



HAL
open science

Algorithmes d'approximation parcimonieuse inspirés d'Orthogonal Least Squares pour les problèmes inverses

Charles Soussen

► **To cite this version:**

Charles Soussen. Algorithmes d'approximation parcimonieuse inspirés d'Orthogonal Least Squares pour les problèmes inverses. Traitement du signal et de l'image [eess.SP]. Université de Lorraine, 2013. tel-00924578v2

HAL Id: tel-00924578

<https://theses.hal.science/tel-00924578v2>

Submitted on 8 Feb 2014

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.



Algorithmes d'approximation parcimonieuse inspirés d'Orthogonal Least Squares pour les problèmes inverses

Mémoire de recherche

présenté publiquement le 28 novembre 2013, en vue de l'obtention de

l'Habilitation à Diriger des Recherches de l'Université de Lorraine

(mention Automatique, Traitement du Signal et des Images, Génie Informatique)

par

Charles SOUSSEN

Maître de Conférences à l'Université de Lorraine

Docteur de l'Université Paris-Sud

Composition du jury :

<i>Président :</i>	Cédric RICHARD	Professeur à l'Université de Nice-Sophia Antipolis
<i>Rapporteurs :</i>	Christian JUTTEN	Professeur à l'Université Joseph Fourier, Grenoble
	Gabriel PEYRÉ	CR CNRS, Cérémade, Université Paris-Dauphine
	Pierre VANDERGHEYNST	Professeur associé à l'École Polytechnique Fédérale de Lausanne
<i>Examineurs :</i>	Hoai An LE THI	Professeur à l'Université de Lorraine
	Christian MUSTIN	DR CNRS. LIEC, Université de Lorraine
	Alain RICHARD	Professeur à l'Université de Lorraine

Centre de Recherche en Automatique de Nancy

CRAN, UMR 7039, Université de Lorraine, CNRS

Faculté des Sciences et Technologies, BP 70239, 54506 Vandœuvre Cedex

Tél: 03 83 68 44 71 — Fax: 03 83 68 44 62 — Courriel: Charles.Soussen@univ-lorraine.fr

Mis en page avec la classe thloria.

Remerciements

Avant tout, je tiens à remercier Didier Wolf et Alain Richard, actuel et ancien directeurs du CRAN pour leur accueil au laboratoire. Je remercie Walter Blondel et Muriel Barberi, co-responsables du département SBS du CRAN, pour leur animation scientifique, leur disponibilité et leur soutien.

Je remercie Alain Richard pour avoir accepté de parrainer ce travail, pour sa disponibilité constante malgré ses responsabilités, pour sa curiosité, son recul et ses conseils avisés. Au delà de l'HdR, merci d'avoir été à mon écoute à des moments clés depuis mon arrivée au CRAN en 2005.

Je remercie Christian Jutten, Gabriel Peyré et Pierre Vanderghenst pour s'être intéressés à mon travail et avoir accepté d'en être les rapporteurs. Je voudrais tout particulièrement souligner leur grande disponibilité. Je remercie vivement Christian Jutten d'avoir annulé une réunion parisienne et l'avoir effectuée par visio-conférence pour pouvoir participer à la fois à la thèse de Simon Henrot et à ma soutenance d'HdR. J'ai beaucoup apprécié que suite à un impondérable, Pierre Vanderghenst ait spontanément proposé de faire le trajet Lausanne-Nancy en voiture pour être présent physiquement à ma soutenance. Enfin, je tiens à remercier Gabriel Peyré pour m'avoir invité à faire un séminaire au Cérémade, pour son accueil chaleureux et celui de son équipe, et les échanges que nous avons eu à cette occasion.

Je remercie vivement Cédric Richard pour avoir présidé le jury et accepté, malgré un emploi du temps chargé, de rester deux jours à Nancy pour participer à la thèse de Simon Henrot et à mon HdR.

Je remercie Hoai An Le Thi pour s'être intéressée à mon travail et pour m'avoir invité à une soutenance de thèse à Metz prochainement. J'espère sincèrement que ces échanges déboucheront sur une collaboration fructueuse.

Je remercie Christian Mustin pour s'être intéressé à ce travail sur une thématique pourtant « exotique ». Merci de m'avoir intégré dans le projet ANR HÆSPRI, pour ta gentillesse, et pour les discussions que nous avons eues depuis 2009. Je dois souligner ta curiosité scientifique remarquable et ta démarche « statistique » et rigoureuse, sans concession avec les analyses superficielles.

Je tiens à remercier tout particulièrement Jérôme Idier pour son amitié, son soutien constant, ses conseils avisés, sa rigueur scientifique, et pour toutes les bonnes pistes qu'il m'a suggéré de suivre et que nous avons partagées. Je lui dois beaucoup, et ce manuscrit aussi. Je suis très reconnaissant à l'IRCCyN et à son directeur, Michel Malabre, de m'avoir fait confiance en m'accueillant en délégation CNRS pendant l'année universitaire 2010-2011. Je remercie l'ensemble de l'équipe ADTSI pour leur accueil si chaleureux au cours de mes déplacements à Nantes, et en particulier les permanents : Sébastien Bourguignon, Eric Le Carpentier, Marie-Françoise Lucas et Saïd Moussaoui.

Je remercie Cédric Herzet et Rémi Gribonval pour notre collaboration aussi fructueuse, pour le dynamisme qu'ils m'ont communiqué, leur rigueur scientifique, et leur état d'esprit remarquable. C'est une chance de vous avoir rencontré. Sans vous, les chapitres 3 et 4.2 ne seraient pas ce qu'ils sont.

Je remercie David Brie pour m'avoir accueilli au sein du G.T. IRIS du CRAN, pour avoir eu l'intuition de projets novateurs et de m'y avoir associé. Je remercie El-Hadi Djermoune et Sebastian Miron pour nos échanges scientifiques et les moments amicaux partagés aux époques du G.T. IRIS et de Simul. Je remercie Christian Daul et Walter Blondel pour m'avoir intégré dans le G.T. IPS et sur le site de l'ENSEM, et pour les bons moments que nous avons partagé.

Je remercie Fabien Gaboriaud et Grégory Francius pour avoir initié le projet sur la microscopie

pie AFM, et tous les autres acteurs du projet Bioforce : Angéline Razafitianamaharavo, David Brie, Stéphanie Grandemange, Claire Barbieux et enfin Rémi Pannequin pour son investissement substantiel dans le projet de développement logiciel.

Merci à Vincent Mazet pour avoir été l'instigateur du projet SpectroDec, pour son dynamisme et son enthousiasme. J'espère que nous aurons à nouveau l'occasion de travailler ensemble.

Je remercie tous les collègues du CRAN que j'ai fréquenté sur les deux sites de la faculté des sciences et technologies (FST) et de l'ENSEM, avec une pensée particulière pour les automaticiens du 4ème étage de la FST et du couloir jaune de l'ENSEM. Je remercie Jean-Christophe Ponsart et Frédéric Hamelin pour nos nombreuses discussions et pour leurs conseils avisés. Merci à Marc Jungers pour m'avoir fait partager son expérience de l'HdR quelques mois plus tôt (et de ses méandres administratifs...), pour sa disponibilité et son aide logistique très appréciables le jour J, malgré le décalage horaire.

J'associe naturellement à ces remerciements les docteurs que j'ai co-encadré : Junbo Duan, Achraf Ben Hamadou et Simon Henrot qui sont des acteurs importants de la recherche. Et parmi ceux que je n'ai pas encadré, une pensée particulière pour Xijing Guo, que j'ai eu la chance de connaître lors de son trop court séjour à Nancy.

Je remercie tout particulièrement Sabine Hureau pour s'être occupée de mes missions, pour sa disponibilité, sa bonne humeur et pour m'avoir souvent dépanné. Merci à Christelle Kondratow pour avoir si bien géré le projet Bioforce, et à Olivier Cervellin pour sa réactivité et ses dépannages informatiques.

Je remercie enfin tous les anciens du Groupe Problèmes Inverses du L2S pour les relations que nous avons entretenues. Une mention spéciale pour Christian Heinrich et Hervé Carfantan pour leur relecture détaillée du chapitre 2 de ce manuscrit et quelques suggestions bibliographiques très pertinentes. Merci à Myriam Nouvel pour s'être déplacée depuis Paris pour assister à ma soutenance, et pour les belles photos qu'elle a prises.

Je remercie mes amis Dominique (pour s'être aussi déplacée!), David, Julien et Christophe, Christophe et Julien, Stéphane K. et surtout Stéphane O. pour avoir supporté mes états d'âmes pendant si longtemps.

Merci au groupe du jeudi soir de la Cie « Ca Respire Encore » (Daniel, Bérengère, Philippe, Marie-Hélène, Sylvain, Sylvia, Marie C., Marie E., François, Alain, Mathilde, Michèle) et de l'Amature à Nantes sans qui je n'aurais jamais connu Koltès, Pinter, Tchekhov, Mouawad, Pommerat, Lagarce, *etc.* et pour m'avoir fait oublier mes soucis. Merci pour tous ces moments forts partagés sur scène et en dehors.

Merci enfin à mes parents pour leur soutien indéfectible.

Table des matières

Organisation du document

xiii

Partie I Présentation générale

<i>Curriculum Vitæ</i> détaillé	3
1 Etat civil	3
2 Situation actuelle	3
3 Titres universitaires	3
4 Parcours	3
5 Activité d'enseignement	4
5.1 Enseignements dispensés	4
5.2 Responsabilités et mise en place d'enseignements	4
5.3 Responsabilité des stages	5
5.4 Vacations externes	5
6 Activité d'encadrement	5
6.1 Encadrement doctoral	5
6.2 Encadrement de stagiaires de Master	6
7 Contrats de recherche	6
7.1 Analyse d'images force volume	6
7.2 Bioforce : analyse d'images PeakForce Mapping de tissus biologiques .	7
7.3 HÆSPRI : analyse hyperspectrale des interactions bactérie-minéral .	7
7.4 SpectroDec : décomposition parcimonieuse de signaux spectroscopiques	7
8 Activités de valorisation et de transfert	7
8.1 Logiciel d'analyse d'images force-volume	7
8.2 Mosaïquage d'images endoscopiques de la vessie	8
8.3 Logiciel déposé : optimisation locale sous Matlab	8
8.4 Code source en accès libre	8
8.5 Contrat industriel	8

9	Principaux collaborateurs extérieurs au CRAN	8
10	Activités liées à l'administration	8
11	Activités liées à la recherche	9
12	Liste de publications	9
Synthèse générale de mon activité de recherche		15
1	Reconstruction tomographique (2001-2007)	15
1.1	Travail de thèse : reconstruction d'un objet polyédrique	15
1.2	Modèle volumique parcimonieux	16
2	Problèmes inverses et traitement de signaux tridimensionnels	17
2.1	Microscopie AFM	17
2.2	Microscopie confocale de fluorescence	19
2.3	Conception d'algorithmes de restauration de signaux parcimonieux	22
3	Analyse en k itérations des algorithmes OMP et OLS	25
3.1	Contexte et cheminement du projet	25
3.2	Contributions	26
4	Recalage et reconstruction d'images 3D	27

Partie II Algorithmes d'approximation parcimonieuse pour les problèmes inverses

Avant-propos	33
Liste des notations	35

<p>Chapitre 1 Algorithmes gloutons bidirectionnels pour la minimisation de critères mixtes ℓ_2-ℓ_0 37</p>

1.1	Approximation parcimonieuse de signaux	37
1.2	Problèmes inverses mal conditionnés	37
1.3	Problème des moindres carrés sous contrainte ℓ_0	38
1.4	Algorithmes de poursuite gloutons orthogonaux	40
1.5	Formulation bayésienne basée sur le modèle Bernoulli-gaussien	43
1.6	Minimisation mixte ℓ_2 - ℓ_0 pour un niveau de parcimonie donné	44
1.6.1	Méthode de descente pour l'optimisation combinatoire	44
1.6.2	Adaptation de l'algorithme SMLR : algorithme Single Best Replacement	45
1.6.3	SBR en tant qu'algorithme de minimisation locale	46

1.6.4	Exemple de la déconvolution impulsionnelle	47
1.6.5	Exemple de la segmentation d'un signal	49
1.7	Minimisation mixte ℓ_2 - ℓ_0 pour des niveaux de parcimonie variés	50
1.7.1	Connexion avec l'optimisation bi-objectif	51
1.7.2	Idées sous-jacentes aux algorithmes proposés	53
1.7.3	Algorithme Continuation Single Best Replacement	54
1.7.4	Algorithme de reconstruction du chemin de régularisation ℓ_0	58
1.7.5	Evaluation numérique des algorithmes proposés	61
1.8	Conclusion	66

Chapitre 2

Analyse d'images force-volume en microscopie AFM 67

2.1	Introduction à la microscopie de force atomique	68
2.1.1	Caractérisation de cellules biologiques	68
2.1.2	Fonctionnement d'un microscope AFM	68
2.1.3	Modèles physiques par morceaux	72
2.2	Traitement de données en spectroscopie AFM	75
2.2.1	Objectifs et méthodologie de traitement	75
2.2.2	Segmentation d'une courbe de force par approximation parcimonieuse	76
2.3	Approximation parcimonieuse pour la segmentation	78
2.3.1	Reformulation de la détection conjointe de discontinuités	78
2.3.2	Construction du dictionnaire	79
2.3.3	Sélection de variables scalaires	80
2.3.4	Sélection de variables vectorielles	80
2.3.5	Difficultés techniques et aspects implémentation	81
2.4	Traitement de données réelles	82
2.4.1	Segmentation d'une courbe de retrait : comparaison de stratégies	82
2.4.2	Traitement complet d'une courbe de retrait	84
2.4.3	Traitement d'une image force-volume dans la phase d'approche	85
2.5	Conclusion	85

Chapitre 3

Analyse en k itérations des algorithmes OMP et OLS 87

3.1	Motivations	87
3.2	Analyse en k itérations	88
3.2.1	Définition	88
3.2.2	Extension bruitée	89

3.2.3	Hypothèses de travail	89
3.3	Différence entre OMP et OLS : considérations géométriques	89
3.4	Analyse au pire cas	92
3.5	Reconstruction exacte par OLS	93
3.6	Reconstruction exacte par OMP et OLS à partir d'une connaissance partielle du support	93
3.6.1	Conditions pour un support initial \mathcal{Q} fixé	94
3.6.2	Conditions pour une itération q fixée	94
3.6.3	Cas particulier $q = 0$	95
3.7	Evaluation empirique des conditions partielles de reconstruction	95
3.8	Garantie de non-reconstruction pour OMP	97
3.9	Transition de phase pour la non-reconstruction	99
3.10	Conclusions	100

