme indiqué dans le Tableau 4.4.

Facette	**C**	**E**	**A**
Indice du plan n_z	30	50	70
Position longitudinale (m)	-0,21	0	0,20
Amplitude propre (dB)	-40	-45	-50
Phase propre (rad)	$\pi/4$	$\pi/2$	$\pi/3$

TABLEAU 4.4 – Description de la carte synthétique « **CEA** »

Une représentation par iso-surfaces de $20\log_{10}|\boldsymbol{a}^v|$ (la SER exprimée de dBm^2 de la carte synthétique « **CEA** ») est donnée Figure 4.2 ainsi qu'une coupe longitudinale selon $(\hat{\boldsymbol{x}}\hat{\boldsymbol{z}})$. Nous insistons sur le fait que l'objet « **CEA** » est une carte de rétro-diffusion synthétique qui contient seulement trois facettes spéculaires. Ainsi, il ne s'agit pas d'une cible physique pour laquelle des rebonds multiples entre les facettes apparaitraient.



(a) (b)

FIGURE 4.2 – $20\log_{10}|\boldsymbol{a}^v|$ (a) Iso-surfaces. (b) Coupe longitudinale.

Données simulées

Les données sont simulées avec le modèle d'observation (1.18) (p. 40) pour la bande fréquentielle et les bandes angulaires d'observation indiquées dans le Tableau 4.5. Ceci représente environ $7,7 \cdot 10^5$ données.

f	θ	φ
8 GHz : 5 MHz : 12 GHz	-15 ° : 1° : 15 °	-15 ° : 1° : 15 °

TABLEAU 4.5 – Bandes fréquentielle et angulaires pour la simulation des données de la carte synthétique « **CEA** »

Un bruit blanc gaussien centré additif est ajouté aux données simulées. Le RSB vrai en entrée est de -7.7 dB. Ce RSB est choisi volontairement inférieur à celui constaté lors de mesures en chambre anéchoïque (≈ 20 dB) afin d'éprouver la robustesse de la méthode au bruit.

Paramètres des méthodes PFA et SPRITE

Les caractéristiques de la grille spatiale sont indiquées dans le Tableau 4.6 et l'origine du repère est au milieu de la carte. Ceci représente 10^6 inconnues.

$N_x \times N_y \times N_z$	Δ_x	Δ_y	Δ_z
$100 \times 100 \times 100$	1 m	1 m	1,1 m

TABLEAU 4.6 – Grille spatiale pour l'estimation de la carte synthétique « **CEA** »

Le Tableau 4.7 détaille les paramètres de la méthode SPRITE. Le support spatial est choisi égal à l'ensemble du domaine imagé (le problème est non contraint).

μ	λ	β	ρ_{P}	ρ_{D}	$\rho_{\mathcal{C}}$	ε
10	100	150	50	$5 \cdot 10^5$	$5 \cdot 10^5$	10^{-4}

TABLEAU 4.7 – Paramètres de la méthode SPRITE pour la carte synthétique « **CEA** »

Résultats

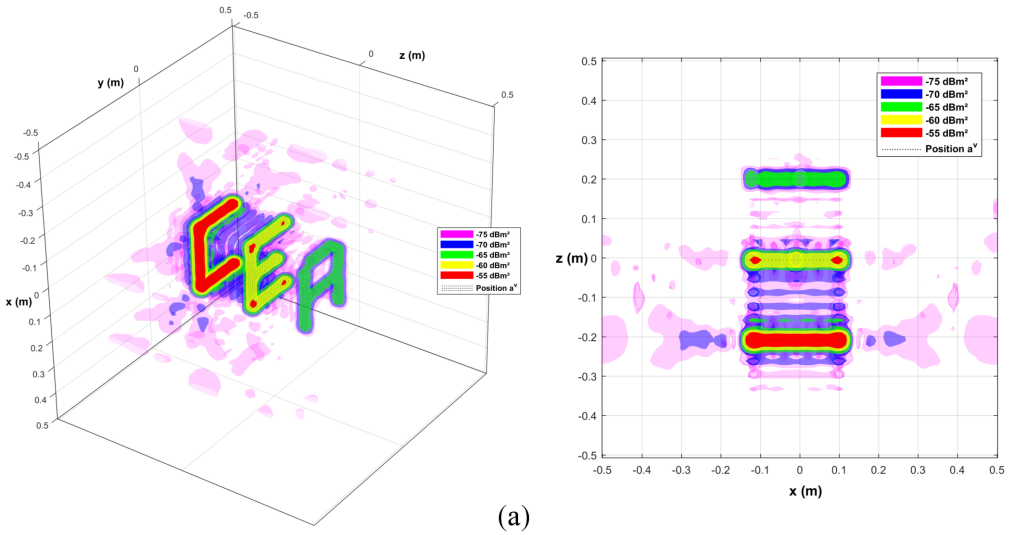
La Figure 4.3 présente les cartes construites avec :

(a) la méthode PFA.

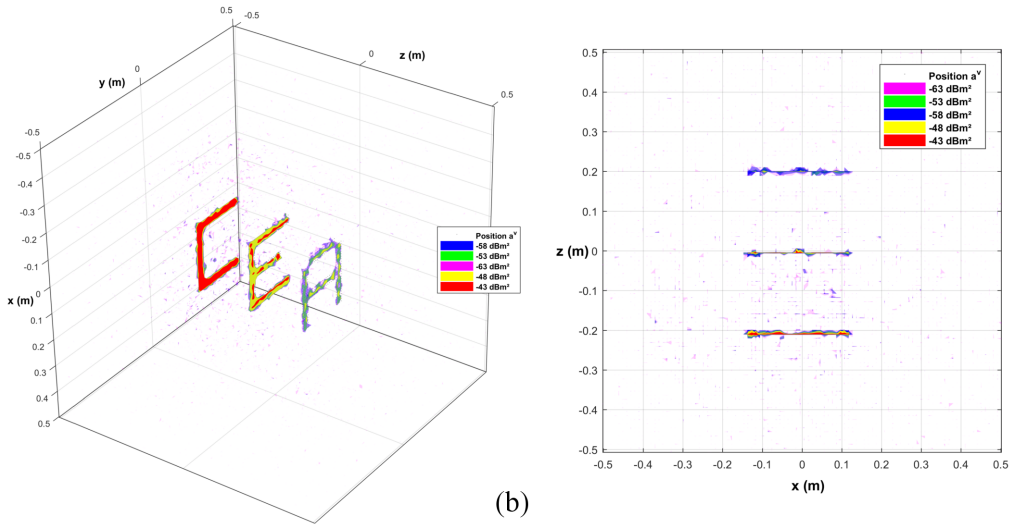
(b) La méthode parcimonieuse ℓ_1 proposée par [Austin *et al.*, 2011]. Le critère consiste en un terme d'attache aux données $\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2$, une pénalité $\|\boldsymbol{a}\|_1$ (*cf.* p.58) et un hyperparamètre μ contrôlant le compromis. Ici $\mu = 0,01$ et l'optimisation est assurée avec l'ADMM.

(c) La méthode SPRITE.

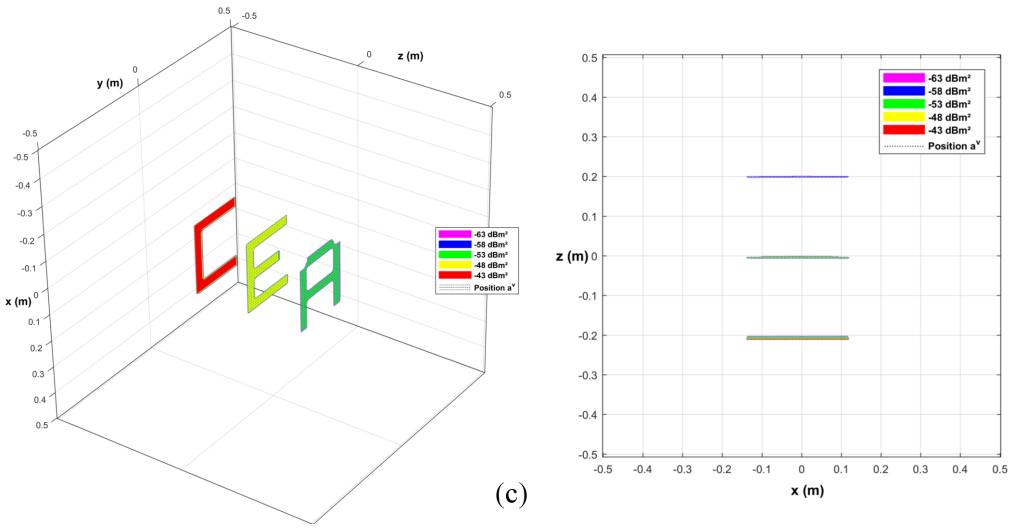
Sur la figure Figure 4.4, nous comparons la SER (exprimée en dBm^2) des coupes transverses de $\boldsymbol{a}^v$, $\hat{\boldsymbol{a}}_{\mathrm{PFA}}$ et $\hat{\boldsymbol{a}}_{\mathrm{SPRITE}}$, pour les plans n_z voisins de $n_z = 30$ (contenant la facette « **C** »), $n_z = 50$ (« **E** ») et $n_z = 70$ (« **A** »).



(a)



(b)



(c)

FIGURE 4.3 – (a) PFA. (b) Méthode parcimonieuse ℓ_1. (c) SPRITE.

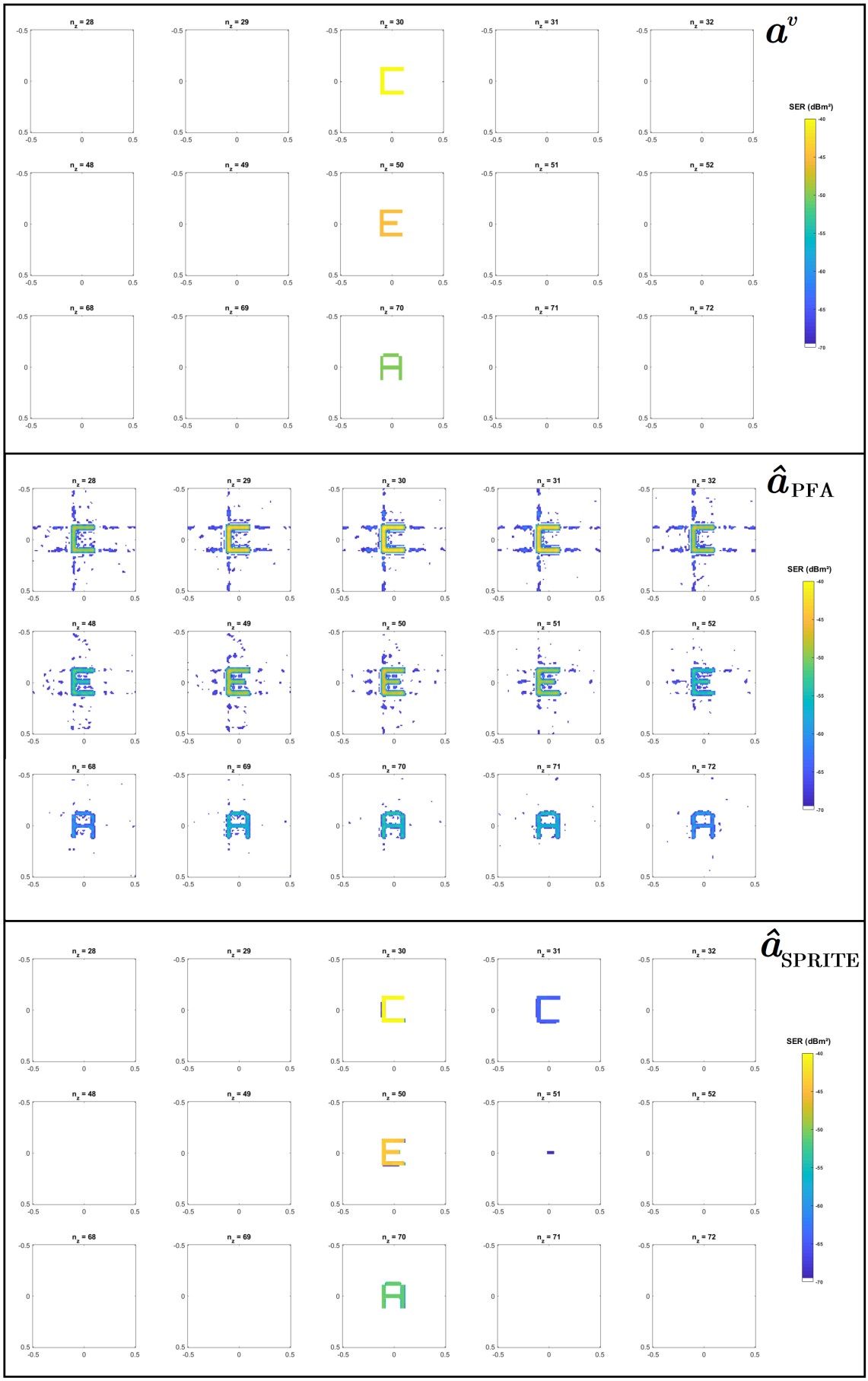


FIGURE 4.4 – Coupes transverses de a^v, $\hat{a}_{\text{PFA}}$ et $\hat{a}_{\text{SPRITE}}$ pour les plans $n_z = \{[\![28, 32]\!], [\![48, 52]\!], [\![68, 72]\!]\}$

La résolution est considérablement accrue avec la méthode SPRITE comparativement à la méthode conventionnelle PFA, qui présente de nombreux lobes secondaires de niveau élevé et de larges lobes principaux. La méthode parcimonieuse fait quant à elle apparaitre de nombreux points brillants en lieu et place des facettes brillantes et l'estimation des niveaux est médiocre. Ainsi, les facettes spéculaires « **C** », « **E** » et « **A** » sont plus précisément localisées et leur niveau est mieux estimé avec SPRITE.

La différence de niveau observé sur la carte PFA (près de 10 dBm^2 en dessous des valeurs de $\boldsymbol{a}^v$) s'explique par le fait que la distribution spatiale des contributeurs est plus diffuse que celle de la méthode SPRITE.

Le temps de calcul pour produire la carte de SER estimée avec la méthode SPRITE est de 3 minutes et 49 secondes.

Analyse fréquentielle

Dans cette section, sur la Figure 4.5, nous comparons le contenu fréquentiel de la carte synthétique à celui de la carte estimée avec la méthode PFA, la méthode parcimonieuse ℓ_1 proposée par [Austin *et al.*, 2011] et enfin à celui de la carte estimée avec la méthode SPRITE. L'axe des abscisses correspond à $\sqrt{k_x^2 + k_y^2 + k_z^2}$ c'est-à-dire la distance d'un point du k-space à l'origine et l'axe des ordonnées est la SER associée à ce point.

Comparativement à la méthode PFA, nous constatons que le contenu fréquentiel de la carte estimée avec la méthode SPRITE est le plus fidèle à celui de la carte synthétique. En particulier, le k-space est mieux complété et les fréquences non observées sont mieux restituées, notamment les hautes et basses fréquences spatiales (en dehors de la portion de k-space observée).

La méthode parcimonieuse ℓ_1 proposée par [Austin *et al.*, 2011] complète également le k-space. Cependant, la SER est sur-estimée dans les hautes et basses fréquences spatiales et elle ne correspond pas à celle de la carte synthétique. Ceci s'explique par le fait que la carte estimée est constituée d'un ensemble de points brillants, or un signal correspondant à une distribution de Dirac a un contenu fréquentiel blanc, *i.e.* chaque fréquence spatiale est présente avec une intensité identique.

Enfin, notons que dans la portion de k-space observée, le contenu fréquentiel de la carte estimée avec la méthode PFA est le plus proche de celui de la carte synthétique. Ceci s'explique par le fait que le PFA correspond à une régularisation au sens des moindres carrés à norme minimale (comme nous l'avons montré dans la **Section 1.3.2.1**, p. 54). Des trois méthodes c'est donc celle qui minimise le plus l'écart entre la sortie du modèle et les données observées.

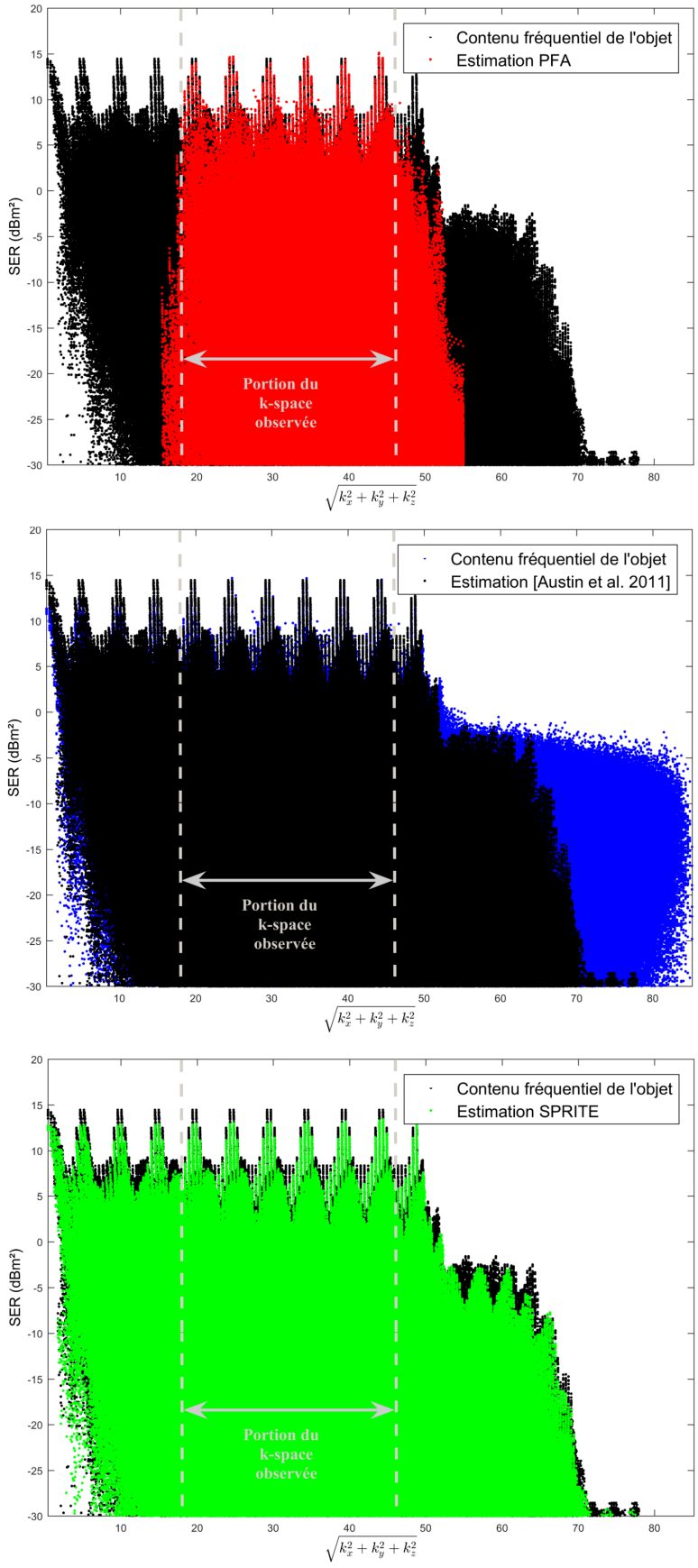


FIGURE 4.5 – Analyse fréquentielle des cartes estimées avec les méthodes (a) PFA, (b) parcimonieuse ℓ_1, (c) SPRITE

Fonctionnement de SPRITE au cours des itérations

La Figure 4.6 présente l'évolution du critère pénalisé $\mathcal{J}(\boldsymbol{a}^{(k)})$ (3.10) et du lagrangien augmenté $\mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}(\boldsymbol{a}^{(k)}, \boldsymbol{v}_P^{(k)}, \boldsymbol{v}_{Dx}^{(k)}, \boldsymbol{v}_{Dy}^{(k)}, \boldsymbol{v}_\mathcal{C}^{(k)}, \boldsymbol{u}_P^{(k)}, \boldsymbol{u}_{Dx}^{(k)}, \boldsymbol{u}_{Dy}^{(k)}, \boldsymbol{u}_\mathcal{C}^{(k)})$ (3.13) au cours des itérations k.

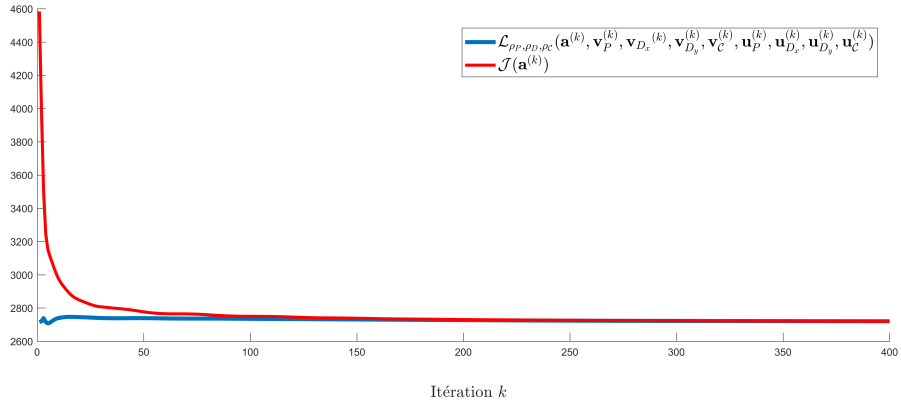


FIGURE 4.6 – Évolution de $\mathcal{J}$ et $\mathcal{L}$ au cours des itérations

Premièrement, nous constatons que l'algorithme converge en une centaine d'itérations. Deuxièmement, et conformément à l'étude théorique sur le fonctionnement de l'ADMM menée **Section 2.2.4.4** (p. 83), nous remarquons sur la Figure 4.6 que l'évolution de $\mathcal{J}$ n'est pas monotone. Ceci s'explique par le fait que l'ADMM n'est pas un algorithme de descente du critère. Cependant, le critère converge bien asymptotiquement vers son minimum. De la même manière, l'évolution du lagrangien augmenté n'est pas monotone. En effet, l'ADMM n'est pas un algorithme de descente du lagrangien. Cependant le lagrangien augmenté converge bien vers la valeur qu'il atteint à son point selle et qui est la valeur du primal et du dual optimal car la dualité est forte. Les résidus primaux et duaux tendent bien tous deux asymptotiquement vers 0.

4.1.2 Robustesse au bruit

Pour apprécier et comparer la robustesse des méthodes au bruit nous évaluons la qualité des cartes reconstruites à partir de données entachées d'un bruit blanc gaussien centré additif. Il s'agit d'une hypothèse communément faite en radar [Borden, 1992; Mensa, 1981].

Dans le cas d'une carte de rétro-diffusion synthétique telle que la carte « **CEA** » présentée avant, $\boldsymbol{a}^v$ est connue. Pour quantifier l'erreur d'estimation des méthodes comparées, nous choisissons la REQM (Racine carrée de l'Erreur Quadratique Moyenne) qui est définie par :

$$\text{REQM} = \sqrt{\frac{\|\hat{\boldsymbol{a}}_\mathcal{M} - \boldsymbol{a}^v\|_2^2}{N}} \tag{4.1}$$

où $\hat{\boldsymbol{a}}_\mathcal{M}$ est la carte estimée selon la méthode $\mathcal{M} \in \{\text{PFA, SPRITE}\}$.

La Figure 4.8 présente l'évolution de la REQM des cartes $\hat{\boldsymbol{a}}_{\text{PFA}}$ et $\hat{\boldsymbol{a}}_{\text{SPRITE}}$ avec les paramètres donnés par le Tableau 4.7, pour différentes variances de bruit et donc différents RSB.

Sur la Figure 4.8, nous constatons que la REQM des cartes estimées avec la méthode PFA est plate de l'ordre de $1,9 \cdot 10^{-4}$ jusqu'à un RSB de 0 dB. Elle augmente ensuite rapidement pour des RSB inférieurs. La REQM des cartes estimées avec la méthode SPRITE est quant à elle plus petite, de l'ordre de $1,3 \cdot 10^{-4}$. Elle reste plate jusqu'à un RSB plus faible de l'ordre -10 dB. Nous pouvons en conclure que la méthode SPRITE est plus robuste au bruit que le PFA : elle permet d'estimer des cartes à partir de données bruitées (sous l'hypothèse d'un bruit blanc gaussien centré additif) tout en conservant une qualité d'estimation raisonnable.

4.1.3 Influence des hyperparamètres et calibration

4.1.3.1 Influence de l'hyperparamètre μ

Nous rappelons que l'hyperparamètre μ permet de contrôler le compromis entre le terme d'attache aux données $\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2$ et la pénalité $\|\boldsymbol{P}\boldsymbol{a}\|_1$ qui favorise une projection de la carte sur l'axe $\hat{\boldsymbol{z}}$ essentiellement nulle sauf pour un nombre limité de plans n_z. Sur la Figure 4.7 nous affichons :

- en ordonnée, $20\log_{10}|\boldsymbol{P}\boldsymbol{a}|$, le logarithme du module de la projection sur l'axe $\hat{\boldsymbol{z}}$ des cartes : $\boldsymbol{a}^v$ (★) et $\hat{\boldsymbol{a}}_{\text{SPRITE}}$ estimées par la méthode SPRITE avec $\mu = 0$ (–), $\mu = 10$ (–) et $\mu = 100$ (–), les autres paramètres étant fixés comme indiqué dans le Tableau 4.7,
- en abscisse, le numéro des plans $n_z = 0, \cdots, N_z - 1$ de discrétisation de l'axe $\hat{\boldsymbol{z}}$.

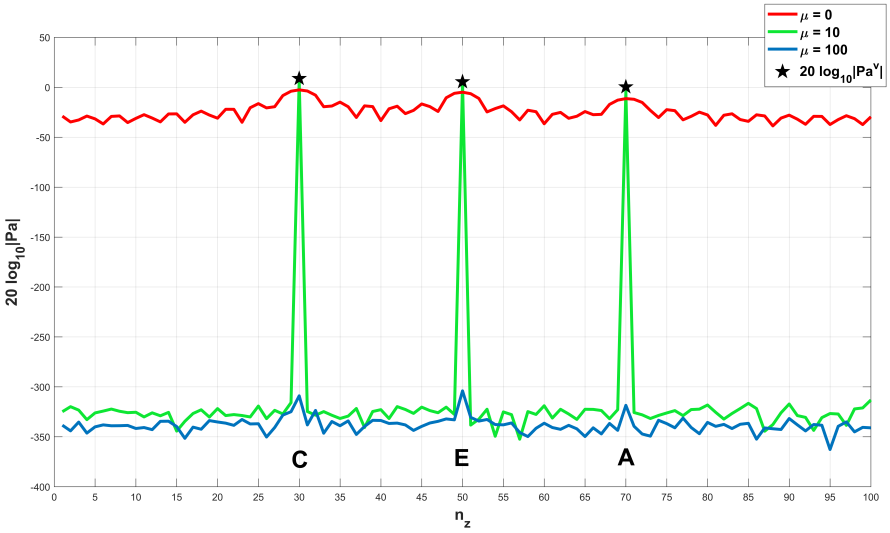


FIGURE 4.7 – Influence du paramètre μ (pour $\lambda = 100, \beta = 150$ fixés)

Comme attendu, nous constatons que plus μ est grand, plus la projection de $\hat{\boldsymbol{a}}_{\text{SPRITE}}$ sur l'axe $\hat{\boldsymbol{z}}$ est parcimonieuse. Plus explicitement, nous notons que pour :

- $\mu = 0$, la projection présente de larges lobes principaux et de nombreux lobes secondaires avec des niveaux élevés,
- $\mu = 100$, $20\log_{10}|\boldsymbol{P}\hat{\boldsymbol{a}}| \leqslant -300$ dBm^2 pour chaque plan n_z. Ceci signifie que la projection est quasiment nulle et donc qu'elle est trop parcimonieuse.
- $\mu = 10$, qui est la valeur adéquate ici, la projection de $\hat{\boldsymbol{a}}_{\text{SPRITE}}$ est très proche de celle de la carte vraie $\boldsymbol{a}^v$.

Nous en concluons que μ permet bien de contrôler la parcimonie de la projection de la carte sur l'axe $\hat{\boldsymbol{z}}$.

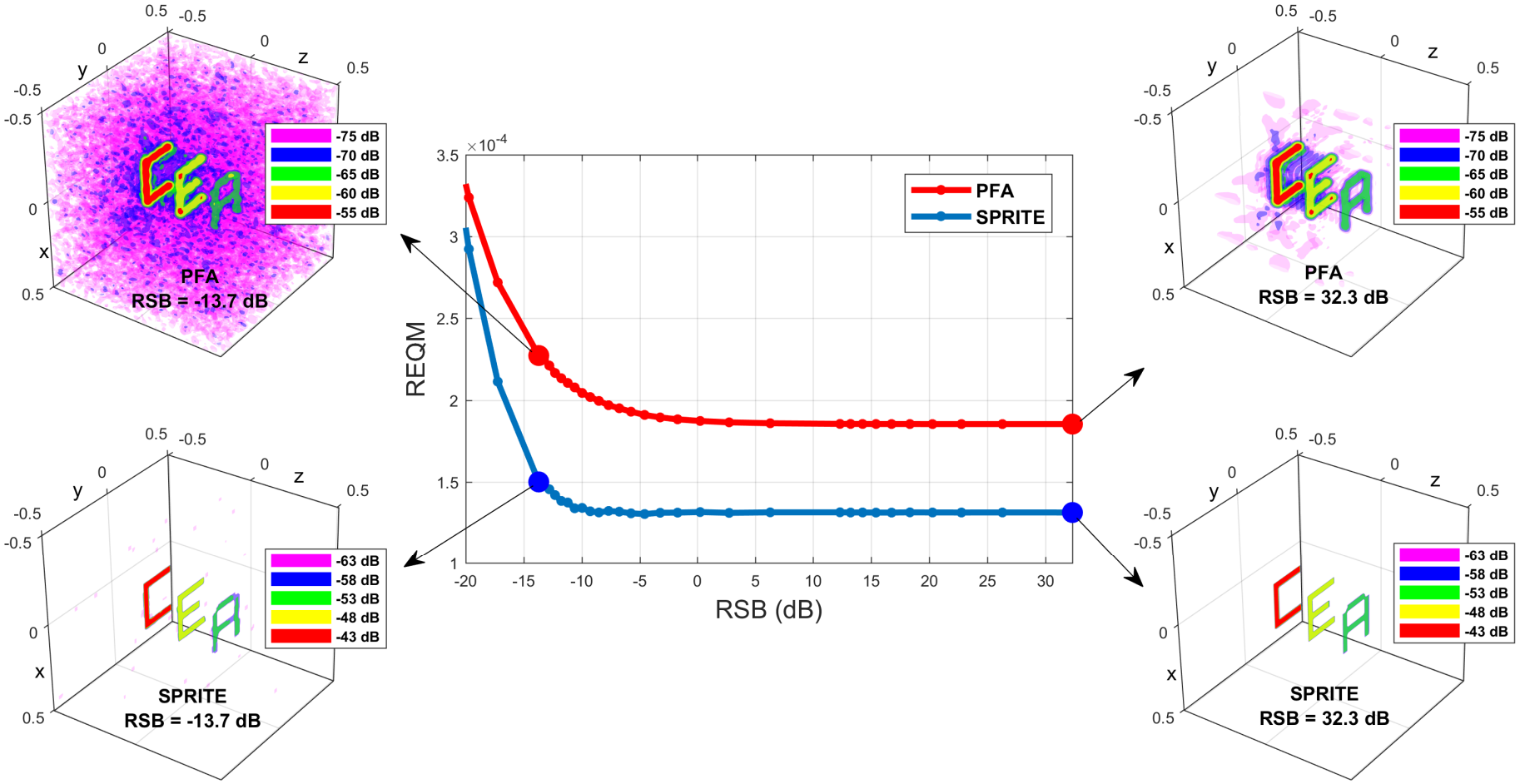


FIGURE 4.8 – Évolution de la REQM selon le RSB

4.1.3.2 Influence de l'hyperparamètre λ

Nous rappelons que l'hyperparamètre λ permet de contrôler le compromis entre le terme d'attache aux données $\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2$ et les pénalités $\|\boldsymbol{D}_x\boldsymbol{a}\|_1$ et $\|\boldsymbol{D}_y\boldsymbol{a}\|_1$ qui favorisent une carte constante par morceaux, chaque morceau correspondant à une facette spéculaire. Sur le Tableau 4.8 nous affichons la SER (exprimée en dBm^2) du plan $n_z = 30$ (qui contient la lettre « **C** ») des cartes $\boldsymbol{a}^v$ et $\hat{\boldsymbol{a}}_{\text{SPRITE}}$ estimées avec $\lambda = \{0, 10, 100\}$. Les autres paramètres de la méthode SPRITE sont fixés comme indiqué dans le Tableau 4.7.

Comme escompté, nous remarquons que plus λ augmente, plus les facettes spéculaires connexes sur lesquelles le coefficient de rétro-diffusion est constant apparaissent. Plus spécifiquement, nous notons que pour :

- $\lambda = 0$ la SER estimée présente de nombreuses zones hétérogènes avec des niveaux sensiblement différents de ceux de $\boldsymbol{a}^v$.
- $\lambda = 10$, le nombre de facettes spéculaires connexes est réduit comparativement à l'estimation pour $\lambda = 0$. Il est cependant toujours trop important comme nous le constatons dans les zones normalement nulles en dehors de la facette « **C** ».
- $\lambda = 100$, qui est la valeur adéquate ici, le niveau de SER estimé pour la facette spéculaire « **C** » est très proche de celui de la carte « vraie ».

Nous en concluons que λ permet bien de contrôler, sur chaque plan n_z, le nombre de facettes spéculaires connexes où le coefficient de rétro-diffusion est constant. En pratique nous constatons que les cartes évoluent de manière significative quand les paramètres λ et/ou μ varient d'un facteur 10.

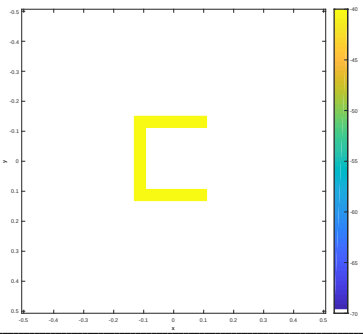
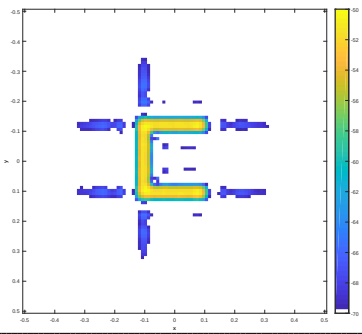
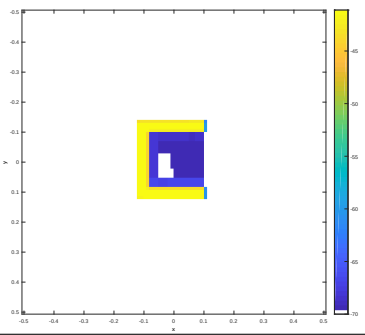
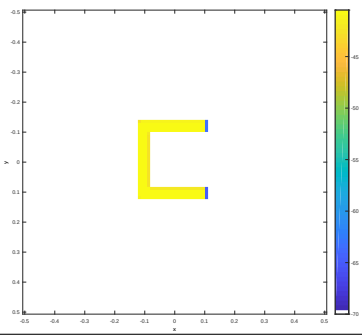
$\boldsymbol{a}^v$	$\hat{\boldsymbol{a}}_{\text{SPRITE}}$, $\lambda = 0$
	
	
$\hat{\boldsymbol{a}}_{\text{SPRITE}}$, $\lambda = 10$	$\hat{\boldsymbol{a}}_{\text{SPRITE}}$, $\lambda = 100$

TABLEAU 4.8 – Influence du paramètre λ (pour $\mu = 10, \beta = 150$ fixés)

4.1.3.3 Calibration des hyperparamètres

Comme expliqué **Section 3.4.3** (p. 108) et comme souligné juste avant, le choix des hyperparamètres est crucial car il détermine la carte $\hat{\boldsymbol{a}}_{\text{SPRITE}}$ estimée. Cependant, à ce jour et à notre connaissance, il n'existe pas de consensus sur une méthode permettant d'estimer conjointement les hyperparamètres et la carte de rétro-diffusion pour des problèmes similaires à celui que nous étudions.