Chapitre 4

Quelques perspectives à court et moyen terme

103

4.1	Développement d'algorithmes d'approximation parcimonieuse	103
4.1.1	Parcimonie et contrainte de positivité	104
4.1.2	Algorithmes plus efficaces que SBR pour la minimisation $\ell_2\text{-}\ell_0$	107
4.1.3	Algorithmes hybrides pour la décomposition en motifs paramétriques	108
4.1.4	Décomposition conjointe de signaux en motifs	109
4.2	Analyse d'algorithmes d'approximation parcimonieuse	110
4.2.1	Bilan et positionnement par rapport à la littérature	110
4.2.2	Discrimination OMP <i>vs</i> OLS	111
4.2.3	Forme faible des conditions ERC-Oxx	113
4.2.4	Analyse approfondie des conditions de non-reconstruction	114
4.2.5	Analyse d'algorithmes bidirectionnels	114
4.3	Analyse de l'invasivité de cellules cancéreuses par microscopie AFM	116
4.3.1	Bilan et contexte	116
4.3.2	Analyse de cellules cancéreuses	117
4.3.3	Traitement du signal	118
4.3.4	Liens avec d'autres projets pour le diagnostic du cancer	120

Annexe A Intérêt de la microscopie AFM pour le diagnostic de tumeurs	121
---	------------

Index	123
--------------	------------

Table des figures

1	Courbes de force en microscopie AFM	18
2	Microscopie confocale et imagerie hyperspectrale	20
3	Prototype de système à lumière structurée	29
1.1	Déconvolution impulsionnelle	47
1.2	Compromis temps de calcul <i>vs</i> performance	48
1.3	Comparaison d'algorithmes pour la segmentation d'un signal	50
1.4	Minimisation de critères ℓ_2 - ℓ_0 en tant que problème d'optimisation bi-objectif	52
1.5	Courbe ℓ_0 caractérisant le chemin de régularisation ℓ_2 - ℓ_0	53
1.6	Interprétation graphique des algorithmes SBR et CSBR.	55
1.7	Paramétrisation du chemin de régularisation ℓ_2 - ℓ_0 estimé	59
1.8	Algorithme ℓ_0 -PT : interprétation graphique	60
1.9	Comparaison CSBR <i>vs</i> SBR	62
1.10	Comparaison des méthodes de continuation	63
1.11	Fonctionnements empiriques de CSBR et ℓ_0 -PT	64
2.1	Spectroscopie de force	69
2.2	Courbe de force schématique	70
2.3	Courbes de forces expérimentales relatives à des bactéries	71
2.4	Imagerie force-volume	72
2.5	Modèles par morceaux pour les courbes de force	73
2.6	Modèles FJC et WLC pour l'étirement de polymères	74
2.7	Dictionnaire polynomial	80
2.8	Lissage par morceaux d'une courbe de force	83
2.9	Traitement complet d'une courbe de retrait	84
2.10	Traitement d'une image force-volume dans la phase d'approche	85
3.1	Illustration graphique des algorithmes OMP et OLS	91
3.2	Comparaison des conditions ERC-OMP/OLS pour un dictionnaire convolutif	96
3.3	Garantie de non-reconstruction d'un support par OMP	99
3.4	Transitions de phase pour OMP, <i>basis pursuit</i> et OLS	101
4.1	Séquence de spectres	108
4.2	Déconvolution de données acoustiques simulées	112
4.3	Modèle de courbe de force pour une cellule biologique	118
4.4	Prise en compte de la géométrie de la pointe AFM	120

Liste des tableaux

1.1	Algorithme <i>Matching Pursuit</i>	41
1.2	Algorithme <i>Orthogonal Matching Pursuit</i>	42
1.3	Algorithme <i>Orthogonal Least Squares</i>	42
1.4	Algorithme <i>Single Best Replacement</i>	45
1.5	Algorithme SBR adapté en vue des appels dans <i>Continuation SBR</i>	57
1.6	Algorithme <i>Continuation SBR</i>	58
1.7	Algorithme ℓ_0 <i>regularization path track</i>	61
3.1	Présentation unifiée des algorithmes OMP et OLS	92

Organisation du document

Le travail présenté dans ce manuscrit a été effectué au Centre de Recherche en Automatique de Nancy (CRAN, UMR 7039, CNRS, Université de Lorraine) depuis ma mutation au CRAN en septembre 2005. Mes projets de recherche ont été principalement menés au sein du groupe thématique *Identification, Restauration, Images et Signaux* (G.T. IRIS) jusqu'en 2012 et mineurement dans le groupe *Ingénierie Pour la Santé* (G.T. IPS), puis entièrement dans le département *Santé, Biologie, Signal* (SBS) créé en janvier 2013 qui regroupe, pour le traitement du signal et de l'image, les chercheurs provenant des deux anciens groupes thématiques.

Les projets de recherche que j'ai menés s'inscrivent dans les domaines des **problèmes inverses** en traitement du signal et des images, de l'**approximation parcimonieuse**, du **traitement d'images hyperspectrales** et de la **reconstruction d'images 3D** en endoscopie à lumière structurée. Dans ce manuscrit, j'ai choisi de privilégier mon travail sur le développement et l'analyse d'algorithmes d'approximation parcimonieuse pour les problèmes inverses. C'est un sujet sur lequel je me suis plus particulièrement investi, notamment lors de mon séjour de délégation CNRS à l'IRCCyN (Nantes) durant l'année universitaire 2010-2011.

Ce document est organisé en deux parties. La première comporte mon *curriculum vitae* détaillé et une synthèse générale de mon activité de recherche. La deuxième partie est dédiée aux algorithmes d'approximation parcimonieuse. Elle comporte quatre chapitres.

Chapitre 1

Le premier chapitre aborde le problème d'approximation parcimonieuse comme la minimisation d'un critère mixte appelé « $\ell_2\text{-}\ell_0$ ». Ce critère est la somme pondérée de deux termes décrivant l'erreur d'approximation (terme ℓ_2) et le degré de parcimonie du modèle (terme ℓ_0). Ce problème d'optimisation est connu pour être NP-complet ; cependant nous proposons de développer des algorithmes sous-optimaux efficaces et de montrer leur pertinence pour les problèmes inverses caractérisés par un dictionnaire très mal conditionné, dont les (certains) atomes sont extrêmement corrélés. De nombreux algorithmes ont été proposés dans la littérature mais un grand nombre d'entre eux ne donnent pas satisfaction pour des dictionnaires mal conditionnés. Nous nous focalisons d'abord sur les algorithmes dits « gloutons » monodirectionnels qui sélectionnent un à un des atomes du dictionnaire pour raffiner l'approximation des données. Le point de départ de ce travail est le constat empirique que l'algorithme *Orthogonal Least Squares* (OLS) est mieux adapté que l'algorithme *Orthogonal Matching Pursuit* (OMP) pour traiter des problèmes inverses mal conditionnés, avec un coût calculatoire certes plus important. En collaboration avec l'IRCCyN, nous avons proposé des extensions bidirectionnelles d'OLS (algorithmes itératifs qui procèdent non seulement à des ajouts mais aussi à des retraits d'atomes) permettant d'une part, d'améliorer sensiblement les performances d'OLS pour les problèmes inverses et d'autre part, d'estimer un chemin de régularisation pour le problème $\ell_2\text{-}\ell_0$, c'est-à-dire de

réaliser conjointement des estimations à des degrés de parcimonie variables.

Chapitre 2

Le deuxième chapitre est une application en microscopie de force atomique (AFM) pour l'imagerie à l'échelle nanométrique de bactéries présentes dans l'environnement. Ce travail a pour cadre une collaboration avec le Laboratoire de Chimie Physique et Microbiologie pour l'Environnement (LCPME, UMR 7564, CNRS, Université de Lorraine). La microscopie AFM est une technologie récente permettant de caractériser les propriétés physico-chimiques d'objets et en particulier d'échantillons biologiques à l'échelle nanométrique. Dans ce chapitre, je souligne la pertinence des méthodes d'approximation parcimonieuse pour segmenter les signaux mesurés en microscopie AFM, c'est-à-dire rechercher les points de discontinuité (sauts, changement de pente ou de courbure) du signal. La solution proposée repose sur le lissage par morceaux du signal et la détermination automatique des morceaux, délimités par deux points de discontinuité consécutifs. Cette solution est obtenue en appliquant les algorithmes d'approximation parcimonieuse du chapitre 1 à un dictionnaire particulier (polynomial) dont chaque atome correspond à l'activation d'un point de discontinuité en un échantillon du signal. Appliqué aux courbes de force mesurées en microscopie AFM, ce traitement permet de détecter de manière adaptative les régions du signal où les modèles physiques d'interaction entre nano-objets sont valides, et finalement de reconstruire des paramètres d'intérêt en ajustant les modèles physiques dans ces régions par moindres carrés. Cela conduit à cartographier (imagerie 2D) un ensemble de paramètres électrostatiques et bio-mécaniques avec une résolution spatiale nanométrique.

Chapitre 3

Le troisième chapitre étudie les propriétés de reconstruction exacte de supports par les algorithmes OMP et OLS. C'est un travail entrepris lors de mon séjour de délégation CNRS à l'IRCCyN (Nantes) durant l'année universitaire 2010-2011, en collaboration avec l'INRIA Rennes. L'objectif initial était d'analyser les algorithmes bidirectionnels du chapitre 1, définis en tant qu'extensions d'OLS. Nous nous sommes aperçus qu'aucune analyse de reconstruction exacte par OLS n'était disponible dans la littérature. Nous nous sommes donc focalisés sur ce problème en laissant en suspens les extensions d'OLS. Nous avons établi une condition nécessaire et suffisante de reconstruction du support inconnu en k itérations, où k est la taille du support. Cette condition s'avère être identique à la condition ERC (pour *Exact Recovery Condition*) pour OMP. Dans le but de distinguer les deux algorithmes et de confirmer le meilleur comportement empirique d'OLS pour des dictionnaires corrélés (chapitre 1), nous avons proposé une analyse plus poussée à l'itération $q + 1$. Son principe est de supposer que les q premières itérations des algorithmes ont toutes fourni de « vrais atomes » (faisant partie du support inconnu), et d'établir des conditions garantissant que les $k - q$ prochaines itérations vont également sélectionner de vrais atomes, permettant ainsi la reconstruction exacte du support. Cette analyse technique valide le meilleur comportement d'OLS par rapport à OMP. Elle permet par ailleurs d'établir une condition suffisante de « non-reconstruction » par OMP garantissant que certains supports ne peuvent pas être reconstruits *quelquesoit* la combinaison linéaire des atomes du support. Inversement, le phénomène de non-reconstruction n'existe pas avec OLS : tout support est atteignable à partir d'une combinaison spécifique des atomes indicés par le support.

Chapitre 4

Le chapitre 4 dresse quelques perspectives à court et moyen terme en lien avec les autres chapitres de la partie II :

1. Le développement de nouveaux algorithmes d'approximation parcimonieuse pour trois problèmes différents. Les premiers algorithmes envisagés sont des raffinements des algorithmes basés OLS présentés dans ce manuscrit pour des problèmes inverses qui mettent en jeu des dictionnaires structurés. En exploitant la structure du problème, il semble en effet possible de réduire le temps de calcul des algorithmes à performance égale, voire en améliorant leur performance. Le deuxième problème est l'approximation parcimonieuse utilisant la norme ℓ_0 sous contraintes de positivité. Enfin, le troisième problème est la décomposition parcimonieuse d'une séquence (par exemple temporelle) de signaux sur un même dictionnaire en imposant des contraintes de proximité entre les décompositions de signaux temporellement voisins.
2. La poursuite du travail sur l'analyse des algorithmes gloutons, l'un de nos objectifs principaux étant de fournir des outils d'analyse d'algorithmes bidirectionnels.
3. Du point de vue applicatif, l'analyse de cellules biologiques cancéreuses par microscopie AFM. L'objectif est d'utiliser la microscopie AFM comme outil de diagnostic du caractère métastatique d'une tumeur. Ce projet est pluridisciplinaire (biologie, physique des interactions, traitement du signal) et soutenu plus particulièrement par le département SBS du CRAN qui intègre des chercheurs en biologie, neurosciences et traitement du signal.

Première partie

Présentation générale

Curriculum Vitæ détaillé

1 Etat civil

Charles SOUSSEN
CRAN, Université de Lorraine
Faculté des Sciences et Technologies
54506 Vandœuvre Cedex
Tél : 03 83 68 44 71
Courriel : Charles.Soussen@univ-lorraine.fr
www.cran.univ-lorraine.fr/perso/charles.soussen/

Né le 24 novembre 1972, 41 ans
Nationalité française

2 Situation actuelle

Maître de Conférences en section 61 à l'Université de Lorraine, affecté au CRAN depuis septembre 2005.

- Enseignement dispensé dans la Licence Sciences Pour l'Ingénieur (SPI) et dans le Master Ingénierie des Systèmes Complexes (ISC) de l'Université de Lorraine : traitement du signal et des images, génie informatique.
- Recherche au CRAN : problèmes inverses, approximation parcimonieuse, traitement d'images hyperspectrales, reconstruction 3D en vision par ordinateur. Applications : microscopie de force atomique, microscopie de fluorescence, endoscopie à lumière structurée.

3 Titres universitaires

1997–00 : Doctorat en traitement d'images au Laboratoire des Signaux et Systèmes (Orsay).
Thèse soutenue en décembre 2000. Sujet : *Reconstruction 3D d'un objet compact en tomographie*. Directeur de thèse : Ali MOHAMMAD–DJAFARI.

Jury : Jean-Louis COATRIEUX (président), Bernard CHALMOND et Line GARNERO (rapporteurs), Jean-Marc DINTEN et Kenneth SAUER (examineurs).

1995–96 : DEA de mathématiques appliquées à l'Université Joseph Fourier (Grenoble). Major.

1993–96 : Diplômé de l'ENSIMAG, mention bien. Spécialisation en mathématiques appliquées.

4 Parcours

2010–11 : Séjour de délégation CNRS d'un an à l'IRCCyN (UMR CNRS 6597, Nantes). *Approximation parcimonieuse et problèmes inverses : minimisation de critères mixtes « l_2 - l_0 »*.

- 2005–13 : Maître de Conférences à l’Université Henri Poincaré (UHP, Nancy I) puis à l’Université de Lorraine, créée en 2012, qui regroupe notamment l’UHP et l’Institut National Polytechnique de Lorraine (INPL).
- 2001–05 : Maître de Conférences à l’Université de Paris-Sud (UPS, Orsay). Recherche au LIMSI-CNRS (UPR CNRS 3251). Reconstruction d’images 3D.
- 2000–01 : Attaché Temporaire à l’Enseignement et la Recherche (ATER) à l’UPS. Recherche dans le Groupe Problèmes Inverses du Laboratoire des Signaux et Systèmes.

5 Activité d’enseignement

5.1 Enseignements dispensés

En tant que Maître de Conférences, j’enseigne depuis l’année 2001, et à la faculté des Sciences et Technologies de Nancy depuis 2005. Mon enseignement est essentiellement dispensé dans la Licence Sciences Pour l’Ingénieur (SPI) et le Master Ingénierie des Systèmes Complexes (ISC). Il se répartit entre le traitement du signal et des images (50%) et le génie informatique (50%) :

- *Codage de l’information* (L2) : échantillonnage et quantification d’un signal, codage entropique, codage de Huffman, codage avec perte.
- *Traitement d’image* (M1) : analyse élémentaire (histogramme, seuillage), échantillonnage, quantification, transformée de Fourier, filtrage linéaire, notion de séparabilité, introduction au filtrage non linéaire et à l’interpolation d’image.
- *Traitement du signal* (M1) : analyse de données multidimensionnelles, analyse par composantes principales.
- *Introduction à l’optimisation globale* (M2 orientation recherche, cours commun avec l’Ecole Nationale Supérieure d’Electricité et de Mécanique, ENSEM) : méthode du crible, optimisation par intervalle, algorithmes évolutifs (essais particuliers, algorithmes génétiques), méthodes de type Monte Carlo, recuit simulé.
- *Algorithmique et programmation en langage C* (L3).
- *Génie informatique et base de données* (L3) : modélisation d’une base de donnée, programmation orientée objet, mini-projet de commande d’un robot avec des bibliothèques écrites en C++.
- *Systèmes temps-réel* (M2) : mise en place de TP illustrant le fonctionnement du système temps-réel RTAI sous Linux.

5.2 Responsabilités et mise en place d’enseignements

- Responsable de l’Unité d’Enseignement (UE) *Génie informatique* en L3 ; mise en place des TP avec L. Baghli.
- Mise en place de l’UE *Signal et Image* du M1 ISC avec A. Richard (partie *Signal*) ; responsabilité de la partie *Image* de cette UE ;
- Mise en place de l’UE *Optimisation* avec P. Riedinger (partie *Optimisation locale*) ; responsabilité de la partie *Optimisation globale* de cette UE ;
- Mise en place avec E.-H. Djermoune des TP illustrant le fonctionnement des systèmes temps-réel sous Linux.

5.3 Responsabilité des stages

Dans le cadre de la mise en place du Master Ingénierie Système (ancêtre du Master ISC) en 2005, j'ai eu la charge de repenser la gestion des stages, le Master étant issu de la fusion de plusieurs DESS qui avaient chacun leur propre fonctionnement. En concertation avec les responsables de spécialité du Master (Automatique, Productique, Réseaux, Electronique et Electrotechnique), j'ai proposé une organisation commune et rationnelle des stages.

Entre 2006 et 2010, j'ai été responsable des stages au sein du Master 2 ISC : mise en place administrative de la recherche de stage, signature des conventions de stage, gestion des visites de stage, préparation et organisation des soutenances de stage.

5.4 Vacances externes

Entre 2006 et 2008, j'ai été invité à enseigner une introduction au traitement d'images (filtrage et analyse de Fourier) à l'Ecole Nationale Supérieure des Technologies et Industries du Bois (ENSTIB, Epinal).

Depuis 2008, j'interviens en troisième année de l'Ecole Centrale de Nantes pour enseigner une introduction aux algorithmes d'optimisation globale avec une application en traitement du signal et des images (14 heures équivalent TD). L'objectif de ce cours est de présenter un panorama des différentes techniques employées dans le domaine de l'optimisation continue et discrète afin de faciliter et guider les futurs utilisateurs dans le choix d'une méthode appropriée.

6 Activité d'encadrement

6.1 Encadrement doctoral

Thèses passées au CRAN.

1. **Achraf Ben Hamadou** (2007-2011) : *Contribution à la cartographie 3D des parois internes de la vessie par cystoscopie à vision active.*
 - Allocataire de recherche (bourse ministérielle) puis ATER à l'UHP. Thèse débutée en octobre 2007, soutenue le 19 septembre 2011.
 - Directeur : C. Daul (CRAN, 40 %), co-encadrants : **C. Soussen (40 %)**, W. Blondel (CRAN, 20 %).
 - Qualifié aux fonctions de Maître de Conférences en 2012.
 - Post-doctorant à l'École des Ponts ParisTech en 2011-2012.
 - Situation depuis octobre 2013 : ingénieur R & D en vision par ordinateur chez Valéo.

La thèse d'A. Ben Hamadou a donné lieu à un article de revue internationale [R2]¹, trois publications dans des conférences internationales [C6, C7, C11] et une conférence nationale [CN6].

2. **Junbo Duan** (2007-2010) : *Restauration et séparation de signaux polynômiaux par morceaux. Application à la microscopie de force atomique.*
 - Boursier Université Henri Poincaré/Région Lorraine. Thèse débutée en octobre 2007, soutenue le 15 novembre 2010.
 - Directeur : D. Brie (CRAN, 30 %), co-encadrants : **C. Soussen (50 %)**, J. Idier (IRC-CyN, Nantes, 20 %).
 - Situation depuis 2013 : enseignant-chercheur permanent à l'Université de Xi'An Jiatong, Xi'An, Chine.

1. Les références internes se rapportent à la liste de publications page 9.

La thèse de J. Duan a donné lieu à trois publications dans des revues internationales [R5, R8, R9], une publication de conférence internationale [C12], une conférence nationale [CN5], plusieurs communications sans acte dont une présentation au GDR ISIS [COM6, COM7, COM8].

3. **Simon Henrot** (2010-2013) : *Déconvolution et séparation d'images hyperspectrales. Applications en microscopies*.
 - Boursier CNRS/Région Lorraine (ANR blanc HÆSPRI).
 - Thèse débutée en octobre 2010, soutenue le 27 novembre 2013.
 - Directeur : D. Brie (50 %), co-encadrant : **C. Soussen (50 %)**.
 - Situation actuelle : ATER à l'Université de Lorraine.

La thèse de S. Henrot a donné lieu à une publication de revue internationale [R4], deux publications de conférence internationale [C2, C3], deux conférences nationales [CN2, CN3] et deux présentations au GDR ISIS [COM1, COM3].

6.2 Encadrement de stagiaires de Master

1. Cyril Kessler (stage de Master 1 Sciences de la Vie et de la Santé). *Analyse de spectres de forces atomiques*, CRAN, Université Henri Poincaré, Nancy, juin 2007.
Co-encadrement à 70 % avec David Brie.
2. Fabrice Caland (stage de Master 1 Ingénierie des Systèmes). *Analyse d'images hyperspectrales de fluorescence par méthode de séparation de sources non négatives*. CRAN, Université Henri Poincaré, Nancy, juin 2008.
Co-encadrement à 50 % avec Sebastian Miron.
3. Simon Henrot (stage de Master 2 Ingénierie des Systèmes Complexes).
Restauration d'images de fluorescence en microscopie confocale : application à l'imagerie de bio-capteurs bactériens, juin 2010, CRAN, Université Henri Poincaré.
Co-encadrement à 50 % avec David Brie.
4. Salmane Lahdachi (stage de deuxième année de l'Ecole des Mines de Nancy – équivalent Master 1). *Modélisation et analyse quantitative des images force-volume de cellules*, sep. 2012, CRAN, Université de Lorraine.
Co-encadrement à 50 % avec David Brie.

7 Contrats de recherche

7.1 Analyse d'images force volume

- Financement : PEPS CNRS, BQR UHP, Région Lorraine.
- Laboratoires partenaires : CRAN, LCPME (Nancy), IRCCyN (Nantes).
- Durée : 3 ans (2007-2010).
- Responsables : D. Brie, **C. Soussen**.
- Autres chercheurs : J. Duan (CRAN), G. Francius (LCPME), J. Idier (IRCCyN).

Je suis porteur de deux demandes de moyen ayant permis la mise en place de cette activité et le co-financement de la thèse de Junbo Duan :

1. Porteur d'un Projet Exploratoire Pluridisciplinaire du CNRS (PEPS) en 2008 : « Développement de méthodes d'analyse d'images force-volume ».
2. Porteur d'un projet BQR UHP-Région Lorraine pour le co-financement de la thèse de Junbo Duan (période 2007-2010).

Ce projet est détaillé dans le chapitre 2.

7.2 Bioforce : analyse d'images PeakForce Mapping de tissus biologiques

- Financement : PEPS conjoint CNRS-Région Lorraine.
- **Porteur : Charles Soussen** (CRAN).
- Durée : 1 an (2013-2014).
- Laboratoires partenaires : CRAN, LCPME (Nancy), LIEC (Nancy).
- Autres chercheurs : D. Brie, S. Grandemange, C. Barbieux (CRAN), G. Francius (LCPME), A. Razafitianamaharavo (LIEC).
- Ingénieur de recherche : R. Pannequin (CRAN).

Ce projet est dans la continuité du projet sur l'imagerie force-volume mais le contexte est très différent de par son orientation vers l'imagerie de cellules biologiques. Le consortium intègre une biologiste du CRAN (S. Grandemange) et le LIEC comme nouveau partenaire. Ce projet est résumé de façon détaillée dans le chapitre 4 (§ 4.3).

7.3 HÆSPRI : analyse hyperspectrale des interactions bactérie-minéral

- Financement : ANR.
- Laboratoires partenaires : LIMOS (actuel LIEC), LEM (actuel LIEC), LCPME, CRAN.
- Durée : 4 ans (2009-2013).
- Porteur : Christian Mustin (LIMOS).
- Responsable CRAN : D. Brie.
- Autres chercheurs CRAN : S. Miron, **C. Soussen**, F. Caland, S. Henrot.

7.4 SpectroDec : décomposition parcimonieuse de signaux spectroscopiques

- Financement : PEPS CNRS.
- Porteur : Vincent Mazet (LSIIT nouvellement ICube, Strasbourg).
- Durée : 2 ans (2012-2013).
- Laboratoires partenaires : ICube, CRAN, Observatoire Astronomique de Strasbourg, Laboratoire Francis Perrin (Saclay).
- Chercheurs CRAN : **C. Soussen**, E.-H. Djermoune.

8 Activités de valorisation et de transfert

8.1 Logiciel d'analyse d'images force-volume

Avec le LCPME et l'IRCCyN, nous avons proposé une méthode pionnière d'analyse automatique des images force-volume acquises en microscopie AFM [R9]. Contrairement à beaucoup d'approches *ad hoc* de la littérature, la méthode nécessite très peu de réglages empiriques et permet de reconstruire des images 2D représentant la cartographie de paramètres physico-chimiques à partir de données expérimentales.

Je suis fortement impliqué dans ce projet de valorisation logicielle. Le développement des méthodes de traitement du signal est assuré par le CRAN. En particulier, la réalisation d'une interface graphique conviviale est assurée par Rémi Pannequin, ingénieur de recherche au CRAN. La phase de test du logiciel avec divers jeux de données réelles est réalisée par le LCPME ainsi que les interactions avec des partenaires industriels susceptibles d'être acheteurs du logiciel.

8.2 Mosaïquage d'images endoscopiques de la vessie

Ce projet, porté par Christian Daul, fait suite à la thèse d'Achraf Ben Hamadou. Il concerne la valorisation des prototypes d'endoscope (cystoscope) conçus au CRAN et des méthodes de traitement d'images vidéo-endoscopiques visant à recalibrer et reconstruire des surfaces 3D mosaïquées représentant l'intérieur de la vessie.

8.3 Logiciel déposé : optimisation locale sous Matlab

Le logiciel [LOG1] a été développé conjointement par les anciens membres du Groupe Problèmes Inverses du Laboratoire des Signaux et Systèmes. Il se présente sous la forme d'une boîte à outil permettant de résoudre des problèmes d'optimisation locale avec quatre stratégies de direction de descente et trois politiques concernant l'adaptation du pas de descente.

8.4 Code source en accès libre

Une implémentation efficace de l'algorithme d'approximation parcimonieuse Single Best Replacement (SBR) [R8] est disponible sur le site web des IEEE, en document attaché à l'article : <http://dx.doi.org/10.1109/TSP.2011.2160633> (rubrique Multimedia).

8.5 Contrat industriel

J'ai participé à un contrat [CTR1] entre l'IRCCyN (Nantes) et l'équipe « Traitement avancé de l'information » (TAI, EdF R & D, Chatou) d'EdF, portant sur la reconstruction 3D d'images binaires en tomographie à rayons X pour le contrôle non destructif des matériaux.

9 Principaux collaborateurs extérieurs au CRAN

- Jérôme Idier (IRCCyN) : reconstruction tomographique pour le contrôle non destructif des matériaux (2005-2008), approximation parcimonieuse et problèmes inverses : minimisation de critères mixtes « ℓ_2 - ℓ_0 » (depuis 2007).
- Cédric Herzet, Rémi Gribonval (INRIA Rennes) : analyse d'algorithmes d'approximation parcimonieuse de signaux (depuis 2011).
- Grégory Francius (LCPME, Nancy) : analyse d'images force-volume en microscopie de force atomique (depuis 2009).
- Christian Mustin (LIEC, Nancy) : imagerie hyperspectrale de bio-capteurs bactériens (depuis 2009).
- Vincent Mazet (ICube, Strasbourg) : approches parcimonieuses pour la décomposition de signaux spectroscopiques (depuis 2012).

10 Activités liées à l'administration

- Membre élu de la commission de spécialistes de la section 61 de l'Université Paris-Sud (Orsay) entre 2003 et 2005.
- Participation à un comité de sélection à l'Université Henri Poincaré (Nancy 1, actuelle Université de Lorraine) pour le poste MCF 698 en 2010.
- Responsable des stages au sein du Master Ingénierie des Systèmes Complexes de l'UHP entre 2006 et 2010.

11 Activités liées à la recherche

- Révision d'articles pour les revues *IEEE Trans. Signal Processing*, *Signal Processing*, *Inverse Problems*, *Applied Optics*, *International Journal of Biomedical Imaging* et la revue française *Traitement du Signal* ;
- Révision d'articles pour les conférences internationales *IEEE ICASSP*, *IEEE ICIP*, *Eusipco* et *SPARS* et pour la conférence nationale *Gretsi*.
- Participation au comité de suivi de thèse de Clément Gilavert (IRCCyN, Nantes), co-directeurs : J. Idier et S. Moussaoui. Thèse débutée en octobre 2011.