Pour faciliter le choix des hyperparamètres et éviter de procéder par tâtonnement, nous proposons une méthode de calibration. Dans un premier temps, nous simulons des données pour un objet synthétique caractéristique d'une famille d'objets réels (notamment en terme de niveau de SER). Ces données sont ensuite bruitées de façon à ce que le RSB soit représentatif des mesures réelles. Les cartes $\hat{\boldsymbol{a}}_{\text{SPRITE}}$ sont alors calculées pour différentes combinaisons de valeurs d'hyperparamètres qui varient sur une grille. La combinaison retenue est alors celle qui minimise la REQM, c'est à dire celle pour laquelle la carte estimée est la plus proche de l'objet synthétique connu.

Par souci de simplicité, nous illustrons cette démarche dans un contexte 1D sur la Figure 4.9. Dans cet exemple, la carte estimée la plus proche de l'objet synthétique (★) est celle correspondant à $\mu = 3,2 \cdot 10^{-1}$ (–). Ce sera donc la valeur retenue quand il s'agira de traiter des objets réels similaires à cet objet synthétique.

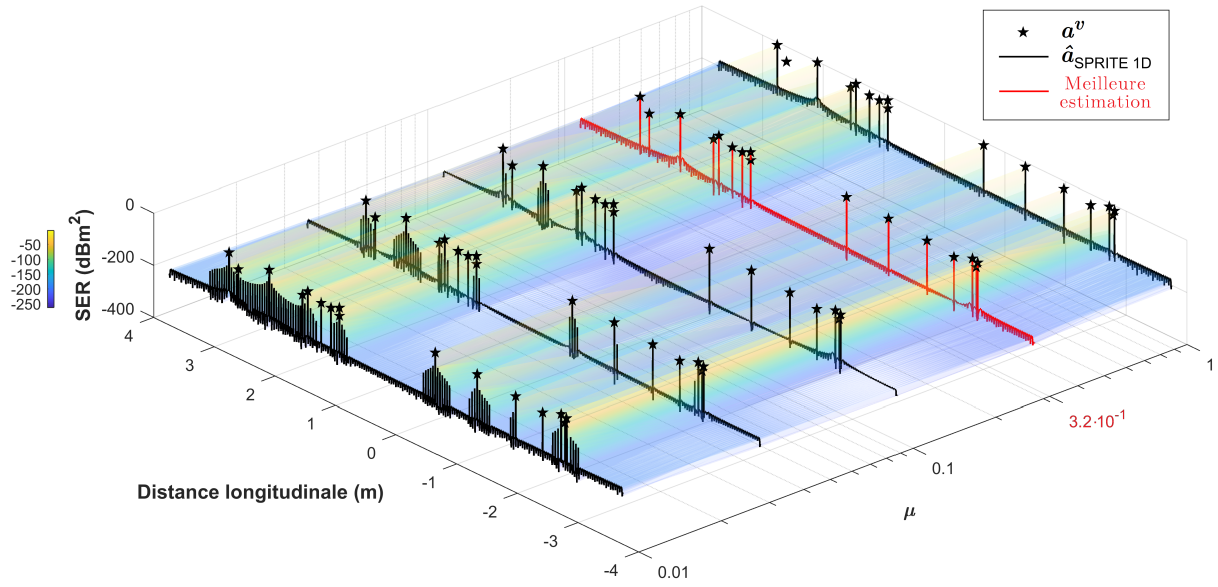


FIGURE 4.9 – Exemple de calibration de l'hyperparamètre en 1D

Dans un second temps, cette démarche est répétée pour d'autres familles d'objets dans le but de construire un « abaque » applicable à de nombreux cas réels.

Nous évoquons également dans les perspectives des méthodes récentes qui reposent sur des stratégies d'inférence bayésienne et qui fournissent un cadre plus approprié à l'estimation conjointe des hyperparamètres.

4.2 L'analyse de SER

Dans la section précédente, la méthode SPRITE a été validée dans un contexte d'imagerie radar mono et tri-dimensionnelle pour des cartes de rétro-diffusion synthétiques. Dans cette section, nous montrons l'apport de la méthode SPRITE 1D pour l'analyse de SER des cibles imagées.

REMARQUE

> Nous avons sélectionné les figures jugées pertinentes pour illustrer nos propos. D'autres figures, notamment des visualisations additionnelles pour les objets 3D, sont données dans l'**Annexe C**.

4.2.1 Imagerie d'un drone

Nous nous intéressons à l'estimation de la carte de rétro-diffusion et à l'analyse de SER 3D d'une maquette numérique d'un drone développé par l'ONERA à l'occasion des ateliers de travail électromagnétique 2014 et 2016 de l'ISAE [Simon et Soudais, 2014].

Le drone mesure 3,077 m de long et 2,873 m d'envergure. Il est décliné en deux versions. La version 2014 où le drone est entièrement CEP et la version 2016, représentée Figure 4.10 où le nez ainsi que les raccords avec la trappe de fuselage ont été recouverts de RAM de permittivité relative $\varepsilon_r = 1,5 - 0,1\jmath$ et de perméabilité relative $\mu_r = 2,5 - 1,8\jmath$.

La particularité de cette application est que les données sont simulées avec un code de calcul EM 3D développé au CEA-CESTA.

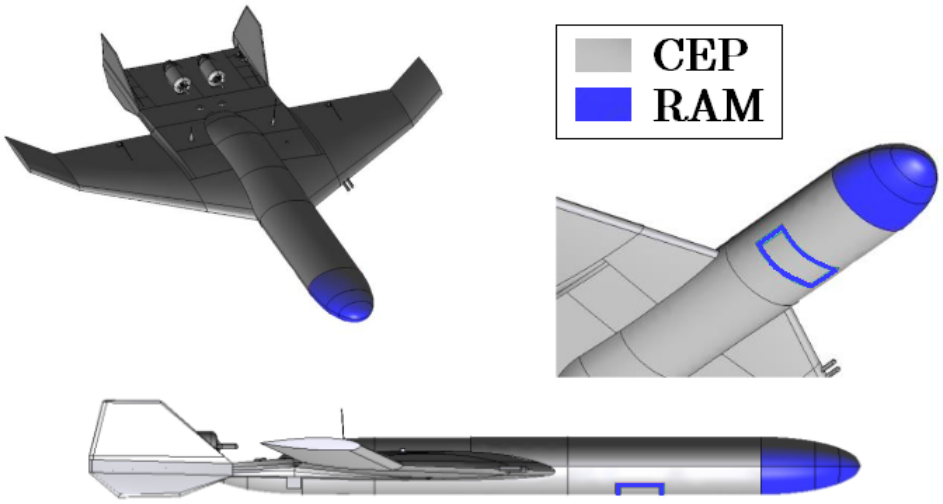


FIGURE 4.10 – Modèle CAO du drone version 2016

Données simulées

Les données sont calculées en bande S à partir de modèles CAO du drone, avec un code de calcul EM 3D développé au CEA [Augonnet *et al.*, 2015]. Ce dernier repose sur une méthode d'équations intégrales surfaciques dans le domaine fréquentiel. Il permet de traiter des objets constitués de parties diélectriques et conductrices. Il est par ailleurs massivement parallèle et exploite les accélérateurs de calcul tels que les cartes graphiques (GPU).

La bande fréquentielle et des bandes angulaires de calcul en polarisation HH sont indiquées dans le Tableau 4.9. Ceci représente environ $8,6 \cdot 10^4$ données.

f	θ	φ
2,5 GHz : 30 MHz : 4 GHz	-20 ° : 1° : 20 °	-20 ° : 1° : 20 °

TABLEAU 4.9 – Bandes fréquentielle et angulaires pour le drone

Un bruit blanc gaussien centré additif est ajouté aux données calculées. Le RSB vrai en entrée est de 15 dB, ce qui est de l'ordre de grandeur de données réelles mesurées.

Paramètres des méthodes PFA et SPRITE

Les caractéristiques de la grille spatiale sont indiquées dans le Tableau 4.10 et l'origine du repère est au milieu de la carte. Ceci représente environ $6,3 \cdot 10^7$ inconnues.

$N_x \times N_y \times N_z$	Δ_x	Δ_y	Δ_z
$350 \times 350 \times 512$	3 m	3 m	4,5 m

TABLEAU 4.10 – Grille spatiale pour le drone

Enfin, le Tableau 4.11 détaille les paramètres de la méthode SPRITE déterminés d'après le processus de calibration expliqué en **Section 4.1.3.3**.

μ	λ	β	ρ_{P}	ρ_{D}	$\rho_{\mathcal{C}}$	ε
10	120	100	256	$3,1 \cdot 10^7$	$3,1 \cdot 10^7$	10^{-4}

TABLEAU 4.11 – Paramètres de la méthode SPRITE pour le drone

Sans connaissances *a priori* sur l'extension spatiale des objets de rétro-diffusion, le support spatial est choisi égal à l'ensemble du domaine et le problème n'est donc pas contraint.

Résultats

La Figure 4.11 présente les cartes du drone version 2016 (exprimées par iso-surfaces en dBm^2), estimées par les méthodes PFA et SPRITE.

Nous remarquons que la carte obtenue avec la méthode SPRITE est plus résolue et nous permet de mieux localiser et de mettre en évidence certains mécanismes de rétro-diffusion qui n'apparaissent pas clairement avec la méthode conventionnelle. Nous pouvons ainsi distinguer les interactions suivantes (certaines sont visibles sur la Figure C.1 et la Figure C.2 (p. 154)).

- L'interaction avec les réacteurs est celle qui concourt le plus à la SER globale de la cible. En effet, l'étage de compression des réacteurs est une turbine constituée de nombreuses aubes. Quand le drone est illuminé dans l'axe (ce qui est le cas du vecteur d'onde central ici), ces dernières rétro-diffusent l'onde vers le radar à la manière d'une facette spéculaire [Knott *et al.*, 2004].
- La diffraction sur le bord de fuite des ailes et l'empennage.
- Le spéculaire sur le nez du drone.
- La diffraction sur les raccords de la trappe de fuselage.
- La diffraction sur les différentes antennes que nous détaillons ensuite.

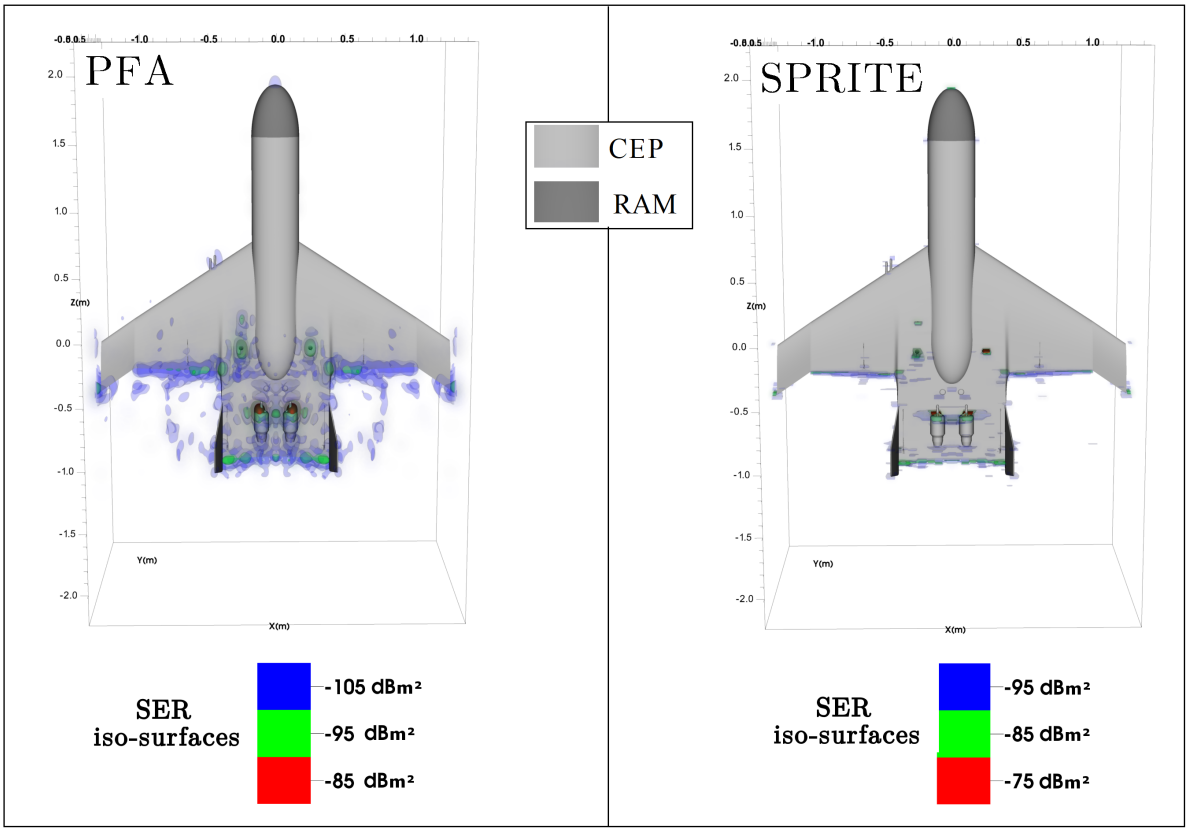


FIGURE 4.11 – Cartes de rétro-diffusion du drone version 2016.

Sur la Figure 4.12, nous nous focalisons sur les mécanismes de rétro-diffusion intervenant au niveau des antennes du drone version 2016 (sachant que ces mécanismes sont similaires pour la version 2014). Nous remarquons encore une fois que les interactions avec les antennes sont plus précisément localisées avec la méthode que nous proposons. Nous constatons également que la méthode SPRITE permet de séparer les deux sondes sur le bord d'attaque de l'aile alors qu'elles sont difficilement distinguables avec la méthode PFA.

La Figure 4.13 permet de comparer l'interaction avec la trappe de fuselage, pour les cartes des versions 2014 et 2016 produites avec la méthode SPRITE. Nous constatons l'influence du RAM appliqué sur le nez et les raccords de la trappe. En effet, le spéculaire sur le nez $\mathbf{N}_{14}$ et la diffraction sur les raccords de la trappe $\mathbf{T}_{14}$ du drone CEP version 2014 ont des

niveaux de 10 à 20 dBm^2 supérieurs à ceux du nez (N$_{16}$) et de la trappe (T$_{16}$) de la version 2016 avec RAM. Par ailleurs, contrairement à la méthode conventionnelle PFA, la méthode SPRITE permet de mettre en évidence les discontinuités de matériaux (D$_{16}$) sur le nez de la version 2016 du drone.

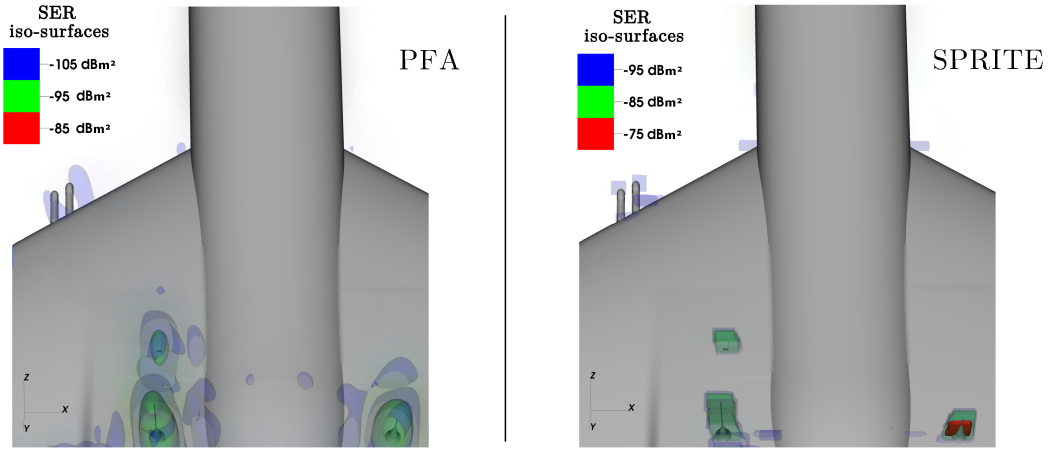


FIGURE 4.12 – Détails des interactions avec les antennes.
Comparaison entre la méthode PFA et SPRITE.

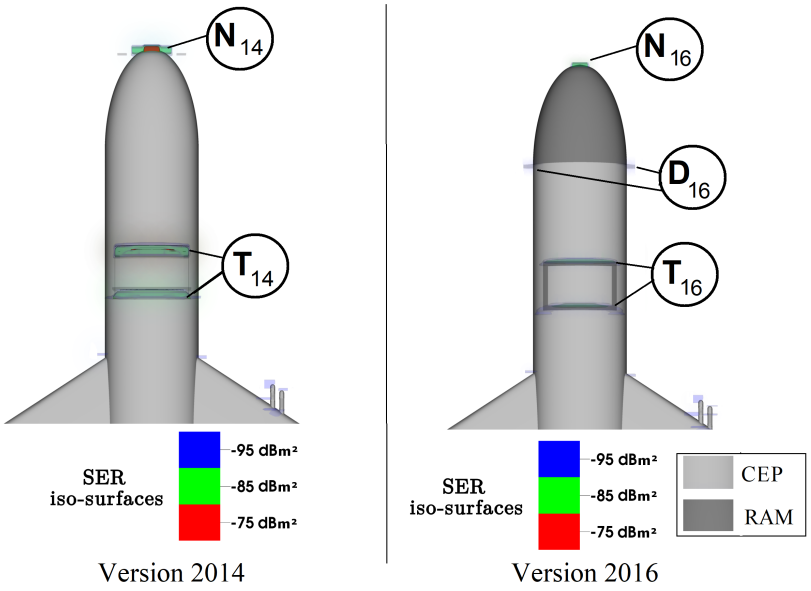


FIGURE 4.13 – Détails de l'interaction avec la trappe sur le fuselage.
Comparaison entre la version 2014 et 2016 (méthode SPRITE).

Le temps de calcul pour produire les cartes de SER estimées avec la méthode SPRITE pour les versions 2014 et 2016 est en moyenne de 8 minutes.

4.2.2 Imagerie d'un cône avec pastilles

Nous montrons tout d'abord l'apport de la méthode SPRITE pour l'analyse de la SER mono-dimensionnelle d'un cône avec pastilles mesuré avec l'Arche 3D, à angles de gisement et de roulis fixés à 0°(*i.e.* l'antenne est à la verticale) et pour une excursion fréquentielle.

La SER du cône avec pastilles est ensuite analysée en 3D à l'aide des cartes de rétro-diffusion estimées par la méthode SPRITE . En particulier, nous montrons que la considération de contraintes de support permet d'accroitre sensiblement la résolution transverse et améliore finalement l'analyse de SER 3D de la cible.

4.2.2.1 Analyse 1D de SER

La cible mesurée est un cône tronqué métallique CEP sur lequel trois pastilles métalliques sont collées.

Les caractéristiques du cône sont indiquées dans le Tableau 4.12.

Hauteur	**Diamètre de la base**	**Ouverture**
550 mm	150 mm	14,2 °

TABLEAU 4.12 – Caractéristiques du cône tronqué

Les positions des pastilles sur le cône sont quant à elles détaillées dans le Tableau 4.13.

Pastilles	②	③	④
Distance à la base du cône	400 mm	250 mm	115 mm
Angle de roulis	135 °	-45 °	0 °

TABLEAU 4.13 – Positions des pastilles sur le cône

Pour la mesure, la cible est placée sur un mât de support cylindrique en plexiglas dont les caractéristiques sont indiquées dans le Tableau 4.14.

Hauteur	**Diamètre**	**Constante diélectrique** ε_r
992 mm	40 mm	2,65

TABLEAU 4.14 – Caractéristiques du mât de support

La Figure 4.14 présente une photographie de la cible ainsi qu'un schéma du dispositif de mesure.

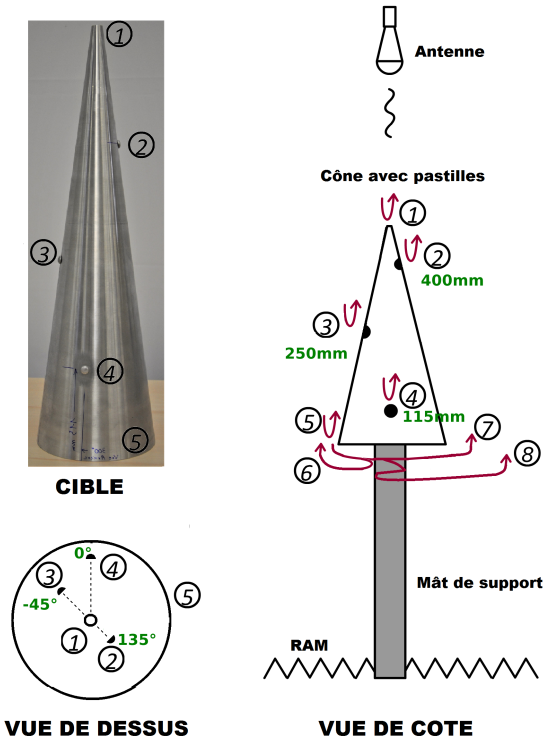


FIGURE 4.14 – Photographie de la cible et schéma du dispositif de mesure.

Données mesurées

Nous nous intéressons à l'analyse en bande X de la SER mono-dimensionnelle du cône avec pastilles. Les données sont mesurées avec l'Arche 3D pour une bande fréquentielle de 8 GHz à 12 GHz, par pas de 3,91 MHz. Ceci représente 1024 données. Les données sont mesurées en polarisation HH et VV.

Le processus de mesure consiste en différentes étapes successives qui sont explicitées dans l'**Annexe A** et rappelées ci-après.

- Une soustraction de chambre à vide est opérée afin de limiter les échos parasites.
- Un étalonnage par substitution avec une sphère métallique CEP de 300 mm de diamètre, dont la SER théorique est calculée par séries de Mie [Knott *et al.*, 2004].
- Un filtrage adaptatif entre -0,5 m et 0,4 m est appliqué afin de supprimer le TOS ainsi que les échos résiduels provenant par exemple d'interactions avec les murs ou le sol, qui pourraient affecter le signal utile.

Paramètres de la méthode conventionnelle et de SPRITE 1D

Les caractéristiques de la grille spatiale sont indiquées dans le Tableau 4.15 et l'origine du repère est au milieu de la carte.

N_z	Δ_z
4096	1,25 m

TABLEAU 4.15 – Grille spatiale 1D pour le cône avec pastilles

Enfin, le Tableau 4.16 détaille les paramètres de la méthode SPRITE 1D. L'hyperparamètre μ est choisi d'après le processus de calibration expliqué en **Section 4.1.3.3**. Le support spatial est choisi égal à l'ensemble du domaine imagé Δ_z.

μ	ρ_P	ε
3	2048	10^{-5}

TABLEAU 4.16 – Paramètres de la méthode SPRITE 1D pour le cône avec pastilles

Résultats

La Figure 4.15 présente la SER estimée selon la distance longitudinale, par la méthode conventionnelle et par la méthode SPRITE 1D que nous proposons, pour les polarisations HH et VV. Notons que la forme de la cible est superposée pour une meilleure analyse de la SER mais qu'elle ne participe en aucun cas au processus d'inversion.

Avec la méthode SPRITE 1D, les principaux contributeurs sont mieux localisés : la pointe tronquée du cône (1), la diffraction sur les pastilles métalliques (2), (3), (4) et la base du cône (5). Derrière la cible, notre méthode permet de mettre en évidence les pics (6), (7) et (8) qui ne sont pas distinguables avec l'approche conventionnelle par TFRI. Après investigations reposant sur des calculs de SER avec un solveur Maxwell harmonique [Stupfel *et al.*, 1991], il s'avère que ces pics résultent d'interactions entre l'onde rampante et le mât de support. Plus précisément, et comme indiqué par les flèches rouges sur la Figure 4.14 :

- le pic (6), situé à 55 mm de la base du cône, correspond à la réflexion directe de l'onde rampante du cône sur le mât,
- le pic (7), positionné à 87,6 mm de la base du cône, correspond au parcours de l'onde rampante à travers le mât. En effet, sa position est cohérente du retard de propagation de l'onde dans le plexiglas dont l'index de réfraction est $n = \sqrt{\varepsilon_r} = 1.63$.
- le pic (8), positionné à 120 mm de la base du cône, correspond à une double réflexion de l'onde rampante à l'intérieur du mât.

Notons que les calculs de SER mettent également en évidence des contributeurs d'amplitudes plus faibles derrière le dernier pic (8). Ils n'apparaissent pas sur la Figure 4.15 car ils ont été filtrés lors du processus de mesure.

Nous constatons également que, dans la partie centrale, la SER associée aux pastilles (2), (3) et (4) dépend de la polarisation de l'onde. Á roulis 0 °, cela s'explique par leurs positions angulaires respectives, indiquées dans le Tableau 4.13. En effet, la pastille (4) est espacée de 135°de la pastille (2) et de 45°de la pastille (3) alors que les pastilles (2) et (3) sont séparées de 180°. Ainsi, la pastille (4) ne signe pas en polarisation VV. Les pastilles (2) et (3) signent toujours de la même manière, quelle que soit la polarisation de l'onde, car elles sont diamétralement opposées. A l'inverse, les pics (1), (5), (6), (7) et (8) ne varient pas avec la polarisation parce que le cône et la mât de support sont axisymétriques et qu'il n'y a pas d'interactions importantes entre les pastilles.

Finalement, ces résultats montrent que la méthode SPRITE 1D permet de produire des cartes avec une résolution accrue qui facilite grandement l'analyse des mécanismes de rétrodiffusion.

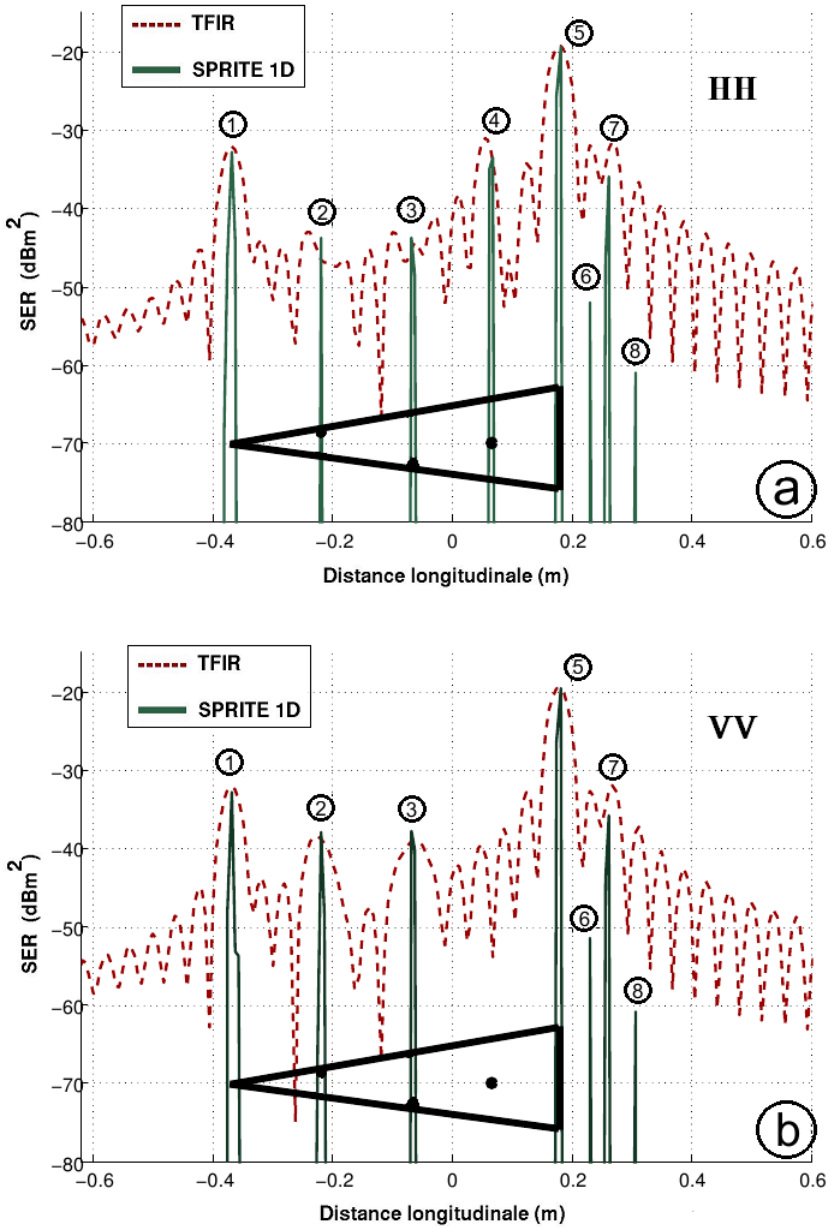


FIGURE 4.15 – SER estimées par TFRI et SPRITE 1D selon la distance longitudinale
(a) Polarisation HH. (b) Polarisation VV.

Le temps de calcul pour produire les cartes estimées par la méthode SPRITE 1D et présentées Figure 4.15 est en moyenne de 1 seconde.

4.2.2.2 Analyse 3D de SER

Données mesurées

Les données sont mesurées en bande X avec l'Arche 3D pour une polarisation HH. Le dispositif de mesure, illustré Figure 4.14, et les étapes de calibration du processus de mesure ont été décrits précédemment en **Section 4.2.2.1**. La bande fréquentielle et les bandes angulaires de mesures sont indiquées dans le Tableau 4.17. Ceci représente environ $4,6 \cdot 10^5$ données.

f	θ	φ
8 GHz : 3,9 MHz : 12 GHz	-20 ° : 1° : 20 °	-20 ° : 4° : 20 °

TABLEAU 4.17 – Bandes fréquentielle et angulaires pour le cône avec pastilles

Paramètres des méthodes PFA et SPRITE

Les caractéristiques de la grille spatiale sont indiquées dans le Tableau 4.18 et l'origine du repère est au milieu de la carte. Ceci représente environ $1,7 \cdot 10^7$ inconnues.

$N_x \times N_y \times N_z$	Δ_x	Δ_y	Δ_z
$256 \times 256 \times 256$	0,3 m	0,3 m	1 m

TABLEAU 4.18 – Grille spatiale pour le cone

Le Tableau 4.19 détaille les paramètres de la méthode SPRITE déterminés d'après le processus de calibration expliqué dans la **Section 4.1.3.3**.

μ	λ	β	ρ_{P}	ρ_{D}	$\rho_{\mathcal{C}}$	ε
5	10	100	128	$8,4 \cdot 10^6$	$8,4 \cdot 10^6$	10^{-4}

TABLEAU 4.19 – Paramètres de la méthode SPRITE pour le cone

Pour souligner l'intérêt de l'intégration d'une contrainte de support justifiée, nous comparons les cartes estimées par la méthode SPRITE pour deux choix de support spatial :

- le support spatial est choisi égal à l'ensemble du domaine, ce qui ne contraint pas le problème.
- le support spatial, en dehors duquel la solution est contrainte à 0, est défini relativement aux contours de la cible physique tout en tenant compte de l'épaisseur des pastilles et d'éventuels rampants à l'arrière de cette dernière. Notons que l'axisymétrie du cône est exploitée pour définir ce support spatial, que nous appelons « pelure de cône », comme illustré sur la Figure 4.16.

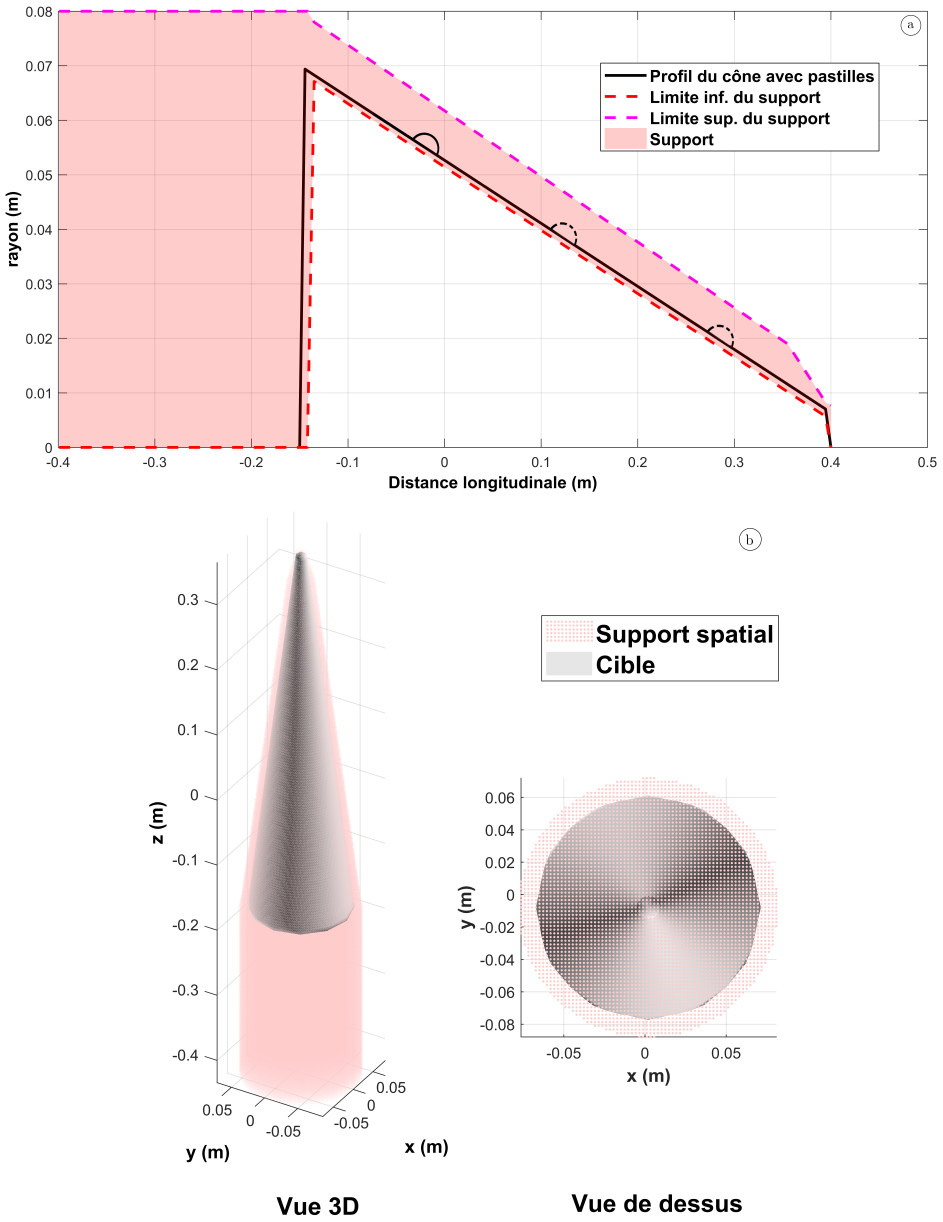


FIGURE 4.16 – (a) Définition du support spatial à partir du profil de la cible. (b) Représentation spatiale du support « pelure de cône ».

Résultats

La Figure 4.17 présente les cartes (exprimées par iso-surfaces en dBm^2) estimées par les méthodes PFA et SPRITE avec et sans contrainte de support. D'autres vues sont données sur la Figure C.3 dans l'**Annexe C** (p. 156).