12 Liste de publications

Thèse de doctorat

- [T1] **C. Soussen**, « Reconstruction 3D d'un objet compact en tomographie », Thèse de doctorat, Université de Paris-Sud, Orsay, déc. 2000.

Articles de revue

- [R1] C. Herzet, **C. Soussen**, J. Idier et R. Gribonval, « Exact recovery conditions for sparse representations with partial support information », *IEEE Trans. Information Theory*, vol. 59, n° 11, pp. 7509-7524, nov. 2013.
- [R2] A. Ben Hamadou, **C. Soussen**, C. Daul, W. Blondel et D. Wolf, « Flexible calibration of structured-light systems projecting point patterns », *Comput. Vision Image Understanding*, vol. 117, n° 10, pp. 1468-1481, oct. 2013.
- [R3] **C. Soussen**, R. Gribonval, J. Idier et C. Herzet, « Joint k -step analysis of orthogonal matching pursuit and orthogonal least squares », *IEEE Trans. Information Theory*, vol. 59, n° 5, pp. 3158-3174, mai 2013.
- [R4] S. Henrot, **C. Soussen** et D. Brie, « Fast positive deconvolution of hyperspectral images », *IEEE Trans. Image Processing*, vol. 22, n° 2, pp. 828-833, fév. 2013.
- [R5] J. Duan, **C. Soussen**, D. Brie, J. Idier et Y.-P. Wang, « On LARS/homotopy equivalence conditions for over-determined LASSO », *IEEE Signal Processing Letters*, vol. 19, n°12, pp. 894-897, déc. 2012.
- [R6] S. Sahnoun, E.-H. Djermoune, **C. Soussen** et D. Brie, « Sparse multidimensional modal analysis using a multigrid dictionary refinement », *Eurasip Journal on Advances in Signal Processing*, vol. 2012, n°60, pp. 1-10, mars 2012.
- [R7] **C. Soussen**, J. Idier, E. Carcreff, L. Simon et C. Potel, « Ultrasonic non destructive testing based on sparse deconvolution », *Journal of Physics : Conference Series*, vol. 353, n°012018, pp. 1-10, mars 2012.
- [R8] **C. Soussen**, J. Idier, D. Brie et J. Duan, « From Bernoulli-Gaussian deconvolution to sparse signal restoration », *IEEE Trans. Signal Processing*, vol. 59, n°10, pp. 4572-4584, oct. 2011.

- [R9] P. Polyakov, **C. Soussen**, J. Duan, J. Duval, D. Brie et G. Francius, « Automated force volume image processing for biological samples », *PLoS ONE*, vol. 6, n°4, pp. e18887, avr. 2011.
- [R10] **C. Soussen** et J. Idier, « Reconstruction of three-dimensional localized objects from limited angle x-ray projections : an approach based on sparsity and multigrid image representation », *J. Electr. Imag.*, vol. 17, n°3, pp. 033011, août 2008.
- [R11] **C. Soussen** et A. Mohammad-Djafari, « Polygonal and polyhedral contour reconstruction in computed tomography », *IEEE Trans. Image Processing*, vol. 13, n°11, pp. 1507–1523, nov. 2004.

Chapitres d’ouvrage

- [CH1] **C. Soussen**, D. Brie, G. Francius et J. Idier, « Approximation parcimonieuse pour l’analyse d’images force-volume en microscopie de force atomique », in *Méthodes d’inversion appliquées au traitement du signal et de l’image*, J.-F. Giovannelli et J. Idier, Eds., chapitre 2, pp. 53–77. Hermès, nov. 2013.
- [CH2] **C. Soussen** et A. Mohammad-Djafari, « Reconstruction des surfaces en tomographie », in *Problèmes inverses en imagerie et en vision*, A. Mohammad-Djafari, Ed., chapitre 8, pp. 287–319. Traité IC2, Série traitement du signal et de l’image, Hermès, Paris, sep. 2009.
- *Version en anglais* : **C. Soussen** et A. Mohammad-Djafari, « Surface reconstruction from tomography data », in *Inverse Problems in Vision and 3D Tomography*, A. Mohammad-Djafari, Ed., chapitre 8, pp. 289–322. Série Digital Signal and Image processing, Wiley, New York, déc. 2009.
- [CH3] A. Mohammad-Djafari et **C. Soussen**, « Compact object reconstruction », in *Discrete Tomography : Foundations, Algorithms and Applications*, G. T. Herman et A. Kuba, Eds., chapitre 14, pp. 317–342. Birkhäuser, Boston, nov. 1999.

Actes de congrès internationaux avec comité de lecture

- [C1] **C. Soussen**, C. Herzet, J. Idier et R. Gribonval, « Unrecoverable subsets by OMP and Basis Pursuit », in *Signal Processing with Adaptive Sparse Structured Representations (SPARS workshop)*, Lausanne, juil. 2013, pp. 1–1.
- [C2] S. Henrot, S. Moussaoui, **C. Soussen** et D. Brie, « Edge-preserving nonnegative hyperspectral image restoration », in *Proc. IEEE ICASSP*, Vancouver, mai 2013, pp. 1622-1625.
- [C3] S. Henrot, **C. Soussen** et D. Brie, « Fast deconvolution of large fluorescence hyperspectral images », in *Proc. IEEE WHISPERS*, Lisbonne, juin 2011, pp. 1-4.
- [C4] S. Sahnoun, E.-H. Djermoune, **C. Soussen** et D. Brie, « Sparse multiresolution modal estimation », in *Proc. IEEE SSP*, Nice, juin 2011, pp. 309-312.
- [C5] S. Bourguignon, **C. Soussen**, H. Carfantan et J. Idier, « Sparse deconvolution : Comparison of statistical and deterministic approaches », in *Proc. IEEE SSP*, Nice, juin 2011, pp. 317-320.
- [C6] A. Ben Hamadou, **C. Soussen**, C. Daul, W. Blondel et D. Wolf, « Flexible projector calibration for active stereoscopic systems », in *Proc. IEEE ICIP*, Hong Kong, sep. 2010, pp. 4241-4244.

- [C7] A. Ben Hamadou, C. Daul, **C. Soussen**, A. Rekik et W. Blondel, « A novel 3D surface construction approach : Application to three-dimensional endoscopic data », in *Proc. IEEE ICIP*, Hong Kong, sep. 2010, pp. 4425-4428.
- [C8] C. Daul, W. Blondel, A. Ben Hamadou, R. Miranda-Luna, **C. Soussen**, D. Wolf et F. Guillemin, « From 2D towards 3D cartography of hollow organs », in *Proc. 7-th Int. Conf. Elec. Eng. (CCE 2010), Computing Science and Automatic Control*, Tuxtla Gutiérrez, Chiapas, Mexique, sep. 2010, pp. 285-293.
- [C9] M. Sauget, R. Laurent, J. Henriët, M. Salomon, R. Gschwind, S. Contassot-Vivier, L. Makovicka et **C. Soussen**, « Efficient domain decomposition for a neural network learning algorithm, used for the dose evaluation in external radiotherapy », in *Artificial Neural Networks - ICANN*, Lecture Notes in Computer Science, K. Diamantaras, W. Duch et L. Iliadis, Eds, avr. 2010, pp. 261-266.
- [C10] **C. Soussen**, S. Miron, F. Caland, D. Brie, P. Billard et C. Mustin, « Complexity reduction by convex cone detection for unmixing hyperspectral images of bacterial biosensors », in *Actes de la 17^e conf. EUSIPCO*, Glasgow, UK, août 2009, pp. 1938-1942.
- [C11] A. Ben Hamadou, **C. Soussen**, W. Blondel, C. Daul et D. Wolf, « Comparative study of image registration techniques for bladder video-endoscopy », in *European Conferences on Biomedical Optics (ECBO)*, Munich, Allemagne, juin 2009, vol. 7371, pp. 737118-737118-7.
- [C12] J. Duan, **C. Soussen**, D. Brie et J. Idier, « A continuation approach to estimate a solution path of mixed L2-L0 minimization problems », in *Signal Processing with Adaptive Sparse Structured Representations (SPARS workshop)*, Saint-Malo, avr. 2009, pp. 1-6.
- [C13] **C. Soussen**, D. Brie, F. Gaboriaud et C. Kessler, « Modeling of force-volume images in atomic force microscopy », in *Proc. IEEE ISBI*, Paris, mai 2008, pp. 1605-1608.
- [C14] **C. Soussen** et J. Idier, « 3D reconstruction of localized objects from radiographs and based on multiresolution and sparsity », in *Proc. IEEE ICIP*, Gênes, Italie, sep. 2005, vol. III, pp. 744-747.
- [C15] **C. Soussen** et A. Mohammad-Djafari, « Closed surface reconstruction in X-ray tomography », in *Proc. IEEE ICIP*, Thessalonique, Grèce, oct. 2001, vol. 1, pp. 718-721.
- [C16] **C. Soussen** et A. Mohammad-Djafari, « Contour-based models for 3D binary reconstruction in X-ray tomography », in *Bayesian Inference and Maximum Entropy Methods (MaxEnt Workshops)*, A. Mohammad-Djafari, Ed, Gif-sur-Yvette, juil. 2000, pp. 543-554.
- [C17] **C. Soussen** et A. Mohammad-Djafari, « Multiresolution approach to the estimation of the shape of a 3D compact object from its radiographic data », in *Proc. SPIE, Mathematical Modeling, Bayesian Estimation, and Inverse Problems*, F. Prêteux, A. Mohammad-Djafari et E. R. Dougherty, Eds, Denver, CO, USA, juil. 1999, pp. 150-160.

Actes de congrès nationaux avec comité de lecture

- [CN1] V. Mazet, **C. Soussen** et E.-H. Djermoune, « Décomposition de spectres en motifs paramétriques par approximation parcimonieuse », in *Actes 24^e coll. GRETSI*, Brest, sep. 2013.

Actes de congrès nationaux avec comité de lecture

- [CN2] S. Henrot, **C. Soussen**, S. Moussaoui et D. Brie, « Restauration positive d'images hyperspectrales avec préservation des contours », in *Actes 24^e coll. GRETSI*, Brest, sep. 2013.
- [CN3] S. Henrot, **C. Soussen** et D. Brie, « Déconvolution rapide d'images hyperspectrales de fluorescence », in *Actes 23^e coll. GRETSI*, Bordeaux, sep. 2011.
- [CN4] S. Sahnoun, E.-H. Djermoune, **C. Soussen** et D. Brie, « Analyse modale bidimensionnelle par approximation parcimonieuse et multirésolution », in *Actes 23^e coll. GRETSI*, Bordeaux, sep. 2011.
- [CN5] J. Duan, **C. Soussen**, D. Brie et J. Idier, « Détection conjointe de discontinuités d'ordres différents dans un signal par minimisation de critère L2-L0 », in *Actes 22^e coll. GRETSI*, Dijon, sep. 2009.
- [CN6] A. Ben Hamadou, **C. Soussen**, W. Blondel, C. Daul et D. Wolf, « Mosaïquage d'images cystoscopiques : étude comparative de méthodes de recalage », in *Actes 22^e coll. GRETSI*, Dijon, sep. 2009.
- [CN7] **C. Soussen** et A. Mohammad-Djafari, « Reconstruction de la surface fermée d'un objet compact en tomographie », in *Actes 18^e coll. GRETSI*, Toulouse, sep. 2001.
- [CN8] **C. Soussen** et A. Mohammad-Djafari, « Approche multirésolution pour la reconstruction 3D de défaut en tomographie X », in *Actes 17^e coll. GRETSI*, Vannes, sep. 1999.

Communications sans acte

- [COM1] S. Henrot, **C. Soussen**, S. Moussaoui et D. Brie, « Restauration positive d'images hyperspectrales avec préservation des contours », in *GDR ISIS*, journée « Optimisation convexe sous contraintes », Paris, oct. 2013.
- [COM2] **C. Soussen**, D. Brie et G. Francius, « Analyse d'échantillons biologiques en microscopie AFM », in *GDR ISIS*, journée « Imagerie biologique », Rennes, nov. 2011.
- [COM3] S. Henrot, **C. Soussen**, D. Brie et C. Mustin, « Déconvolution d'images hyperspectrales de fluorescence en microscopie confocale sous contrainte de positivité », in *GDR ISIS*, journée « Imagerie biologique », Rennes, nov. 2011.
- [COM4] **C. Soussen**, S. Miron, F. Caland, D. Brie, P. Billard et C. Mustin, « Réduction de complexité pour la séparation de sources en imagerie hyperspectrale : application à l'imagerie de biocapteurs bactériens », in *GDR ISIS*, journée « Contraintes de non-négativité en traitement du signal et des images », Paris, jan. 2011.
- [COM5] **C. Soussen**, J. Idier, D. Brie, J. Duan et G. Francius, « Algorithmes d'approximation parcimonieuse basés sur la minimisation de critères mixtes ℓ_2 - ℓ_0 », in *Séminaires de la Fédération Charles Hermite^a*, journée « Problèmes Inverses », Nancy, mars 2010.

^a. Fédération lorraine qui regroupe les chercheurs en mathématique, informatique et automatique des laboratoires IECL, LORIA et CRAN.

- [COM6] **C. Soussen**, J. Duan, D. Brie et F. Gaboriaud, « Problèmes d'inversion de données en imagerie de force atomique », in *GDR ISIS*, journée « Problèmes Inverses et Imagerie », Paris, juin 2008.
- [COM7] **C. Soussen**, J. Duan, D. Brie et F. Gaboriaud, « Modeling of force-volume images in atomic force microscopy », in *AFM Workshop. Biological interfaces : from the model membrane to microbial cells*, Nancy, oct. 2008.
- [COM8] J. Duan, **C. Soussen**, D. Brie, J. Idier et F. Gaboriaud, « Sparse representation of a force spectrum and a force-volume image », in *AFM workshop. Biological interfaces : from the model membrane to microbial cells*, Nancy, oct. 2008.

Contrat industriel

- [CTR1] M. Allain, **C. Soussen** et J. Idier, « Détection et dimensionnement de défauts dans un volume par gammagraphie », Rapport de contrat EDF–Ecole centrale de Nantes, sep. 2005.