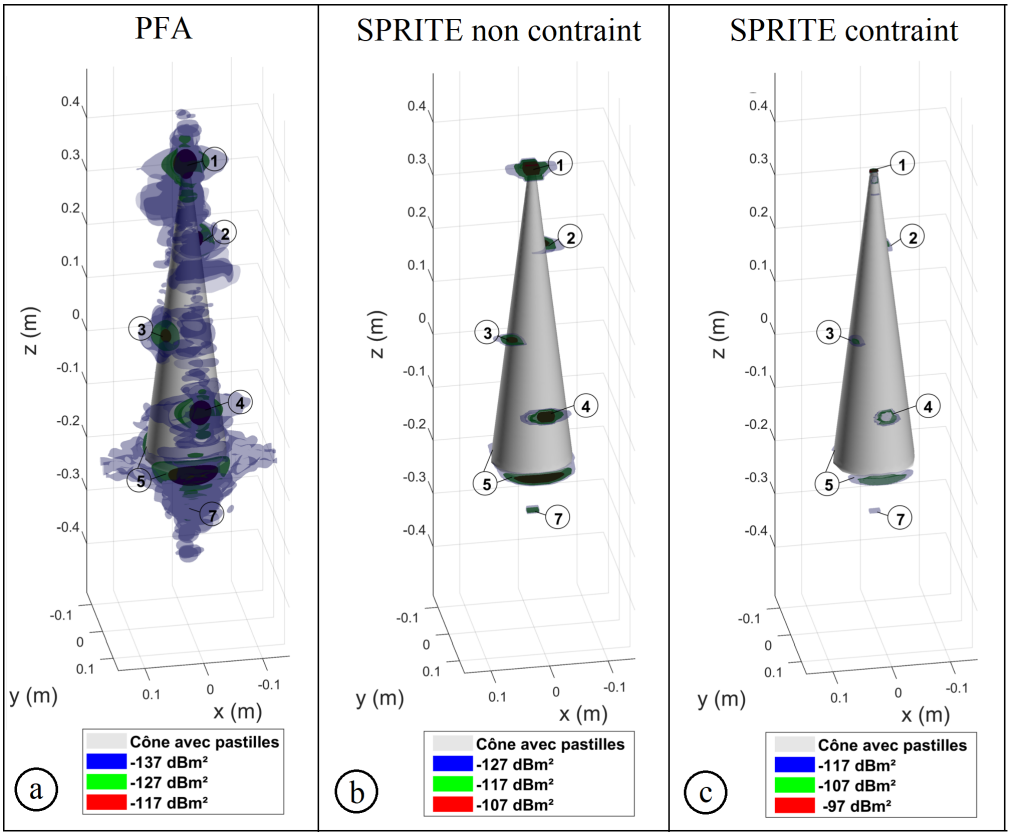


FIGURE 4.17 – Cartes de rétro-diffusion du cône avec pastilles.
(a) PFA. (b) SPRITE non contraint. (c) SPRITE contraint.

L'amélioration de la résolution facilite l'analyse de SER 3D et nous permet de mettre en évidence les principaux contributeurs et mécanismes de rétro-diffusion. Les résultats sont en cohérence avec l'analyse 1D réalisée dans la précédente section. En effet, nous pouvons visualiser très précisément :

- L'interaction avec la pointe tronquée du cône (1).
- Les interactions avec les pastilles (2), (3) et (4). La vue développée des cartes estimées, illustrée sur la Figure C.4 dans l'**Annexe C** (p. 157) , nous permet de constater que ces dernières sont nettement mieux localisées avec la méthode SPRITE qu'avec la méthode PFA. Nous remarquons également qu'en polarisation HH, la pastille (4) signe plus que les deux autres. Ceci est en cohérence avec l'analyse 1D réalisée avant.
- L'interaction avec la base du cône (5).
- L'interaction entre l'onde rampante et le mât de support en plexiglas (7). Cette dernière apparaît derrière la cible à 87.6 mm de la base du cône pour les raisons mentionnées dans l'analyse 1D. Notons que les autres interactions du rampant avec le mât (6) et (8) n'apparaissent pas sur la Figure 4.17 car leur niveau est inférieur aux iso-surfaces représentées.

Par ailleurs, nous constatons que la considération de contraintes de support spatial dans la méthode SPRITE permet d'accroitre la résolution, notamment transverse. Ceci est particulièrement visible sur la Figure 4.17 au niveau de l'interaction avec la pointe tronquée du

cône qui est plus précisément localisée. Il est donc intéressant de considérer des contraintes de support quand l'extension spatiale de l'objet de rétro-diffusion peut être définie.

Le temps de calcul pour produire les cartes de SER estimées avec la méthode SPRITE est de 5 minutes et 48 secondes pour le cas non contraint et de 6 minutes et 16 secondes pour le cas contraint.

4.3 Le contrôle 3D de SER

Nous montrons maintenant l'intérêt de la SPRITE pour le contrôle 3D de SER. La SER d'une maquette de lanceur multi-étages, appelée « lanceur MX-14 » et développée au CEA, est étudiée. Il s'agit en particulier de contrôler les différences entre les cartes de rétro-diffusion de trois configurations de cette cible. Ces dernières sont présentées Figure 4.18 et détaillées ci-dessous :

- $\#1$ l'objet nominal CEP d'une hauteur de 1,3 m.
- $\#2$ l'objet nominal avec une rainure (R), située à 67 cm de la pointe du lanceur (P), dans la partie basse au dessus des ailerons.
- $\#3$ l'objet nominal avec la même rainure que la configuration $\#2$ et un méplat, sur la partie haute du lanceur, de 2 cm de diamètre et dont le centre est situé à 14 cm de la pointe.

Dans la suite nous appelons « écarts à la configuration nominale » la rainure et le méplat.

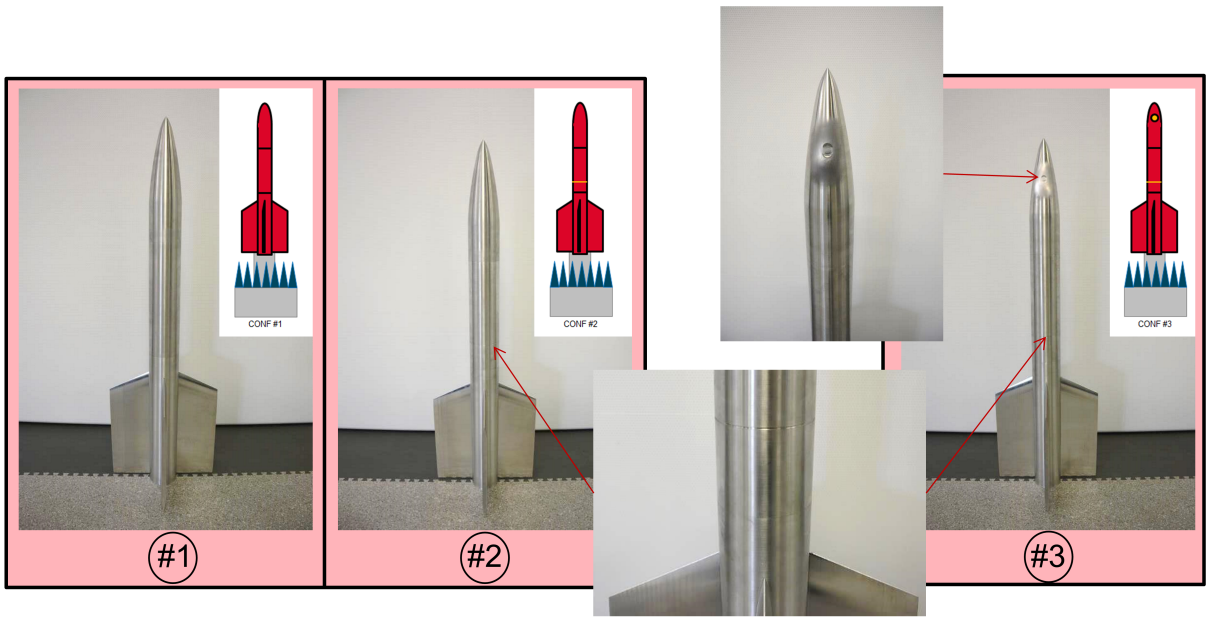


FIGURE 4.18 – Configurations $\#1$, $\#2$ et $\#3$ du lanceur MX-14

Données mesurées

Comme illustré sur la photographie du dispositif de mesure Figure 4.19, le lanceur MX-14 est placé sur un mât de support en polystyrène cylindrique de constante diélectrique $\epsilon_r \approx 1,2$.

Les données sont mesurées en bande X avec l'Arche 3D, pour une polarisation VV. Les étapes de calibration du processus de mesure sont similaires à celles opérées pour la mesure du cône avec pastilles et sont présentées dans la section **Section 4.2.2.1**. Un filtrage distance adaptatif est réalisé entre -0,55 m et 1 m pour éliminer le TOS et les échos parasites.

FIGURE 4.19 – Dispositif de mesure de l'Arche 3D pour le lanceur MX-14

La bande fréquentielle et les bandes angulaires de mesures sont indiquées dans le Tableau 4.20. Ceci représente environ $4,6 \cdot 10^5$ données.

f	θ	φ
8 GHz : 3,9 MHz : 12 GHz	-20 ° : 1° : 20 °	-20 ° : 4° : 20 °

TABLEAU 4.20 – Bandes fréquentielle et angulaires pour le lanceur MX-14

Paramètres des méthodes PFA et SPRITE

Les caractéristiques de la grille spatiale sont indiquées dans le Tableau 4.21 et l'origine du repère est au milieu de la carte. Ceci représente environ $6,7 \cdot 10^7$ inconnues.

$N_x \times N_y \times N_z$	Δ_x	Δ_y	Δ_z
$256 \times 256 \times 512$	0,5 m	0,5 m	2 m

TABLEAU 4.21 – Grille spatiale pour le lanceur MX-14

Enfin, le Tableau 4.22 détaille les paramètres de la méthode SPRITE déterminés d'après le processus de calibration (*cf.* **Section 4.1.3.3**).

μ	λ	β	ρ_{P}	ρ_{D}	$\rho_{\mathcal{C}}$	ε
5	15	50	256	$3,4 \cdot 10^{7}$	$3,4 \cdot 10^{7}$	10^{-5}

TABLEAU 4.22 – Paramètres de la méthode SPRITE pour le lanceur MX-14

En l'absence de connaissances *a priori* sur l'extension spatial des objets de rétro-diffusion, nous choisissons un support spatial égal au domaine imagé (le problème n'est pas contraint).

Résultats

La Figure 4.20 présente, pour les trois configurations et selon deux angles de vue, les cartes de rétro-diffusion estimées par la méthode SPRITE. Une comparaison avec les cartes estimées par la méthode PFA est données sur la Figure C.5 dans l'**Annexe C** (p. 153).

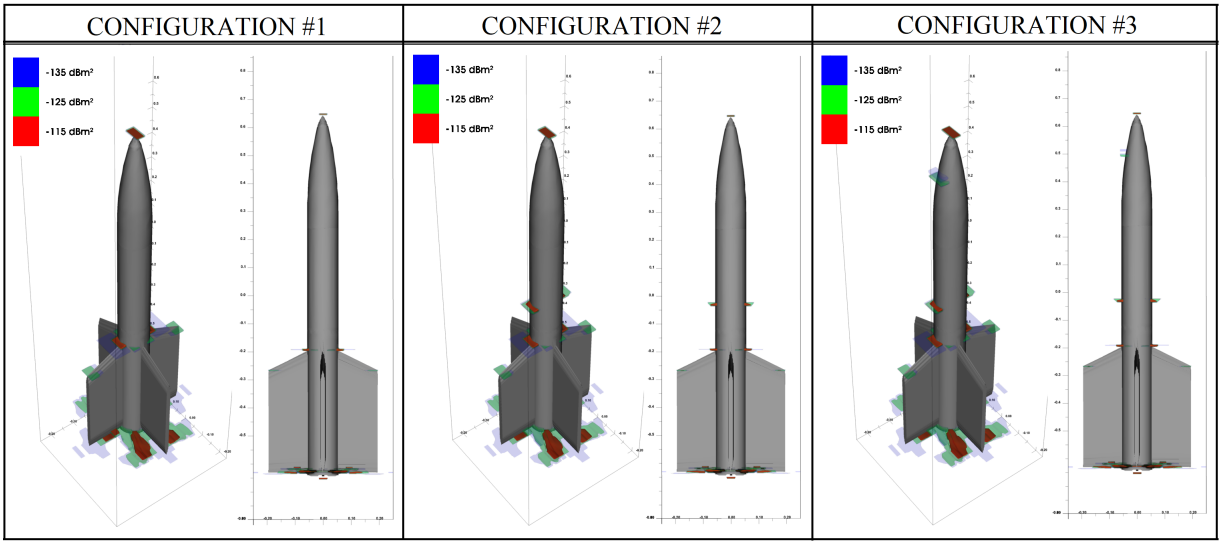


FIGURE 4.20 – Contrôle des cartes de rétro-diffusion des trois configurations du lanceur MX-14

La méthode SPRITE permet de discerner très distinctement la rainure et le méplat sur les configurations #2 et #3. Il est alors possible d'envisager des méthodes de traitement pour localiser les écarts à la configuration nominale (ces dernières n'ont pas été développées dans ce travail). Par ailleurs, comme illustré sur la Figure 4.21, la méthode SPRITE permet aussi de caractériser précisément ces écarts à la configuration nominale. Par exemple, nous pouvons distinguer les interactions sur les bords supérieur et inférieur du méplat (notés respectivement (M$_1$) et (M$_2$) sur la Figure 4.21) alors qu'elles ne sont pas séparables avec l'approche conventionnelle. Nous remarquons que le bord inférieur du méplat (M$_2$) signe plus que le bord supérieur (M$_1$). Ceci s'explique par le fait que dans ce dispositif de mesure (faible excursion angulaire autour de $\theta = 0^{\circ}$), le bord inférieur apparait comme une petite facette spéculaire qui rétro-diffuse l'onde alors que le bord supérieur apparait comme une arrête sur laquelle l'onde est diffractée.

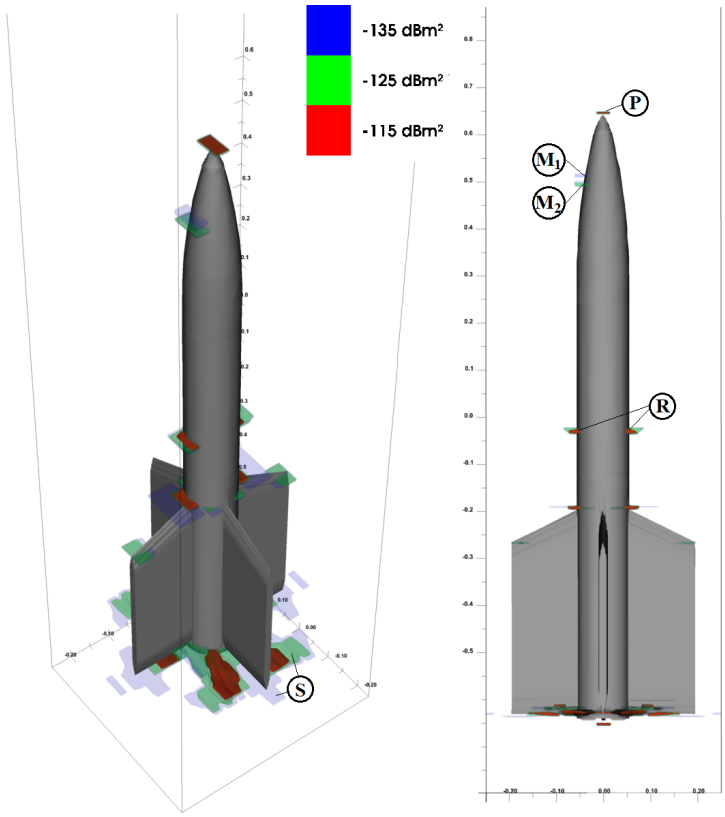


FIGURE 4.21 – Vue détaillée de la carte 3D de la configuration #3 du lanceur MX-14.

Le Tableau 4.23 compare les positions de la rainure et du méplat relevées sur la carte 3D à leur position réelle sur la cible. Relativement aux dimensions du lanceur (1,3 m de hauteur), les écarts entre les positions réelles et les positions estimées sont très faibles ($< 0,01\%$).

	Cible	Carte 3D	Ecart
Distance entre (M1) et (P)	13 cm	12,6 cm	4 mm
Distance entre (M2) et (P)	15 cm	14,4 cm	6 mm
Diamètre du méplat	2 cm	1,8 cm	2 mm
Distance entre (R) et (P)	67 cm	66,1 cm	9 mm

TABLEAU 4.23 – Comparaison des positions du méplat et de la rainure sur la cible et sur la carte 3D SPRITE

REMARQUE

Le lanceur est enfoncé dans le support en polystyrène (comme illustré sur la Figure 4.19). Le filtrage distance réalisé lors processus de calibration ne permet pas de supprimer l'interaction entre l'onde et le support. La section plane du mât de support apparait donc comme une facette spéculaire (S), comme illustré sur la Figure 4.21).

Conclusion du Chapitre 4

Dans ce chapitre, nous avons tout d'abord validé la méthode SPRITE sur des données synthétiques. Nous avons montré qu'elle permet d'accroitre drastiquement la résolution des cartes estimées comparativement à la méthode conventionnelle PFA. Ces résultats confirment l'étude menée dans le chapitre précédent. Nous avons également étudié empiriquement l'influence des hyperparamètres et nous avons proposé une méthode de calibration.

Dans un second temps, nous avons évalué la méthode SPRITE sur des données réelles. A travers divers exemples, nous avons montré qu'elle permet d'analyser plus finement les mécanismes de rétro-diffusion et de contrôler précisément la SER de cibles qui diffèrent d'une configuration nominale.

CONCLUSIONS GÉNÉRALES ET PERSPECTIVES

Conclusions

Dans le cadre de ce travail de thèse, nous nous sommes intéressés à la thématique de l'imagerie 3D HR de SER. La contribution centrale de cette thèse consiste en la proposition et l'évaluation d'une nouvelle méthode HR pour l'imagerie 3D de SER, dénommée SPRITE. Cette méthode vient en rupture avec les méthodes classiques reposant sur un regridding des données dans le k-space et une TFRI 3D qui conduisent à des résultats de résolution limitée. Elle est de plus en rupture avec les dernières méthodes de l'état de l'art qui favorisent des cartes avec un faible nombre de points brillants, ce qui ne correspond pas à la réalité des interactions onde-cible. Nous rappelons brièvement les différentes parties que nous avons développées dans ce document ainsi que le cheminement qui nous a conduit à la méthode SPRITE.

Dans la première partie du document, nous avons rappelé la notion de SER. Nous avons présenté les principaux mécanismes de rétro-diffusion et d'interactions onde-cible qui concourent à la SER en hautes fréquences. Il s'agit des spéculaires qui s'expriment comme des facettes brillantes, des ondes propagées et rampantes, des interactions multiples ou encore des phénomènes de diffraction. Nous avons introduit la notion d'*analyse de SER* qui consiste à localiser et à caractériser précisément les principaux contributeurs à la SER d'une cible ainsi que la notion de *contrôle de SER* qui consiste à s'assurer que la SER est conforme à un certain niveau de spécifications. À travers une série d'exemples nous avons montré que ces deux notions sont des éléments clés pour de nombreuses applications radar. Nous avons ensuite présenté le processus de mesure de l'hologramme de SER ainsi que le modèle d'observation liant la SER mesurée à l'objet de rétro-diffusion à estimer. Nous avons alors expliqué que l'imagerie 3D de SER est un outil adapté pour analyser et contrôler la SER. Nous avons montré qu'il s'agit d'un problème de synthèse de Fourier non inversible et qui plus est posé en grande dimension. Nous avons alors présenté les méthodes conventionnelles d'imagerie PFA et RPF, qui consistent en une TFRI après reformatage des données avec bourrage de zéro ; nous avons souligné leurs limites en terme de résolution. Afin d'améliorer la résolution des cartes de rétro-diffusion estimées nous avons été amenés à étudier de près la notion d'*approche problème inverse* avec une attention particulière sur la *régularisation parcimonieuse et contrainte*. Celle-ci fournit un cadre rigoureux pour développer des méthodes HR. Dans ce formalisme, nous avons présenté quelques méthodes HR qui favorisent la formation de cartes contenant un faible nombre de points brillants.

Dans la deuxième partie du document, nous nous sommes plus spécifiquement intéressés aux méthodes de régularisation parcimonieuse qui permettent de promouvoir des cartes avec un faible nombre d'éléments. Nous avons premièrement rappelé que la minimisation

de critères composites avec un terme d'attache aux données et des pénalités à norme ℓ_1 favorise la parcimonie tout en respectant les données mesurées. Nous avons alors présenté quelques algorithmes d'optimisation de critères non différentiables et contraints que nous avons classé en quatre grandes familles : les méthodes de lissage des non différentiabilités, les méthodes de sous-gradients et de sous-gradients projetés, les méthodes proximales et les méthodes primales-duales. Nous nous sommes en particulier focalisés sur l'algorithme primal-dual ADMM.

Dans la troisième partie du document, nous avons présenté la contribution centrale de cette thèse : le développement de la méthode SPRITE. Elle s'appuie sur la construction d'un critère pénalisé et contraint afin de prendre en compte les connaissances *a priori* sur la solution recherchée. La minimisation de ce critère composite permet de favoriser la parcimonie de la projection de l'objet sur l'axe longitudinal colinéaire au vecteur d'onde ainsi que des facettes connexes sur lesquelles le coefficient de rétro-diffusion est constant. L'optimisation en grande dimension est assurée par une adaptation de l'ADMM au problème d'imagerie qui mène à des calculs efficaces et particulièrement simples à l'aide de TFR.

Enfin, dans la dernière partie du document, nous avons tout d'abord validé la méthode SPRITE sur des données synthétiques. Nous avons montré qu'il s'agit d'une méthode performante qui permet d'accroitre drastiquement la résolution des cartes estimées comparativement à la méthode conventionnelle PFA tout en conservant des temps de calcul limités. Nous avons également étudié empiriquement l'influence des hyperparamètres et nous avons également proposé une méthode de calibration des hyperparamètres. Dans un second temps, nous avons évalué la méthode SPRITE sur des données réelles et nous avons montré qu'elle permet d'analyser plus finement les mécanismes de rétro-diffusion et de contrôler précisément la SER de cibles qui diffèrent d'une configuration nominale.

Perspectives

Nous listons maintenant quelques perspectives qui nous semblent intéressantes à court et moyen terme.

Acquisitions comprimées

La notion d'*acquisitions comprimées* (ou *compressive sensing* en anglais) [Donoho, 2006] consiste à estimer un signal à partir d'un faible nombre d'acquisitions. Des premiers résultats empiriques montrent que contrairement au PFA, la méthode SPRITE donne la possibilité de diminuer considérablement le nombre d'acquisitions nécessaires pour produire des cartes de qualité donnée. Cet aspect est particulièrement intéressant en ce qui concerne la diminution des durées d'acquisition de l'hologramme de SER.

Quantification d'incertitude

Nous avons indiqué dans la **Section 1.2.3** (p. 36) qu'une erreur de mesure et de modélisation est prise en compte dans le modèle d'observation (1.18). Il pourrait alors être intéressant de quantifier l'incertitude associée aux mesures et au choix du modèle et d'étudier la propagation de ces incertitudes sur les cartes estimées. Ceci permettrait notamment d'associer des intervalles d'incertitude aux cartes produites.

Méthode non supervisée

Comme expliqué dans la section **Section 3.4.3** (p. 108), le choix des hyperparamètres μ, λ et ν de la méthode SPRITE est crucial car il détermine le compromis entre le terme d'attache aux données et chaque pénalité. Il détermine donc la solution estimée. Cependant ce réglage est complexe pour des critères impliquant des normes ℓ_1 et la question de la détermination de ces hyperparamètres est un problème encore ouvert.

Dans notre dans notre cas, ces hyperparamètres sont réglés par calibration comme expliqué dans la section **Section 4.1.3.3** (p. 124).

Une perspective intéressante serait d'estimer ces hyperparamètres en même temps que la carte $\hat{\boldsymbol{a}}$ ce qui permettrait d'aboutir à une méthode non-supervisée. Dans cette esprit, nous évoquons ci-après quelques pistes, qui n'ont pas été étudiées dans ce travail de thèse, mais qui pourraient s'avérer pertinentes.

De nombreuses méthodes ont été proposées pour des régularisations de Tikhonov (*i.e.* avec des pénalités ℓ_2). Parmi les plus connues nous pouvons citer la L-curve [Hansen, 2000] ou la validation croisée généralisée [Golub *et al.*, 1979]. Le formalisme bayésien présenté dans l'**Annexe B** offre plus de souplesse quand à l'estimation conjointe des hyperparamètres et de l'objet d'intérêt [Idier, 2001, Chapitre 8]. Dans le cas d'une régularisation de Tikhonov, les stratégies reposent généralement sur des *a priori* gaussien. Dans cet esprit nous pouvons citer les algorithmes d'échantillonnage de Markov Chain Monte Carlo (MCMC), voir par exemple [Orieux *et al.*, 2013], ou les approches Variational Bayes Approach (VBA), voir par exemple [Ormerod *et al.*, 2017]. Cependant il ne semble pas y avoir de consensus sur la méthode à utiliser.

En ce qui concerne les régularisations avec pénalités ℓ_1, le nombre de méthodes étudiées dans la littérature est plus restreint. A titre d'exemple, nous pouvons citer la méthode déterministe SURE (Stein's Unbiased Risk Estimate) (voir par exemple [Weller *et al.*, 2014]) ou encore une fois des techniques reposant sur des stratégies bayésiennes (avec *a priori* laplacien) (par exemple [Frecon *et al.*, 2017; O'Hara et Sillanpää, 2009]).

Pénalité avec pseudo-norme ℓ_0

Comme expliqué dans la **Section 2.1.2** (p. 62), la minimisation de critère avec des pénalités ℓ_0 est un problème combinatoire complexe. C'est pourquoi la pseudo-norme ℓ_0 est généralement relaxée par une norme ℓ_1.

Néanmoins, il pourrait être intéressant de construire et minimiser des critères avec des pénalités ℓ_0 pour renforcer la promotion de la parcimonie. Dans cet esprit, [Storath *et al.*, 2014] ont par exemple proposé une méthode pour résoudre un problème de minimisation appelé *iPotts* (pour Inverse Potts Problem). Le critère à minimiser est similaire à celui de la

variation totale présenté dans la **Section 2.1.2** à ceci près que la pénalité en norme ℓ_1 est remplacée par une pseudo-norme ℓ_0. L'optimisation est assurée par l'ADMM et l'étape de mise-à-jour de l'objet (qui fait apparaitre la minimisation d'une fonctionnelle non convexe) est assurée par une méthode de graph-cuts. Dans cet esprit nous pouvons citer les travaux de [Giovannelli et Vacar, 2017; Mohammad-Djafari, 2008; Tupin, 2007].

ANNEXE A

PROCESSUS DE CALIBRATION DES MESURES DE SER

La Figure A.1, illustre les différents signaux parasites intervenant dans le modèle du signal complexe mesuré par l'ARV :

$$S_{Mes.} = G\left(S_{Cible} + S_{Envir.} + S_{Coupl.}\right) + B \tag{A.1}$$

- $S_{Mes.}$ est le signal mesuré à la sortie du récepteur,
- G est le gain complexe de la chaîne radar,
- S_{Cible} est le signal rétro-diffusé par la cible (homogène à la racine de la SER),
- $S_{Envir.}$ représente les échos parasites dus à l'environnement (chambre, positionneur) appelés *fouillis* ainsi que la signature propre du TOS de l'antenne,
- $S_{Coupl.}$ est le couplage entre l'objet et l'environnement. Il s'agit du terme le plus difficile à identifier et à éliminer,
- B est le bruit thermique de l'instrumentation. Il est négligeable devant les autres signaux compte tenu de la bande passante du filtre de réception de l'ARV qui varie entre quelques dizaines de Hz et quelques centaines de kHz autour de la porteuse [Garat et Morvan, 1993].

Notons que ces signaux dépendent du vecteur d'onde $\boldsymbol{k}$ (*i.e.* de la fréquence d'analyse f et des angles d'incidence θ et φ).

Nous détaillons ci-après les étapes de *soustraction vectorielle* et d'*étalonnage* [Garat *et al.*, 1993].

La soustraction vectorielle

Pour éliminer le terme $S_{Envir.}$ nous effectuons une soustraction vectorielle dont un schéma de principe est présenté en Figure A.2. Cette opération consiste à effectuer préalablement une mesure de la chambre à vide sans la cible $S_{Ch.\ Vide\ Cible}$ (dans la même configuration que la mesure de la cible), puis à la soustraire à la mesure du signal $S_{Mes.}$.

$$S_{Cible} \approx S_{Mes.} - S_{Ch.\ Vide\ Cible} \tag{A.2}$$

La différence entre ces deux signaux complexes permet d'obtenir (aux couplages entre l'objet et l'environnement près) le signal S_{cible} (homogène à la SER de la cible) en s'affranchissant de

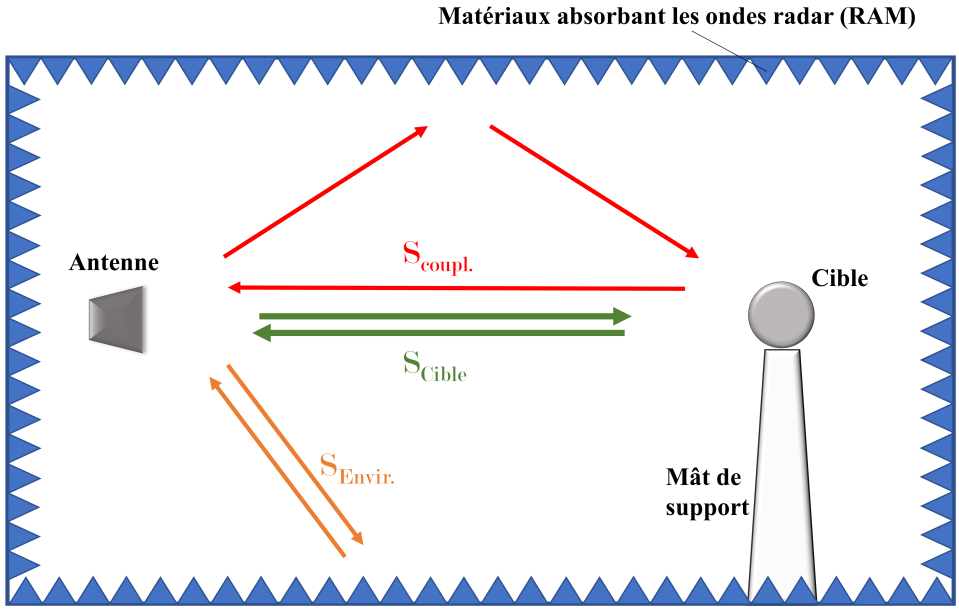


FIGURE A.1 – Mesure de SER : signaux utile et parasites

la contribution de l'enceinte de mesure. Une telle opération est d'autant plus efficace que les conditions dans lesquelles les deux mesures sont effectuées sont proches. Ceci exige une très bonne stabilité thermique et mécanique du système de mesure. Néanmoins, cette méthode ne permet pas d'éliminer les couplages parasites entre la chambre et l'objet $S_{Coupl.}$ car ces derniers sont absents lors de la mesure de la chambre à vide.

L'étalonnage

Dans l'équation (A.1) nous constatons que le signal mesuré par l'instrumentation $S_{Mes.}$ est proportionnel à S_{Cible} (et donc à la SER de la cible), à un coefficient près G : le gain complexe de la chaîne de mesure. Ce coefficient prend en compte les gains des voies d'émission et de réception, le gain de l'antenne ainsi que l'atténuation due à la propagation en espace libre entre l'antenne et la zone de mesure. Notons que la modélisation employée ici néglige les couplages et présuppose un coefficient multiplicatif indépendant de la cible sous test. Dans ce cas, pour étalonner les mesures de SER, il suffit de mesurer un objet dont la SER est théoriquement connue et d'appliquer la formule suivante :

$$S_{\text{Cible Etalonné}} = \frac{S_{\text{Total Mes.}} - S_{\text{Ch. Vide Cible}}}{S_{\text{Etalon Mes}} - S_{\text{Ch. Vide Etalon}}} \times S_{\text{Etalon Théo.}} \tag{A.3}$$

où :

- $S_{\text{Cible Etalonné}}$ est le signal étalonné de la cible sous test,
- S_{Cible} est le signal mesuré à la sortie du récepteur pour un étalon de référence,
- $S_{\text{Ch. Vide Cible Etalon}}$ est le signal mesuré pour la chambre vide dans la même configuration que la mesure de l'étalon,
- $S_{\text{Etalon Théo.}}$ est la SER théorique de l'étalon.

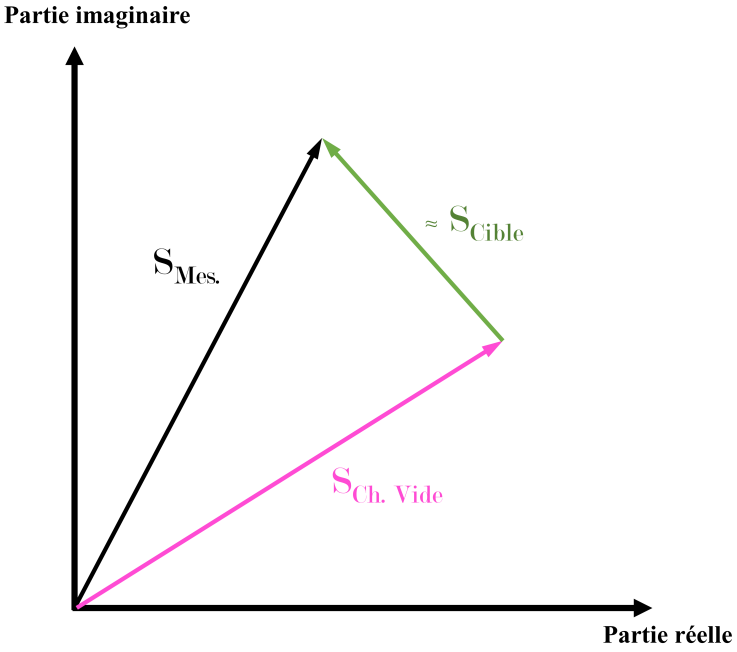


FIGURE A.2 – Schéma de principe de la soustraction vectorielle

En général, l'étalon choisi est une sphère métallique parfaitement conductrice, de diamètre connu, et dont la SER théorique est calculée analytiquement par série de Mie [Knott *et al.*, 2004].

L'opération de calibration permet également de ramener la référence de phase des mesures au centre de l'étalon. Ainsi toutes les mesures ont une référence de phase commune. Par ailleurs, afin d'éliminer les échos parasites en dehors de la zone d'intérêt et le TOS de l'antenne, il est possible d'appliquer des techniques de filtrage spatial [Stutzman et Thiele, 2013].

Enfin, si le contexte champ lointain du modèle d'observation, présenté dans la section **Section 1.2.3**, n'est pas respecté (*i.e.* les fronts d'onde ne sont pas plans mais sphériques), il est possible d'appliquer des corrections champ proche-champ lointain appelées *correction de sphéricité*. Ces dernières reposent généralement sur la *méthode de la lentille* qui dérive elle-même du principe d'Huygens-Fresnel [Guillemot *et al.*, 1997; Odendaal et Joubert, 1996].

ANNEXE B

INFÉRENCE BAYÉSIENNE

Lien entre la régularisation déterministe et l'approche bayésienne

La régularisation d'un problème consiste à prendre en compte des informations *a priori* sur la solution recherchée. Dans un cadre déterministe, cela peut se traduire par la construction d'un critère composite composé d'un terme d'attache aux données et d'une ou plusieurs pénalités qui favorisent l'information *a priori*. Ce critère est ensuite minimiser pour trouver une solution au problème. Cette idée générale peut s'interpréter dans le cadre de l'inférence bayésienne. Pour cela, il faut disposer des éléments suivants.