Logiciel déposé

- [LOG1] S. Brette, H. Carfantan, J.-F. Giovannelli, T. Martin, J.-F. Bercher, C. Heinrich, J. Idier et **C. Soussen**, « Gradient à Pas Adaptatif avec Corrections, une mise en œuvre Matlab : GPAC.m », déclaration d'invention n°DI-0258-01, Logibox APP n° 21961, IDDN.FR.001.070033.000.S.P.2006.000.10800, fév. 2006.

Rapports techniques

- [RT1] S. Henrot, **C. Soussen**, M. Dossot et D. Brie, « Does deblurring improve geometrical hyperspectral unmixing? », accepté pour publication dans *IEEE Trans. Image Processing*, pp. 1-14, déc. 2013.
- [RT2] **C. Soussen**, J. Idier, J. Duan et D. Brie, « ℓ_2 - ℓ_0 regularization path tracking algorithms », Rapport interne, Centre de Recherche en Automatique de Nancy, sep. 2013 (*en cours de finalisation*).
- [RT3] A. Ben Hamadou, C. Daul, **C. Soussen**, W. Blondel et F. Guillemin, « Construction of large 3-D field of views of the internal bladder wall surface : a proof of concept », Rapport interne, CRAN, pp. 1-11, déc. 2013.

Synthèse générale de mon activité de recherche

Je présente dans ce chapitre une synthèse de mes thématiques de recherche depuis que je suis Maître de Conférences en fournissant davantage de précisions sur mon activité qui n'est pas couverte dans la deuxième partie du manuscrit.

Mon activité de recherche s'est orientée jusqu'à présent vers le traitement du signal et des images dans différents contextes applicatifs. Je possède une bonne expérience des **problèmes d'inversion de données** dans le contexte où les données disponibles sont en nombre limité et partiellement informatives sur la quantité à estimer (problèmes inverses mal posés). Pour ce type de problèmes, la prise en compte d'informations *a priori* sur les inconnues est indispensable, et s'effectue par le biais des techniques de régularisation. Je me suis intéressé aux **algorithmes d'approximation parcimonieuse**, à la reconstruction d'images 2D/3D, et aux problèmes de **séparation de sources** et de **recalage d'images**.

Je commence par présenter mon activité antérieure à mon arrivée au CRAN (section 1) puis mon activité menée au sein des groupes thématiques *Identification, Restauration, Images et Signaux* (G.T. IRIS : section 2) et *Ingénierie Pour la Santé* (G.T. IPS : section 4) dans l'organisation du CRAN pour la période 2005-2012. Ces activités sont à présent totalement intégrées dans le département *Santé, Biologie, Signal* (SBS) créé en 2013. La section 3 résume une activité amorcée lors de mon séjour de délégation CNRS à Nantes en 2011 en collaboration avec l'IRCCyN puis l'INRIA Rennes, portant sur l'analyse théorique d'algorithmes gloutons pour l'approximation parcimonieuse. Les sections 1, 2 et 3 sont scientifiquement liées. Ces travaux se sont déroulés dans l'ordre chronologique mais les contributions de chaque section sont distinctes. La section 4 est indépendante. Les travaux qui y sont présentés se sont déroulés en parallèle.

1 Reconstruction tomographique (2001-2007)

Cette activité concerne la reconstruction d'images 3D à partir de projections limitées pour le contrôle non destructif (CND) des matériaux. C'est le prolongement de mon travail de thèse avec des méthodes néanmoins différentes. Elle a été menée au LIMSI entre 2001 et 2005 et poursuivie lors de mes deux premières années au CRAN. Pour pouvoir présenter cette activité, je résume brièvement mon travail de thèse.

1.1 Travail de thèse : reconstruction d'un objet polyédrique

Ma thèse menée dans le Groupe Problèmes Inverses du Laboratoire des Signaux et Systèmes portait sur la reconstruction 3D d'un objet binaire en tomographie à rayons X à partir d'un faible nombre de projections pour des applications en CND où l'objet 3D représente un défaut

présent dans un métal. Décrire l'objet binaire inconnu par une surface 3D maillée (modèle polyédrique) permet de modéliser l'image de façon parcimonieuse, et ainsi de réduire le nombre de paramètres à estimer. La reconstruction de la surface à partir des données projectives est alors abordée comme une méthode de contour actif déformable. C'est une modélisation courante en segmentation d'images mais plus rarement utilisée en reconstruction tomographique. La contribution principale de ma thèse était d'adapter les méthodes de contours déformables à la reconstruction tomographique pour des données limitées en nombre et en angles. Cela m'a conduit à suivre une méthodologie de type « problèmes inverses » en régularisant le problème pour compenser son caractère mal posé, et estimer les sommets du polyèdre au sens du maximum *a posteriori* (MAP). Le critère à optimiser est généralement non convexe car l'opérateur de projection dépend des paramètres inconnus de façon non linéaire. L'optimisation s'avère complexe et une partie importante de ma thèse consistait à proposer des algorithmes locaux efficaces.

Références associées : [R11, CH2].

1.2 Modèle volumique parcimonieux

Ce travail est le point de départ de ma collaboration avec Jérôme Idier (IRCCyN, UMR CNRS 6597, Nantes) et de notre réflexion sur l'utilisation de la parcimonie pour les problèmes inverses mal conditionnés (*cf.* § 2). Le projet sur la reconstruction tomographique impliquait également Marc Allain, alors post-doctorant à l'IRCCyN, et l'équipe « Traitement avancé de l'information » (TAI, EdF R & D, Chatou) d'EdF dans le cadre du contrat [CTR1].

La thématique est proche de celle de ma thèse. Il s'agissait de reconstruire une image 3D binaire représentant un *ou plusieurs* défauts (valeur 1) inclus dans un fond homogène (valeur 0) à partir de projections limitées en nombre et en angles. Nous avons ici opté pour un *modèle volumique* de l'image : l'image est paramétrée par un ensemble de voxels dont la valeur est estimée à partir des données [DIN 90]. Ces voxels constituent une matrice creuse : les voxels du fond de l'image (le métal) ont une valeur nulle et les voxels non nuls correspondent aux défauts recherchés dans la structure métallique. La modélisation est dite parcimonieuse car les défauts occupent généralement un volume limité dans le métal. Il s'agit donc de détecter de façon adaptative l'ensemble des voxels appartenant au fond de l'image, et ainsi de les exclure de la paramétrisation de l'image. Nous avons formulé la reconstruction d'image au sens du MAP où l'information *a priori* favorisant la parcimonie des voxels prend la forme d'une pénalisation ℓ_1 , et on impose de plus la positivité des voxels. L'algorithme d'optimisation proposé, de type *Single Site Update* [BRE 96, IDI 01] permet de reconstruire des images effectivement parcimonieuses, comportant un grand nombre de voxels nuls (méthode appelée mono-échelle).

Une implémentation efficace de cette méthode repose sur une stratégie multi-échelle, qui reconstruit une séquence d'images creuses à des échelles de plus en plus fines où seule la première image à l'échelle la plus grossière est codée par une matrice pleine. Cette stratégie permet 1) de détecter à chaque échelle la liste des voxels nuls, 2) de mettre à jour la paramétrisation « creuse » de l'image en les excluant, et 3) de fournir une solution initiale pertinente à l'échelle suivante pour accélérer la convergence de l'algorithme d'optimisation. Par cette implémentation, on améliore très largement les performances de la méthode mono-échelle (modélisation « pleine » de l'image par N^3 voxels) à la fois en termes de temps de calcul et d'occupation mémoire.

Références associées : [R10, C14, CTR1].

2 Problèmes inverses et traitement de signaux tridimensionnels

Dans le groupe thématique IRIS du CRAN (jusqu'en 2012) et dans l'actuel département SBS, je suis principalement rattaché à l'équipe de recherche de David Brie, qui s'intéresse aux problèmes inverses en traitement du signal et des images et aux techniques de régularisation. Ces techniques couvrent un vaste champ d'applications et de problématiques comme le débruitage, la déconvolution, l'analyse spectrale et la séparation de sources. Depuis quelques années, l'équipe s'intéresse à des problèmes pour lesquels les données ou les inconnues sont multidimensionnelles. L'inversion de données requiert alors des algorithmes qui intègrent le fait que les données sont en grand nombre et que chaque dimension apporte une information de nature différente. Dans ce contexte, je me consacre plus particulièrement à l'approximation parcimonieuse et à la séparation de sources pour l'analyse d'images tridimensionnelles.

Je commence par introduire les deux principales applications qui ont motivé mes recherches. Puis je décris les problèmes de traitement du signal qui en découlent et les investissements méthodologiques pour résoudre ces problèmes (section 2.3). La microscopie de force atomique (AFM pour *Atomic Force Microscopy*) a donné lieu à la thèse de Junbo Duan, co-encadrée avec David Brie et Jérôme Idier. La microscopie confocale de fluorescence est le sujet de thèse de Simon Henrot, co-encadrée avec David Brie. Le projet sur la microscopie de force atomique sera plus amplement détaillé dans la deuxième partie du manuscrit.

Le terme générique « signaux tridimensionnels » désigne une collection de signaux mesurés sur une grille spatiale (x, y) . En imagerie confocale, il s'agit d'images hyperspectrales (chaque signal mesuré pour un pixel (x, y) est un spectre qui dépend de la longueur d'onde λ) mais en microscopie AFM, les trois dimensions (x, y, z) sont spatiales.

2.1 Microscopie AFM

Imagerie force-volume

Le traitement des données obtenues en microscopie AFM est une collaboration entre le CRAN et le LCPME (UMR 7564, CNRS, Université de Lorraine). Elle concerne l'étude des forces inter-atomiques et intermoléculaires aux interfaces solide/solution aqueuse. Un microscope AFM fournit des mesures de force à l'échelle nanométrique indépendamment de la nature des échantillons (biologique, organique, minérale) et des images tridimensionnelles dites *images force-volume*. Au cours des dernières années, la microscopie AFM a suscité un engouement important dans des domaines aussi divers que la biologie, la microbiologie, l'industrie pharmaceutique et la médecine clinique car cette modalité rend possible la détection de forces inter-atomiques de l'ordre du nano-Newton et permet la reconstruction d'images avec une résolution spatiale de l'ordre du nanomètre.

L'objectif du projet entre le CRAN et le LCPME est de proposer des outils avancés de traitement adaptés aux images force-volume. Nous avons en effet constaté que les outils d'analyse disponibles sont soit peu sophistiqués, soit non automatiques (ils requièrent des réglages empiriques pour chaque signal à traiter). Les signaux traités sont :

1. des courbes de force. Une courbe de force $z \mapsto f(z)$ est un signal 1D qui représente la force d'interaction entre deux nano-objets en fonction de leur distance z . Les objets en interaction sont la pointe nanométrique du microscope (pointe AFM) et l'échantillon à imager, par exemple une bactérie ou une cellule biologique : *cf.* fig. 1(a). On approche progressivement la pointe vers l'échantillon jusqu'à atteindre le contact et compresser mécaniquement l'échantillon. La force d'interaction est mesurée pour chaque distance z , ce

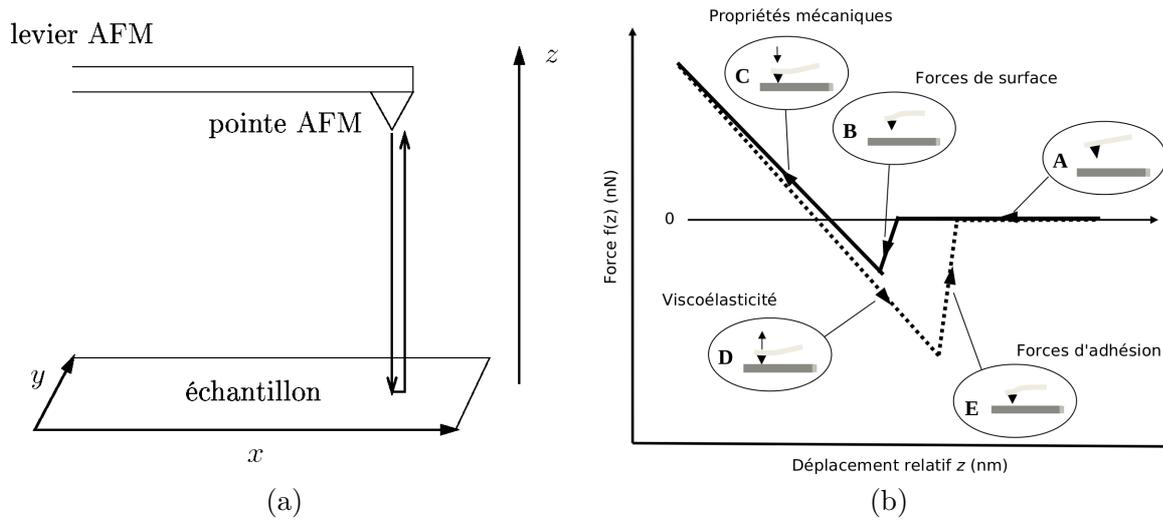


FIGURE 1 – (a) Mesure d’une courbe de force en un point (x, y) ; (b) Modèle par morceaux d’une courbe de force : signaux d’approche (trait plein) et de retrait (en pointillés).

qui constitue une courbe d’approche. Dans une courbe d’approche, on distingue plusieurs régions (ou types d’interaction) avant contact, au moment du contact et après contact. Inversement, le retrait progressif de la pointe à partir d’une position de contact avec l’échantillon constitue une courbe de retrait. Ce sont ces deux courbes (fig. 1(b)) que l’on propose d’analyser en utilisant des modèles physiques. Leur analyse consiste à estimer des paramètres physiques à partir des données expérimentales par inversion de données.

- des images force-volume. Une image force-volume est une collection de courbes de forces acquises successivement en effectuant un balayage spatial (x, y) de la surface de l’échantillon. Ce sont donc des images 3D du type $f(x, y, z)$. Analyser une image force-volume permet d’établir une cartographie spatiale (suivant (x, y)) des paramètres physiques.

Les échantillons étudiés dans ce projet sont la bactérie *Escherichia coli* présente dans le corps humain dont certaines formes pathogènes sont responsables d’infections intestinales, et la bactérie *Pseudomonas fluorescens* présente dans les réseaux de distribution d’eau et responsable de la formation des biofilms, eux-mêmes à l’origine de certaines maladies nosocomiales.

Traitement de données

Le traitement de données est composé de deux axes décrits ci-dessous. Le premier axe est bien abouti et le deuxième reste exploratoire.