- $p(\boldsymbol{a})$: la densité de probabilité *a priori* sur l'objet recherché,
- $p(\boldsymbol{\sigma}|\boldsymbol{a})$: le modèle d'observation appelé appelé *vraisemblance*.

L'estimateur du *Maximum A Posteriori* (MAP) est défini par :

$$\hat{\boldsymbol{a}}_{\text{MAP}} = \underset{\boldsymbol{a}}{\operatorname{argmax}}\, p(\boldsymbol{a}|\boldsymbol{\sigma}) \tag{B.1}$$

où $p(\boldsymbol{a}|\boldsymbol{\sigma})$ est appelé *densité a posteriori*.

D'après la règle de Bayes nous avons :

$$\hat{\boldsymbol{a}}_{\text{MAP}} = \underset{\boldsymbol{a}}{\operatorname{argmax}}\, \frac{p(\boldsymbol{\sigma}|\boldsymbol{a})p(\boldsymbol{a})}{p(\boldsymbol{\sigma})} \tag{B.2}$$

REMARQUE

1. Le terme $p(\boldsymbol{\sigma})$ peut quant à lui être occulté car il ne joue qu'un rôle de normalisation.
2. La maximisation de la densité *a posteriori* et équivalente à la maximisation de son logarithme (du fait de la monotonie du logarithme).

Nous pouvons donc écrire la relation suivante :

$$\begin{aligned} \hat{\boldsymbol{a}}_{\text{MAP}} &= \underset{\boldsymbol{a}}{\operatorname{argmax}} \Big\{ p(\boldsymbol{\sigma}|\boldsymbol{a})p(\boldsymbol{a}) \Big\} \\ &= \underset{\boldsymbol{a}}{\operatorname{argmax}} \Big\{ \ln p(\boldsymbol{\sigma}|\boldsymbol{a}) + \ln p(\boldsymbol{a}) \Big\} \end{aligned} \tag{B.3}$$

Et finalement, nous avons :

$$\hat{\boldsymbol{a}}_{\text{MAP}} = \underset{\boldsymbol{a}}{\operatorname{argmin}} \left\{ -\ln p(\boldsymbol{\sigma}|\boldsymbol{a}) - \ln p(\boldsymbol{a}) \right\} \tag{B.4}$$

Le terme $-\ln p(\boldsymbol{\sigma}|\boldsymbol{a})$ est l'anti-log-vraisemblance et traduit l'adéquation aux données. Le terme $\ln p(\boldsymbol{a})$ traduit quant à lui les connaissances *a priori* introduites sur la solution recherchée. Il est donc possible d'envisager l'estimateur MAP comme une régularisation de l'estimateur du maximum de vraisemblance (avec égalité si la densité *a priori* $p(\boldsymbol{a})$ est uniforme).

Application pour la méthode SPRITE

Pour trouver une expression pour la vraisemblance nous considérons le modèle d'observation (1.18) et nous supposons que le bruit $\boldsymbol{b}$ est un bruit blanc centré gaussien complexe, dont les composantes sont indépendantes et identiquement distribuées (i.i.d.), et dont la densité s'écrit :

$$p(\boldsymbol{b}) \propto \exp -\frac{\|\boldsymbol{b}\|_2^2}{2\varrho^2} \tag{B.5}$$

où la variance ϱ^2 est proportionnelle à la puissance du bruit. Ce choix de modèle de bruit est couramment utilisé pour représenter les mesures radar [Borden, 1992; Mensa, 1981]. La vraisemblance s'écrit donc :

$$p(\boldsymbol{\sigma}|\boldsymbol{a}) \propto \exp -\frac{1}{2\varrho^2} \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 \tag{B.6}$$

Nous choisissons la densité *a priori* suivante :

$$p(\boldsymbol{a}) \propto \exp -\eta\Psi(\boldsymbol{a}) \tag{B.7}$$

où

$$\Psi(\boldsymbol{a}) = \mu \|\boldsymbol{P}\boldsymbol{a}\|_1 + \lambda \|\boldsymbol{D}_x\boldsymbol{a}\|_1 + \lambda \|\boldsymbol{D}_y\boldsymbol{a}\|_1 + \frac{\nu}{2} \|\boldsymbol{a}\|_2^2$$

est la somme des pénalités qui apparaissent dans le critère (3.10) de la méthode SPRITE . Notons que la densité *a priori* $p(\boldsymbol{a})$ (B.7) est proportionnelle à un produit de densités relevant de lois de Laplace (pour les termes avec une norme ℓ_1) et d'une densité relevant d'une loi normale (pour les termes en norme ℓ_2).

Finalement, en exploitant (B.4) nous avons :

$$\hat{\boldsymbol{a}}_{\text{MAP}} = \underset{\boldsymbol{a}}{\operatorname{argmin}} \left\{ \frac{1}{2\varrho^2} \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \eta\Psi(\boldsymbol{a}) \right\} \tag{B.8}$$

En posant $\eta = 1/(\varrho^2)$ et en substituant dans (B.8) nous avons :

$$\hat{\boldsymbol{a}}_{\text{MAP}} = \underset{\boldsymbol{a}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \mu \|\boldsymbol{P}\boldsymbol{a}\|_1 + \lambda \|\boldsymbol{D}_x\boldsymbol{a}\|_1 + \lambda \|\boldsymbol{D}_y\boldsymbol{a}\|_1 + \frac{\nu}{2} \|\boldsymbol{a}\|_2^2 \right\} \tag{B.9}$$

REMARQUE

L'estimation du MAP $\hat{a}_{\text{MAP}}$ (B.9) est le minimiseur du critère non contraint de la méthode SPRITE (3.11).

ANNEXE C

FIGURES ADDITIONNELLES

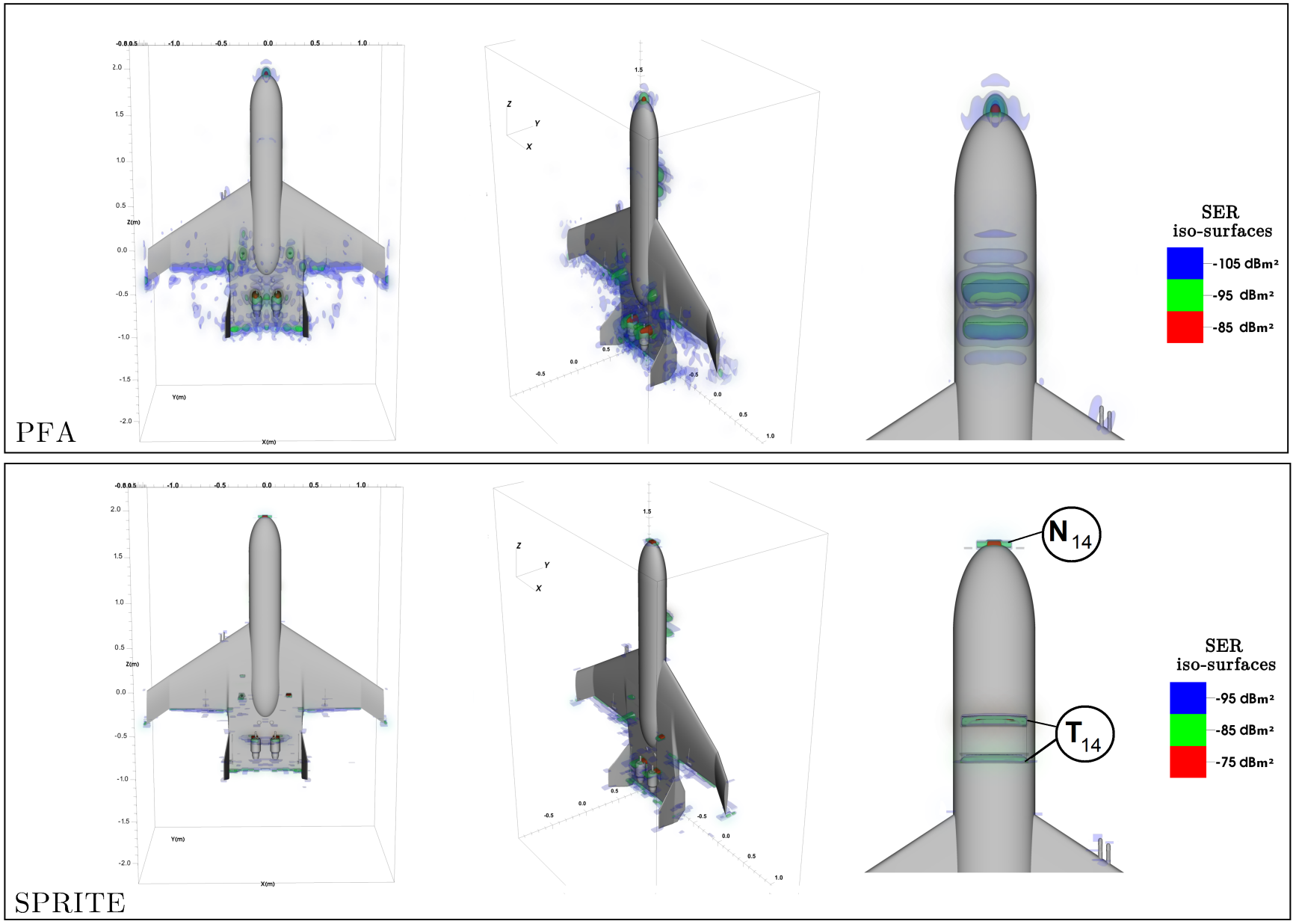


FIGURE C.1 – Carte de rétro-diffusion du drone version 2014

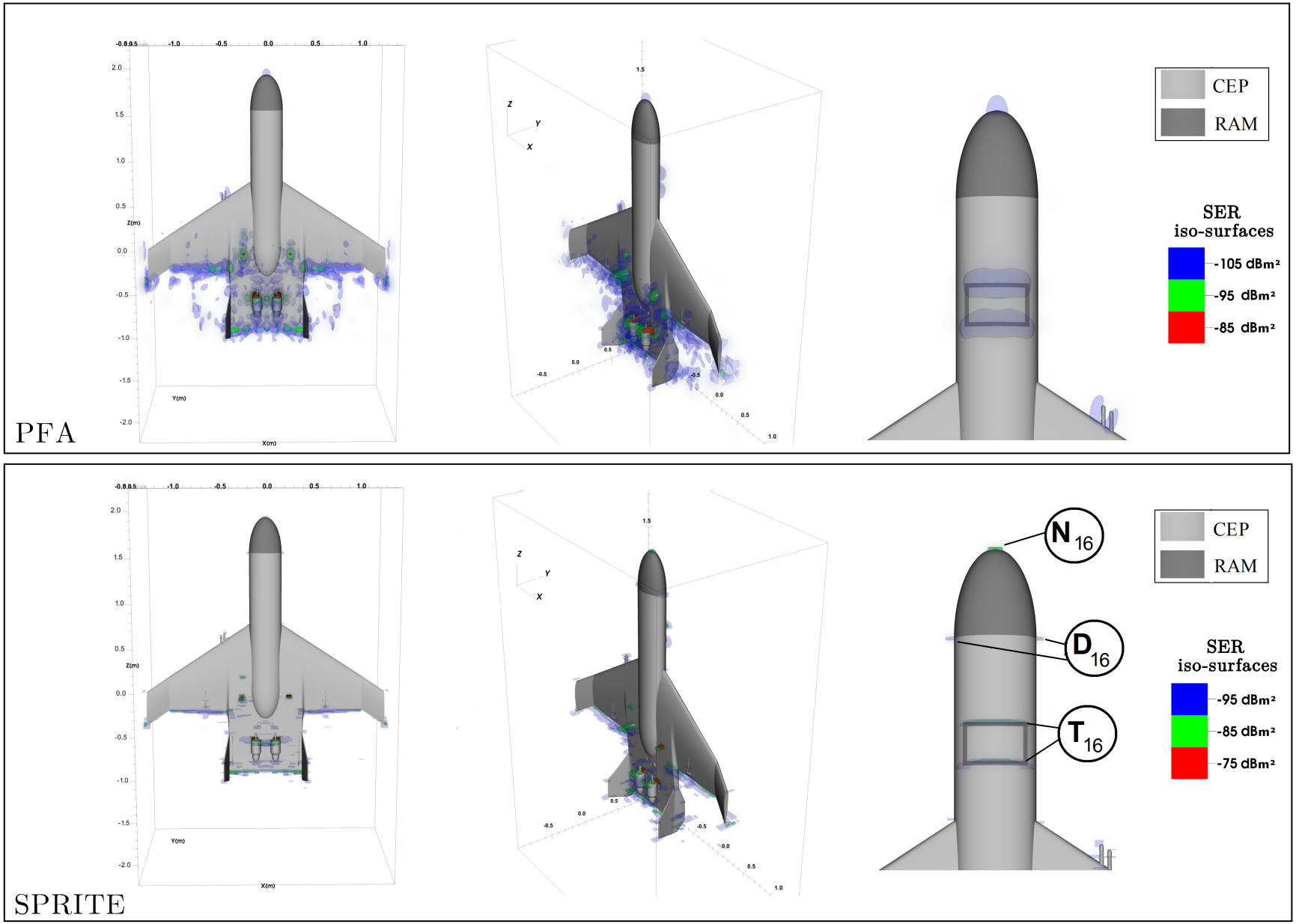


FIGURE C.2 – Carte de rétro-diffusion du drone version 2016

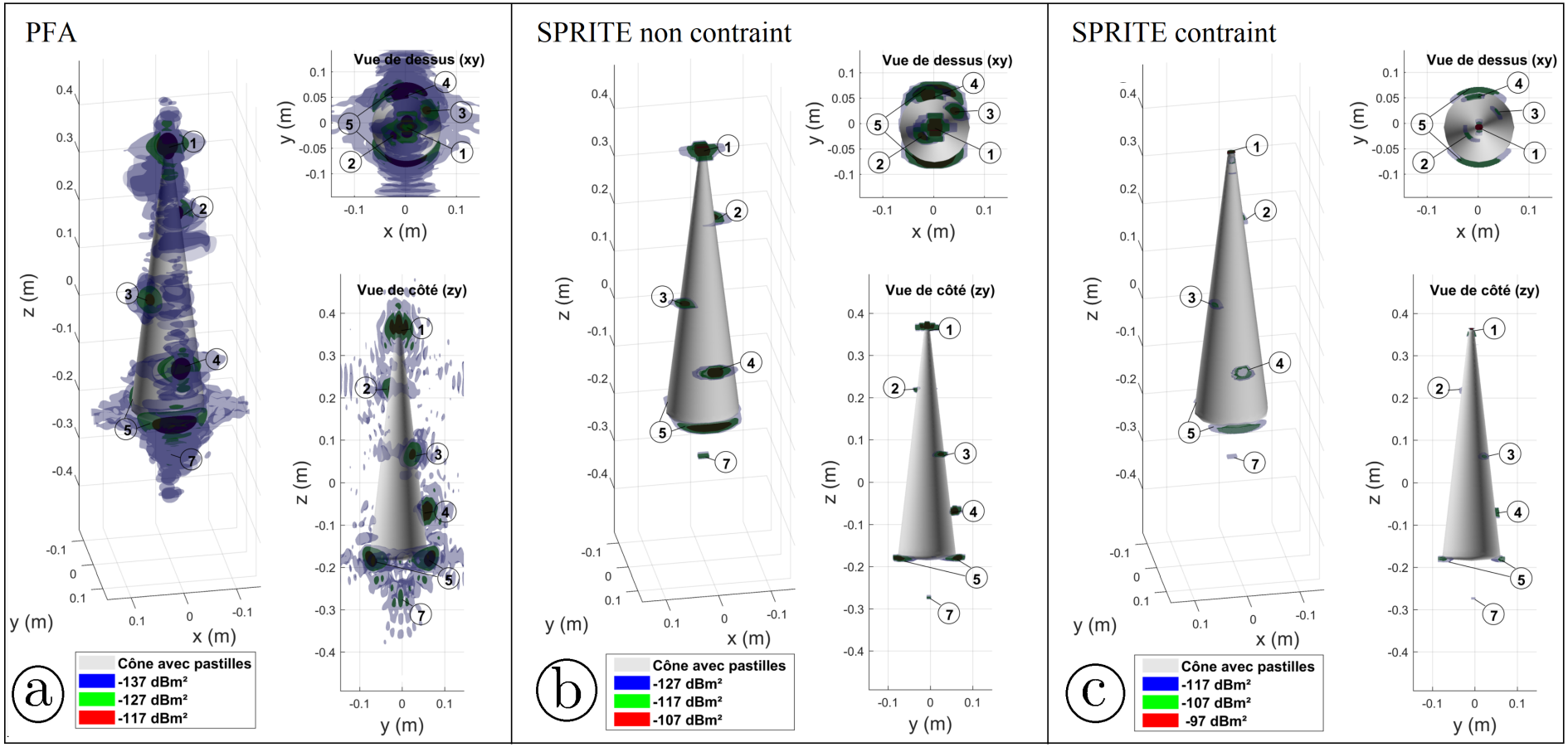


FIGURE C.3 – Comparaisons des cartes estimées (polarisation HH).
(a) PFA. (b) SPRITE non contraint. (c) SPRITE contraint.

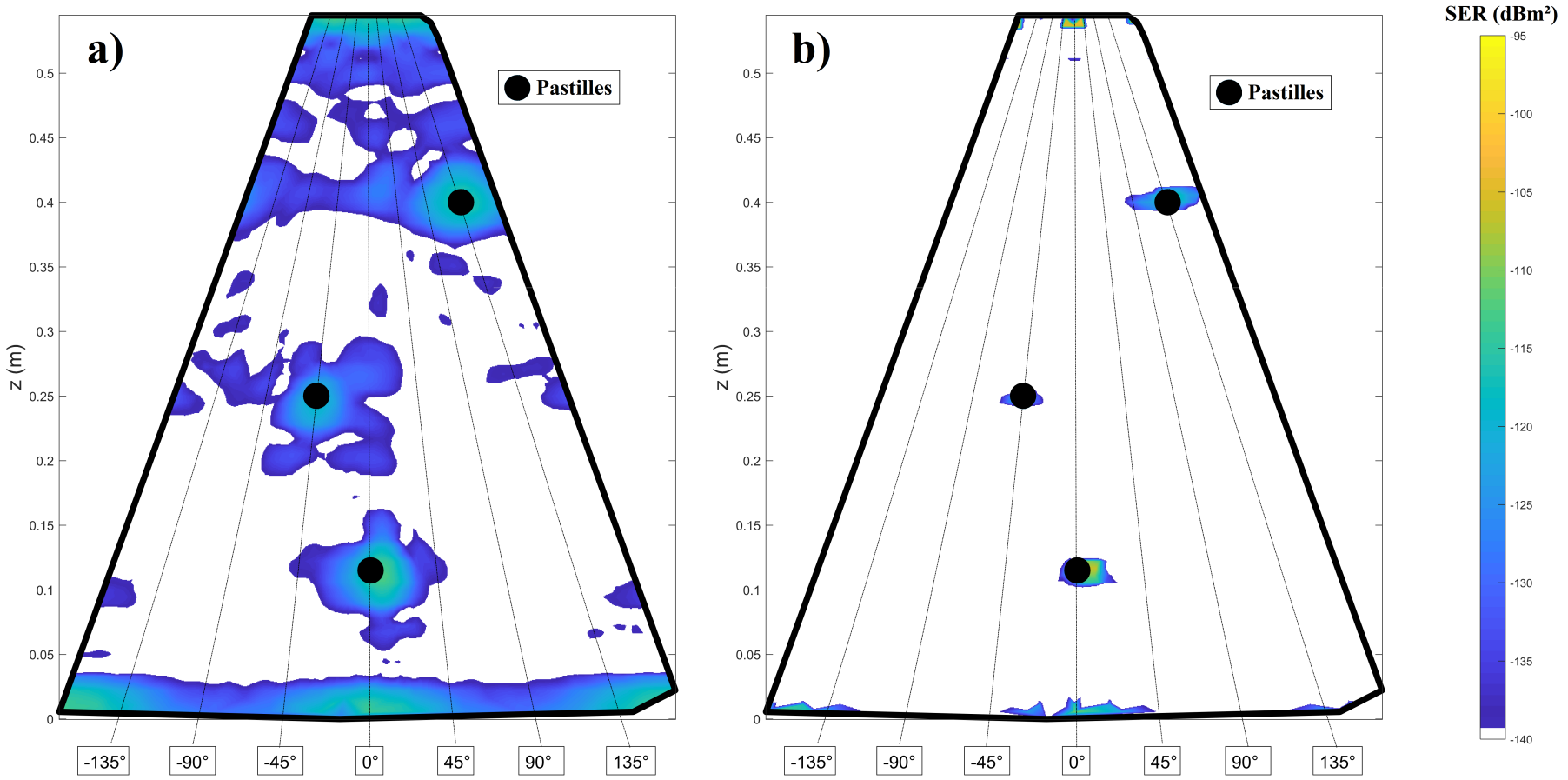


FIGURE C.4 – Vue développée des cartes de rétro-diffusion estimées avec les méthodes :
a) PFA. b) SPRITE contraint.

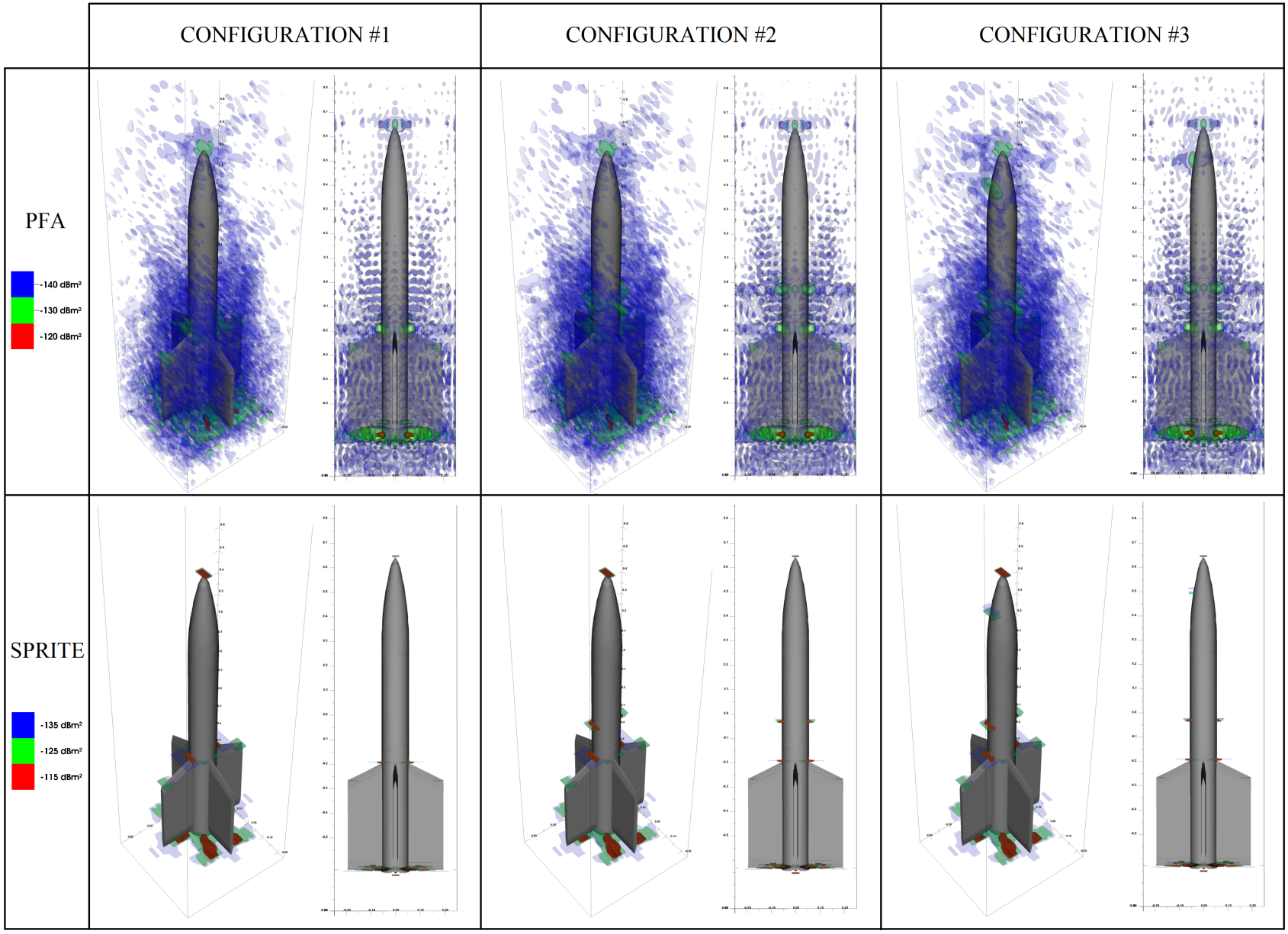


FIGURE C.5 – Contrôle des cartes de rétro-diffusion des trois configurations du lanceur MX-14

ANNEXE D

DÉMONSTRATIONS DES FORMULES PROPOSÉES DANS LE DOCUMENT

D.1 Preuves Chapitre 1

▶ **Preuve réécriture de $\boldsymbol{H}$** (1.22)

Expression de $\sigma(\boldsymbol{k}_{m_x m_y m_z})$

Nous notons $\boldsymbol{k}_{m_x m_y m_z} = k_x^{[m_x]}\hat{\boldsymbol{x}} + k_y^{[m_y]}\hat{\boldsymbol{y}} + k_z^{[m_z]}\hat{\boldsymbol{z}}$ le vecteur d'onde associé au point $\left[k_x^{[m_x]}, k_y^{[m_y]}, k_z^{[m_z]}\right]$.

La SER complexe $\sigma(\boldsymbol{k}_{m_x m_y m_z})$ associée au point $\left[k_x^{[m_x]}, k_y^{[m_y]}, k_z^{[m_z]}\right]$ de la grille du k-space est alors donnée par :

$$\begin{aligned}\sigma(\boldsymbol{k}_{m_x m_y m_z}) &= \sum_{n_x=0}^{N_x-1}\sum_{n_y=0}^{N_y-1}\sum_{n_z=0}^{N_z-1} a(n_x,n_y,n_z)\exp\left\{-2\jmath\pi\boldsymbol{k}_{m_x m_y m_z}\cdot\boldsymbol{r}_{n_x n_y n_z}\right\} \\ &= \alpha\Psi\sum_{n_x=0}^{N_x-1}\sum_{n_y=0}^{N_y-1}\sum_{n_z=0}^{N_z-1}\left[a(n_x,n_y,n_z)\Phi\,\mathrm{e}^{-2\jmath\pi(\delta k_x\delta_x m_x n_x+\delta k_y\delta_y m_y n_y+\delta k_z\delta_z m_z n_z)}\right]\end{aligned} \tag{D.1}$$

avec :

- $\alpha = \exp\left\{-2\jmath\pi(k_x^{[0]}x^{[0]} + k_y^{[0]}y^{[0]} + k_z^{[0]}z^{[0]})\right\}$
- $\Psi = \exp\left\{-2\jmath\pi(x^{[0]}\delta k_x m_x + y^{[0]}\delta k_y m_y + z^{[0]}\delta k_z m_z)\right\}$
- $\Phi = \exp\left\{-2\jmath\pi(k_x^{[0]}\delta_x n_x + \pi k_y^{[0]}\delta_y n_y + k_z^{[0]}\delta_z n_z)\right\}$

Conditions de non repliement

Par définition, la TFDI de l'objet d'intérêt a est périodique dans le domaine spatiale (selon les trois directions spatiales). Par ailleurs, l'échantillonnage du k-space conditionne la dimension du champ de vu dans le domaine spatial comme illustré Figure D.1. Ainsi, si le champ de

vue ne contient pas l'objet d'intérêt dans sa totalité, il y a recouvrement spatial (ou *aliasing* en anglais) [Nyquist, 1928; Shannon, 1949]. En lien avec le théorème de Nyquist-Shannon, [Mezrich, 1995] et [Mahafza et Elsherbeni, 2004] définissent les conditions limites (D.2) pour qu'il n'y ait pas de recouvrement spatial.

$$\begin{cases} \delta k_x \delta_x \leqslant 1/(N_x - 1) \\ \delta k_y \delta_y \leqslant 1/(N_y - 1) \\ \delta k_z \delta_z \leqslant 1/(N_z - 1) \end{cases} \tag{D.2}$$

A pas d'échantillonnage du k-space donnés, afin de bénéficier des plus grandes extensions spatiales possibles (selon les trois directions) et pour faire apparaitre une TFD 3D dans (D.1) nous posons :

$$\begin{cases} \delta k_x \delta_x = 1/N_x \\ \delta k_y \delta_y = 1/N_y \\ \delta k_z \delta_z = 1/N_z \end{cases} \tag{D.3}$$

qui respecte bien les conditions définies dans (D.2).

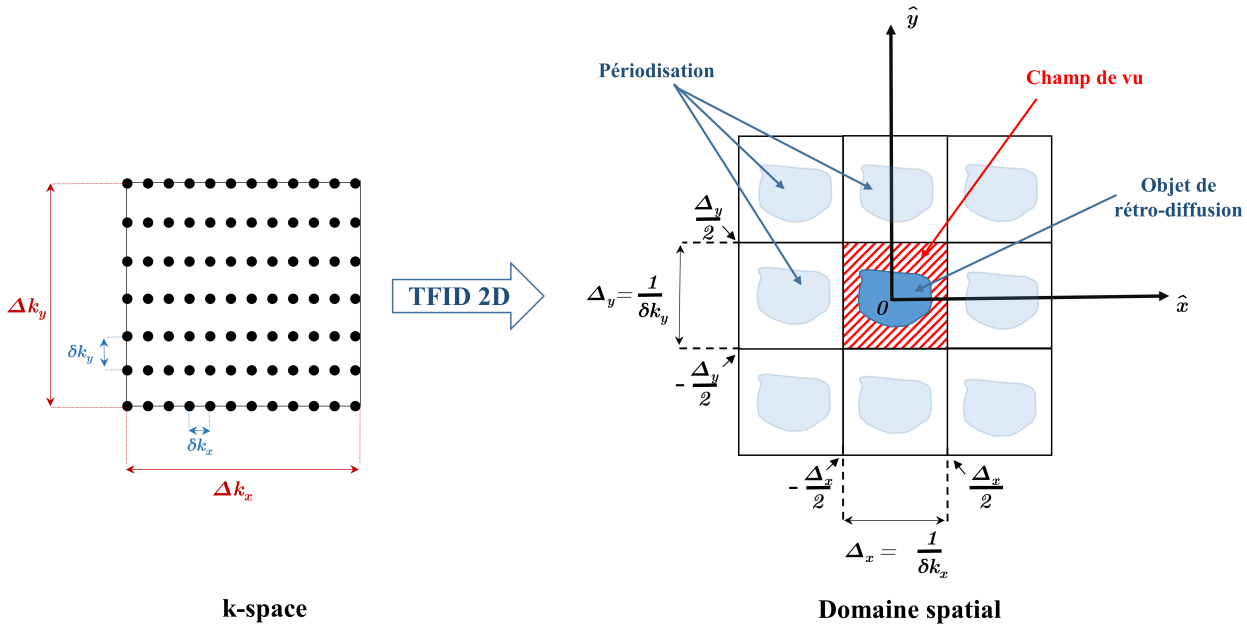


FIGURE D.1 – Exemple 2D de la périodisation de l'objet de rétro-diffusion dans le domaine spatial induite par l'échantillonnage du k-space.

Dès lors, (D.1) devient :

$$\sigma(\boldsymbol{k}_{m_x m_y m_z}) = \alpha\Psi \sum_{n_x=0}^{N_x-1} \sum_{n_y=0}^{N_y-1} \sum_{n_z=0}^{N_z-1} \left[a(n_x, n_y, n_z)\Phi \exp\left\{ -2\jmath\pi \left(\frac{m_x n_x}{N_x} + \frac{m_y n_y}{N_y} + \frac{m_z n_z}{N_z} \right) \right\} \right] \tag{D.4}$$

où nous reconnaissons une TFD 3D de l'objet de rétro-diffusion a.

Mise en forme matricielle

Nous pouvons alors réécrire la relation linéaire (D.4) sous la forme matricielle suivante :

$$\boxed{\boldsymbol{H} = \alpha\sqrt{N}\,\boldsymbol{\Delta}_{\Psi}\boldsymbol{S}\boldsymbol{F}_{3D}\boldsymbol{\Delta}_{\Phi}} \tag{D.5}$$

avec

- $\alpha = \exp\left\{-2\jmath\pi(k_x^{[0]}x^{[0]} + k_y^{[0]}y^{[0]} + k_z^{[0]}z^{[0]})\right\}$ est un coefficient complexe exprimant le déphasage à l'origine des grilles du k-space et spatiale.
- $N = N_xN_yN_z$ est le nombre d'éléments du vecteur $\boldsymbol{a}$ c'est-à-dire le nombre de voxels de la carte.
- $\boldsymbol{S}$ est une matrice de sélection binaire qui permet de sélectionner uniquement les coefficients de rétro-diffusion associés aux points de la grille du k-space affectés lors de l'étape de regridding.
- $\boldsymbol{F}_{3D} = \boldsymbol{F}_{N_x} \otimes \boldsymbol{F}_{N_y} \otimes \boldsymbol{F}_{N_z}$ est la matrice de TFD 3D [1] où

$$\boldsymbol{F}_{\aleph} = \frac{1}{\sqrt{\aleph}}\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{\aleph-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(\aleph-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{\aleph-1} & \omega^{2(\aleph-1)} & \cdots & \omega^{(\aleph-1)(\aleph-1)} \end{pmatrix}$$

  est la matrice de TFD normalisée et

$$\omega = \mathrm{e}^{-2\jmath\pi/\aleph}$$

  la racine $\aleph^{\text{ième}}$ de l'unité.
- $\boldsymbol{\Delta}_{\Psi}$ est une matrice diagonale de déphasage dans le k-space. Le $(m_xm_ym_z)^{\text{ième}}$ élément de sa diagonale est $\Psi = \exp\left\{-2\jmath\pi(x^{[0]}\delta k_x m_x + y^{[0]}\delta k_y m_y + z^{[0]}\delta k_z m_z)\right\}$.
- $\boldsymbol{\Delta}_{\Phi} = \boldsymbol{\Delta}_{\Phi_y} \otimes \boldsymbol{\Delta}_{\Phi_x} \otimes \boldsymbol{\Delta}_{\Phi_z}$ est une matrice diagonale de déphasage dans le domaine spatial où :
  - $\boldsymbol{\Delta}_{\Phi_x} = \mathrm{diag}\left\{\exp\left\{-2\jmath\pi k_x^{[0]}\delta_x n_x\right\},\ n_x = 0\cdots N_x - 1\right\}$
  - $\boldsymbol{\Delta}_{\Phi_y} = \mathrm{diag}\left\{\exp\left\{-2\jmath\pi k_y^{[0]}\delta_y n_y\right\},\ n_y = 0\cdots N_y - 1\right\}$
  - $\boldsymbol{\Delta}_{\Phi_z} = \mathrm{diag}\left\{\exp\left\{-2\jmath\pi k_z^{[0]}\delta_z n_z\right\},\ n_z = 0\cdots N_z - 1\right\}$

1. $\otimes$ est le produit de Kronecker

► Preuve reformatage données

Nous supposons que les données sont reformatées avec le regridding plus proche voisin lacunaire et que plusieurs données ont été affectées à un même point de la grille régulière du k-space. Nous montrons que ce point prend la moyenne des données affectées. Nous ajoutons qu'il s'agit d'une approximation permettant de mener les calculs par TFD. En effet, rigoureusement, il faudrait pondérer par le nombre de points affectés pour tenir compte de la variance du bruit plus faible pour les données reformatées.

Nous pouvons noter :

$$\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 = \sum_{\substack{l=1 \\ l/\{p_j\}}}^{N_{\text{obs}}} (\sigma_l - \boldsymbol{h}_l\boldsymbol{a})^\dagger(\sigma_l - \boldsymbol{h}_l\boldsymbol{a}) + \sum_{j=1}^{N_{\text{a}}} (\sigma_{p_j} - \boldsymbol{h}_{p_j}\boldsymbol{a})^\dagger(\sigma_{p_j} - \boldsymbol{h}_{p_j}\boldsymbol{a}) \tag{D.6}$$

où $\boldsymbol{h}_l$ sont les lignes de la matrice $\boldsymbol{H}$, N_{a} le nombre de données affectées à un même point de la grille régulière du k-space. Les lignes $\boldsymbol{h}_{p_j}$ sont donc des lignes semblables dans la matrice $\boldsymbol{H}$. Pour tout j, nous notons $\boldsymbol{h}_{p_j} = \boldsymbol{h}$.

L'expression (D.6) s'écrit donc :

$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 &= \sum_{\substack{l=1 \\ l/\{p_j\}}}^{N_{\text{obs}}} \left[(\sigma_l - \boldsymbol{h}_l\boldsymbol{a})^\dagger(\sigma_l - \boldsymbol{h}_l\boldsymbol{a})\right] + \sum_{j=1}^{N_{\text{a}}} \left[(\sigma_{p_j} - \boldsymbol{h}\boldsymbol{a})^\dagger(\sigma_{p_j} - \boldsymbol{h}\boldsymbol{a})\right] \\
&= \sum_{\substack{l=1 \\ l/\{p_j\}}}^{N_{\text{obs}}} \left[(\sigma_l - \boldsymbol{h}_l\boldsymbol{a})^\dagger(\sigma_l - \boldsymbol{h}_l\boldsymbol{a})\right] + \sum_{j=1}^{N_{\text{a}}} \left[\sigma_{p_j}^\dagger\sigma_{p_j} - 2\,\text{Re}\left\{\sigma_{p_j}^\dagger \boldsymbol{h}\boldsymbol{a}\right\} + \boldsymbol{a}^\dagger\boldsymbol{h}^\dagger\boldsymbol{h}\boldsymbol{a}\right] \\
&= \sum_{\substack{l=1 \\ l/\{p_j\}}}^{N_{\text{obs}}} \left[(\sigma_l - \boldsymbol{h}_l\boldsymbol{a})^\dagger(\sigma_l - \boldsymbol{h}_l\boldsymbol{a})\right] + \sum_{j=1}^{N_{\text{a}}} \left[\sigma_{p_j}^\dagger\sigma_{p_j}\right] - 2\,\text{Re}\sum_{j=1}^{N_{\text{a}}}\left\{\sigma_{p_j}^\dagger \boldsymbol{h}\boldsymbol{a}\right\} + N_{\text{a}}\boldsymbol{a}^\dagger\boldsymbol{h}^\dagger\boldsymbol{h}\boldsymbol{a}
\end{aligned} \tag{D.7}$$

Nous notons $\bar{\sigma}$ la moyenne des σ_{p_j} qui s'écrit :

$$\bar{\sigma} = \frac{1}{N_{\text{a}}}\sum_{j=1}^{N_{\text{a}}} \sigma_{p_j} \tag{D.8}$$

Et il vient donc :

$$\bar{\sigma}^\dagger\bar{\sigma} = \frac{1}{N_{\text{a}}^2}\sum_{j=1}^{N_{\text{a}}} \sigma_{p_j}^\dagger\sigma_{p_j} + \beta \tag{D.9}$$

où $\beta = 2\,\text{Re}\left\{\sigma_{p_1}^\dagger(\sigma_{p_2} + \cdots + \sigma_{p_{N_{\text{a}}}}) + \sigma_{p_2}^\dagger(\sigma_{p_3} + \cdots + \sigma_{p_{N_{\text{a}}}}) + \cdots + \sigma_{p_{N_{\text{a}}-1}}^\dagger\sigma_{p_{N_{\text{a}}}}\right\}$
Par ailleurs, nous pouvons ré-écrire (D.7) en y introduisant (D.8) :

$$\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 = \sum_{\substack{l=1 \\ l/\{p_j\}}}^{N_{\text{obs}}} \left[(\sigma_l - \boldsymbol{h}_l\boldsymbol{a})^\dagger(\sigma_l - \boldsymbol{h}_l\boldsymbol{a})\right] + \sum_{j=1}^{N_{\text{a}}} \left[\sigma_{p_j}^\dagger\sigma_{p_j}\right] - 2N_{\text{a}}\,\text{Re}\left\{\bar{\sigma}_{p_j}^\dagger \boldsymbol{h}\boldsymbol{a}\right\} + N_{\text{a}}\boldsymbol{a}^\dagger\boldsymbol{h}^\dagger\boldsymbol{h}\boldsymbol{a} \tag{D.10}$$

Par ailleurs, nous avons :

$$\begin{aligned}\|\boldsymbol{\sigma}-\boldsymbol{H}\boldsymbol{a}\|_2^2 &= \sum_{\substack{l=1\\ l/\{p_j\}}}^{N_{\text{obs}}}\left[(\sigma_l-\boldsymbol{h}_l\boldsymbol{a})^\dagger(\sigma_l-\boldsymbol{h}_l\boldsymbol{a})\right]+N_{\text{a}}^2(\bar{\sigma}^\dagger\bar{\sigma}-\beta)-2N_{\text{a}}\,\text{Re}\left\{\bar{\sigma}_{p_j}^\dagger\boldsymbol{h}\boldsymbol{a}\right\}+N_{\text{a}}\boldsymbol{a}^\dagger\boldsymbol{h}^\dagger\boldsymbol{h}\boldsymbol{a}\\ &= \sum_{\substack{l=1\\ l/\{p_j\}}}^{N_{\text{obs}}}\left[(\sigma_l-\boldsymbol{h}_l\boldsymbol{a})^\dagger(\sigma_l-\boldsymbol{h}_l\boldsymbol{a})\right]\\ &\qquad +N_{\text{a}}\Big[\bar{\sigma}^\dagger\bar{\sigma}-2\,\text{Re}\left\{\bar{\sigma}_{p_j}^\dagger\boldsymbol{h}\boldsymbol{a}\right\}+\boldsymbol{a}^\dagger\boldsymbol{h}^\dagger\boldsymbol{h}\boldsymbol{a}+\bar{\sigma}^\dagger\bar{\sigma}(N_{\text{a}}-1)\Big]-N_{\text{a}}\beta\end{aligned} \tag{D.11}$$

En réarrangeant et en compactant (D.11) nous obtenons finalement :

$$\|\boldsymbol{\sigma}-\boldsymbol{H}\boldsymbol{a}\|_2^2=\sum_{\substack{l=1\\ l/\{p_j\}}}^{N_{\text{obs}}}\left[(\sigma_l-\boldsymbol{h}_l\boldsymbol{a})^\dagger(\sigma_l-\boldsymbol{h}_l\boldsymbol{a})\right]+N_{\text{a}}\left[(\bar{\sigma}-\boldsymbol{h}\boldsymbol{a})^\dagger(\bar{\sigma}-\boldsymbol{h}\boldsymbol{a})\right]+\beta' \tag{D.12}$$

où $\beta'=N_{\text{a}}(N_{\text{a}}-1)\bar{\sigma}^\dagger\bar{\sigma}-N_{\text{a}}\beta$

REMARQUE

Lors de la minimisation du terme $\|\boldsymbol{\sigma}-\boldsymbol{H}\boldsymbol{a}\|_2^2$ par rapport à $\boldsymbol{a}$, la constante β' n'intervient pas.

Nous avons donc montré que si plusieurs données sont affectées à un même point de la grille régulière du k-space, ce dernier a pour valeur la moyenne des données affectées. Rigoureusement, il faudrait également pondérer d'un facteur N_{a} puisque si le bruit est indépendant et identiquement distribué la variance sur $\bar{\sigma}$ est N_{a} fois plus faible que la variance sur σ. Néanmoins, pour pouvoir mener les calculs par TFD nous omettons cette pondération. Cette hypothèse se justifie en pratique car si la grille est suffisamment discrétisée le nombre de points affectés N_{a} tend vers 1.

REMARQUE

Une preuve par récurrence moins lourde est également envisageable.

► Preuve solution des moindres carrés à norme minimale (PFA) (1.32)

Nous rappelons que la solution au sens des moindres carrés à norme minimale s'écrit :

$$\hat{\boldsymbol{a}}=\boldsymbol{H}^\dagger(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\boldsymbol{\sigma} \tag{D.13}$$

et qu'après reformatage, la matrice d'observation $\boldsymbol{H}$ s'écrit :

$$\boldsymbol{H}=\alpha\sqrt{N}\,\boldsymbol{\Delta}_\Psi\boldsymbol{S}\boldsymbol{F}_{3D}\boldsymbol{\Delta}_\Phi \tag{D.14}$$

Ainsi, en remplaçant (D.14) dans (D.13) il vient :

$$\begin{aligned}\hat{\boldsymbol{a}} &= \alpha^*\sqrt{N}\,\boldsymbol{\Delta}_\Phi^*\boldsymbol{F}_{3D}^\dagger\boldsymbol{S}^{\mathrm{t}}\boldsymbol{\Delta}_\Psi^*\left[\left(\alpha\sqrt{N}\,\boldsymbol{\Delta}_\Psi\boldsymbol{S}\boldsymbol{F}_{3D}\boldsymbol{\Delta}_\Phi\right)\left(\alpha^*\sqrt{N}\,\boldsymbol{\Delta}_\Phi^*\boldsymbol{F}_{3D}^\dagger\boldsymbol{S}^{\mathrm{t}}\boldsymbol{\Delta}_\Psi^*\right)\right]^{-1}\breve{\boldsymbol{\sigma}} \\ &= \alpha^*\sqrt{N}\,\boldsymbol{\Delta}_\Phi^*\boldsymbol{F}_{3D}^\dagger\boldsymbol{S}^{\mathrm{t}}\boldsymbol{\Delta}_\Psi^*\left[\alpha\alpha^* N\boldsymbol{\Delta}_\Psi\boldsymbol{S}\boldsymbol{F}_{3D}\boldsymbol{\Delta}_\Phi\boldsymbol{\Delta}_\Phi^*\boldsymbol{F}_{3D}^\dagger\boldsymbol{S}^{\mathrm{t}}\boldsymbol{\Delta}_\Psi^*\right]^{-1}\breve{\boldsymbol{\sigma}} \\ &= \alpha^*\sqrt{N}\,\boldsymbol{\Delta}_\Phi^*\boldsymbol{F}_{3D}^\dagger\boldsymbol{S}^{\mathrm{t}}\boldsymbol{\Delta}_\Psi^*\left[N\boldsymbol{I}_{N_{\text{obs}}}\right]^{-1}\breve{\boldsymbol{\sigma}}\end{aligned} \tag{D.15}$$

car :

- $\alpha\alpha^* = |\alpha|^2 = 1$
- $\boldsymbol{\Delta}_\Phi\boldsymbol{\Delta}_\Phi^* = \boldsymbol{I}_N$
- $\boldsymbol{F}_{3D}\boldsymbol{F}_{3D}^\dagger = \boldsymbol{I}_N$
- $\boldsymbol{S}\boldsymbol{S}^{\mathrm{t}} = \boldsymbol{I}_{N_{\text{obs}}}$
- $\boldsymbol{\Delta}_\Psi\boldsymbol{\Delta}_\Psi^* = \boldsymbol{I}_{N_{\text{obs}}}$

et finalement il vient :

$$\boxed{\hat{\boldsymbol{a}} = \alpha^*\frac{1}{\sqrt{N}}\boldsymbol{\Delta}_\Phi^*\boldsymbol{F}_{3D}^\dagger\boldsymbol{S}^t\boldsymbol{\Delta}_\Psi^*\breve{\boldsymbol{\sigma}}} \tag{D.16}$$

D.2 Preuves Chapitre 2

► Preuve fonction duale (2.16)

En partant de la définition de la fonction duale et en se servant de la définition de la conjuguée convexe **Annexe E** nous avons :

$$\begin{aligned}\Theta(\boldsymbol{\gamma}) &\triangleq \inf_{\boldsymbol{a}} \mathcal{L}(\boldsymbol{a},\boldsymbol{\gamma})\\ &= \inf_{\boldsymbol{a}}\big(f(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big)\\ &= -\sup_{\boldsymbol{a}}\big(-f(\boldsymbol{a}) - \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{a} + \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{c}\big)\\ &= -\sup_{\boldsymbol{a}}\big(-(\boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma})^{\mathrm{t}}\boldsymbol{a} - f(\boldsymbol{a})\big) - \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{c}\\ &= -f^{\star}(-\boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma}) - \boldsymbol{c}^{\mathrm{t}}\boldsymbol{\gamma}\end{aligned}$$

► Preuve concavité de la fonction duale (2.16)

Nous voulons montrer que $\Theta(\boldsymbol{\gamma}) = \inf_{\boldsymbol{a}}\big(f(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big)$ est concave sur $\mathcal{U}$.

Soit $\theta \in [0,1],\ \forall(\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2) \in (\mathcal{U}\times\mathcal{U})$:

$$\begin{aligned}\Theta\big(\theta\boldsymbol{\gamma}_1 + (1-\theta)\boldsymbol{\gamma}_2\big) &= \inf_{\boldsymbol{a}}\big(f(\boldsymbol{a}) + (\theta\boldsymbol{\gamma}_1 + (1-\theta)\boldsymbol{\gamma}_2)^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big)\\ &= \inf_{\boldsymbol{a}}\big(\theta\big[f(\boldsymbol{a}) + \boldsymbol{\gamma}_1{}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big] + (1-\theta)\big[f(\boldsymbol{a}) + \boldsymbol{\gamma}_2{}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big]\big)\\ \Rightarrow \Theta\big(\theta\boldsymbol{\gamma}_1 + (1-\theta)\boldsymbol{\gamma}_2\big) &\geqslant \inf_{\boldsymbol{a}}\big(\theta\big[f(\boldsymbol{a}) + \boldsymbol{\gamma}_1{}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big]\big) + \inf_{\boldsymbol{a}}\big((1-\theta)\big[f(\boldsymbol{a}) + \boldsymbol{\gamma}_2{}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big]\big)\\ &\geqslant \theta\inf_{\boldsymbol{a}}\big(f(\boldsymbol{a}) + \boldsymbol{\gamma}_1{}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big) + (1-\theta)\inf_{\boldsymbol{a}}\big(f(\boldsymbol{a}) + \boldsymbol{\gamma}_2{}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{c})\big)\\ &\geqslant \theta\Theta(\boldsymbol{\gamma}_1) + (1-\theta)\Theta(\boldsymbol{\gamma}_2)\end{aligned}$$

Θ est donc concave sur $\mathcal{U}$.

► Preuve Min-Max inégalité (E.3)

Soient $\mathbb{W} \subseteq \mathbb{R}^n$ et $\mathbb{Z} \subseteq \mathbb{R}^m$ deux sous-ensembles et soient $g : \mathbb{Z} \to \mathbb{R}$ et $f : \mathbb{W}\times\mathbb{Z} \to \mathbb{R}$ deux fonctions telles que :

$$g(\boldsymbol{z}) \triangleq \inf_{\boldsymbol{w}\in\mathbb{W}} f(\boldsymbol{w},\boldsymbol{z})$$

Nous avons alors, $\forall(\boldsymbol{w},\boldsymbol{z}) \in \mathbb{W}\times\mathbb{Z}$:

$$\begin{aligned} g(\boldsymbol{z}) &\leqslant f(\boldsymbol{w},\boldsymbol{z}) \\ \Rightarrow \sup_{\boldsymbol{z}\in\mathbb{Z}} g(\boldsymbol{z}) &\leqslant \sup_{\boldsymbol{z}\in\mathbb{Z}} f(\boldsymbol{w},\boldsymbol{z}) \\ \Rightarrow \sup_{\boldsymbol{z}\in\mathbb{Z}} \inf_{\boldsymbol{w}\in\mathbb{W}} f(\boldsymbol{w},\boldsymbol{z}) &\leqslant \sup_{\boldsymbol{z}\in\mathbb{Z}} f(\boldsymbol{w},\boldsymbol{z}) \end{aligned}$$

Et finalement il vient bien :

$$\boxed{\sup_{\boldsymbol{z}\in\mathbb{Z}} \inf_{\boldsymbol{w}\in\mathbb{W}} f(\boldsymbol{w},\boldsymbol{z}) \leqslant \inf_{\boldsymbol{w}\in\mathbb{W}} \sup_{\boldsymbol{z}\in\mathbb{Z}} f(\boldsymbol{w},\boldsymbol{z})}$$

► **Preuve résidu contrainte égalité** (2.23)

Cette preuve repose sur l'utilisation de $f^\star$ la conjuguée convexe de f rappelée **Annexe E**.

Soit f une fonction strictement convexe et dérivable. Nous avons :

$$\begin{aligned} \widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}} &= \underset{\boldsymbol{a}}{\operatorname{argmin}}(\boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{a} - f(\boldsymbol{a})) \\ &\Leftrightarrow \underset{\boldsymbol{a}}{\nabla}\left(\boldsymbol{\gamma}^{\mathrm{t}}\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}\right) - \underset{\boldsymbol{a}}{\nabla} f(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}) = 0 \\ &\Leftrightarrow \boldsymbol{\gamma} = \underset{\boldsymbol{a}}{\nabla} f(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}) \end{aligned}$$

Tout d'abord, montrons que :

$$f \text{ strictement convexe et dérivable} \Rightarrow \underset{\boldsymbol{\gamma}}{\nabla} f^\star(\boldsymbol{\gamma}) = \underset{\boldsymbol{a}}{\operatorname{argmax}}\left(\boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{a} - f(\boldsymbol{a})\right) = \widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}$$

Autrement dit, il faut prouver que :

$$f^\star(\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}) - f^\star(\boldsymbol{\gamma}) = \Delta\boldsymbol{\gamma}^{\mathrm{t}}\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}} + \underset{\|\Delta\boldsymbol{\gamma}\|\to 0}{o}\|\Delta\boldsymbol{\gamma}\|$$

$$\begin{aligned} f^\star(\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}) - f^\star(\boldsymbol{\gamma}) &= \sup_{\boldsymbol{a}}\left((\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma})^{\mathrm{t}}\boldsymbol{a} - f(\boldsymbol{a})\right) - \sup_{\boldsymbol{a}}\left(\boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{a} - f(\boldsymbol{a})\right) \\ &= (\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma})^{\mathrm{t}}\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}} - f(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}}) - \boldsymbol{\gamma}^{\mathrm{t}}\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}} + f(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}) \\ &= \boldsymbol{\gamma}^{\mathrm{t}}(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}} + \widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}) + \Delta\boldsymbol{\gamma}^{\mathrm{t}}\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}} + \left(f(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}) - f(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}})\right) \end{aligned}$$

Or

$$\begin{aligned} f(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}) - f(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}}) &= \left(\underset{\boldsymbol{a}}{\nabla} f^\star(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}})\right)^{\mathrm{t}}(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}}) + \underset{\|\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}-\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}}\|\to 0}{o}\|\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}}\| \\ &= (\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma})^{\mathrm{t}}(\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}}) + \underset{\|\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}}-\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}}\|\to 0}{o}\|\widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{a}}_{\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}}\| \end{aligned}$$

Et en développant dans l'expression précédente il vient bien :

$$f^\star(\boldsymbol{\gamma}+\Delta\boldsymbol{\gamma}) - f^\star(\boldsymbol{\gamma}) = \Delta\boldsymbol{\gamma}^{\mathrm{t}}\widetilde{\boldsymbol{a}}_\gamma + \underset{\|\Delta\boldsymbol{\gamma}\|\to 0}{\mathcal{O}}\|\Delta\boldsymbol{\gamma}\|$$

Désormais, montrons que l'expression du résidu pour la contrainte d'égalité s'écrit :

$$\boxed{\underset{\boldsymbol{\gamma}}{\nabla}\Theta(\boldsymbol{\gamma}) = \boldsymbol{A}\boldsymbol{a}^+ - \boldsymbol{c}}$$

Soit :

$$h(\boldsymbol{\gamma}) = -\boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma} \Rightarrow \underset{\boldsymbol{\gamma}}{\nabla} h(\boldsymbol{\gamma}) = -\boldsymbol{A}$$

Alors :

$$\begin{aligned}
\underset{\boldsymbol{\gamma}}{\nabla}\Theta(\boldsymbol{\gamma}) &= \underset{\boldsymbol{\gamma}}{\nabla}\inf_{\boldsymbol{a}}\mathcal{L}(\boldsymbol{a},\boldsymbol{\gamma}) \\
&= \underset{\boldsymbol{\gamma}}{\nabla}\big(-f^\star(-\boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma}) - \boldsymbol{c}^{\mathrm{t}}\boldsymbol{\gamma}\big) \\
&= \underset{\boldsymbol{\gamma}}{\nabla}\big(-f^\star\circ h(\boldsymbol{\gamma})\big) - \boldsymbol{c} \\
&= -\underset{\boldsymbol{\gamma}}{\nabla}h(\boldsymbol{\gamma})\underset{h(\boldsymbol{\gamma})}{\nabla}f^\star(h(\boldsymbol{\gamma})) - \boldsymbol{c} \\
&= \boldsymbol{A}\operatorname*{argmax}_{\boldsymbol{a}}\big(-\boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{a} - f(\boldsymbol{a})\big) - \boldsymbol{c} \\
&= \boldsymbol{A}\operatorname*{argmin}_{\boldsymbol{a}}\big(f(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{a}\big) - \boldsymbol{c} \\
&= \boldsymbol{A}\operatorname*{argmin}_{\boldsymbol{a}}\big(f(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c})\big) - \boldsymbol{c} \\
&= \boldsymbol{A}\operatorname*{argmin}_{\boldsymbol{a}}\mathcal{L}(\boldsymbol{a},\boldsymbol{\gamma}) - \boldsymbol{c}
\end{aligned}$$

et en posant $\boldsymbol{a}^+ = \operatorname*{argmin}_{\boldsymbol{a}}\mathcal{L}(\boldsymbol{a},\boldsymbol{\gamma})$ la preuve est achevée.

► Mise à jour de a dans l'algorithme ADL dans le cas d'une fonction objectif f quadratique strictement convexe (2.25)

Nous écrivons le lagrangien :

$$\begin{aligned}
\mathcal{L}(\boldsymbol{a},\boldsymbol{\gamma}) &= f(\boldsymbol{a}) + \boldsymbol{\gamma}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c}) \\
&= \boldsymbol{a}^{\mathrm{t}}\boldsymbol{P}\boldsymbol{a} + \boldsymbol{q}^{\mathrm{t}}\boldsymbol{a} + r + \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{a} - \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{c} \\
&= \boldsymbol{a}^{\mathrm{t}}\boldsymbol{P}\boldsymbol{a} + (\boldsymbol{q}^{\mathrm{t}} + \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{A})\boldsymbol{a} + \underbrace{r - \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{c}}_{=\mathrm{cste}_{\boldsymbol{a}}}
\end{aligned}$$

$\mathcal{L}(\boldsymbol{a},\boldsymbol{\gamma})$ est donc une fonction quadratique en $\boldsymbol{a}$, strictement convexe (car $\boldsymbol{P}$ est définie positive). Nous annulons donc le gradient pour trouver le minimiseur $\widetilde{\boldsymbol{a}}$:

$$\begin{aligned}
\underset{\boldsymbol{a}}{\nabla}\mathcal{L}(\boldsymbol{a},\boldsymbol{\gamma}) &= 0 \\
\Rightarrow 2\boldsymbol{P}\widetilde{\boldsymbol{a}} + (\boldsymbol{q}^{\mathrm{t}} + \boldsymbol{\gamma}^{\mathrm{t}}\boldsymbol{A})^{\mathrm{t}} &= 0 \\
\Rightarrow \widetilde{\boldsymbol{a}} = \frac{1}{2}\boldsymbol{P}^{-1}(-\boldsymbol{q} - \boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma})
\end{aligned}$$

la mise à jour de $\boldsymbol{a}$ dans l'**Algorithme 5** s'écrit donc :

$$\boxed{\boldsymbol{a}^{(k+1)} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\, \mathcal{L}(\boldsymbol{a}, \boldsymbol{\gamma}^{(k)}) = \frac{1}{2}\boldsymbol{P}^{-1}(-\boldsymbol{q} - \boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma}^{(k)})}$$

► **Minimiseurs du problème primal et dual théoriques pour un objectif quadratique** (2.26)

La fonction objectif f est strictement convexe et il s'agit d'une contrainte d'égalité affine donc d'après les conditions KKT et le théorème de Slater la dualité est forte et il existe un couple $(\widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{\gamma}})$ qui est point selle du lagrangien.

Nous avons :

$$\widetilde{\boldsymbol{a}} = \frac{1}{2}\boldsymbol{P}^{-1}(-\boldsymbol{q} - \boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}})$$
$$\text{et } \Theta(\widetilde{\boldsymbol{\gamma}}) = f(\widetilde{\boldsymbol{a}}) + \widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}(\boldsymbol{A}\widetilde{\boldsymbol{a}} - \boldsymbol{c})$$

Dans un premier temps nous calculons $f(\widetilde{\boldsymbol{a}})$:

$$\begin{aligned} f(\widetilde{\boldsymbol{a}}) &= \frac{1}{4}\left[\left(\boldsymbol{P}^{-1}(-\boldsymbol{q} - \boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}})\right)^{\mathrm{t}}\boldsymbol{P}\left(\boldsymbol{P}^{-1}(-\boldsymbol{q} - \boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}})\right)\right] + \frac{1}{2}\boldsymbol{q}^{\mathrm{t}}\boldsymbol{P}^{-1}(-\boldsymbol{q} - \boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma}) + r \\ &= \frac{1}{4}\widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma} - \frac{1}{4}\boldsymbol{q}^{\mathrm{t}}\boldsymbol{P}^{-1}\boldsymbol{q} \end{aligned}$$

Nous calculons alors le gradient de f par rapport à la variable $\boldsymbol{\gamma}$:

$$\nabla_{\boldsymbol{\gamma}} f(\widetilde{\boldsymbol{a}}) = \frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}}$$

Nous calculons dans un second temps :

$$\begin{aligned} \widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}(\boldsymbol{A}\widetilde{\boldsymbol{a}} - \boldsymbol{c}) &= \widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}(\frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}(-\boldsymbol{q} - \boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}}) - \boldsymbol{c}) \\ &= -\frac{1}{2}\widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}} - \frac{1}{2}\widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{q} - \widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}\boldsymbol{c} \end{aligned}$$

En calculant cette fois le gradient de $\left(\widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}(\boldsymbol{A}\widetilde{\boldsymbol{a}} - \boldsymbol{c})\right)$ par rapport à $\boldsymbol{\gamma}$ il vient :

$$\nabla_{\boldsymbol{\gamma}}\left(\widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}(\boldsymbol{A}\widetilde{\boldsymbol{a}} - \boldsymbol{c})\right) = -\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}} - \frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{q} - \boldsymbol{c}$$

Ainsi, en reportant dans l'expression du gradient de Θ par rapport à $\boldsymbol{\gamma}$, nous avons :

$$\begin{aligned} \nabla_{\boldsymbol{\gamma}}\Theta(\widetilde{\boldsymbol{\gamma}}) &= \frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}} - \boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}} - \frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{q} - \boldsymbol{c} \\ &= -\frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\widetilde{\boldsymbol{\gamma}} - \frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{q} - \boldsymbol{c} \end{aligned}$$

Or, comme Θ est strictement concave (car f est strictement convexe), $\widetilde{\boldsymbol{\gamma}}$ annule son gradient :

$$\nabla_{\boldsymbol{\gamma}}\Theta(\widetilde{\boldsymbol{\gamma}}) = 0 \Longrightarrow \boxed{\widetilde{\boldsymbol{\gamma}} = -2\Big(\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}}\Big)^{-1}\Big(\frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{q} + \boldsymbol{c}\Big)}$$

Et finalement, en reportant dans l'expression de $\widetilde{\boldsymbol{a}}$ nous avons :

$$\boxed{\widetilde{\boldsymbol{a}} = \frac{1}{2}\boldsymbol{P}^{-1}\Big[-\boldsymbol{q} + 2\boldsymbol{A}^{\mathrm{t}}\Big((\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{A}^{\mathrm{t}})^{-1}(\frac{1}{2}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{q} + \boldsymbol{c})\Big)\Big]}$$

Ce qui conclut la preuve.

► **Convergence de la méthode des multiplicateurs de Lagrange** **Algorithme 7**

Nous montrons tout d'abord que :

$$\boxed{\widetilde{p} - p^{(k+1)} \leqslant \widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})} \tag{D.17}$$

Par hypothèse f est convexe, fermée et propre et nous considérons des contraintes d'égalité affines. D'après le théorème de Slater et les conditions KKT, la dualité est forte et il existe un couple $(\widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{\gamma}})$ point selle de $\mathcal{L}_0$. Par définition du point selle Figure E, nous avons :

$$\begin{aligned}
&\mathcal{L}_0(\widetilde{\boldsymbol{a}}, \widetilde{\boldsymbol{\gamma}}) \leqslant \mathcal{L}_0(\boldsymbol{a}^{(k+1)}, \widetilde{\boldsymbol{\gamma}}) \\
\Rightarrow \underbrace{f(\widetilde{\boldsymbol{a}})}_{\widetilde{p}} + \widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}\underbrace{(\boldsymbol{A}\widetilde{\boldsymbol{a}} - \boldsymbol{c})}_{0} &\leqslant \underbrace{f(\boldsymbol{a}^{(k+1)})}_{p^{(k+1)}} + \widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c}) \\
\Rightarrow \widetilde{p} - p^{(k+1)} &\leqslant \widetilde{\boldsymbol{\gamma}}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})
\end{aligned}$$

Nous montrons ensuite que :

$$\boxed{p^{(k+1)} - \widetilde{p} \leqslant -(\boldsymbol{\gamma}^{(k+1)})^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})}$$

Soit $f \in \mathrm{Conv}(\mathrm{dom}\, f)$. Par définition $\boldsymbol{a}^{(k+1)}$ minimise $\mathcal{L}_\rho(\boldsymbol{a}, \boldsymbol{\gamma}^{(k)})$. Nous obtenons alors la condition KKT de stationnarité (ie. condition d'optimalité au premier ordre) :

$$\begin{aligned}
0 &\in \partial_{\boldsymbol{a}}\mathcal{L}_\rho(\boldsymbol{a}^{(k+1)}, \boldsymbol{\gamma}^{(k)}) \\
0 &\in \partial_{\boldsymbol{a}} f(\boldsymbol{a}^{(k+1)}) + \boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma}^{(k)} + \rho\boldsymbol{A}^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})
\end{aligned}$$

or $\boldsymbol{\gamma}^{(k+1)} = \boldsymbol{\gamma}^{(k)} + \rho(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})$, donc

$$0 \in \underset{\boldsymbol{a}}{\partial}\mathcal{L}_\rho(\boldsymbol{a}^{(k+1)}, \boldsymbol{\gamma}^{(k)}) \Rightarrow 0 \in f(\boldsymbol{a}^{(k+1)}) + \boldsymbol{A}^{\mathrm{t}}\boldsymbol{\gamma}^{(k+1)}$$

donc $\forall \boldsymbol{a}, \boldsymbol{a}^{(k+1)}$ minimise $f(\boldsymbol{a}) + (\boldsymbol{\gamma}^{(k+1)})^{\mathrm{t}}\boldsymbol{A}\boldsymbol{a}$ et donc aussi $f(\boldsymbol{a}) + (\boldsymbol{\gamma}^{(k+1)})^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a} - \boldsymbol{c})$. Ainsi :

$$\underbrace{f(\boldsymbol{a}^{(k+1)})}_{p^{(k+1)}} + (\boldsymbol{\gamma}^{(k+1)})^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c}) \leqslant \underbrace{f(\widetilde{\boldsymbol{a}})}_{\widetilde{p}} + (\boldsymbol{\gamma}^{(k+1)})^{\mathrm{t}}\underbrace{(\boldsymbol{A}\widetilde{\boldsymbol{a}} - \boldsymbol{c})}_{0}$$

$$\Rightarrow p^{(k+1)} - \widetilde{p} \leqslant -(\boldsymbol{\gamma}^{(k+1)})^{\mathrm{t}}(\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c})$$

Lorsque $k \to \infty$, la contrainte est respectée et donc le résidu pour la contrainte d'égalité $\boldsymbol{r}^{(k+1)} = (\boldsymbol{A}\boldsymbol{a}^{(k+1)} - \boldsymbol{c}) \to 0$. Ainsi, d'après les deux inégalités il vient :
$\widetilde{p} - p^{(k+1)} \leqslant 0$ et $p^{(k+1)} - \widetilde{p} \leqslant 0 \Rightarrow p^{(k+1)} = \widetilde{p}$, ce qui prouve la convergence.

D.3 Preuves Chapitre 3

► Mises-à-jour de $\boldsymbol{a}$, $\boldsymbol{v}_P$, $\boldsymbol{v}_x$, $\boldsymbol{v}_y$et $\boldsymbol{v}_\mathcal{C}$ Section 3.3.2

Dans la suite, pour travailler avec des variables complexes nous utilisons la définition du gradient d'une fonction à variables complexes et à valeur réelle donnée par [Haykin, 1996, eq. (8.8)] et [Petersen et Pedersen, 2012, eq. (232)] :

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = \frac{\partial f(\boldsymbol{x})}{\partial \mathrm{Re}\{\boldsymbol{x}\}} + \jmath \frac{\partial f(\boldsymbol{x})}{\partial \mathrm{Im}\{\boldsymbol{x}\}} \tag{D.18}$$

◊ Mise-à-jour de $\boldsymbol{a}$

Le lagrangien augmenté $\mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}$ (3.13) est une fonction quadratique de $\boldsymbol{a}$. Nous calculons son gradient par rapport à $\boldsymbol{a}$ (dans chaque étape nous éliminons les termes ne dépendant pas de $\boldsymbol{a}$ car leur gradient est nul) :

$$\begin{aligned}
\nabla_{\boldsymbol{a}} \mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}} = \nabla_{\boldsymbol{a}} \Bigg[& \frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + \frac{\nu}{2}\|\boldsymbol{a}\|_2^2 \\
& + \frac{\rho_P}{2}\|\boldsymbol{P}\boldsymbol{a} - \boldsymbol{v}_P + \boldsymbol{u}_P\|_2^2 + \frac{\rho_D}{2}\|\boldsymbol{D}_x\boldsymbol{a} - \boldsymbol{v}_x + \boldsymbol{u}_x\|_2^2 \\
& + \frac{\rho_D}{2}\|\boldsymbol{D}_y\boldsymbol{a} - \boldsymbol{v}_y + \boldsymbol{u}_y\|_2^2 + \frac{\rho_\mathcal{C}}{2}\|\boldsymbol{a} - \boldsymbol{v}_\mathcal{C} + \boldsymbol{u}_\mathcal{C}\|_2^2 \Bigg] \\
= \nabla_{\boldsymbol{a}} \Bigg[& \frac{1}{2}\left(\boldsymbol{a}^\dagger\boldsymbol{H}^\dagger\boldsymbol{H}\boldsymbol{a} - 2\mathrm{Re}\{\boldsymbol{\sigma}^\dagger\boldsymbol{H}\boldsymbol{a}\}\right) + \frac{\nu}{2}\left(\boldsymbol{a}^\dagger\boldsymbol{a}\right) \\
& + \frac{\rho_P}{2}\left(\boldsymbol{a}^\dagger\boldsymbol{P}^\dagger\boldsymbol{P}\boldsymbol{a} - 2\mathrm{Re}\left\{\boldsymbol{v}_P^\dagger\boldsymbol{P}\boldsymbol{a}\right\} + 2\mathrm{Re}\left\{\boldsymbol{u}_P^\dagger\boldsymbol{P}\boldsymbol{a}\right\}\right) \\
& + \frac{\rho_D}{2}\left(\boldsymbol{a}^\dagger\boldsymbol{D}_x^\dagger\boldsymbol{D}_x\boldsymbol{a} - 2\mathrm{Re}\left\{\boldsymbol{v}_x{}^\dagger\boldsymbol{D}_x\boldsymbol{a}\right\} + 2\mathrm{Re}\left\{\boldsymbol{u}_x{}^\dagger\boldsymbol{D}_x\boldsymbol{a}\right\}\right) \\
& + \frac{\rho_D}{2}\left(\boldsymbol{a}^\dagger\boldsymbol{D}_y^\dagger\boldsymbol{D}_y\boldsymbol{a} - 2\mathrm{Re}\left\{\boldsymbol{v}_y{}^\dagger\boldsymbol{D}_y\boldsymbol{a}\right\} + 2\mathrm{Re}\left\{\boldsymbol{u}_y{}^\dagger\boldsymbol{D}_y\boldsymbol{a}\right\}\right) \\
& + \frac{\rho_\mathcal{C}}{2}\left(\boldsymbol{a}^\dagger\boldsymbol{a} - 2\mathrm{Re}\left\{\boldsymbol{v}_\mathcal{C}{}^\dagger\boldsymbol{a}\right\} + 2\mathrm{Re}\left\{\boldsymbol{u}_\mathcal{C}^\dagger\boldsymbol{a}\right\}\right) \\
= & \ \boldsymbol{G}\boldsymbol{a} - \boldsymbol{d}
\end{aligned} \tag{D.19}$$

avec :
$\boldsymbol{G} = \boldsymbol{H}^\dagger\boldsymbol{H} + \rho_P\boldsymbol{P}^\dagger\boldsymbol{P} + \rho_D(\boldsymbol{D}_x^\dagger\boldsymbol{D}_x + \boldsymbol{D}_y^\dagger\boldsymbol{D}_y) + (\nu + \rho_\mathcal{C})\boldsymbol{I}_N$
$\boldsymbol{d} = \boldsymbol{H}^\dagger\boldsymbol{\sigma} + \rho_P\boldsymbol{P}^\dagger(\boldsymbol{v}_P - \boldsymbol{u}_P) + \rho_D\big(\boldsymbol{D}_x^\dagger(\boldsymbol{v}_x - \boldsymbol{u}_x) + \boldsymbol{D}_y^\dagger(\boldsymbol{v}_y - \boldsymbol{u}_y)\big) + \rho_\mathcal{C}(\boldsymbol{v}_\mathcal{C} - \boldsymbol{u}_\mathcal{C})$
Le gradient est donc une fonction affine de $\boldsymbol{a}$.