Inversion de données. Il s’agit d’estimer des paramètres physiques à partir d’une courbe de force expérimentale. Une courbe de force $z \mapsto f(z)$ est modélisée « par morceaux » : chaque morceau est relatif à un type d’interaction particulier (pas de contact entre les deux nano-objets, zone de contact, puis interaction mécanique après contact). La difficulté de l’identification du modèle repose sur la recherche des morceaux dont les extrémités sont des points de discontinuité de la courbe de force ou de ses dérivées. Cette recherche revient à segmenter la courbe de force en détectant notamment les sauts du signal et ses changements de pente. Une fois le signal segmenté, il s’agit d’ajuster les modèles physiques paramétriques dans chaque morceau par moindres carrés. L’algorithme de lissage par morceaux proposé est décrit dans le chapitre 2.

Couplé à une estimation des paramètres physiques sur chaque morceau par moindres carrés, cet algorithme fournit des résultats quantitatifs très précis. Cette méthode a donné lieu à la publication de revue [R9].

Séparation de sources retardées. Une image force-volume peut être analysée en traitant indépendamment chaque courbe de force qui la constitue par la méthode décrite ci-dessus. Cela conduit à la reconstruction d'un ensemble d'images 2D qui représentent chacune une propriété physico-chimique de l'échantillon (topographie, élasticité, *etc.*). C'est la méthode que nous avons publiée dans [R9].

Nous avons en parallèle envisagé deux modélisations plus fines des images force-volume [C13] qui reposent sur un modèle convolutif lorsque la pointe AFM n'est plus supposée ponctuelle, et un modèle de mélanges de sources [DUA 10, Chap. 5]. L'idée sous-jacente au modèle de mélange de sources est que la pointe AFM n'est plus en interaction avec un seul point (x, y) de la surface de l'échantillon mais avec une surface locale autour de ce point. Cela conduit à un problème de séparation de sources à partir de mélanges retardés du type $f(x, y, z) = \sum_{k=1}^p a_k(x, y, z) f_k(z - z_k(x, y))$. Dans cette équation, $f(x, y, z)$ représente la courbe de force mesurée pour la position (x, y) , les signaux sources $f_k(z)$ modélisent des interactions élémentaires, et les coefficients $a_k(x, y)$ caractérisent la décomposition du mélange $f(x, y, z)$ en terme de forces d'interaction élémentaires. Enfin, les retards $z_k(x, y)$ s'interprètent comme la topographie de l'échantillon en (x, y) . Nous avons montré la pertinence de ce modèle de mélange [C13] et abordé le problème de la séparation de sources sous l'hypothèse de parcimonie des signaux sources : une source est décrite par un signal polynomial par morceaux avec peu de morceaux. La procédure de séparation de sources se résume par la segmentation de chaque courbe de force, puis la mise en correspondance des points de discontinuités trouvés, correspondant aux extrémités des morceaux. Une première solution algorithmique a été proposée dans la thèse de Junbo Duan [DUA 10, Chap. 5] et mérite d'être plus amplement développée.

Références associées : [R8, R9, CH1, C12, C13, CN5, COM2, COM5, COM6, COM7, COM8].

2.2 Microscopie confocale de fluorescence

Ce projet a pour cadre une collaboration entre le CRAN et le Laboratoire Interdisciplinaire des Environnements Continentaux (LIEC, UMR CNRS 7360, Nancy) ayant pour objectif le développement d'outils originaux de traitement des images hyperspectrales acquises en microscopie confocale. Ce travail s'inscrit dans le projet ANR HÆSPRI (« *Hyperspectral Analysis and Enhanced Surface Probing of Representative bacteria-mineral Interaction* », 2009-2013) porté par Christian Mustin (LIEC).

Contrairement au projet sur la microscopie AFM qui permet d'imager *une* bactérie ou *une* cellule biologique à l'échelle de l'atome, la microscopie confocale fournit des images d'*une population* de cellules pour comprendre les interactions des cellules avec leur environnement. L'objectif est donc de fournir une description statistique du comportement d'une population à l'échelle macroscopique. Les deux points de vue sont complémentaires : l'analyse d'une cellule biologique à l'échelle nanométrique permet de comprendre les mécanismes responsables des processus d'infection de la cellule alors que le point de vue macroscopique met en évidence le comportement statistique de différents types de cellules.

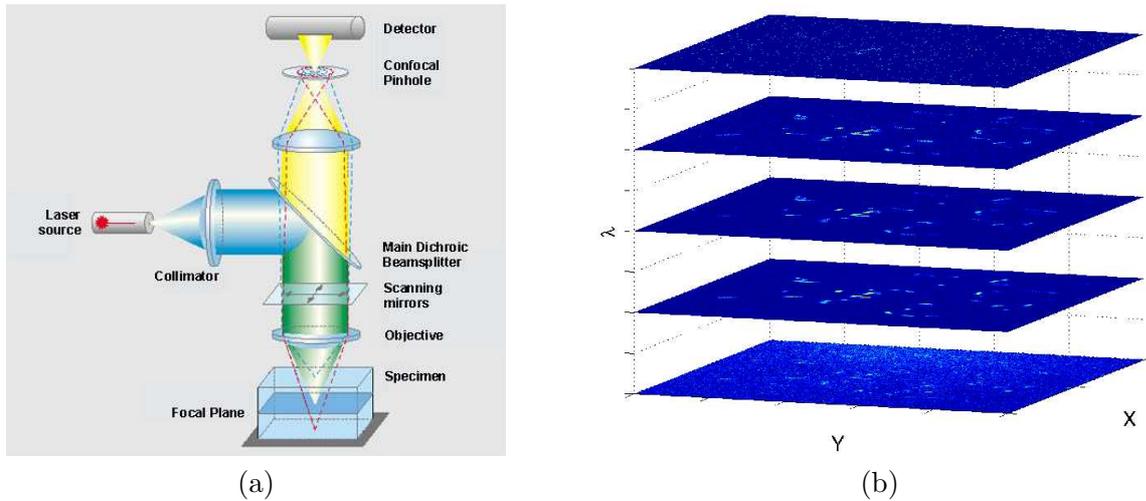


FIGURE 2 – Microscopie confocale et imagerie hyperspectrale. (a) Schéma de fonctionnement d’un microscope confocal (mesures de fluorescence). (b) Acquisition d’une pile d’images hyperspectrales $f(x, y, \lambda_i)$ dans le même plan focal. Les images sont de taille 512×512 pixels et la taille d’un pixel est de $100 \times 100 \text{ nm}^2$.

Imagerie de bio-capteurs bactériens

Le contexte du projet est l’étude de bio-capteurs bactériens dans les sols. Ce sont des bactéries génétiquement modifiées qui émettent des signaux de fluorescence lorsqu’elles sont excitées par un laser. Ces signaux sont liés à l’interaction de la bactérie avec son environnement : ils diffèrent selon que la bactérie est en contact avec une source de pollution ou non. Leur analyse a pour but de détecter et d’identifier les polluants présents dans les sols.

Pour pouvoir modéliser la réponse fonctionnelle adaptative des bactéries à leur environnement minéral, le LIEC possède à la fois des micro-spectromètres performants (spectrofluoromètre, microscope confocal) capables d’acquérir des images multivariées ou multimodales (espace, longueur d’onde, concentration, température, *etc.*) résolues à l’échelle d’une cellule, et des outils de microbiologie et de biologie moléculaire permettant de sélectionner des agents microbiens particulièrement actifs des sols et de les modifier génétiquement. Les bio-capteurs environnementaux sont obtenus par insertion de gènes rapporteurs codant des émetteurs lumineux ou fluorescents [LEV 02, LAR 05]. Grâce à ces outils, il est possible d’entreprendre des analyses spectrales de suspensions de bio-capteurs. Ils jouent le rôle de bactéries « rapporteuses » interagissant avec des grains minéraux de petite taille dans des conditions variables. L’émission polychromique de lumière résulte des multiples interactions entre les différentes bactéries et leur milieu environnemental.

Un microscope confocal est un microscope optique qui permet d’acquérir des images de fluorescence $g(x, y, \lambda)$ à différentes longueurs d’ondes λ . Ces images représentent un ensemble de bactéries génétiquement modifiées (bio-capteurs), illuminé par un rayonnement laser. Le schéma de fonctionnement d’un microscope confocal est illustré sur la figure 2. Une lumière laser monochromatique est envoyée sur l’échantillon en positionnant le plan focal à un certain niveau de profondeur. Pour un niveau de profondeur z donné et en chaque pixel (x, y) , l’interaction de la lumière incidente avec l’échantillon produit l’émission d’un signal ou *spectre de fluorescence* $\lambda \mapsto g(x, y, \lambda)$ pour des longueurs d’ondes λ supérieures à la longueur d’onde excitatrice. Le miroir dichroïque permet de séparer la longueur d’onde incidente des longueurs d’ondes liées aux

rayons émis. Ce système d'acquisition peut donc produire des *images hyperspectrales* 3D $g(x, y, \lambda)$ ou 4D $g(x, y, z, \lambda)$ en faisant varier la profondeur z du plan focal. Les données peuvent être vues comme une pile d'images (*cf.* fig. 2(b)) ou comme un ensemble de spectres 1D $\lambda \mapsto g(x, y, \lambda)$ pour un grand nombre de positions spatiales (x, y) ou (x, y, z) .

Séparation et restauration d'images hyperspectrales

Déconvolution d'images. Lors de l'acquisition, les images présentent des dégradations (convolution ou flou) dues à l'appareil de mesure [BIG 04, PAN 09]. La déconvolution consiste à enlever le flou pour restaurer des images nettes à chacune des longueurs d'ondes [SAR 06, DUP 09]. Les données $g(x, y, \lambda)$ sont modélisées par :

$$g(x, y, \lambda) = h(x, y; \lambda) \underset{(x,y)}{*} f(x, y, \lambda), \quad (1)$$

où h est la réponse impulsionnelle du microscope et pour chaque longueur d'onde λ , $f(x, y, \lambda)$ désigne l'image nette inconnue.

Au lieu de restaurer séparément chacune des images hyperspectrales $f(x, y, \lambda)$, nous avons proposé dans la thèse de Simon Henrot de traiter ces problèmes *de façon conjointe* (quelquesoit λ) [R4, C2] en imposant d'une part la positivité de l'image f et d'autre part une régularisation inter-images, *i.e.*, l'image restaurée à une longueur d'onde λ prend en compte l'image $g(x, y, \lambda)$ mais aussi les données aux longueurs d'ondes adjacentes $g(x, y, \lambda \pm \delta\lambda)$. Nous avons proposé un algorithme rapide adapté aux images hyperspectrales de grandes dimensions en imposant une régularisation de Tikhonov [R4]. Puis, nous avons étendu cet algorithme à une régularisation préservant les contours spatiaux dans l'image $f(x, y, \lambda)$ [C2]. Ces algorithmes sont basés sur des transformées de Fourier 2D (transformée de Fourier de chaque image spatiale $g(x, y, \lambda)$ pour λ fixé) et des calculs rapides dans le domaine de Fourier spatial.

Séparation de sources. Pour comprendre l'interaction de micro-organismes dans un milieu naturel particulier (par exemple, l'exposition de bactéries à un métal toxique), nous avons proposé d'analyser les images hyperspectrales par une procédure de séparation de sources. Un mélange correspond à un pixel, et le nombre de mélanges peut typiquement atteindre 512^2 . Pour un ensemble de bactéries génétiquement modifiées, un signal source décrit soit la fluorescence d'une bactérie non modifiée génétiquement, soit celle de l'un des gènes ajoutés. En un pixel (x, y) , le signal de fluorescence $\lambda \mapsto f(x, y, \lambda)$ s'exprime comme le mélange bilinéaire de signaux sources $s_k(\lambda)$:

$$f(x, y, \lambda) = \sum_{k=1}^p a_k(x, y) s_k(\lambda), \quad (2)$$

où $a_k(x, y)$ représente le poids de la k -ème source à l'intérieur du pixel et p désigne le nombre de sources présentes. Les poids $a_k(x, y)$ et les signaux sources $s_k(\lambda)$ sont tous à valeurs positives ou nulles.

La spécificité du problème de séparation de sources est liée au très grand nombre de mélanges. Pour pouvoir appliquer les algorithmes de séparation de sources positives en un temps de calcul raisonnable, il faut procéder à une réduction de dimensionnalité, par exemple par une procédure de sélection de pixels. Nous avons proposé une solution [C10] qui couple une procédure de sélection de mélanges et l'utilisation de l'algorithme *Bayesian Positive Source Separation* conçu au CRAN [MOU 06].

Déconvolution d'images et séparation de sources

Dans thèse de Simon Henrot, nous avons abordé la séparation aveugle de sources en considérant l'influence d'une étape préliminaire de déconvolution sur les résultats de séparation. Nous nous sommes intéressés aux conditions d'identifiabilité des signaux sources qui reposent sur le modèle géométrique de simplexe de volume minimal. Pour le modèle de mélange positif de sources (2) sans effet convolutif, la condition (suffisante) classique de « pixels purs » [BIO 12] stipule que chaque source doit être observable à l'état pur dans l'un des mélanges ($f(x, y, \lambda) = a_k(x, y)s_k(\lambda)$) pour garantir l'identifiabilité. Géométriquement, chaque sommet du simplexe de volume minimal est atteint. Cette condition très restrictive peut néanmoins être relaxée en une condition portant la présence de mélanges de sources, situés sur les facettes du simplexe [RT1].

Pour des données convoluées $g(x, y, \lambda)$ générées par les modèles (1) et (2), l'analyse du problème d'identifiabilité montre que la convolution (1) par une réponse impulsionnelle $h(x, y; \lambda)$ à valeurs positives induit une contraction des points (mélanges de sources) à l'intérieur du simplexe de volume minimal. En d'autres termes, le caractère mal posé du problème de séparation de sources à partir des données brutes $g(x, y, \lambda)$ est accru. Pratiquer une déconvolution des images $g(x, y, \lambda)$ en amont de la séparation de sources est donc très pertinent [RT1]. Ces résultats théoriques ont été validés par le traitement de données réelles en microscopie Raman, fournies par Manuel Dossot (LCPME).

Ces travaux ouvrent la voie au développement de méthodes conjointes de déconvolution et de séparation de sources. Le défaut d'une approche en deux temps (déconvolution puis séparation) est d'une part une propagation d'erreurs : des résultats de déconvolution imparfaits sont utilisés en entrée de la procédure de séparation. D'autre part, lorsqu'on régularise le problème de déconvolution (1), on impose des contraintes sur les images $f(x, y, \lambda)$ liées au mélange de sources alors qu'en réalité, les images que l'on souhaite estimer sont les images $a_k(x, y)$. Pour toutes ces raisons, notre projet est de concevoir un algorithme qui réalise simultanément la restauration des images hyperspectrales et la séparation de sources positives. La difficulté de ce problème vient du fait que la réponse impulsionnelle $h(x, y; \lambda)$ dépend des dimensions spatiales x et y , mais aussi de la longueur d'onde λ à laquelle l'image est acquise. En intégrant les modèles (1) et (2), on s'aperçoit que les données convoluées

$$g(x, y, \lambda) = \sum_k [h(x, y; \lambda) \underset{(x,y)}{*} a_k(x, y)] s_k(\lambda) \quad (3)$$

ne s'expriment pas comme un mélange bilinéaire des sources s_k , car le terme $h(x, y; \lambda) \underset{(x,y)}{*} a_k(x, y)$ dépend de λ . L'hypothèse de mélange instantané de signaux sources n'est donc plus valable.