Nous calculons alors le Hessien de $\mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}$:

$$\nabla^2_{\boldsymbol{a}} \mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}} = \boldsymbol{G} = \boldsymbol{H}^\dagger\boldsymbol{H} + \rho_P\boldsymbol{P}^\dagger\boldsymbol{P} + \rho_D(\boldsymbol{D}_x^\dagger\boldsymbol{D}_x + \boldsymbol{D}_y^\dagger\boldsymbol{D}_y) + (\nu + \rho_\mathcal{C})\boldsymbol{I}_N \tag{D.20}$$

Le Hessien est donc constant.

La matrice hermitienne $\boldsymbol{H}^\dagger\boldsymbol{H}$ est semi-définie positive car pour tout $\boldsymbol{a}$ nous avons :

$$\boldsymbol{a}^\dagger\boldsymbol{H}^\dagger\boldsymbol{H}\boldsymbol{a} = (\boldsymbol{H}\boldsymbol{a})^\dagger\boldsymbol{H}\boldsymbol{a} = \|\boldsymbol{H}\boldsymbol{a}\|_2^2 \geqslant 0 \tag{D.21}$$

Avec le même raisonnement, nous montrons que les matrices $\boldsymbol{P}^\dagger\boldsymbol{P}$, $\boldsymbol{D}_x^\dagger\boldsymbol{D}_x$, $\boldsymbol{D}_y^\dagger\boldsymbol{D}_y$ (et a fortiori $\boldsymbol{I}_N$) sont semi-définies positives.

En tant que somme de matrices semi-définies positives, la matrice $\nabla^2_{\boldsymbol{a}}\mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}$ l'est également. Par ailleurs, $(\nu+\rho_\mathcal{C}) > 0$ (car par définition ν est positif et $\rho_\mathcal{C}$ est strictement positif). Le Hessien est donc une matrice définie positive.

Nous en déduisons que :

le lagrangien augmenté $\mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}$ est une fonction quadratique strictement convexe de $\boldsymbol{a}$.

Nous pouvons en conclure qu'un minimiseur de $\mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}$ annule son gradient. A chaque mise-à-jour de $\boldsymbol{a}$ dans **Algorithme 9** nous avons donc :

$$\begin{aligned}\boldsymbol{a}^{(k+1)} &= \operatorname*{argmin}_{\boldsymbol{a}} \mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}(\boldsymbol{a}, \boldsymbol{v}_P^{(k)}, \boldsymbol{v}_x^{(k)}, \boldsymbol{v}_y^{(k)}, \boldsymbol{v}_\mathcal{C}^{(k)}, \boldsymbol{u}_P^{(k)}, \boldsymbol{u}_x^{(k)}, \boldsymbol{u}_y^{(k)}, \boldsymbol{u}_\mathcal{C}^{(k)}) \\ &\Leftrightarrow\ \boldsymbol{G}\boldsymbol{a}^{(k+1)} - \boldsymbol{d}^{(k)} = \boldsymbol{0}\end{aligned} \tag{D.22}$$

où :
$$\boldsymbol{d}^{(k)} = \Big[\boldsymbol{H}^\dagger\boldsymbol{\sigma} + \rho_P\boldsymbol{P}^\dagger(\boldsymbol{v}_P^{(k)} - \boldsymbol{u}_P^{(k)}) + \rho_D\big(\boldsymbol{D}_x^\dagger(\boldsymbol{v}_x^{(k)} - \boldsymbol{u}_x^{(k)}) + \boldsymbol{D}_y^\dagger(\boldsymbol{v}_y^{(k)} - \boldsymbol{u}_y^{(k)})\big) + \rho_\mathcal{C}(\boldsymbol{v}_\mathcal{C}^{(k)} - \boldsymbol{u}_\mathcal{C}^{(k)})\Big]$$

Et finalement, comme $\boldsymbol{G}$ est inversible, nous déduisons de (D.19) et (D.22) :

$$\boxed{\boldsymbol{a}^{(k+1)} = \boldsymbol{G}^{-1}\boldsymbol{d}^{(k)}} \tag{D.23}$$

Ce qui termine la preuve de la mise-à-jour de $\boldsymbol{a}$.

◊ **Mise-à-jour de $\boldsymbol{v}_P$, $\boldsymbol{v}_x$, et $\boldsymbol{v}_y$**

Dans **Algorithme 9**, la mise-à-jour de $\boldsymbol{v}_P$ s'écrit :

$$\begin{aligned}\boldsymbol{v}_P^{(k+1)} &= \operatorname*{argmin}_{\boldsymbol{v}_P} \mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P, \boldsymbol{v}_x^{(k)}, \boldsymbol{v}_y^{(k)}, \boldsymbol{v}_\mathcal{C}^{(k)}, \boldsymbol{u}_P^{(k)}, \boldsymbol{u}_x^{(k)}, \boldsymbol{u}_y^{(k)}, \boldsymbol{u}_\mathcal{C}^{(k)}) \\ &= \operatorname*{argmin}_{\boldsymbol{v}_P} \left\{\mu\|\boldsymbol{v}_P\|_1 + \frac{\rho_P}{2}\left\|\boldsymbol{P}\boldsymbol{a}^{(k+1)} - \boldsymbol{v}_P + \boldsymbol{u}_P{}^{(k)}\right\|_2^2\right\} \\ &= \operatorname{prox}_{\mu\|\boldsymbol{v}_P\|_1}\left(\boldsymbol{P}\boldsymbol{a}^{(k+1)} + \boldsymbol{u}_P^{(k)}\right) \\ &= \mathrm{S}_{\mu/\rho_P}\left(\boldsymbol{P}\boldsymbol{a}^{(k+1)} + \boldsymbol{u}_P^{(k)}\right)\end{aligned} \tag{D.24}$$

où $\text{prox}_{\mu\|\boldsymbol{v}_P\|_1}$ est l'opérateur proximal de la norme ℓ_1 et S_{μ/ρ_P} est l'opérateur de seuillage doux de seuil μ/ρ_P qui sont définis dans l'**Annexe E**. Le seuillage doux consiste à mettre à 0 les composantes dont le module est inférieur au seuil et à pondérer les autres. Cette opération s'effectue composante par composante.

Les preuves de mise-à-jour des variables $\boldsymbol{v}_x$, et $\boldsymbol{v}_y$ étant similaires à celle de la variable $\boldsymbol{v}_P$, nous ne les détaillons pas.

◊ **Mise-à-jour de $\boldsymbol{v}_\mathcal{C}$**

Dans **Algorithme 9**, la mise-à-jour de $\boldsymbol{v}_\mathcal{C}$ s'écrit :

$$\begin{aligned}
\boldsymbol{v}_\mathcal{C}^{(k+1)} &= \underset{\boldsymbol{v}_\mathcal{C}}{\text{argmin}}\, \mathcal{L}_{\rho_P,\rho_D,\rho_\mathcal{C}}(\boldsymbol{a}^{(k+1)}, \boldsymbol{v}_P^{(k+1)}, \boldsymbol{v}_x^{(k+1)}, \boldsymbol{v}_y^{(k+1)}, \boldsymbol{v}_\mathcal{C}, \boldsymbol{u}_P^{(k)}, \boldsymbol{u}_x^{(k)}, \boldsymbol{u}_y^{(k)}, \boldsymbol{u}_\mathcal{C}^{(k)}) \\
&= \underset{\boldsymbol{v}_\mathcal{C}}{\text{argmin}} \left\{ \mathcal{I}_\mathcal{C}(\boldsymbol{v}_\mathcal{C}) + \frac{\rho_\mathcal{C}}{2} \left\| \boldsymbol{a}^{(k+1)} - \boldsymbol{v}_\mathcal{C} + \boldsymbol{u}_\mathcal{C}^{(k)} \right\|_2^2 \right\} \\
&= \text{prox}_{\mathcal{I}_\mathcal{C}(\boldsymbol{v}_\mathcal{C})}\left(\boldsymbol{a}^{(k+1)} + \boldsymbol{u}_\mathcal{C}^{(k)}\right) \\
&= \mathcal{P}_\mathcal{C}(\boldsymbol{a}^{(k+1)} + \boldsymbol{u}_\mathcal{C}^{(k)})
\end{aligned} \tag{D.25}$$

où $\text{prox}_{\mathcal{I}_\mathcal{C}(\boldsymbol{v}_\mathcal{C})}$ est l'opérateur proximal de la fonction indicatrice $\mathcal{I}_\mathcal{C}$ et $\mathcal{P}_\mathcal{C}$ est l'opérateur de projection euclidienne sur l'ensemble $\mathcal{C}$ qui sont définis **Annexe E**. Dans notre cas, la projection consiste simplement à mettre à 0 les composantes en dehors du support spatial et à conserver les autres. Cette opération s'effectue composante par composante.

► Diagonalisation des matrices $\boldsymbol{H}^\dagger\boldsymbol{H}$, $\boldsymbol{P}^\dagger\boldsymbol{P}$, $\boldsymbol{D}_x^\dagger\boldsymbol{D}_x$, $\boldsymbol{D}_y^\dagger\boldsymbol{D}_y$ et $\boldsymbol{G}$ Section 3.3.3

◊ **Diagonalisation de $\boldsymbol{H}^\dagger\boldsymbol{H}$**

Nous rappelons que si les données sont sur une grille cartésienne régulière du k-space alors la matrice $\boldsymbol{H}$ s'écrit :

$$\boldsymbol{H} = \alpha\sqrt{N}\, \boldsymbol{\Delta}_\Psi \boldsymbol{S} \boldsymbol{F}_{3D} \boldsymbol{\Delta}_\Phi \tag{D.26}$$

Les termes intervenant dans l'expression de $\boldsymbol{H}$ sont explicités dans la **Section 1.2.3.4** (p. 43).

Nous déduisons de (D.26) que :

$$\begin{aligned}
\boldsymbol{H}^\dagger\boldsymbol{H} &= |\alpha|^2 N \boldsymbol{\Delta}_\Phi^* \boldsymbol{F}_{3D}^\dagger \boldsymbol{S}^t \boldsymbol{\Delta}_\Psi^* \boldsymbol{\Delta}_\Psi \boldsymbol{S} \boldsymbol{F}_{3D} \boldsymbol{\Delta}_\Phi \\
&= N \boldsymbol{\Delta}_\Phi^* \boldsymbol{F}_{3D}^\dagger \boldsymbol{S}^t \boldsymbol{S} \boldsymbol{F}_{3D} \boldsymbol{\Delta}_\Phi
\end{aligned} \tag{D.27}$$

car $|\alpha|^2 = 1$ et $\boldsymbol{\Delta}_\Psi^* \boldsymbol{\Delta}_\Psi = \boldsymbol{I}_{\text{Nobs}}$.

Nous nous intéressons maintenant au terme :

$$\boldsymbol{F}_{3D}\boldsymbol{H}^{\dagger}\boldsymbol{H}\boldsymbol{F}_{3D}^{\dagger} = N\boldsymbol{F}_{3D}\boldsymbol{\Delta}_{\Phi}^{*}\boldsymbol{F}_{3D}^{\dagger}\boldsymbol{S}^{t}\boldsymbol{S}\boldsymbol{F}_{3D}\boldsymbol{\Delta}_{\Phi}\boldsymbol{F}_{3D}^{\dagger} \tag{D.28}$$

Pour cela, nous étudions un premier temps $\boldsymbol{F}_{3D}\boldsymbol{\Delta}_{\Phi}^{*}\boldsymbol{F}_{3D}^{\dagger}$:

$$\begin{aligned}\boldsymbol{F}_{3D}\boldsymbol{\Delta}_{\Phi}^{*}\boldsymbol{F}_{3D}^{\dagger} &= \left(\boldsymbol{F}_{N_x} \otimes \boldsymbol{F}_{N_y} \otimes \boldsymbol{F}_{N_z}\right)\left(\boldsymbol{\Delta}_{\Phi_x}^{*} \otimes \boldsymbol{\Delta}_{\Phi_y}^{*} \otimes \boldsymbol{\Delta}_{\Phi_z}^{*}\right)\left(\boldsymbol{F}_{N_x}^{\dagger} \otimes \boldsymbol{F}_{N_y}^{\dagger} \otimes \boldsymbol{F}_{N_z}^{\dagger}\right)\\ &= \left(\boldsymbol{F}_{N_x}\boldsymbol{\Delta}_{\Phi_x}^{*}\boldsymbol{F}_{N_x}^{\dagger}\right) \otimes \left(\boldsymbol{F}_{N_y}\boldsymbol{\Delta}_{\Phi_y}^{*}\boldsymbol{F}_{N_y}^{\dagger}\right) \otimes \left(\boldsymbol{F}_{N_z}\boldsymbol{\Delta}_{\Phi_z}^{*}\boldsymbol{F}_{N_z}^{\dagger}\right)\\ &= \boldsymbol{\Pi}_x \otimes \boldsymbol{\Pi}_y \otimes \boldsymbol{\Pi}_z\\ &= \boldsymbol{\Pi}_{\text{3D}}\end{aligned} \tag{D.29}$$

Notons que les matrices $\boldsymbol{\Pi}_x = \boldsymbol{F}_{N_x}\boldsymbol{\Delta}_{\Phi_x}^{*}\boldsymbol{F}_{N_x}^{\dagger}$, $\boldsymbol{\Pi}_y = \boldsymbol{F}_{N_y}\boldsymbol{\Delta}_{\Phi_y}^{*}\boldsymbol{F}_{N_y}^{\dagger}$ et $\boldsymbol{\Pi}_z = \boldsymbol{F}_{N_z}\boldsymbol{\Delta}_{\Phi_z}^{*}\boldsymbol{F}_{N_z}^{\dagger}$ sont circulantes car elles sont diagonales dans l'espace de Fourier. De plus, et comme prouvé dans la **Section D.3** (p. 178), si la condition $k_x^{[0]}/\delta k_x = P_x \in \mathbb{Z}$ (resp. $k_y^{[0]}/\delta k_y = P_y \in \mathbb{Z}$, $k_z^{[0]}/\delta k_z = P_z \in \mathbb{Z}$) est vérifiée alors $\boldsymbol{\Pi}_x$ (resp. $\boldsymbol{\Pi}_y$, $\boldsymbol{\Pi}_z$) est une matrice binaire de décalage circulaire.

Nous reportons (D.29) dans (D.28) et il vient :

$$\boldsymbol{F}_{3D}\boldsymbol{H}^{\dagger}\boldsymbol{H}\boldsymbol{F}_{3D}^{\dagger} = N\boldsymbol{\Pi}_{\text{3D}}\boldsymbol{S}^{t}\boldsymbol{S}\boldsymbol{\Pi}_{\text{3D}}^{\dagger} \tag{D.30}$$

Et nous notons :

$$\boldsymbol{\Lambda}_{\boldsymbol{H}^{\dagger}\boldsymbol{H}} = N\boldsymbol{\Pi}_{\text{3D}}\boldsymbol{S}^{t}\boldsymbol{S}\boldsymbol{\Pi}_{\text{3D}}^{\dagger} \tag{D.31}$$

$\boldsymbol{S}^{t}\boldsymbol{S}$ est une matrice de projection canonique : elle est diagonale de 0 et de 1 et son application annule les composantes aux fréquences spatiales non observées.

Nous en déduisons que si $\boldsymbol{\Pi}_x, \boldsymbol{\Pi}_y$ et $\boldsymbol{\Pi}_z$ sont des matrices binaires de décalage circulaire (*i.e.* si $(P_x, P_y, P_z) \in \mathbb{Z}^3$), alors $\boldsymbol{\Lambda}_{\boldsymbol{H}^{\dagger}\boldsymbol{H}}$ est une matrice diagonale dont les éléments diagonaux correspondent à un décalage circulaire des éléments diagonaux de $\boldsymbol{S}^{t}\boldsymbol{S}$. Finalement, en multipliant (D.30) à gauche par $\boldsymbol{F}_{3D}^{\dagger}$ et à droite par $\boldsymbol{F}_{3D}$ il vient :

$$\boxed{\boldsymbol{H}^{\dagger}\boldsymbol{H} = \boldsymbol{F}_{3D}^{\dagger}\boldsymbol{\Lambda}_{\boldsymbol{H}^{\dagger}\boldsymbol{H}}\boldsymbol{F}_{3D}} \tag{D.32}$$

Nous avons donc montré que :

$\boldsymbol{H}^{\dagger}\boldsymbol{H}$ est diagonalisable dans le domaine de Fourier

Ce qui signifie également que :

$\boldsymbol{H}^{\dagger}\boldsymbol{H}$ est une matrice 3-niveaux circulante

◊ **Diagonalisation de $\boldsymbol{P}^\dagger\boldsymbol{P}$**

Nous rappelons ci-dessous l'expression de la matrice $\boldsymbol{P}$ définie **Section 3.2.1** :

$$\boldsymbol{P} = \underbrace{\left(\begin{array}{c|c|c|c} 1 \text{ —— } 1 & 0 \text{ —— } 0 & & 0 \text{ —— } 0 \\ 0 \text{ —— } 0 & 1 \text{ —— } 1 & & 0 \text{ —— } 0 \\ \vdots \quad\quad \vdots & \vdots \quad\quad \vdots & \cdots & \vdots \quad\quad \vdots \\ 0 \text{ —— } 0 & 0 \text{ —— } 0 & & 1 \text{ —— } 1 \\ 0 \text{ —— } 0 & 0 \text{ —— } 0 & & 0 \text{ —— } 1 \end{array}\right)}_{N_xN_y} \Bigg\} N_z$$

Nous avons donc :

$$\boldsymbol{P}^\dagger\boldsymbol{P} = \mathbb{1}_{N_xN_y} \otimes \boldsymbol{I}_{N_z} = \mathbb{1}_{N_x} \otimes \mathbb{1}_{N_y} \otimes \boldsymbol{I}_{N_z} = \begin{pmatrix} \mathbb{1}_{N_xN_y} & \mathbf{O}_{N_xN_y} & \cdots & \mathbf{O}_{N_xN_y} \\ \mathbf{O}_{N_xN_y} & \mathbb{1}_{N_xN_y} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{O}_{N_xN_y} \\ \mathbf{O}_{N_xN_y} & \cdots & \mathbf{O}_{N_xN_y} & \mathbb{1}_{N_xN_y} \end{pmatrix} \tag{D.33}$$

où $\mathbb{1}_N$ est une matrice de taille $N \times N$ remplie de 1.

Nous nous intéressons désormais au terme $\boldsymbol{F}_{3D}\boldsymbol{P}^\dagger\boldsymbol{P}\boldsymbol{F}_{3D}^\dagger$:

$$\begin{aligned} \boldsymbol{F}_{3D}\boldsymbol{P}^\dagger\boldsymbol{P}\boldsymbol{F}_{3D}^\dagger &= \left(\boldsymbol{F}_{N_x} \otimes \boldsymbol{F}_{N_y} \otimes \boldsymbol{F}_{N_z}\right)\left(\mathbb{1}_{N_x} \otimes \mathbb{1}_{N_y} \otimes \boldsymbol{I}_{N_z}\right)\left(\boldsymbol{F}_{N_x}^\dagger \otimes \boldsymbol{F}_{N_y}^\dagger \otimes \boldsymbol{F}_{N_z}^\dagger\right) \\ &= \left(\boldsymbol{F}_{N_x}\mathbb{1}_{N_x}\boldsymbol{F}_{N_x}^\dagger\right) \otimes \left(\boldsymbol{F}_{N_y}\mathbb{1}_{N_y}\boldsymbol{F}_{N_y}^\dagger\right) \otimes \left(\boldsymbol{F}_{N_z}\boldsymbol{F}_{N_z}^\dagger\right) \\ &= \left(\boldsymbol{F}_{N_x}\mathbb{1}_{N_x}\boldsymbol{F}_{N_x}^\dagger\right) \otimes \left(\boldsymbol{F}_{N_y}\mathbb{1}_{N_y}\boldsymbol{F}_{N_y}^\dagger\right) \otimes \boldsymbol{I}_{N_z} \end{aligned} \tag{D.34}$$

Nous notons :

$$\boldsymbol{\Lambda}_{\boldsymbol{P}^\dagger\boldsymbol{P}} = \left(\boldsymbol{F}_{N_x}\mathbb{1}_{N_x}\boldsymbol{F}_{N_x}^\dagger\right) \otimes \left(\boldsymbol{F}_{N_y}\mathbb{1}_{N_y}\boldsymbol{F}_{N_y}^\dagger\right) \otimes \boldsymbol{I}_{N_z} \tag{D.35}$$

Dans (D.35), les matrices $\boldsymbol{F}_{N_x}\mathbb{1}_{N_x}\boldsymbol{F}_{N_x}^\dagger$ et $\boldsymbol{F}_{N_y}\mathbb{1}_{N_y}\boldsymbol{F}_{N_y}^\dagger$ sont diagonales car :

$$\boldsymbol{F}_N\mathbb{1}_N\boldsymbol{F}_N^\dagger = \begin{pmatrix} N & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \tag{D.36}$$

En tant que produit de Kronecker de matrices diagonales nous en déduisons que $\boldsymbol{\Lambda}_{\boldsymbol{P}^\dagger\boldsymbol{P}}$ est diagonale. Finalement, en reportant (D.35) dans (D.34) et en multipliant (D.34) à gauche par $\boldsymbol{F}_{3D}^\dagger$ et à droite par $\boldsymbol{F}_{3D}$ il vient :

$$\boxed{\boldsymbol{P}^\dagger\boldsymbol{P} = \boldsymbol{F}_{3D}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{P}^\dagger\boldsymbol{P}}\boldsymbol{F}_{3D}} \tag{D.37}$$

Nous avons donc montré que :

$\boxed{\boldsymbol{P}^\dagger \boldsymbol{P} \text{ est diagonalisable dans le domaine de Fourier}}$

Ce qui signifie également que :

$\boxed{\boldsymbol{P}^\dagger \boldsymbol{P} \text{ est une matrice 3-niveaux circulante}}$

◊ **Diagonalisation de $\boldsymbol{D}_x^\dagger \boldsymbol{D}_x$ et $\boldsymbol{D}_y^\dagger \boldsymbol{D}_y$**

Nous rappelons ci-dessous l'expression de la matrice $\boldsymbol{D}_x$ définie **Section 3.2.1** :

$$\boldsymbol{D}_x = \boldsymbol{D}_x^{[1]} \otimes \boldsymbol{I}_{N_z} \tag{D.38}$$

où le bloc intermédiaire (bloc-circulant) $\boldsymbol{D}_x^{[1]}$ est donné par :

$$\boldsymbol{D}_x^{[1]} = \boldsymbol{D}_x^{[0]} \otimes \boldsymbol{I}_{N_y} = \begin{pmatrix} \boldsymbol{D}_x^{[0]} & \mathbf{O}_{N_y} & \cdots & \mathbf{O}_{N_y} \\ \mathbf{O}_{N_y} & \boldsymbol{D}_x^{[0]} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{O}_{N_y} \\ \mathbf{O}_{N_y} & \cdots & \mathbf{O}_{N_y} & \boldsymbol{D}_x^{[0]} \end{pmatrix} \tag{D.39}$$

et où le bloc élémentaire (circulant) $\boldsymbol{D}_x^{[0]}$ est donné par :

$$\boldsymbol{D}_x^{[0]} = \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 1 & & & -1 \end{pmatrix} \tag{D.40}$$

Tout d'abord nous remarquons que la matrice $\left(\boldsymbol{D}_x^{[0]}\right)^\dagger \boldsymbol{D}_x^{[0]}$ est circulante. En effet, $\boldsymbol{D}_x^{[0]}$ est circulante par définition (et donc $\boldsymbol{D}_x^{[0]\dagger}$ l'est également) et nous pouvons donc diagonaliser ces deux matrices dans le domaine fréquentiel de sorte que :

$$\begin{aligned} \left(\boldsymbol{D}_x^{[0]}\right)^\dagger \boldsymbol{D}_x^{[0]} &= \boldsymbol{F}_{N_x}^\dagger \boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}} \boldsymbol{F}_{N_x} \boldsymbol{F}_{N_x}^\dagger \boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]}} \boldsymbol{F}_{N_x} \\ &= \boldsymbol{F}_{N_x}^\dagger \boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}} \boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]}} \boldsymbol{F}_{N_x} \end{aligned} \tag{D.41}$$

où $\boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]}}$ et $\boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}}$ sont des matrices diagonales. Nous notons :

$$\boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}\boldsymbol{D}_x^{[0]}} = \boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}}\boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]}} \tag{D.42}$$

En reportant (D.42) dans (D.41) il vient :

$$\left(\boldsymbol{D}_x^{[0]}\right)^\dagger \boldsymbol{D}_x^{[0]} = \boldsymbol{F}_{N_x}^\dagger \boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}\boldsymbol{D}_x^{[0]}} \boldsymbol{F}_{N_x} \tag{D.43}$$

Comme $\boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}\boldsymbol{D}_x^{[0]}}$ est une matrice diagonale (en tant que produit de deux matrices diagonales), nous avons montré que $\left(\boldsymbol{D}_x^{[0]}\right)^\dagger \boldsymbol{D}_x^{[0]}$ est diagonalisable dans le domaine fréquentiel et donc que c'est une matrice circulante.

Nous nous intéressons désormais au terme $\boldsymbol{F}_{3D}\boldsymbol{D}_x^\dagger\boldsymbol{D}_x\boldsymbol{F}_{3D}^\dagger$. En exploitant (D.38), (D.39) et (D.43) il vient :

$$\begin{aligned}
\boldsymbol{F}_{3D}\boldsymbol{D}_x^\dagger\boldsymbol{D}_x\boldsymbol{F}_{3D}^\dagger &= \boldsymbol{F}_{3D}\left(\left(\boldsymbol{D}_x^{[1]}\right)^\dagger \otimes \boldsymbol{I}_{N_z}\right)\left(\boldsymbol{D}_x^{[1]} \otimes \boldsymbol{I}_{N_z}\right)\boldsymbol{F}_{3D}^\dagger \\
&= \boldsymbol{F}_{3D}\left(\boldsymbol{D}_x^{[0]\dagger} \otimes \boldsymbol{I}_{N_y} \otimes \boldsymbol{I}_{N_z}\right)\left(\boldsymbol{D}_x^{[0]} \otimes \boldsymbol{I}_{N_y} \otimes \boldsymbol{I}_{N_z}\right)\boldsymbol{F}_{3D}^\dagger \\
&= \boldsymbol{F}_{3D}\left(\boldsymbol{D}_x^{[0]\dagger}\boldsymbol{D}_x^{[0]} \otimes \boldsymbol{I}_{N_y} \otimes \boldsymbol{I}_{N_z}\right)\boldsymbol{F}_{3D}^\dagger \\
&= \boldsymbol{F}_{3D}\left(\boldsymbol{F}_{N_x}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}\boldsymbol{D}_x^{[0]}}\boldsymbol{F}_{N_x} \otimes \boldsymbol{F}_{N_y}^\dagger\boldsymbol{F}_{N_y} \otimes \boldsymbol{F}_{N_z}^\dagger\boldsymbol{F}_{N_z}\right)\boldsymbol{F}_{3D}^\dagger \\
&= \boldsymbol{F}_{3D}\boldsymbol{F}_{3D}^\dagger\left(\boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}\boldsymbol{D}_x^{[0]}} \otimes \boldsymbol{I}_{N_y} \otimes \boldsymbol{I}_{N_z}\right)\boldsymbol{F}_{3D}\boldsymbol{F}_{3D}^\dagger
\end{aligned} \tag{D.44}$$

Nous notons :

$$\boldsymbol{\Lambda}_{\boldsymbol{D}_x^\dagger\boldsymbol{D}_x} = \boldsymbol{\Lambda}_{\boldsymbol{D}_x^{[0]\dagger}\boldsymbol{D}_x^{[0]}} \otimes \boldsymbol{I}_{N_y} \otimes \boldsymbol{I}_{N_z} \tag{D.45}$$

Et finalement en reportant (D.45) dans (D.44) et en multipliant (D.44) à gauche par $\boldsymbol{F}_{3D}^\dagger$ et à droite par $\boldsymbol{F}_{3D}$ il vient :

$$\boxed{\boldsymbol{D}_x^\dagger\boldsymbol{D}_x = \boldsymbol{F}_{3D}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{D}_x^\dagger\boldsymbol{D}_x}\boldsymbol{F}_{3D}} \tag{D.46}$$

Nous avons donc montré que :

$$\boxed{\boldsymbol{D}_x^\dagger\boldsymbol{D}_x \text{ est diagonalisable dans le domaine de Fourier}}$$

Ce qui signifie également que :

$$\boxed{\boldsymbol{D}_x^\dagger\boldsymbol{D}_x \text{ est une matrice 3-niveaux circulante}}$$

Le raisonnement pour montrer que $\boldsymbol{D}_y^\dagger\boldsymbol{D}_y$ est diagonalisable dans le domaine fréquentiel et donc 3-niveaux circulante est analogue et nous avons :

$$\boxed{\boldsymbol{D}_y^\dagger\boldsymbol{D}_y = \boldsymbol{F}_{3D}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{D}_y^\dagger\boldsymbol{D}_y}\boldsymbol{F}_{3D}} \tag{D.47}$$

REMARQUE

Ce résultat est évident en utilisant les propriétés de le TFD. En effet, les opérations de différence en ligne et en colonne correspondent à des filtrages périodiques dans le domaine spatial, c'est à dire des produits de convolution circulaire qui sont donc diagonalisables dans le domaine de Fourier.

◊ **Diagonalisation de $\boldsymbol{G}$**

Nous rappelons ci-dessous l'expression de la matrice $\boldsymbol{G}$ définie **Section 3.3.2** :

$$\boldsymbol{G} = \left[\boldsymbol{H}^\dagger\boldsymbol{H} + \rho_P\boldsymbol{P}^\dagger\boldsymbol{P} + \rho_D(\boldsymbol{D}_x^\dagger\boldsymbol{D}_x + \boldsymbol{D}_y^\dagger\boldsymbol{D}_y) + (\nu + \rho_{\mathcal{C}})\boldsymbol{I}_{N_xN_yN_z}\right] \tag{D.48}$$

En exploitant les résultats (D.32), (D.37), (D.46) et (D.47) nous avons :

$$\boldsymbol{G} = \boldsymbol{F}_{3D}^\dagger\left[\boldsymbol{\Lambda}_{\boldsymbol{H}^\dagger\boldsymbol{H}} + \rho_P\boldsymbol{\Lambda}_{\boldsymbol{P}^\dagger\boldsymbol{P}} + \rho_D(\boldsymbol{\Lambda}_{\boldsymbol{D}_x{}^\dagger\boldsymbol{D}_x} + \boldsymbol{\Lambda}_{\boldsymbol{D}_y{}^\dagger\boldsymbol{D}_y}) + (\nu + \rho_{\mathcal{C}})\boldsymbol{I}_{N_xN_yN_z}\right]\boldsymbol{F}_{3D} \tag{D.49}$$

Nous notons :

$$\boldsymbol{\Lambda}_{\boldsymbol{G}} = \left[\boldsymbol{\Lambda}_{\boldsymbol{H}^\dagger\boldsymbol{H}} + \rho_P\boldsymbol{\Lambda}_{\boldsymbol{P}^\dagger\boldsymbol{P}} + \rho_D(\boldsymbol{\Lambda}_{\boldsymbol{D}_x{}^\dagger\boldsymbol{D}_x} + \boldsymbol{\Lambda}_{\boldsymbol{D}_y{}^\dagger\boldsymbol{D}_y}) + (\nu + \rho_{\mathcal{C}})\boldsymbol{I}_{N_xN_yN_z}\right] \tag{D.50}$$

En tant que somme de matrices diagonales, $\boldsymbol{\Lambda}_{\boldsymbol{G}}$ est une matrice diagonale. Et finalement en reportant (D.50) dans (D.49) il vient :

$$\boxed{\boldsymbol{G} = \boldsymbol{F}_{3D}^\dagger\boldsymbol{\Lambda}_{\boldsymbol{G}}\boldsymbol{F}_{3D}} \tag{D.51}$$

Nous avons donc montré que :

$\boldsymbol{G}$ est diagonalisable dans le domaine de Fourier

Ce qui signifie également que :

$\boldsymbol{G}$ est une matrice 3-niveaux circulante

► **Preuve $\boldsymbol{\Pi}$**

Notation

Soit $N \neq 0$. Nous notons :

$$\boldsymbol{\Pi} = \boldsymbol{F}_N\boldsymbol{\Delta}_\Phi^*\boldsymbol{F}_N^{-1} = (\boldsymbol{F}_N^{-1}\boldsymbol{\Delta}_\Phi\boldsymbol{F}_N)^\dagger \tag{D.52}$$

Il vient alors :

$$\boldsymbol{\Pi} = \boldsymbol{W}_{N,N}\boldsymbol{\Delta}_\Phi^*\boldsymbol{W}_{N,N}^{-1} \tag{D.53}$$

car, par définition, $\boldsymbol{W}_{N,N} = \sqrt{N}\boldsymbol{F}_N$ et $\boldsymbol{W}_{N,N}^{-1} = \dfrac{1}{\sqrt{N}}\boldsymbol{F}_N^{-1}$.

Expression de $[\mathbf{\Pi}]_{l,c}$

Soient $(l,c) \in [\![1,N]\!]^2$. Nous cherchons premièrement à exprimer $[\mathbf{\Pi}]_{l,c}$ qui est l'élément de la l-ième ligne et de la c-ième colonne de la matrice $\mathbf{\Pi}$.