Références associées : [R4, C2, C3, C10, CN2, CN3, COM3, COM4, RT1].

2.3 Conception d'algorithmes de restauration de signaux parcimonieux

Restauration de signaux parcimonieux

Je résume à présent les principales contributions méthodologiques qui seront détaillées dans la partie II du manuscrit. Elles concernent les problèmes inverses linéaires formulés par :

$$\mathbf{y} = \mathbf{H} \underbrace{\mathbf{W}\mathbf{x}}_{\mathbf{u}} + \mathbf{b} \quad (4)$$

où \mathbf{y} représente le vecteur des observations, $\mathbf{u} = \mathbf{W}\mathbf{x}$ est le signal ou l'image à restaurer, \mathbf{H} est la matrice d'observation qui modélise le processus physique d'acquisition des données, et le

« bruit » \mathbf{b} regroupe le bruit de mesure et les erreurs induites par le choix du modèle linéaire ($\mathbf{H}\mathbf{u}$). Une régularisation parcimonieuse consiste à imposer que \mathbf{u} prend la forme $\mathbf{u} = \mathbf{W}\mathbf{x}$ où \mathbf{x} est un vecteur parcimonieux, *i.e.*, contenant un très faible nombre d'éléments x_i non nuls. L'objectif est finalement d'estimer \mathbf{x} , et implicitement de retrouver le support de \mathbf{x} à partir de \mathbf{y} et de la matrice \mathbf{HW} . Ce problème est aussi connu sous le nom de *sélection de variables* en régression statistique [MIL 02] où une *variable* est une coordonnée x_i du vecteur \mathbf{x} .

Approximation parcimonieuse

L'approximation parcimonieuse a pour but de représenter un signal observé \mathbf{y} comme une combinaison linéaire d'un nombre minimal de signaux élémentaires, appelés atomes. Les atomes sont choisis dans un dictionnaire qui est une matrice regroupant tous les atomes possibles. En notant \mathbf{A} le dictionnaire, il s'agit de trouver une approximation $\mathbf{y} \approx \mathbf{A}\mathbf{x}$, où \mathbf{x} est vecteur parcimonieux.

La restauration de signaux parcimonieux dans le cadre des problèmes inverses et l'approximation parcimonieuse peuvent sembler très similaires puisque \mathbf{A} s'identifie avec \mathbf{HW} (²). La différence réside dans le fait qu'en restauration parcimonieuse, le dictionnaire est imposé par la physique du problème alors qu'en approximation parcimonieuse, il est choisi de façon à représenter les données \mathbf{y} avec un nombre limité d'atomes. Pour les applications mentionnées précédemment, le dictionnaire \mathbf{A} est formé de la façon suivante : en reconstruction tomographique, $\mathbf{A} = \mathbf{H}$ est la transformée en rayons X d'une image 3D, et l'image $\mathbf{u} = \mathbf{x}$ est supposée parcimonieuse dans la base canonique. Pour la segmentation de signaux en microscopie AFM, \mathbf{H} est la matrice unité et le signal est supposé homogène par morceaux : il possède une représentation parcimonieuse dans un dictionnaire polynomial \mathbf{W} qui joue le rôle de la matrice \mathbf{A} , comme nous le verrons un peu plus loin. Le modèle (4) n'a en revanche pas été utilisé dans le contexte de l'imagerie hyperspectrale.

Dictionnaires mal conditionnés

Le dénominateur commun entre les dictionnaires étudiés dans ce manuscrit est qu'ils sont mal conditionnés. La conception de nouveaux algorithmes est motivée par le constat que certains algorithmes populaires dans le domaine de l'approximation parcimonieuse (inspirés du domaine de l'échantillonnage compressé où le dictionnaire est quasi-orthogonal) trouvent des limites lorsque le dictionnaire est une matrice mal conditionnée.

Optimisation mixte ℓ_2 - ℓ_0 sous formes contrainte et pénalisée

Le problème d'approximation parcimonieuse est abordé comme l'optimisation d'un critère mixte $\ell_2 - \ell_0$ où le terme ℓ_2 se rapporte à l'erreur quadratique d'approximation $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ et le coût ℓ_0 , noté $\|\mathbf{x}\|_0$, est le degré de parcimonie de \mathbf{x} défini comme le nombre d'éléments non nuls de \mathbf{x} .

L'approximation parcimonieuse est traditionnellement formulée comme un problème d'optimisation sous contrainte de parcimonie :

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.c.} \quad \|\mathbf{x}\|_0 \leq k \quad (5)$$

2. J'adopte dans ce manuscrit l'approche dite « synthèse » pour la parcimonie où \mathbf{H} et \mathbf{W} sont regroupés dans une seule matrice. Une alternative est l'approche « analyse » qui découple matrice d'observation et matrice de régularisation [ELA 07] ; voir la section 4.2.1 du chapitre 4 pour plus de précisions.

ou comme

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.c.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \varepsilon \quad (6)$$

suisant que l'on possède une connaissance d'une borne k sur le nombre maximal de coefficients non nuls ou d'une borne ε sur l'erreur d'approximation souhaitée, qui s'apparente à la variance du bruit de mesure \mathbf{b} .

L'approche bayésienne pour les problèmes inverses repose sur la définition d'une distribution de probabilité *a priori* pour le vecteur parcimonieux [R8]. Un choix adapté est le modèle Bernoulli-gaussien [KOR 82, MEN 83, CHA 96, TUR 11] : le support de \mathbf{x} suit la loi de Bernoulli et les amplitudes non nulles x_i sont distribuées suivant une loi gaussienne. Nous avons considéré le modèle Bernoulli-gaussien limite quand la variance de la gaussienne tend vers $+\infty$, c'est-à-dire sans imposer d'information *a priori* sur la valeur des amplitudes. Nous avons montré [R8] qu'en supposant la loi du bruit gaussienne, l'estimation de \mathbf{x} au sens du maximum *a posteriori* (MAP) conduit à minimiser le critère pénalisé :

$$\mathcal{J}(\mathbf{x}; \lambda) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0, \quad (7)$$

où l'hyperparamètre λ contrôle le degré de parcimonie de la solution.

Du point vue optimisation, notons que les formulations contraintes (5) et (6) ne sont pas équivalentes à la formulation non contrainte (7) [NIK 13] (*cf.* chapitre 1, § 1.7).

Algorithmes utilisant la pseudo-norme ℓ_0

Nous avons d'abord utilisé des algorithmes gloutons *monodirectionnels*³ adaptés aux formes contraintes (5) ou (6). Leur principe est d'ajouter à chaque itération un élément au support de \mathbf{x} et de mettre à jour les amplitudes x_i par projection orthogonale. Il s'agit des algorithmes *Orthogonal Matching Pursuit* (OMP) [PAT 93, DAV 94] et *Orthogonal Least Squares* (OLS) [CHE 89].

De plus, nous nous sommes intéressés à la minimisation du critère pénalisé (7). **Nous avons conçu de nouveaux algorithmes gloutons *bidirectionnels***⁴ permettant de :

- minimiser $\mathcal{J}(\mathbf{x}; \lambda)$ à λ fixé ;
- minimiser $\mathcal{J}(\mathbf{x}; \lambda)$ pour un continuum de valeurs de λ , et ainsi fournir des solutions à différents niveaux de parcimonie λ .

Ces algorithmes, nommés *Single Best Replacement* (SBR), *Continuation Single Best Replacement* (CSBR) et ℓ_0 *regularization path track* (ℓ_0 -PT) sont qualifiés de « bidirectionnels » car à chaque itération, le support de \mathbf{x} est mis à jour en incrémentant ou en décrémentant sa taille d'un élément. Ces algorithmes sont inspirés de l'algorithme *Single Most Likely Replacement* (SMLR), initialement proposé pour résoudre des problèmes de déconvolution de signaux impulsionnels modélisés par un processus Bernoulli-gaussien (optimisation de critères proches de la forme (7)) [KOR 82, MEN 83, CHA 96]. Ils ont une structure proche des algorithmes appelés *stepwise regression algorithms* en régression statistique [MIL 02, Chap. 3]. L'efficacité reconnue (depuis les années 1980) de l'algorithme SMLR pour des problèmes de restauration parcimonieuse a été une source de motivation pour l'adapter au contexte de la minimisation ℓ_2 - ℓ_0 et proposer de nouveaux algorithmes efficaces.

Références associées : [R8, C5, C12, COM5, RT2].

3. *forward greedy pursuit algorithms* en anglais.

4. *forward-backward greedy algorithms*, dans le sens « ajout et retrait » d'atomes.

Problèmes traités

Nous avons appliqué et adapté les algorithmes ci-dessus dans différents contextes :

- Le lissage d’un signal par un signal polynomial par morceaux (et implicitement la recherche adaptative des morceaux) avec comme application la microscopie AFM [R9]. La notion de parcimonie est ici liée au nombre de morceaux du signal lissé. Par exemple, un signal constant par morceaux est modélisé par la position des sauts du signal, et le degré de parcimonie est le nombre de morceaux où le signal est constant. Plus généralement, pour un signal polynomial par morceaux, on définit un point de discontinuité comme l’extrémité d’un morceau et l’ordre p du point de discontinuité de la façon suivante : $p = 0$ correspond à un saut du signal, $p = 1$ à un changement de pente, $p = 2$ à un changement de courbure. A p fixé, on introduit le dictionnaire \mathbf{A}^p constitué des fonctions $a_i^p(t) = (t - t_i)_+^p \triangleq [\max(t - t_i, 0)]^p$, de telle sorte que l’apparition d’une discontinuité en n’importe quelle position $t = t_i$ soit rendue possible. Finalement, le dictionnaire

$$\mathbf{A} = [\mathbf{A}^0 | \mathbf{A}^1 | \dots | \mathbf{A}^P] \quad (8)$$

est obtenu en concaténant les dictionnaires \mathbf{A}^p relatifs à chaque ordre. Les algorithmes d’approximation parcimonieuse permettent de détecter les points de discontinuité (*i.e.*, les morceaux) tout en lissant le signal (*cf.* chapitre 2).

- Le contrôle non destructif des matériaux par ultrasons qui conduit à un problème de déconvolution impulsionnelle du type $\mathbf{y} = \mathbf{h} \star \mathbf{x} + \mathbf{b}$ où la réponse impulsionnelle \mathbf{h} correspond à l’ondelette acoustique incidente. Les algorithmes SBR et CSBR permettent d’améliorer sensiblement les performances des méthodes utilisées en contrôle non destructif par ultrasons [R7]. De plus, des implémentations efficaces exploitant la structure convolutive du problème sont disponibles [R8].
- L’estimation des paramètres de signaux modaux (sinusoïdes amorties) bidimensionnels en spectroscopie de Résonance Magnétique Nucléaire bidimensionnelle (RMN 2-D) [R6]. Il s’agit de décomposer un signal donné comme une somme de motifs paramétriques dont le centre et les paramètres de forme sont inconnus. En discrétisant ces paramètres inconnus sur une grille, on forme un dictionnaire paramétrique de grande taille contenant tous les motifs potentiellement présents dans le signal. La solution algorithmique proposée consiste à coupler une méthode d’approximation parcimonieuse avec une analyse multi-échelle pour raffiner le dictionnaire. L’algorithme est initialisé avec un dictionnaire correspondant à une grille grossière, cette dernière est ensuite raffinée de manière adaptative en fonction des atomes sélectionnés par la méthode d’approximation parcimonieuse. L’algorithme d’analyse modale 2-D ainsi obtenu est rapide et très compétitif avec les méthodes concurrentes [R6].

Références associées : [R6, R7, R9, CH1, C4, CN4, CN5].

3 Analyse en k itérations des algorithmes OMP et OLS

3.1 Contexte et cheminement du projet

Ce projet a été initié en 2011 lors de mon séjour de délégation CNRS à l’IRCCyN (Nantes). Avec Jérôme Idier, nous nous sommes posé la question de l’analyse du comportement de l’algorithme SBR que nous avons conçu les années précédentes, en terme de capacité de reconstruction exacte du support inconnu. Comme SBR est une version bidirectionnelle de l’algorithme OLS qui intègre une politique de retraits d’atomes en plus de l’ajout d’un atome à la fois, nous nous

sommes naturellement intéressés à la reconstruction exacte de supports par OLS. Nous avons fait le constat qu’aucun résultat de ce type n’était disponible dans la littérature (d’autres auteurs ont fait le même constat [DAV 12]) bien qu’OLS soit un algorithme ancien, conçu au début des années 1990 [CHE 89, NAT 95] pour la sélection de modèles en statistique. Le fait que l’analyse théorique d’OLS soit un problème ouvert s’explique de deux façons :

1. Dans la littérature récente, de nombreux algorithmes d’approximation parcimonieuse *rapides* ont été proposés et analysés (voir par exemple le livre [ELA 10] pour une revue des algorithmes récents). Dans le domaine de l’échantillonnage compressé, utiliser des algorithmes rapides est nécessaire pour reconstruire un signal en « temps-réel » [RAU 08, FOU 13b]. L’algorithme OLS est sensiblement plus coûteux que d’autres algorithmes classiques (l’algorithme OMP notamment) d’où un intérêt limité pour OLS par les chercheurs du domaine de l’approximation parcimonieuse. Notons également que pour les problèmes d’échantillonnage compressé, les colonnes du dictionnaire sont peu corrélées les unes aux autres. Dans ce contexte, des algorithmes plus simples qu’OLS donnent satisfaction, et utiliser OLS au lieu d’OMP n’apporte pas un gain significatif. Ce n’est pas le cas des problèmes inverses.
2. OLS a une structure qui ressemble à OMP. Les deux algorithmes sont d’ailleurs fréquemment confondus dans la littérature (je renvoie le lecteur au rapport technique [BLU 07] pour un éclairage sur la question). La différence entre les deux algorithmes est que la règle de sélection d’atomes diffère : elle est plus complexe et coûteuse pour OLS. Leur point commun est qu’à chaque itération, les amplitudes \mathbf{x} des atomes dans la décomposition $\mathbf{y} \approx \mathbf{A}\mathbf{x}$ sont calculées en effectuant une projection orthogonale. La structure plus complexe d’OLS explique qu’aucune analyse technique de l’algorithme n’ait été effectuée avant 2012.