Pour cela, nous calculons dans un premier temps :

$$[\boldsymbol{W}_{N,N}\mathbf{\Delta}_{\Phi}^{*}]_{l,c} = \sum_{n=1}^{N} [\boldsymbol{W}_{N,N}]_{l,n} [\mathbf{\Delta}_{\Phi}^{*}]_{n,c} \tag{D.54}$$

Or, en notant $\omega = \exp\{-2\jmath\pi/N\}$, nous avons :

$$[\boldsymbol{W}_{N,N}]_{l,n} = \exp\{-2\jmath\pi(l-1)(n-1)/N\} = \omega^{(l-1)(n-1)}.$$

De plus, $\mathbf{\Delta}_{\Phi}^{*}$ est diagonale et en notant $P = k^{[0]}/\delta k$ nous avons :

$$[\mathbf{\Delta}_{\Phi}^{*}]_{n,c} = \begin{cases} 0 & \text{si } n \neq c \\ \exp\{2\jmath\pi\Phi_{n-1}\} = \exp\{2\jmath\pi P(n-1)/N\} & \text{sinon} \end{cases}.$$

(D.54) s'écrit donc :

$$[\boldsymbol{W}_{N,N}\mathbf{\Delta}_{\Phi}^{*}]_{l,c} = \omega^{(c-1)(l-1-P)} \tag{D.55}$$

Nous calculons désormais :

$$[\mathbf{\Pi}]_{l,c} = \sum_{n=1}^{N} [\boldsymbol{W}_{N,N}\mathbf{\Delta}_{\Phi}^{*}]_{l,n} \left[\boldsymbol{W}_{N,N}^{-1}\right]_{n,c} \tag{D.56}$$

Tout d'abord, nous rappelons que $\boldsymbol{W}_{N,N}^{\dagger} = \sqrt{N}\boldsymbol{F}_{N}^{\dagger} = \sqrt{N}\boldsymbol{F}_{N} = N\boldsymbol{W}_{N,N}^{-1}$ et $\left[\boldsymbol{W}_{N,N}^{\dagger}\right]_{n,c} = \omega^{-(n-1)(c-1)}$. Donc nous avons $\left[\boldsymbol{W}_{N,N}^{-1}\right]_{n,c} = \frac{1}{N}\omega^{-(n-1)(c-1)}$

Finalement il vient :

$$\boxed{[\mathbf{\Pi}]_{l,c} = \frac{1}{N}\sum_{n=0}^{N-1}\left(\omega^{l-P-c}\right)^{n}} \tag{D.57}$$

Preuve que $\mathbf{\Pi}$ est de structure de Tœplitz

En partant de (D.57) et par un calcul direct, nous avons :

$$[\mathbf{\Pi}]_{l+1,c+1} = \frac{1}{N}\sum_{n=0}^{N-1}\left(\omega^{l-P-c}\right)^{n} = [\mathbf{\Pi}]_{l,c} \tag{D.58}$$

$\mathbf{\Pi}$ est donc une matrice à diagonales constantes, autrement dit :

$$\boxed{\mathbf{\Pi} \text{ est de structure de Tœplitz}}$$

Preuve que $\mathbf{\Pi}$ est circulante

Nous avons :

$$\begin{aligned}
[\mathbf{\Pi}]_{l,N} &= \frac{1}{N}\sum_{n=0}^{N-1}\left(\omega^{l-P-N}\right)^n \\
&= \frac{1}{N}\sum_{n=0}^{N-1}\left(\omega^{n(l-P)}\omega^{-nN}\right) \\
&= \frac{1}{N}\sum_{n=0}^{N-1}\omega^{n(l-P)} \text{ (car } n \in \mathbb{N}) \\
&= [\mathbf{\Pi}]_{l+1,1}
\end{aligned} \tag{D.59}$$

Ceci signifie que l'élément de la l-ième ligne de la N-ième colonne (dernière colonne) de $\mathbf{\Pi}$ est le même que l'élément de la première colonne de la $(l+1)$-ième ligne (ligne suivante). Or comme $\mathbf{\Pi}$ est de structure de Tœplitz nous en déduisons que :

$\boxed{\mathbf{\Pi} \text{ est circulante}}$

$\mathbf{\Pi}$ est donc entièrement définie par sa première ligne $l = 1$ (ou sa première colonne $c = 1$).

Preuve que $\mathbf{\Pi}$ est de décalage circulaire

Nous allons désormais montrer que, si $P \in \mathbb{Z}$, alors $\mathbf{\Pi}$ est une matrice de décalage circulaire (i.e. une matrice de décalage binaire circulante qui contient un seul et unique « 1 » par ligne ou par colonne).

Nous notons la condition suivante :

$$(l - P - c)/N \in \mathbb{Z} \tag{Cond}$$

Par l'absurde, nous montrons que $P \notin \mathbb{Z} \Rightarrow (l-P-c) \notin \mathbb{Z} \Rightarrow (l-P-c)/N \notin \mathbb{Z}$ car $(l, c, N) \in \mathbb{N}^3$. Donc une condition nécessaire (mais non suffisante) pour que (Cond) soit vérifiée est que $P \in \mathbb{Z}$.

Nous étudions les deux cas qui peuvent se présenter :

- (Cond) est vérifiée et donc $P \in \mathbb{Z}$, nous avons alors :

  $$(l - P - c)/N \in \mathbb{Z} \;\Leftrightarrow\; (l - P - c) \bmod N = 0$$

  et il vient :

  $$\exp\{-2\jmath\pi(l - P - c)/N\} = 1 \Leftrightarrow \omega^{l-P-c} = 1$$

et donc

$$[\mathbf{\Pi}]_{l,c} = \frac{1}{N}\sum_{n=0}^{N-1}\left(\omega^{l-P-c}\right)^n = 1$$

- (Cond) n'est pas vérifiée, nous avons alors :

$$(l-P-c)/N \notin \mathbb{Z}$$

et il vient :

$$\exp\left\{-2\jmath\pi\left(l-P-c\right)/N\right\} \neq 1 \Leftrightarrow \omega^{l-P-c} \neq 1$$

En tant que somme d'une suite géométrique de raison $\omega^{l-P-c} \neq 1$, pour tout P nous avons :

$$[\mathbf{\Pi}]_{l,c} = \frac{1}{N}\frac{1-\left(\omega^{l-P-c}\right)^N}{1-\omega^{l-P-c}} = \frac{\omega^{\frac{1}{2}(l-P-c)(N-1)}}{N}\frac{\sin\left((l-P-c)\pi\right)}{\sin\left((l-P-c)\pi/N\right)}$$

Si en plus $P \in \mathbb{Z}$ (avec toujours (Cond) non vérifiée) alors $(l-P-c) \in \mathbb{Z}$ (car $(l,p) \in \mathbb{N}^2$) et $\sin\left((l-P-c)\pi\right) = 0$. Ainsi, si (Cond) n'est pas vérifiée mais que $P \in \mathbb{Z}$, $[\mathbf{\Pi}]_{l,c} = 0$.

En résumé, $\mathbf{\Pi}$ est toujours circulante et si :

- $P \notin \mathbb{Z}$ alors $\boxed{[\mathbf{\Pi}]_{l,c} = \frac{\omega^{\frac{1}{2}(l-P-c)(N-1)}}{N}\frac{\sin\left((l-P-c)\pi\right)}{\sin\left((l-P-c)\pi/N\right)}}$
- $P \in \mathbb{Z}$ alors $\boxed{[\mathbf{\Pi}]_{l,c} = \begin{cases} 1 & \text{si } (l-P-c) \bmod N = 0 \\ 0 & \text{sinon} \end{cases}}$

Ceci démontre que $\mathbf{\Pi}$ est une matrice de décalage circulaire pour $P = k^{[0]}/\delta k \in \mathbb{Z}$.

REMARQUE

Ce résultat est également évident en utilisant les propriétés de le TFD. En effet, $\mathbf{\Delta}_\Phi^*$ est une matrice de rampe de phase. La transformée de Fourier de $\mathbf{\Delta}_\Phi^* \boldsymbol{F}_N^{-1}\boldsymbol{v}$ ($\boldsymbol{v}$ est un vecteur) est un $P-$ décalage circulaire des éléments de $\boldsymbol{v}$.

▶ Preuve mise-à-jour de $\boldsymbol{a}$ dans le domaine fréquentiel (3.25)

Nous rappelons qu'à chaque itération de l'**Algorithme 9**, la mise-à-jour de $\boldsymbol{a}$ s'écrit :

$$\boldsymbol{a}^{(k+1)} = \boldsymbol{G}^{-1}\boldsymbol{d}^{(k)} \tag{D.60}$$

et nous avons montré **Section D.3** que :

$$\boldsymbol{G} = \boldsymbol{F}_{3D}^{\dagger}\mathbf{\Lambda}_{\boldsymbol{G}}\boldsymbol{F}_{3D} \tag{D.61}$$

où $\mathbf{\Lambda}_{\boldsymbol{G}}$ est une matrice diagonale.

En reportant (D.61) dans (D.60) il vient :

$$\begin{aligned} \boldsymbol{a}^{(k+1)} &= \left(\boldsymbol{F}_{3D}^{\dagger}\boldsymbol{\Lambda}_{\boldsymbol{G}}\boldsymbol{F}_{3D}\right)^{-1}\boldsymbol{d}^{(k)} \\ &= \boldsymbol{F}_{3D}^{-1}\boldsymbol{\Lambda}_{\boldsymbol{G}}^{-1}\boldsymbol{F}_{3D}\boldsymbol{d}^{(k)} \end{aligned} \tag{D.62}$$

car $\left(\boldsymbol{F}_{3D}^{\dagger}\right)^{-1} = \left(\boldsymbol{F}_{3D}^{-1}\right)^{-1} = \boldsymbol{F}_{3D}$. Pour finir la démonstration nous multiplions (D.62) à gauche par $\boldsymbol{F}_{3D}$ et nous avons :

$$\boxed{\boldsymbol{F}_{3D}\boldsymbol{a}^{(k+1)} = \boldsymbol{\Lambda}_{\boldsymbol{G}}^{-1}\boldsymbol{F}_{3D}\boldsymbol{d}^{(k)}} \tag{D.63}$$

Nous étudions désormais le terme $\boldsymbol{F}_{3D}\boldsymbol{d}^{(k)}$:

$$\begin{aligned} \boldsymbol{F}_{3D}\boldsymbol{d}^{(k)} &= \boldsymbol{F}_{3D}\boldsymbol{H}^{\dagger}\boldsymbol{\sigma} + \boldsymbol{l} \\ &= \alpha N \boldsymbol{F}_{3D}\boldsymbol{\Delta}_{\Phi}^{*}\boldsymbol{F}_{3D}^{\dagger}\boldsymbol{S}^{t}\boldsymbol{\Delta}_{\Psi}^{*}\boldsymbol{\sigma} + \boldsymbol{l} \\ &= \alpha N \boldsymbol{\Pi}_{3\mathrm{D}}\boldsymbol{S}^{t}\boldsymbol{\Delta}_{\Psi}^{*}\boldsymbol{\sigma} + \boldsymbol{l} \end{aligned} \tag{D.64}$$

où le vecteur $\boldsymbol{l}$ s'écrit :
$\boldsymbol{l} = \boldsymbol{F}_{3D}\left[\rho_P \boldsymbol{P}^{\dagger}(\boldsymbol{v}_P^{(k)} - \boldsymbol{u}_P^{(k)}) + \rho_D\left(\boldsymbol{D}_x^{\dagger}(\boldsymbol{v}_x^{(k)} - \boldsymbol{u}_x^{(k)}) + \boldsymbol{D}_y^{\dagger}(\boldsymbol{v}_y^{(k)} - \boldsymbol{u}_y^{(k)})\right) + \rho_{\mathcal{C}}(\boldsymbol{v}_{\mathcal{C}}^{(k)} - \boldsymbol{u}_{\mathcal{C}}^{(k)})\right]$

Finalement en reportant (D.64) dans (D.63) il vient :

$$\boxed{\boldsymbol{F}_{3D}\boldsymbol{a}^{(k+1)} = \alpha N \boldsymbol{\Lambda}_{\boldsymbol{G}}^{-1}\left(\boldsymbol{\Pi}_{3\mathrm{D}}\boldsymbol{S}^{t}\boldsymbol{\Delta}_{\Psi}^{*}\boldsymbol{\sigma} + \boldsymbol{l}\right)} \tag{D.65}$$

ANNEXE E

QUELQUES DÉFINITIONS & THÉORÈMES D'ANALYSE CONVEXE

Nous rappelons ici des résultats bien connus d'analyse convexe qui sont nécessaires à la compréhension des algorithmes présentés dans la **Section 2.2**. Ces définitions sont notamment tirées de [Rockafellar, 1970], [Hiriart-Urruty et Lemaréchal, 2001], [Moreau, 1965] et [Combettes et Pesquet, 2011]. Le lecteur intéressé par les preuves des différentes propositions pourra s'y référer.

Avant tout, nous signalons que dans certaines applications, comme la nôtre, l'objet d'intérêt et les données vivent dans l'espace des complexes. De nombreux travaux d'analyse complexe ont permis de généraliser la plupart des résultats et des propriétés des espaces euclidiens aux espaces hermitiens. Ainsi, et sauf mention contraire, les résultats présentés après sont analogues pour des fonctions à variables réelles ou complexes, l'important étant qu'elles soient à valeurs réelles. Par soucis de simplicité, et sans perte de généralité, les définitions présentées ensuite s'appliquent pour un espace vectoriel $\mathbb{E}$ quelconque (aussi bien sur le corps des réels que des complexes) muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ et d'une norme $\|\cdot\|$. Notons qu'il est aussi possible de se ramener au cas réel en exprimant $\boldsymbol{x} \in \mathbb{C}^N$ comme $(\boldsymbol{x}_{\text{re}}^t, \boldsymbol{x}_{\text{im}}^t)^t \in \mathbb{R}^{2N}$. Cette démarche consiste simplement à former un vecteur composite en empilant $\boldsymbol{x}_{\text{re}}$ (les parties réelles de $\boldsymbol{x}$) au dessus de $\boldsymbol{x}_{\text{im}}$ (les parties imaginaires de de $\boldsymbol{x}$) et à travailler sur un espace plus grand.

Définitions

DÉFINITION : **Domaine effectif**

Le *domaine effectif* d'une fonction $f : \mathbb{E} \to \mathbb{R}$ est défini par :

$$\text{dom } f \triangleq \{\boldsymbol{x} \in \mathbb{E} : f(\boldsymbol{x}) < +\infty\}$$

DÉFINITION : **Fonction convexe**

$f : \mathbb{E} \to \mathbb{R}$ est une fonction *convexe* au sens large si $\forall (x, y) \in \text{dom } f$ et $\forall \theta \in [0, 1]$:

$$f\big(\theta x + (1-\theta)y\big) \leqslant \theta f(x) + (1-\theta) f(y)$$

La fonction est strictement convexe si l'inégalité est stricte.

DÉFINITION : **Fonction propre**

$f : \mathbb{E} \to \mathbb{R}$ est une fonction *propre* si :

$$\text{dom } f \neq \emptyset \text{ et } f(x) > -\infty,\ \forall x \in \mathbb{E}$$

Nous notons $\text{Conv}(\mathbb{E})$ l'ensemble des fonctions définies sur $\mathbb{E}$ à valeurs dans $\mathbb{R} \cup \{-\infty, +\infty\}$ qui sont convexes et propres.

DÉFINITION : **Fonction fermée**

$f : \mathbb{E} \to \mathbb{R}$ est une fonction *fermée* si :

$$\text{les sous-ensembles } \{x \in \text{dom } f \mid f(x) \leqslant \alpha\} \text{ sont fermés } \forall \alpha \in \mathbb{R}$$

Nous notons $\overline{\text{Conv}}(\mathbb{E})$ la partie de $\text{Conv}(\mathbb{E})$ formée des fonctions qui sont également fermées.

DÉFINITION : **Sous-gradient**

Soit $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$, une fonction convexe et propre. $s \in \mathbb{E}$ est un *sous-gradient* de f en $x \in \text{dom } f$ si :

$$\forall\, y \in \mathbb{E},\ f(y) \geqslant f(x) + \langle s, y - x \rangle$$

Nous posons $\langle \boldsymbol{s}, \boldsymbol{y} - \boldsymbol{x} \rangle = \boldsymbol{s}^t(\boldsymbol{y} - \boldsymbol{x})$, le produit scalaire sur $\mathbb{R}^N$ qui est une forme bilinéaire symétrique définie positive.
De la même manière, nous posons $\langle \boldsymbol{s}, \boldsymbol{y} - \boldsymbol{x} \rangle = \boldsymbol{s}^\dagger(\boldsymbol{y} - \boldsymbol{x})$ le produit scalaire sur $\mathbb{C}^N$ qui est une forme hermitienne définie positive.

DÉFINITION : **Sous-différentiel**

L'ensemble des sous-gradients de f en x est appelé *sous-différentiel* de f en x et est noté : $\partial f(x)$. f est dite sous-différentiable en x si $\partial f(x) \neq \emptyset$. Par convention, $\partial f(x) = \emptyset$ si $x \notin \text{dom } f$. Si f est différentiable en x alors $\partial f(x) = \{\nabla f\}$ (son gradient est son unique sous-différentiel). $\widetilde{x}$ est un minimiseur de f (pas nécessairement convexe) si et seulement si f est sous-différentiable en $\widetilde{x}$ et $0 \in \partial f(\widetilde{x})$.

DÉFINITION : **Fonction lipschitzienne**

Soient (A, d_A) et (B, d_B) des espaces métriques[a], $f : A \to B$ une fonction et $M \in \mathbb{R}^+$. f est dite *L-lipschitzienne* :

$$\forall (x, y) \in A^2,\ d_B\left(f(x), f(y)\right) \leq L\, d_A(x, y)$$

a. i.e. des ensembles munis d'une distance d, caractérisé, par exemple, par la norme euclidienne $\|\cdot\|_2 = \sqrt{\langle \cdot, \cdot \rangle}$

DÉFINITION : **Fonction conjuguée ou transformée de Legendre-Fenchel**

Soit $f : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty, +\infty\}$ une fonction (pas nécessairement convexe).
Sa *fonction conjuguée convexe* est la fonction $f^\star : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty, +\infty\}$ définie en $\boldsymbol{s} \in \mathbb{R}^N$ par :

$$f^\star(\boldsymbol{s}) \triangleq \sup_{\boldsymbol{x} \in \mathbb{R}^N} \Big(\langle \boldsymbol{s}, \boldsymbol{x} \rangle - f(\boldsymbol{x}) \Big)$$

DÉFINITION : **Fonction biconjuguée**

La *fonction biconjuguée* notée $f^{\star\star} : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty, +\infty\}$ est définie par :

$$f^{\star\star}(\boldsymbol{x}) \triangleq \sup_{\boldsymbol{s} \in \mathbb{R}^N} \Big(\langle \boldsymbol{s}, \boldsymbol{x} \rangle - f^\star(\boldsymbol{s}) \Big)$$

Pour toute fonction $f : \mathbb{R}^N \to \mathbb{R} \cup \{-\infty, +\infty\}$, $f^{\star\star} \in \overline{\mathrm{Conv}}(\mathbb{R}^N)$

DÉFINITION : **Opérateur proximal**

Soit $f \in \overline{\mathrm{Conv}}(\mathbb{E})$, l'*opérateur proximal* $\mathrm{prox}_f : \mathbb{E} \to \mathbb{E}$ est défini par :

$$\mathrm{prox}_f(\boldsymbol{v}) \triangleq \operatorname*{argmin}_{\boldsymbol{x} \in \mathbb{E}} \left(f(\boldsymbol{x}) + \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{v}\|_2^2 \right)$$

Notons que : $\boldsymbol{x} = \mathrm{prox}_f(\boldsymbol{v}) \Leftrightarrow \boldsymbol{v} - \boldsymbol{x} \in \partial f(\boldsymbol{x})$. D'après la définition du sous-gradient, nous en déduisons que $\widetilde{\boldsymbol{x}}$ est un minimiseur de f si et seulement si $\widetilde{\boldsymbol{x}}$ est un point fixe, i.e. $\widetilde{\boldsymbol{x}} = \mathrm{prox}_f(\widetilde{\boldsymbol{x}})$.

Nous rappelons également deux résultats concernant l'opérateur proximal qui seront utilisés ensuite.

Opérateur proximal de la fonction indicatrice

Soit $\mathcal{C} \subseteq \mathbb{E}$ un sous-ensemble non vide, convexe et fermé de $\mathbb{E}$. Nous appelons fonction indicatrice de $\mathcal{C}$ la fonction définie par :

$$\mathcal{I}_{\mathcal{C}}(x) = \begin{cases} 0 & \text{si } x \in \mathcal{C} \\ +\infty & \text{sinon} \end{cases} \tag{E.1}$$

L'opérateur proximal de $\mathcal{I}_{\mathcal{C}}$ s'écrit alors :

$$\mathrm{prox}_{\mathcal{I}_{\mathcal{C}}}(\boldsymbol{v}) = \operatorname*{argmin}_{x \in \mathcal{C}} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{v}\|_2^2 = \mathcal{PC}(\boldsymbol{v})$$

où $\mathcal{PC}$ est la projecteur euclidien sur $\mathcal{C}$. L'opérateur proximal de la fonction indicatrice est donc une généralisation de la projection.

Par exemple, si $\mathcal{C} = \{\boldsymbol{x} \mid \boldsymbol{l} \preccurlyeq \boldsymbol{x} \preccurlyeq \boldsymbol{u}\}$ défini un support de type « boîte », nous avons :

$$\big(\mathcal{PC}(\boldsymbol{x})\big)_i = \begin{cases} l_i & \text{si } x_i \leqslant l_i \\ x_i & \text{si } l_i \leqslant x_i \leqslant u_i \\ u_i & \text{si } x_i \geqslant u_i \end{cases} \qquad \text{pour } i = 1, \cdots, N \tag{E.2}$$

où N est la taille du vecteur $\boldsymbol{x}$.

Opérateur proximal de la norme ℓ_1

Soit $f(\boldsymbol{x}) = \lambda \|\boldsymbol{x}\|_1$, nous avons :

- $\mathrm{prox}_{\lambda\|\boldsymbol{x}\|_1}(\boldsymbol{v}) = \mathrm{S}_\lambda(\boldsymbol{v}) = \mathrm{signe}(\boldsymbol{v}) \max\{|\boldsymbol{v}| - \lambda, 0\}$ sur l'ensemble des réels
- $\mathrm{prox}_{\lambda\|\boldsymbol{x}\|_1}(\boldsymbol{v}) = \mathrm{S}_\lambda(\boldsymbol{v}) = \dfrac{\max\{|\boldsymbol{v}| - \lambda, 0\}\, \boldsymbol{v}}{\max\{|\boldsymbol{v}| - \lambda, 0\} + \lambda}$ est une extension à l'ensemble des complexes donnée par [Wright *et al.*, 2009] et [Li *et al.*, 2015b].

L'opérateur S_λ est appelé *seuillage doux* (ou *soft thresholding* en anglais). Par ailleurs, $\mathrm{prox}_{\lambda\|\boldsymbol{x}\|_1}(\boldsymbol{v})$ est séparable en $\boldsymbol{v}$, les calculs sont donc menés éléments par éléments. La Figure E.1, illustre le seuillage doux dans le cas d'une variable réelle $v \in \mathbb{R}$ et la Figure E.2 dans le cas d'une variable complexe $x \in \mathbb{C}$.

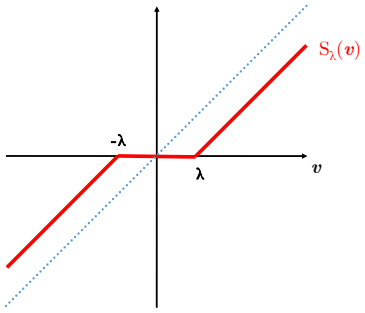


FIGURE E.1 – Seuillage doux pour une variable réelle

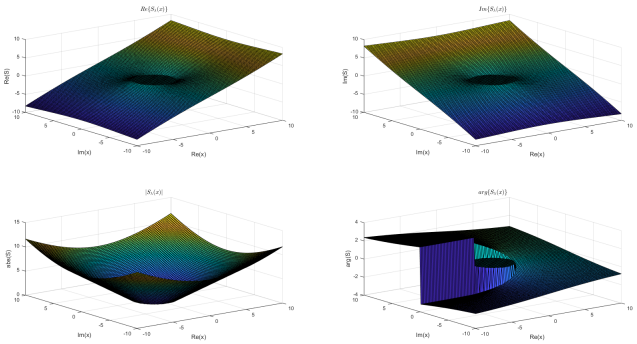


FIGURE E.2 – Seuillage doux pour une variable complexe

DÉFINITION : **Point selle**

On dit que le couple $(\widetilde{\boldsymbol{w}}, \widetilde{\boldsymbol{z}})$ est un point selle de la fonction $f : \mathbb{W} \times \mathbb{Z} \to \mathbb{R}$ si et seulement si :

$$\forall (\boldsymbol{w}, \boldsymbol{z}),\ f(\widetilde{\boldsymbol{w}}, \boldsymbol{z}) \leqslant f(\widetilde{\boldsymbol{w}}, \widetilde{\boldsymbol{z}}) \leqslant f(\boldsymbol{w}, \widetilde{\boldsymbol{z}})$$

Cela signifie que $\widetilde{\boldsymbol{w}}$ minimise $f(\boldsymbol{w}, \widetilde{\boldsymbol{z}})$ sur $\mathbb{W}$ et que $\widetilde{\boldsymbol{z}}$ maximise $f(\widetilde{\boldsymbol{w}}, \boldsymbol{z})$ sur $\mathbb{Z}$:

$$\begin{cases} f(\widetilde{\boldsymbol{w}}, \widetilde{\boldsymbol{z}}) = \inf\limits_{\boldsymbol{w} \in \mathbb{W}} f(\boldsymbol{w}, \widetilde{\boldsymbol{z}}) \\ f(\widetilde{\boldsymbol{w}}, \widetilde{\boldsymbol{z}}) = \sup\limits_{\boldsymbol{z} \in \mathbb{Z}} f(\widetilde{\boldsymbol{w}}, \boldsymbol{z}) \end{cases}$$

Nous avons donc :

$$\inf_{\boldsymbol{w} \in \mathbb{W}} f(\boldsymbol{w}, \widetilde{\boldsymbol{z}}) = \sup_{\boldsymbol{z} \in \mathbb{Z}} f(\widetilde{\boldsymbol{w}}, \boldsymbol{z}) \text{ soit } \inf_{\boldsymbol{w} \in \mathbb{W}} \sup_{\boldsymbol{z} \in \mathbb{Z}} f(\boldsymbol{w}, \boldsymbol{z}) = \sup_{\boldsymbol{z} \in \mathbb{Z}} \inf_{\boldsymbol{w} \in \mathbb{W}} f(\boldsymbol{w}, \boldsymbol{z})$$

Et nous en déduisons que si f satisfait la propriété forte du Min-Max définie par (E.3), alors f admet au moins un *point selle*.

Théorèmes

THÉORÈME : **Min-Max Inégalité**

Soient $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ une fonction et $\mathbb{W} \subseteq \mathbb{R}^n$ et $\mathbb{Z} \subseteq \mathbb{R}^m$ deux sous-ensembles.

$$\sup_{\boldsymbol{z} \in \mathbb{Z}} \inf_{\boldsymbol{w} \in \mathbb{W}} f(\boldsymbol{w}, \boldsymbol{z}) \leqslant \inf_{\boldsymbol{w} \in \mathbb{W}} \sup_{\boldsymbol{z} \in \mathbb{Z}} f(\boldsymbol{w}, \boldsymbol{z})$$

On dit que f satisfait *la propriété forte du Min-Max* si l'égalité suivante est vérifiée :

$$\sup_{\boldsymbol{z} \in \mathbb{Z}} \inf_{\boldsymbol{w} \in \mathbb{W}} f(\boldsymbol{w}, \boldsymbol{z}) = \inf_{\boldsymbol{w} \in \mathbb{W}} \sup_{\boldsymbol{z} \in \mathbb{Z}} f(\boldsymbol{w}, \boldsymbol{z}) \tag{E.3}$$

Une preuve de ce théorème est donnée dans l'**Annexe D** (**Section D.2**).

RÉFÉRENCES

AFONSO, M. V., BIOUCAS-DIAS, J. M. et FIGUEIREDO, M. A. T. (2011). *An Augmented Lagrangian Approach to the Constrained Optimization Formulation of Imaging Inverse Problems*. IEEE Transactions on Image Processing, 20(3):681–695. 86

AUGONNET, C., PUJOLS, A. et SESQUES, M. (2015). *Fast Direct Solver for Integral Equations on Massively Parallel Architectures*. IEEE USNC-URSI, pp. 149–149. 126

AUSTIN, C. D., ERTIN, E. et MOSES, R. L. (2011). *Sparse Signal Methods for 3D Radar Imaging*. IEEE Journal of Selected Topics in Signal Processing, 5(3):408–423. 13, 58, 115, 118

AUSTIN, C. D. et MOSES, R. L. (2008). *Wide-Angle Sparse 3D Synthetic Aperture Radar Imaging for Nonlinear Flight Paths*. IEEE NAECON, pp. 330–336. 58

BALANIS, C. A. (2016). *Antenna Theory : Analysis and Design*. Wiley, 4ième Éd. 19

BARABELL, A. (1983). *Improving the Resolution Performance of Eigenstructure-Based Direction-Finding Algorithms*. Institute of Electrical and Electronics Engineers, 8:336–339. 58

BARTON, D., SKOLNIK, M. I. et SMITH, J. (1998). *IEEE Standard Radar Definitions*. Institute of Electrical and Electronics Engineers. 18

BECHTEL, M. (1969). *Vertically Polarized Radar Backscattering from the Rear of a Cone or Cylinder*. IEEE Transactions on Antennas and Propagation, 17(2):244–246. 31

BECK, A. (2017). *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics. 66

BECK, A. et TEBOULLE, M. (2009a). *Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems*. IEEE Transactions on Image Processing, 18(11):2419–2434. 72, 73

BECK, A. et TEBOULLE, M. (2009b). *A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems*. SIAM Journal on Imaging Sciences, 2(1):183–202. 71, 73

BEN HAMZA, A. et KRIM, H. (2001). *Image Denoising : a Nonlinear Robust Statistical Approach*. IEEE Transactions on Signal Processing, 49(12):3045–3054. 56

BENOUDIBA-CAMPANINI, T., MINVIELLE, P., MASSALOUX, P. et GIOVANNELLI, J.-F. (2017a). *A New Regularization Method for Radar Cross Section Imaging*. IEEE EuCAP, pp. 1419–1423. 107

BENOUDIBA-CAMPANINI, T., MINVIELLE, P., MASSALOUX, P. et GIOVANNELLI, J.-F. (2017b). *Régularisation et Parcimonie pour l'Imagerie Radar Haute Résolution 1D*. JNM 2017. 107

BERTHOUMIEU, Y. (1996). *Contribution à l'Identification de Processus Harmoniques Bidimensionnels - Applications à l'Analyse Haute Résolution de Signaux Radar*. Thèse de doctorat, Université de Bordeaux. 58

BERTSEKAS, D. P. (1975). *Nondifferentiable Optimization via Approximation*. *In Nondifferentiable Optimization*, Vol. 3, pp. 1–25. 66

BHATTACHARYYA, A. K. et SENGUPTA, D. L. (1991). *Radar Cross Section Analysis and Control*. The Artech House radar library. Artech House. 27, 29, 30

BIRCH, G. C., GRIFFIN, J. C. et ERDMAN, M. K. (2015). *UAS Detection, Classification, and Neutralization : Market Survey 2015*. Rapport technique SAND2015-6365, Sandia National Laboratories. 28

BLAKE, A. et ZISSERMAN, A. (1987). *Visual Reconstruction*. The MIT Press series in artificial intelligence. 67

BORDEN, B. (1992). *Maximum Entropy Regularization in Inverse Synthetic Aperture Radar Imagery*. IEEE Transactions on Signal Processing, 40(4):969–973. 96, 120, 150

BORDEN, B. (1999). *Radar Imaging of Airborne Targets : a Primer for Applied Mathematicians and Physicists*. Institute of Physics. 37, 38

BORDEN, B. (2002). *Mathematical Problems in Radar Inverse Scattering*. 18(1):R1–R28. 38

BOURGUIGNON, S., NININ, J., CARFANTAN, H. et MONGEAU, M. (2016). *Exact Sparse Approximation Problems via Mixed-Integer Programming : Formulations and Computational Performance*. IEEE Transactions on Signal Processing, 64(6):1405–1419. 63

BOYD, S. (2010). *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. Foundations and Trends in Machine Learning, 3(1):1–122. 83, 85, 108, 110

BOYD, S. P. et VANDENBERGHE, L. (2004). *Convex Optimization*. Cambridge University Press. 66, 73

BOYLE, J. P. et DYKSTRA, R. L. (1986). *A Method for Finding Projections onto the Intersection of Convex Sets in Hilbert Spaces*. Springer New York, 37:28–47. 83

BRISTOW, H. et LUCEY, S. (2014). *Optimization Methods for Convolutional Sparse Coding*. Computing Research Repository. 109

BRYANT, R., MORAN, M. S., THOMA, D. P., HOLIFIELD COLLINS, C. D., SKIRVIN, S., RAHMAN, M., SLOCUM, K., STARKS, P., BOSCH, D. et GONZALEZ DUGO, M. P. (2007). *Measuring Surface Roughness Height to Parameterize Radar Backscatter Models for Retrieval of Surface Soil Moisture*. IEEE Geoscience and Remote Sensing Letters, 4(1):137–141. 92

BURNSIDE, W. D. et PETERS, L. (1972). *Axial-Radar Cross Section of Finite Cones by the Equivalent-Current Concept with Higher-Order Diffraction*. Radio Science, 7(10):943–948. 31

BURROWS, M. (2004). *Two-Dimensional ESPRIT With Tracking for Radar Imaging and Feature Extraction*. IEEE Transactions on Antennas and Propagation, 52(2):524–532. 58

CANDES, E., ROMBERG, J. et TAO, T. (2006). *Robust Uncertainty Principles : Exact Signal Reconstruction from Highly Incomplete Frequency Information*. IEEE Transactions on Information Theory, 52(2):489–509. 62, 64

CARLAVAN, M., WEISS, P. et BLANC-FÉRAUD, L. (2010). *Régularité et parcimonie pour les problèmes inverses en imagerie. Algorithmes et comparaisons*. Traitement du signal, 27(2):189–219. 56, 69

CETIN, M. (2001). *Feature-enhanced synthetic aperture radar imaging*. Thèse de doctorat, Boston University College of Engineering. 13, 51, 59, 100

CETIN, M. et KARL, W. (1997). *A statistical tomographic approach to synthetic aperture radar image reconstruction*. IEEE Comput. Soc, 1:845–848. 51

CHAARI, L., BATATIA, H. et TOURNERET, J.-Y. (2013). *Sparse Bayesian Regularization Using Bernoulli-Laplacian Priors*. EUSIPCO. 64

CHAMBOLLE, A. (2004). *An Algorithm for Total Variation Minimization and Applications*. Journal of Mathematical Imaging and Vision, Kluwer Academic Publishers, 20(1-2):89–97. 72

CHAMBOLLE, A. et DOSSAL, C. (2015). *On the Convergence of the Iterates of the « Fast Iterative Shrinkage/Thresholding Algorithm »*. Journal of Optimization Theory and Applications, 166(3):968–982. 71

CHASSAY, G. (1982). *Introduction des doubles réflexions dans le modèle des points brillants*. Annales des Télécommunications, 37(11):485–501. 38

CHASSAY, G. (1983). *Représentation ponctuelle de la contribution de la frontière d'ombre d'un objectif radar et son introduction dans le modèle des points brillants*. Annales des Télécommunications, 38(11):460–470. 38

CHEN, M. et CHEN, C.-C. (2014). *RCS Patterns of Pedestrians at 76-77 GHz*. IEEE Antennas and Propagation Magazine, 56(4):252–263. 27, 28

CHEN, S., BILLINGS, S. A. et LUO, W. (1989). *Orthogonal least squares methods and their application to non-linear system identification*. International Journal of Control, 50(5): 1873–1896. 63

CHEN, S. S., DONOHO, D. L. et SAUNDERS, M. A. (1998). *Atomic Decomposition by Basis Pursuit*. SIAM Journal on Scientific Computing, 20(1):33–61. 62, 64

CHEN, V. C. et MARTORELLA, M. (2014). *Inverse Synthetic Aperture Radar Imaging : Principles, Algorithms, and Applications*. Scitech Publishing, IET. 49, 57

CHENEY, M. et BORDEN, B. (2009). *Fundamentals of Radar Imaging*. Num. 79 de CBMS-NSF regional conference series in applied mathematics. Society for Industrial and Applied Mathematics. 38

CIUCIU, P. (2000). *Méthodes Markoviennes en Estimation Spéctrale non Paramétrique - Applications en Imagerie Radar Doppler*. Thèse de doctorat, Université de Bordeaux. 59, 66

CLAERBOUT, J. F. et MUIR, F. (1973). *Robust Modeling With Erratic Data*. GEOPHYSICS, 38(5):826–844. 62

COMBETTES, P. L. et PESQUET, J.-C. (2011). *Proximal Splitting Methods in Signal Processing*. *In Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Vol. 49 de *Springer Optimization and Its Applications*, pp. 185–212. Springer New York. 183

COMBETTES, P. L. et WAJS, V. R. (2005). *Signal Recovery by Proximal Forward-Backward Splitting*. Multiscale Modeling & Simulation, 4(4):1168–1200. 70

CONDAT, L. (2013). *A Direct Algorithm for 1D Total Variation Denoising*. IEEE Signal Processing Letters, 20(11):1054–1057. 64

COOLEY, J. W. et TUKEY, J. W. (1965). *An Algorithm for the Machine Calculation of Complex Fourier Series*. Mathematics of Computation, 19(90):297. 50

COSTA, F., BATATIA, H., CHAARI, L. et TOURNERET, J.-Y. (2015). *Sparse EEG Source Localization Using Bernoulli Laplacian Priors*. IEEE Transactions on Biomedical Engineering, 62(12):2888–2898. 64

CRISPIN, J. W. J. (1968). *Methods of Radar Cross-section Analysis.* Elsevier Science. 29

DAVIS, G. M., MALLAT, S. G. et ZHANG, Z. (1994). *Adaptive time-frequency decompositions with matching pursuit*. p. 402. 63

DE PRONY, G. (1795). *Essai expérimental et analytique : sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alcool à différentes températures*. Journal de l'Ecole Polytechnique, 1(22):24–76. 57

DOERRY, A. W. (2012). *Basics of Polar-Format Algorithm for Processing Synthetic Aperture Radar Images*. Rapport technique SAND2012-3369, Sandia National Laboratories. 41

DONOHO, D. (2006). *Compressed sensing*. IEEE Transactions on Information Theory, 52(4): 1289–1306. 62, 142

DONOHO, D. et HUO, X. (2001). *Uncertainty principles and ideal atomic decomposition*. IEEE Transactions on Information Theory, 47(7):2845–2862. 62

DONOHO, D. L. et LOGAN, B. F. (1992). *Signal Recovery and the Large Sieve*. SIAM Journal on Applied Mathematics, 52(2):577–591. 62

DONOHO, D. L. et STARK, P. B. (1989). *Uncertainty Principles and Signal Recovery*. SIAM Journal on Applied Mathematics, 49(3):906–931. 62

DONOHO, D. L. et TSAIG, Y. (2008). *Fast Solution of L1-Norm Minimization Problems When the Solution May Be Sparse*. IEEE Transactions on Information Theory, 54(11):4789–4812. 64

DOUGLAS, J. et RACHFORD, H. H. (1956). *On the Numerical Solution of Heat Conduction Problems in Two and Three Space Variables*. Transactions of the American Mathematical Society, 82(2):421. 83

DUBOIS, F. (1998). *Reconstruction des Images Tomographiques par Rétroprojection filtrée*. ACOMEN, 4(2). 51

ECKSTEIN, J. et BERTSEKAS, D. P. (1992). *On the Douglas-Rachford Splitting Method and the Proximal Point Algorithm for Maximal Monotone Operators*. Mathematical Programming, 55(1-3):293–318. 85

ELAD, M. et BRUCKSTEIN, A. (2002). *A Generalized Uncertainty Principle and Sparse Representation in Pairs of Bases*. IEEE Transactions on Information Theory, 48(9):2558–2567. 62

ENDER, J. H. (2010). *On Compressive Sensing Applied to Radar*. Signal Processing, 90(5): 1402–1414. 62

EYRAUD, C., GEFFRIN, J.-M., LITMAN, A. et TORTEL, H. (2015). *Complex Permittivity Determination From Far-Field Scattering Patterns*. IEEE Antennas and Wireless Propagation Letters, 14:309–312. 48

EYRAUD, C., GEFFRIN, J.-M., SABOUROUX, P., CHAUMET, P. C., TORTEL, H., GIOVANNINI, H. et LITMAN, A. (2008). *Validation of a 3D Bistatic Microwave Scattering Measurement Setup*. Radio Science, 43(4):n/a–n/a. 35

FORTUNY, J. (2001). *Efficient Algorithms for Three-Dimensional Near-Field Synthetic Aperture Radar Imaging*. Thèse de doctorat, University of Karslruhe, Faculty of Electrical Engineering. 26, 35

FORTUNY, J. et SIEBER, A. (1999). *Three-dimensional synthetic aperture radar imaging of a fir tree : first results*. IEEE Transactions on Geoscience and Remote Sensing, 37(2):1006–1014. 27

FRECON, J., PUSTELNIK, N., DOBIGEON, N., WENDT, H. et ABRY, P. (2017). *Bayesian-driven Criterion to Automatically Select the Regularization Parameter in the l1-Potts Model*. IEEE ICASSP 2017. 143

FRITSCH, P. (1965). *Special Issue on Radar Reflectivity*. Proceedings of the IEEE, 53(8). 29

FUCHS, J.-J. (2004). *Recovery of Exact Sparse Representations in the Presence of Noise*. IEEE ICASSP, 2:II–533–536. 64

FUHS, A. E. (1983). *Radar Cross Section Lectures*. 21, 24, 29, 30

GABAY, D. et MERCIER, B. (1976). *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*. Computers & Mathematics with Applications, 2(1):17–40. 83

GABOR, D. (1946). *Theory of communication. Part 1 : The analysis of information*. Journal of the Institution of Electrical Engineers - Part III : Radio and Communication Engineering, 93(26):429–441. 93

GARAT, J. et MORVAN, S. (1993). *Le Traitement du Signal dans les Mesures de Surface Equivalente Radar*. Revue scientifique et technique de la Direction des applications militaires, (8):75–86. 145

GARAT, J., NAUD, P. et PRULHIERE, J. P. (1993). *Moyen de Mesure de la Surface Equivalente Radar*. Revue scientifique et technique de la Direction des applications militaires, (8):65–74. 49, 145

GEFFRIN, J.-M., EYRAUD, C. et LITMAN, A. (2015). *3-D Imaging of a Microwave Absorber Sample From Microwave Scattered Field Measurements*. IEEE Microwave and Wireless Components Letters, 25(7):472–474. 48

GEMAN, D. et REYNOLDS, G. (1992). *Constrained Restoration and the Recovery of Discontinuities*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(3):367–383. 67

GEMAN, D. et YANG, C. (1995). *Nonlinear Image Recovery With Half-Quadratic Regularization*. IEEE Transactions on Image Processing, 4(7):932–946. 67

GEMAN, S. et GEMAN, D. (1984). *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6(6):721–741. 64

GIOVANNELLI, J.-F. et COULAIS, A. (2005). *Positive Deconvolution for Superimposed Extended Source and Point Sources*. Astronomy & Astrophysics, 439(1):401–412. 65

GIOVANNELLI, J.-F. et IDIER, J. (2013). *Méthodes d'Inversion Appliquées au Traitement du Signal et de l'Image*. Hermès Science : Lavoisier. 53

GIOVANNELLI, J.-F. et VACAR, C. (2017). *Deconvolution-segmentation for textured images. In Signal Processing Conference (EUSIPCO), 2017 25th European*, pp. 191–195. IEEE. 144

GLOWINSKI, R., LIONS, J.-L. et TRÉMOLIÈRES, R. (1976). *Analyse Numérique des Inéquations Variationnelles*. Méthodes mathématiques de l'informatique ; 5. 66

GLOWINSKI, R. et MARROCO, A. (1975). *Sur l'Approximation, par Éléments Finis d'Ordre Un, et la Résolution, par Pénalisation-Dualité d'une Classe de Problèmes de Dirichlet Non Linéaires*. Revue française d'automatique, informatique, recherche opéra- tionnelle., 9(R2):41–76. 83

GOLUB, G. H., HEATH, M. et WAHBA, G. (1979). *Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter*. Technometrics, 21(2):215. 143

GRAFF, C. G. et SIDKY, E. Y. (2015). *Compressive Sensing in Medical Imaging*. Applied Optics, 54(8):C23. 62

GRIBONVAL, R. et NIELSEN, M. (2003). *Sparse Representations in Unions of Bases*. IEEE Transactions on Information Theory, 49(12):3320–3325. 62

GRIFFITHS, H. D. et BAKER, C. J. (2007). *Radar Imaging for Combating Terrorism*. Springer Netherlands, pp. 29–48. 31

GUAY, R. (2016). *Mesurer Dynamiquement la Surface Equivalente Radar de Drones*. Thèse de doctorat, Royal Military College of Canada,. 28

GUILLEMOT, T., POTTIER, E., SAILLARD, J., LEVREL, J.-R. et LEBRETON, E. (1997). *Transformation Champ Proche - Champ Lointain Bidimensionnelle et Monodimensionnelle Appliquée à Des Mesures de SER*. Annales des Télécommunications, 52(1):81–93. 147

HANSEN, P. (2000). *The L-Curve and its Use in the Numerical Treatment of Inverse Problems*. WIT Press, pp. 119–142. 143

HASTIE, T., TIBSHIRANI, R. et WAINWRIGHT, M. (2015). *Statistical Learning with Sparsity : the Lasso and Generalizations*. Num. 143 de Monographs on statistics and applied probability. CRC Press, Taylor & Francis Group. 64

HAWE, S., KLEINSTEUBER, M. et DIEPOLD, K. (2012). *Cartoon-Like Image Reconstruction via Constrained lp-Minimization*. IEEE ICASSP, pp. 717–720. 64

HAYKIN, S. S. (1996). *Adaptive Filter Theory*. Prentice Hall information and system sciences series. 3rd ed Éd. 171

HENROT, S. (2013). *Déconvolution et Séparation d'Images Hyperspectrales en Microscopie*. Thèse de doctorat, Université de Lorraine. 65, 66

HERMAN, M. et STROHMER, T. (2009). *High-Resolution Radar via Compressed Sensing*. IEEE Transactions on Signal Processing, 57(6):2275–2284. 58

HERSCHFELT, A., BIRTCHER, C. R., GUTIERREZ, R. M., RONG, Y., YU, H., BALANIS, C. A. et BLISS, D. W. (2017). *Consumer-Grade Drone Radar Cross-Section and Micro-Doppler Phenomenology*. IEEE RADAR, pp. 0981–0985. 28

HÖGBOM, J. (1974). *Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines*. Astronomy and Astrophysics Supplement Series, 15:417–426. 57

HIRIART-URRUTY, J.-B. et LEMARÉCHAL, C. (2001). *Fundamentals of Convex Analysis*. Grundlehren text editions. Springer. 183

HODGE, D. et SCHULTZ, F. (1965). *The Born Approximation Applied to Electromagnetic Scattering from a Finite Cone*. Proceedings of the IEEE, 53(8):1066–1067. 37

HUBER, P. J. (1996). *Robust Statistical Procedures*. Num. 68 de CBMS-NSF regional conference series in applied mathematics. 2nd ed Éd. 66

HUNT, B. (1975). *Digital Image Processing*. Proceedings of the IEEE, 63(4):693–708. 56

HUNT, B. R. (1973). *The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer*. IEEE Transactions on Computers, C-22(9):805–812. 105

IDIER, J. (2001). *Approche Bayésienne pour les Problèmes Inverses*. Hermès Science Publications. 13, 56, 67, 143

JAY, F. (1984). *IEEE Standard Dictionary of Electrical and Electronics Terms*. Institute of Electrical and Electronics Engineers : Distributed in cooperation with Wiley-Interscience, 3rd ed Éd. 48

KAK, A. C. et SLANEY, M. (2001). *Principles of Computerized Tomographic Imaging*. Society for Industrial and Applied Mathematics. 51

KELLER, J. (1961). *Backscattering from a Finite Cone–Comparison of Theory and Experiment*. IRE Transactions on Antennas and Propagation, 9(4):411–412. 31

KEMPF, T., PEICHL, M., DILL, S. et SUESS, H. (2007). *Microwave Radar Signature Acquisition of Urban Structures*. pp. 68–73. 27

KNOTT, E. F. et SENIOR, T. B. A. (1973). *Second Order Diffraction by a Ring Discontinuity*. University of Michigan. 31

KNOTT, E. F., SHAEFFER, J. F. et TULEY, M. T. (2004). *Radar Cross Section*. SciTech radar and defense series. SciTech Pub, 2nd ed., corr. reprinting Éd. 18, 19, 21, 24, 25, 26, 29, 32, 49, 91, 127, 130, 147

KOLAWOLE, M. O. (2002). *Radar Systems, Peak Detection and Tracking*. Newnes. 52

KORMYLO, J. J. et MENDEL, J. M. (1983). *Maximum-Likelihood Seismic Deconvolution*. IEEE Transactions on Geoscience and Remote Sensing, GE-21(1):72–82. 64

KUHN, H. W. et TUCKER, A. W. (1951). *Nonlinear Programming*. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, pp. 481–492. 76

KUMARESAN, R. et TUFTS, D. (1982). *Estimating the Parameters of Exponentially Damped Sinusoids and Pole-Zero Modeling in Noise*. IEEE Transactions on Acoustics, Speech, and Signal Processing, 30(6):833–840. 57

LANGLEY, K., HAMRAN, S.-E., HOGDA, K. A., STORVOLD, R., BRANDT, O., HAGEN, J. O. et KOHLER, J. (2007). *Use of C-Band Ground Penetrating Radar to Determine Backscatter Sources Within Glaciers*. IEEE Transactions on Geoscience and Remote Sensing, 45(5): 1236–1246. 31

LASSERRE, M., BIDON, S. et LE CHEVALIER, F. (2016). *New Sparse-Promoting Prior for the Estimation of a Radar Scene with Weak and Strong Targets*. IEEE Transactions on Signal Processing, 64(17):4634–4643. 59

LE CHEVALIER, F. et FUGIER-GARREL, C. (1979). *Synthèse Tomographique Appliquée au Radar (S.T.A.R)*. 38

LE CHEVALIER, F. et POUIT, C. (1983). *Imagerie Radar et Synthèse Tomographique*. (76): 119–132. 38

LI, C. J., BHALLA, R. et LING, H. (2015a). *Investigation of the Dynamic Radar Signatures of a Vertical-Axis Wind Turbine*. IEEE Antennas and Wireless Propagation Letters, 14:763–766. 29

LI, C. J. et LING, H. (2017). *An Investigation on the Radar Signatures of Small Consumer Drones*. IEEE Antennas and Wireless Propagation Letters, 16:649–652. 28

LI, L., WANG, X. et WANG, G. (2015b). *Alternating Direction Method of Multipliers for Separable Convex Optimization of Real Functions in Complex Variables*. Mathematical Problems in Engineering, 2015:1–14. 73, 85, 186

LIN, Y., WOHLBERG, B. et VESSELINOV, V. (2017). *ADMM Penalty Parameter Selection with Krylov Subspace Recycling Technique for Sparse Coding*. Proceedings of IEEE. 86

LIONS, P. L. et MERCIER, B. (1979). *Splitting Algorithms for the Sum of Two Nonlinear Operators*. SIAM Journal on Numerical Analysis, 16(6):964–979. 70

LLIBRE, M. (2008). *Résolution de Systèmes Linéaires Moindres Carrés Récursifs et Filtre de Kalman Discret*. Rapport technique DCSD-2008_069-NOT-001-1.0, ISAE et ONERA. 55

LONNQVIST, A., MALLAT, J. et RAISANEN, A. (2006). *Phase-Hologram-Based Compact RCS Test Range at 310 GHz for Scale Models*. IEEE Transactions on Microwave Theory and Techniques, 54(6):2391–2397. 31

LUSTIG, M., DONOHO, D. et PAULY, J. M. (2007). *Sparse MRI : The Application of Compressed Zensing for Rapid MR Imaging*. Magnetic Resonance in Medicine, 58(6):1182–1195. 62

MAHAFZA, B. R. (2013). *Radar Systems Analysis and Design Using MATLAB*. CRC/Taylor & Francis, third edition Éd. 93

MAHAFZA, B. R. et ELSHERBENI, A. Z. (2004). *MATLAB Simulations for Radar Systems Design*. CRC Press/Chapman & Hall. 160

MALLAT, S. et ZHIFENG ZHANG (1993). *Matching Pursuits with Time-Frequency Dictionaries*. IEEE Transactions on Signal Processing, 41(12):3397–3415. 63

MASSALOUX, P., MINVIELLE, P. et GIOVANNELLI, J.-F. (2014). *Indoor 3D Spherical Near Field RCS Measurement Facility : Localization of Scatterers*. IEEE EuCAP, pp. 2765–2769. 33

MENDENHALL, M. R. (1992). *Tactical Missile Aerodynamics. General Topics*. Num. 141 de Progress in Astronautics and Aeronautics. 2. ed Éd. 22

MENSA, D. L. (1981). *High Resolution Radar Cross-Section Imaging*. The Artech House radar library. Artech House. 48, 50, 96, 120, 150

MEZRICH, R. (1995). *A Perspective on K-Space.* Radiology, 195(2):297–315. 160

MINVIELLE, P., MASSALOUX, P. et GIOVANNELLI, J.-F. (2017). *Indoor 3-D Radar Imaging for Low-RCS Analysis*. IEEE Transactions on Aerospace and Electronic Systems, 53(2): 995–1008. 36, 55

MOCH, R., DALLMANN, T. et HEBERLING, D. (2017). *A tomographic approach to polarimetrie radar cross-section imaging*. IEEE EuCAP, pp. 3940–3944. 58

MOHAMMAD-DJAFARI, A. (2008). *Gauss-Markov-Potts priors for images in computer tomography resulting to joint optimal reconstruction and segmentation*. International Journal of Tomography and Statistics, 11:76–92. 144

MOREAU, J.-J. (1965). *Proximité et Dualité dans un Espace Hilbertien*. Bulletin de la Société Mathématique de France, pp. 273–299. 183

MOREIRA, A., PRATS-IRAOLA, P., YOUNIS, M., KRIEGER, G., HAJNSEK, I. et PAPATHANASSIOU, K. P. (2013). *A Tutorial on Synthetic Aperture Radar*. IEEE Geoscience and Remote Sensing Magazine, 1(1):6–43. 92

MUNSON, D., O'BRIEN, J. et JENKINS, W. (1983). *A tomographic formulation of spotlight-mode synthetic aperture radar*. Proceedings of the IEEE, 71(8):917–925. 50, 51

MUNSON, D. et VISENTIN, R. (1989). *A Signal Processing View of Strip-Mapping Synthetic Aperture Radar*. IEEE Transactions on Acoustics, Speech, and Signal Processing, 37(12): 2131–2147. 50

MUSICUS, B. R. (1988). *Levinson and Fast Cholesky Algorithms for Toeplitz and Almost Toeplitz Matrice*. Rapport technique RLE n°538, Research Laboratory of Electronics Massachusetts Institute of Technology. 107

NATARAJAN, B. K. (1995). *Sparse Approximate Solutions to Linear Systems*. SIAM Journal on Computing, 24(2):227–234. 63

NESTEROV, Y. (1983). *A Method of Solving a Convex Programming Poblem with Convergence Rate O(1/k²)*. Soviet Mathematics Doklady, 27(2):372–376. 71, 72

NESTEROV, Y. (2005). *Smooth Minimization of Non-Smooth Functions*. Mathematical Programming, 103(1):127–152. 67

NIKOLOVA, M. (2004). *A Variational Approach to Remove Outliers and Impulse Noise*. Journal of Mathematical Imaging and Vision, 20(1/2):99–120. 56

NITHYA, T. et LEELA LAKSHMI, S. (2014). *Performance of Alternating Direction Method of Multipliers (ADMM) Based on Deconvolving Images with Unknown Boundaries*. 2(8): 5449–5456. 86

NOUVEL, J.-F., HERIQUE, A., KOFMAN, W. et SAFAEINILI, A. (2004). *Radar signal simulation : Surface modeling with the Facet Method : Radar Signal Simulation*. Radio Science, 39(1):n/a–n/a. 22

NYQUIST, H. (1928). *Certain Topics in Telegraph Transmission Theory*. Transactions of the American Institute of Electrical Engineers, 47(2):617–644. 160

ODENDAAL, J. et JOUBERT, J. (1996). *Radar Cross Section Measurements Using Near-Field Radar Imaging*. IEEE Transactions on Instrumentation and Measurement, 45(6):948–954. 147

O'HARA, R. B. et SILLANPÄÄ, M. J. (2009). *A review of Bayesian variable selection methods : what, how and which*. Bayesian Analysis, 4(1):85–117. 143

OLDENBURG, D. W., SCHEUER, T. et LEVY, S. (1983). *Recovery of the Acoustic Impedance from Reflection Seismograms*. GEOPHYSICS, 48(10):1318–1337. 62

ONHON, N. et CETIN, M. (2012). *A Sparsity-Driven Approach for Joint SAR Imaging and Phase Error Correction*. IEEE Transactions on Image Processing, 21(4):2075–2088. 100

ORIEUX, F., GIOVANNELLI, J.-F., RODET, T. et ABERGEL, A. (2013). *Estimating Hyperparameters and Instrument Parameters in Regularized Inversion Illustration for Herschel /SPIRE Map Making*. Astronomy & Astrophysics, 549:A83. 143

ORMEROD, J. T., YOU, C. et MÜLLER, S. (2017). *A Variational Bayes Approach to Variable Selection*. Electronic Journal of Statistics, 11(2):3549–3594. 143

PASTORINO, M. (2015). *Electromagnetic Imaging : Methods and Applications*. IEEE MMS, pp. 1–4. 48

PEICHL, M., KEMPF, T. et DILL, S. (2008). *Radar Signature Analysis of Urban Structures*. IEEE IGARSS, pp. V – 522–V – 525. 27

PENROSE, R. et TODD, J. A. (1955). *A Generalized Inverse for Matrices*. Mathematical Proceedings of the Cambridge Philosophical Society, 51(03):406. 55

PETERSEN, K. B. et PEDERSEN, M. S. (2012). *The Matrix Cookbook*. 171

PISARENKO, V. F. (1973). *The Retrieval of Harmonics from a Covariance Function*. Geophysical Journal International, 33(3):347–366. 57

POLIYAK, B. T. (1987). *Introduction to Optimization*. Translations series in mathematics and engineering. Optimization Software, Publications Division. 68

POTTER, L. C., ERTIN, E., PARKER, J. T. et CETIN, M. (2010). *Sparsity and Compressed Sensing in Radar Imaging*. Proceedings of the IEEE, 98(6):1006–1020. 58

POULALION, G. (1997). *Détection et Éstimation Conjointes en Analyse Spectrale Haute Résolution - Application à l'Analyse de Signaux Issus de Mesures de Surface Équivalente Radar*. Thèse de doctorat, Université de Bordeaux. 58

RAHMAN, M. et KAI-BOR YU (1987). *Total Least Squares Approach for Frequency Estimation Using Linear Prediction*. IEEE Transactions on Acoustics, Speech, and Signal Processing, 35(10):1440–1454. 57

RASHID, L. et BROWN, A. K. (2011). *Radar Cross-Section Analysis of Wind Turbine Blades with Radar Absorbing Materials*. pp. 97–100. 29

REN, X., QIAO, L., QIN, Y. et LI, P. (2016). *Sparse Regularization Based Imaging Method for Inverse Synthetic Aperture Radar*. IEEE PIERS, pp. 4348–4351. 59

RITCHIE, M., FIORANELLI, F., GRIFFITHS, H. et TORVIK, B. (2015). *Micro-Drone RCS Analysis*. IEEE RADAR conf., pp. 452–456. 28

ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. 77, 183

ROS, J.-M. (1994). *Contribution au Développement des Méthodes d'Analyse Spectrale Haute Résolution et Temps-Fréquence pour l'Analyse de Signaux de Mesures de Surface Équivalente Radar (SER) Non Stationnaires*. Thèse de doctorat, Université de Bordeaux. 49

ROY, R., GOLDBURG, M., OTTERSTEN, B., SWINDLEHURST, A. L., VIBERG, M. et KAILATH, T. (1989). *ESPRIT And Uniform Linear Arrays*. p. 370. 58

ROY, R., PAULRAJ, A. et KAILATH, T. (1986). *ESPRIT - A Subspace Rotation Approach to Estimation of Parameters of Cisoids in Noise*. IEEE Transactions on Acoustics, Speech, and Signal Processing, 34(5):1340–1342. 58

RUCK, G. T., BARRICK, D. E., STUART, W. D. et KRICHBAUM, C. K. (1970). *Radar Cross Section Handbook*. 29

RUDIN, L. I., OSHER, S. et FATEMI, E. (1992). *Nonlinear Total Variation Based Noise Removal Algorithms*. Physica D : Nonlinear Phenomena, 60(1-4):259–268. 62, 64

SAHNOUN, S. (2012). *Développement de Méthodes d'Éstimation Modale de Signaux Multidimensionnels. Application à la Spectroscopie RMN Multidimensionnelle.* Thèse de doctorat, Université de Lorraine. 57, 58

SANDERS, T., GELB, A. et PLATTE, R. B. (2017). *Composite SAR Imaging Using Sequential Joint Sparsity*. Journal of Computational Physics, 338:357–370. 100

SBAI, S. M. A. (2013). *Traitement des Signaux Parcimonieux et Applications*. Thèse de doctorat. 62

SCHMIDT, R. (1986). *Multiple Emitter Location and Signal Parameter Estimation*. IEEE Transactions on Antennas and Propagation, 34(3):276–280. 57

SHANG, Y. et SHEN, Z. (2017). *Radar Cross-Section Enhancement Techniques*. IEEE COMPEM, pp. 336–337. 30

SHANG, Y., XIAO, S. et SHEN, Z. (2015). *Edge-On Backscattering Enhancement Based on Quasi-Superdirective Reradiation*. IEEE Antennas and Wireless Propagation Letters, 14:539–542. 30

SHANNON, C. (1949). *Communication in the Presence of Noise*. Proceedings of the IRE, 37(1):10–21. 160

SIDIROPOULOS, N. D., GINI, F., CHELLAPPA, R. et THEODORIDIS, S. (2014). *Communications and Radar Signal Processing*. Academic Press, (volume 2). 48

SIMON, J. et SOUDAIS, P. (2014). *Workshops EM ISAE 2014 et 2016. ONERA*. 125

SITKO, M. (2013). *Imagerie SAR-3D à Données Incomplètes en Mode Polarisé*. Rapport de stage, CEA CESTA, ENSEIRB Matmeca. 42

SKOLNIK, M. I., éditeur (1990). *Radar handbook*. McGraw-Hill, 2nd ed Éd. 29

SLATER, M. (1950). *Lagrange Multipliers Revisited : A Contribution to Non-Linear Programming*. Cowles Foundation For Research In Economics. 76

SOERGEL, U., éditeur (2010). *Radar remote sensing of urban areas*. Num. v. 15 de Remote Sensing and Digital Image Processing. Springer. 27

SOUBIES, E., BLANC-FÉRAUD, L. et AUBERT, G. (2017). *Quelles Relaxations Continues pour le Critère L2-L0 ?* 63

SOUSSEN, C., IDIER, J., BRIE, D. et DUAN, J. (2011). *From Bernoulli–Gaussian Deconvolution to Sparse Signal Restoration*. IEEE Transactions on Signal Processing, 59(10):4572–4584. 63

STANKWITZ, H., DALLAIRE, R. et FIENUP, J. (1995). *Nonlinear Apodization for Sidelobe Control in SAR Imagery*. IEEE Transactions on Aerospace and Electronic Systems, 31(1): 267–279. 52

STONE, W. R. (1989). *Special Issue on Radar Cross Sections of Complex Objects*. IEEE Trans. on Antennas and Propagation, AP-37(5). 29

STORATH, M., WEINMANN, A. et DEMARET, L. (2014). *Jump-Sparse and Sparse Recovery Using Potts Functionals*. IEEE Transactions on Signal Processing, 62(14):3654–3666. 143

STUFF, M., BIANCALANA, M., ARNOLD, G. et GARBARINO, J. (2004). *Imaging Moving Objects in 3D from Single Aperture Synthetic Aperture Radar*. IEEE NRC, pp. 94–98. 57

STUPFEL, B., LE MARTRER, R., BONNEMASON, P. et SCHEURER, B. (1991). *Combined Boundary-Element and Finite-Element Method for Scattering Problem by Axi-Symmetrical Penetrable Objects*. Mathematical and numerical aspects of wave propagation phenomena, pp. 332–341. 131

STUTZMAN, W. L. et THIELE, G. A. (2013). *Antenna Theory and Design*. Wiley, 3rd ed Éd. 147

TAO, S., BOLEY, D. et ZHANG, S. (2015). *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence : Buenos Aires, Argentina, 25-31 July 2015*. 86

TARANTOLA, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics Philadelphia. 13

TAYLOR, H. L., BANKS, S. C. et MCCOY, J. F. (1979). *Deconvolution with the L1-Norm*. GEOPHYSICS, 44(1):39–52. 62

TIBSHIRANI, R. (1996). *Regression Shrinkage and Selection via the Lasso*. Wiley for the Royal Statistical Society, 58(1):267–288. 62

TIBSHIRANI, R. (2011). *Regression Shrinkage and Selection via the Lasso : a Retrospective*. Journal of the Royal Statistical Society : Series B (Statistical Methodology), 73(3):273–282. 64

TIBSHIRANI, R., SAUNDERS, M., ROSSET, S., ZHU, J. et KNIGHT, K. (2005). *Sparsity and Smoothness via the Fused Lasso*. Journal of the Royal Statistical Society : Series B (Statistical Methodology), 67(1):91–108. 64

TIKHONOV, A. N. et ARSENIN, V. I. (1977). *Solutions of Ill-Posed Problems*. Scripta series in mathematics. Winston, Halsted Press. 56

TON, C. (2013). *Radar Cross Section Simulation for Wind Turbines*. Thèse de doctorat, Naval Postgraduate School. 29

TRIA, M. (2005). *Imagerie Radar à Synthèse d'Ouverture (RSO) par Analyse en Ondelettes Continues Multidimensionnelles.* Thèse de doctorat, Université Paris-Sud XI - Faculté des sciences d'Orsay. 39

TSAIG, Y. et DONOHO, D. L. (2006). *Extensions of Compressed Sensing*. Signal Processing, 86(3):549–571. 62

TSAO, J. et STEINBERG, B. (1988). *Reduction of Sidelobe and Speckle Artifacts in Microwave Imaging : the CLEAN Technique*. IEEE Transactions on Antennas and Propagation, 36(4):543–556. 57

TUPIN, F. (2007). *Champs de Markov sur graphes pour le traitement des images radar*. Habilitation à Diriger des Recherches, École nationale supérieure des télécommunications. 144

TUPIN, F., HOUSHMAND, B. et DATCU, M. (2002). *Road Detection in Dense Urban Areas Using SAR Imagery and the Usefulness of Multiple Views*. IEEE Transactions on Geoscience and Remote Sensing, 40(11):2405–2414. 27

UBERALL, H. (1964). *Radar Scattering from Coated Perfect Conductors : Application to the Semi- Infinite Cone and Use of Exact Eikonal*. 68D(6):749–766. 37

URAGO, S., BERTHOD, M. et ZERUBIA, J. (1992). *Restauration d'Image de Contours Incomplets par Modelisation par Champs de Markov*. Rapport de recherche RR-1688, Institut National de Recherche en Informatique et en Automatique (INRIA) Nice. 64

VEGA, D., JENN, D., ANGULO, I. et GUERRA, D. (2016). *Simplified Characterization of Radar Cross Section of Wind Turbines in the Air Surveillance Radars Band*. IEEE EuCAP, pp. 1–5. 29

VIGNAUD, L. (1996). *Imagerie Micro-Ondes des Scènes Instationnaires*. Thèse de doctorat, Université Paris VI. 49

VOGEL, C. R. et OMAN, M. E. (1996). *Iterative Methods for Total Variation Denoising*. SIAM Journal on Scientific Computing, 17(1):227–238. 67

WEISS, P. (2008). *Algorithmes Rapides d'Optimisation Convexe. Applications à la Reconstruction d'Images et à la Détection de Changements.* Thèse de doctorat, Université Nice Sophia Antipolis. 69

WEISS, P., AUBERT, G. et BLANC-FÉRAUD, L. (2006). *Some Applications of L-infinite Constraints in Image Processing*. Rapport de recherche 6115, Institut National de Recherche en Informatique et en Automatique. 56

WEISS, P., BLANC-FÉRAUD, L. et AUBERT, G. (2009). *Efficient Schemes for Total Variation Minimization Under Constraints in Image Processing*. SIAM Journal on Scientific Computing, 31(3):2047–2080. 67

WELLER, D. S., RAMANI, S., NIELSEN, J.-F. et FESSLER, J. A. (2014). *Monte Carlo SURE-Based Parameter Selection for Parallel Magnetic Resonance Imaging Reconstruction*. Magnetic Resonance in Medicine, 71(5):1760–1770. 143

WHITE, M. (1998). *Radar cross-section : measurement, prediction and control*. Electronics & Communication Engineering Journal, 10(4):169–180. 29

WOOD, W. (1998). *An open-cavity RCS enhancement device for airborne applications*. IEEE APS, 4:2190–2193. 30

WRIGHT, S., NOWAK, R. et FIGUEIREDO, M. (2009). *Sparse Reconstruction by Separable Approximation*. IEEE Transactions on Signal Processing, 57(7):2479–2493. 186

XIAO-CHUN, X. (2010). *Application of Compressive Sensing in 3D Radar Imaging*. IEEE ICOSP, pp. 2117–2120. 13, 58

XU, Z., FIGUEIREDO, M. A. T. et GOLDSTEIN, T. (2017). *Adaptive ADMM with Spectral Penalty Parameter Selection*. Aarti Singh & Jerry Zhu PMLR, 54:718–727. 86

YIN, W., OSHER, S., GOLDFARB, D. et DARBON, J. (2008). *Bregman Iterative Algorithms for L1-Minimization with Applications to Compressed Sensing*. SIAM Journal on Imaging Sciences, 1(1):143–168. 83

YU, C., YUAN, M. et LIU, Q. H. (2009). *Reconstruction of 3D Objects from Multi-Frequency Experimental Data with a Fast DBIM-BCGS Method*. Inverse Problems, 25(2):024007. 48

ÖZDEMIR, C. (2012). *Inverse Synthetic Aperture Radar Imaging with MATLAB Algorithms*. Num. 210 de Wiley series in microwave and optical engineering. Wiley. 36, 41, 45, 50, 51, 52

ZHU, S. (2013). *A Bayesian Approach for Inverse Problems in Synthetic Aperture Radar Imaging*. Thèse de doctorat, Université Paris-Sud XI - Faculté des sciences d'Orsay. 13, 59