Nous avons donc entrepris le projet d’analyser OLS en collaboration avec l’INRIA Rennes (Rémi Gribonval et Cédric Herzet), un laboratoire reconnu dans l’analyse théorique des algorithmes de représentation parcimonieuse.

Comme OMP et OLS ont une structure très proche, l’idée de départ était de s’inspirer des conditions de reconstruction exacte pour OMP qui garantissent de reconstruire le support de \mathbf{x} à partir de données non bruitées $\mathbf{y} = \mathbf{A}\mathbf{x}$. Parmi ces conditions, la condition de reconstruction exacte ERC (*Exact Recovery Condition*) proposée par Tropp [TRO 04] est un résultat majeur. Elle s’écrit :

$$\max_{j \notin \mathcal{Q}^*} \|\mathbf{A}_{\mathcal{Q}^*}^\dagger \mathbf{a}_j\|_1 < 1, \quad \text{ERC}(\mathbf{A}, \mathcal{Q}^*)$$

où \mathcal{Q}^* désigne le support inconnu (le support de \mathbf{x}), \mathbf{a}_j sont les « mauvais » atomes, la matrice $\mathbf{A}_{\mathcal{Q}^*}$ rassemble les « vrais » atomes $\{\mathbf{a}_i, i \in \mathcal{Q}^*\}$ et \cdot^\dagger désigne la pseudo-inverse d’une matrice. $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ est une condition suffisante de reconstruction exacte : si $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ est vérifiée, alors le support \mathcal{Q}^* est correctement reconstruit par OMP en $k = \text{Card}[\mathcal{Q}^*]$ itérations à partir de n’importe quelle combinaison linéaire $\mathbf{y} = \mathbf{A}_{\mathcal{Q}^*}\mathbf{t}$ de k atomes indicés par \mathcal{Q}^* . $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ est de plus une condition nécessaire *au pire cas* : si $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ n’est pas vérifiée, alors il existe une combinaison linéaire $\mathbf{y} = \mathbf{A}_{\mathcal{Q}^*}\mathbf{t}$ pour laquelle OMP sélectionne un mauvais atome $j \notin \mathcal{Q}^*$. En fait, le mauvais atome est sélectionné à la première itération.

3.2 Contributions

Nous avons d’abord montré que **la condition ERC est également une condition suffisante et nécessaire au pire cas pour OLS** [R3]. La partie « nécessaire » de ce résultat est

immédiate car la première itération d'OLS s'identifie avec celle d'OMP : comme ERC est une condition nécessaire pour OMP à la première itération, OLS sélectionne forcément un mauvais atome à la première itération pour les contre-exemples où OMP sélectionne un mauvais atome.

Pour pouvoir discriminer le comportement des deux algorithmes, nous avons élaboré une **analyse plus fine au delà de la première itération**. Cette analyse consiste à supposer que les $q < k$ premières itérations se sont déroulées avec succès (q vrais atomes ont été sélectionnés) et de définir des conditions garantissant le succès de la $(q + 1)$ -ème itération. Nous avons :

- proposé des nouvelles conditions, notées $\text{ERC-OMP}(\mathbf{A}, \mathcal{Q}^*, q)$ et $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, q)$, garantissant le succès de la $(q + 1)$ -ème itération d'OMP (respectivement d'OLS).
- montré que ces conditions sont nécessaires dans un certain sens ;
- montré que ces conditions garantissent le succès de toutes les itérations suivantes $q + 1, \dots, k$, ce qui implique que le support \mathcal{Q}^* est correctement reconstruit en k itérations ;
- montré qu'il n'y a pas d'implication systématique, ni dans un sens ni dans l'autre, entre $\text{ERC-OMP}(\mathbf{A}, \mathcal{Q}^*, q)$ et $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, q)$;
- cependant, des évaluations numériques montrent que $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, q)$ est plus souvent vérifiée lorsque le dictionnaire est corrélé, ce qui valide le choix d'OLS dans ce contexte.

Les conditions $\text{ERC-OMP}(\mathbf{A}, \mathcal{Q}^*, q)$ et $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, q)$ sont des conditions suffisantes raffinées mais très difficilement exploitables en pratique car leur évaluation est coûteuse, c'est pourquoi je ne les détaille pas dans ce chapitre. Le lecteur intéressé est renvoyé au chapitre 3. Nous avons établi une condition nécessaire plus forte mais très simple portant sur la cohérence mutuelle du dictionnaire [R1]. Elle est définie comme le produit scalaire maximal entre deux atomes du dictionnaire (que l'on suppose normalisé) :

$$\mu = \max_{i \neq j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle|, \quad (9)$$

et la condition suffisante prend la forme $\mu < 1/(2k - q - 1)$. C'est une condition nécessaire et suffisante au pire cas (pour tous les supports possibles de cardinalité q) pour le succès d'OMP et d'OLS en k itérations lorsque q vrais atomes ont déjà été sélectionnés. Dans le cas $q = 0$, on retrouve la condition classique $\mu < 1/(2k - 1)$ [TRO 04].

Références associées : [R1, R3, C1].

4 Recalage et reconstruction d'images 3D

Cette activité a été amorcée (pour ma part) en 2007 au sein du groupe thématique « Ingénierie Pour la Santé » (G.T. IPS) du CRAN et poursuivie dans le cadre du département SBS. Bien qu'étant principalement affecté dans le G.T. IRIS à mon arrivée au CRAN, j'ai constaté qu'une activité de recherche en reconstruction d'image (mosaïquage d'images) en vision par ordinateur était menée dans le G.T. IPS, par Christian Daul et Walter Blondel, et que des interactions étaient possibles. Nous avons entamé une collaboration dans le cadre de la thèse d'Achraf Ben Hamadou, dont le sujet consistait à proposer de nouvelles méthodes de reconstruction et de recalage de surfaces 3D représentant l'intérieur d'un organe creux à partir d'images endoscopiques.

Le contexte applicatif du projet est la détection du cancer de la vessie à partir d'une séquence d'images vidéo acquises lors d'un examen endoscopique. Le CRAN a développé par le passé [MIR 05, HER 07] des algorithmes permettant de reconstruire une image panoramique qui représente l'intérieur de la surface de la vessie, dépliée en une image 2D. Ce traitement repose sur :

1. le recalage deux à deux des images de la séquence. Il s'agit d'estimer une transformation perspective 2D/2D pour mettre en correspondance chaque couple d'images consécutives.

La mise en correspondance est effectuée en maximisant une mesure de similarité (par exemple l'information mutuelle) entre l'une des images et la transformation perspective de la seconde.

2. la construction d'une image panoramique (aussi appelée *mosaïquage*) résultant de l'assemblage (du collage) des images de la séquence vidéo sur une même carte. Cet assemblage prend en compte toutes les transformations géométriques estimées lors de l'étape de recalage. Le mosaïquage d'images permet de visualiser en 2D l'intérieur de la surface de la vessie. Cette visualisation est néanmoins partielle car l'endoscope ne filme généralement pas la totalité de l'organe mais des zones particulières de sa surface.

L'objectif de la thèse d'Achraf Ben Hamadou était de **construire une représentation 3D de la surface interne de la vessie**. La reconstruction 3D n'est pas envisageable à partir de la séquence d'images acquises lors d'un examen endoscopique standard car la position spatiale de l'endoscope est inconnue, et de plus, l'endoscope se déplace à l'intérieur de la vessie durant l'examen endoscopique avec un mouvement inconnu. Pour pouvoir aborder la reconstruction 3D, nous avons :

1. conçu un **dispositif de vision active** monté en bout d'endoscope. Ce dispositif est capable de fournir en plus de la séquence d'images vidéo, des informations 3D liées à la position de l'endoscope par rapport à la surface de la vessie ;
2. proposé des algorithmes de **reconstruction de surface** (recalage et mosaïquage) à partir des informations partielles recueillies par le dispositif.

Je résume de façon plus précise nos contributions dans les deux prochains paragraphes. Elles ont été pensées dans le cadre de l'application à l'endoscopie médicale, mais elles sont transposables à d'autres applications de la vision active.

Système à lumière structurée préservant les textures. La partie instrumentation du projet consiste à projeter un motif lumineux (lumière structurée) sur la surface de la vessie. Le prototype que nous avons conçu est illustré sur la figure 3. Le système est constitué d'un projecteur à lumière structurée rigidement attaché à la caméra CCD qui acquiert les images 2D. La lumière structurée est générée par une source laser verte et un système d'optique diffractive permettant de transformer le rayon laser incident en 8 nouveaux rayons concentriques (fig. 3(a)). Contrairement à la plupart des systèmes à lumière structurée, ce système préserve la texture des surfaces car le motif projeté est parcimonieux (8 rayons lasers seulement, qui se projettent de façon quasi-ponctuelle en 8 « points laser » dans l'image acquise : cf. fig. 3(d)). La préservation des textures est indispensable en endoscopie médicale car l'objectif est de diagnostiquer la présence d'une tumeur en visualisant la séquence vidéo d'images texturées.

Nous avons proposé une méthode flexible pour calibrer un système projecteur-caméra (aussi appelé *Procam* dans la littérature) comme celui de la figure 3(a) [R2]. Contrairement à d'autres méthodes de la littérature [MAR 02, CHA 03, DIP 05], aucun système de positionnement n'est nécessaire pour effectuer la calibration, et les résultats sont cependant très précis. La géométrie du projecteur est déterminée dans le système de coordonnées de la caméra par une procédure d'optimisation numérique, où le centre optique du projecteur est estimé en plus des droites supportant les rayons laser 3D issues du projecteur. Le système projectif permet finalement de reconstruire les coordonnées 3D de chaque point laser dans le repère de la caméra [R2, C6].

Reconstruction 3D surfacique. En utilisant le prototype d'endoscope à lumière structurée présenté ci-dessus, nous sommes capables d'acquérir une séquence d'*images augmentées*.

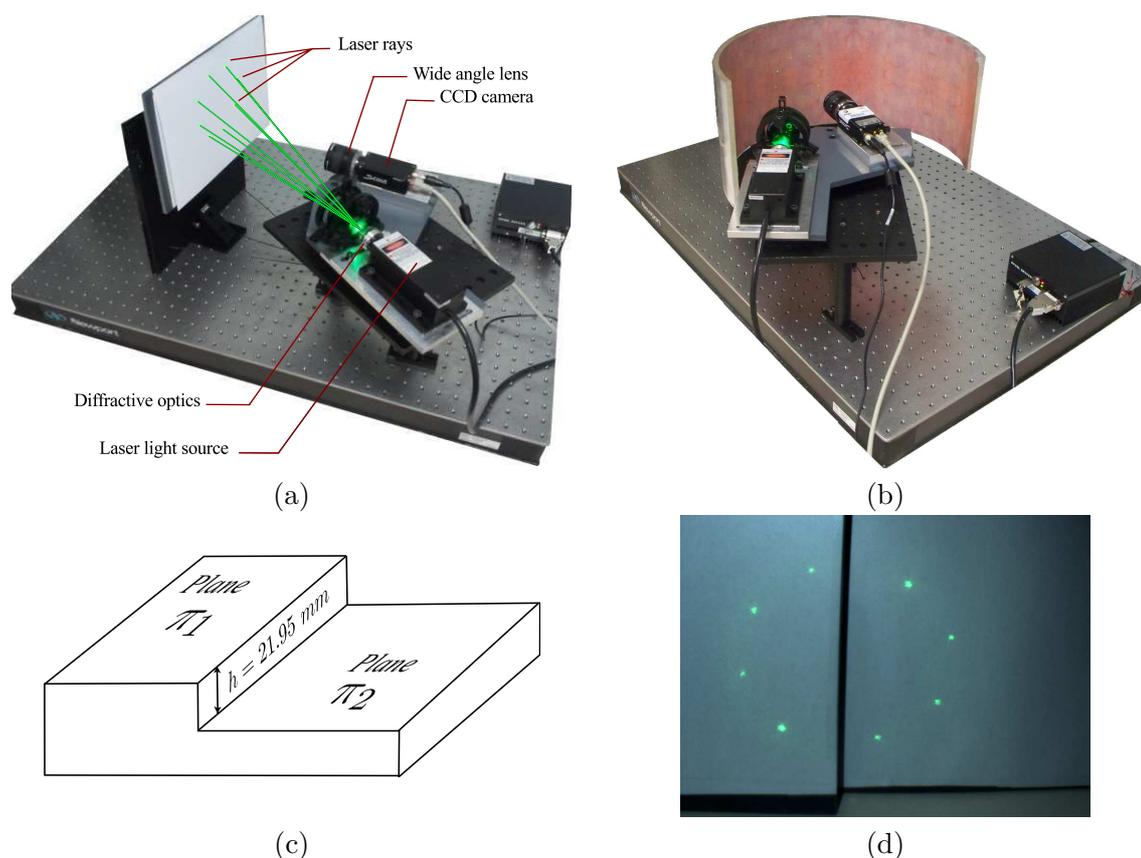


FIGURE 3 – Prototype de système à lumière structurée conçu au CRAN. Le système est constitué d'une caméra CCD et d'un projecteur de lumière structurée. Les 8 rayons laser de la figure (a) montrent la trajectoire des points projetés dans l'espace. Ils ne sont pas visibles en pratique. (b) Acquisition d'une image vidéo d'un fantôme cylindrique. L'image acquise contient, en plus des informations de texture, les 8 points laser projetés. (d) Exemple d'image acquise pour un objet formé de deux plans parallèles (c).

Chaque image augmentée est constituée d'une image couleur standard représentant la texture de la surface, et des 8 points lasers 3D reconstruits dans le repère local de la caméra CCD. La reconstruction surfacique consiste à mettre en correspondance ces points lasers en considérant toutes les images de la séquence vidéo. De façon équivalente, il s'agit d'estimer le mouvement 3D de l'endoscope dans un repère commun unique.

Nous avons proposé **une méthode d'estimation du mouvement du système Procam** [C7]. Le mouvement du système projectif provoque un déplacement des points laser sur la surface inspectée, ce qui rend les méthodes traditionnelles basées sur l'appariement des points homologues inopérantes. L'estimation du mouvement est guidée par le recalage des images 2D acquises à des instants consécutifs. L'information homologue exploitée est la texture de la surface imagée, que l'on retrouve en grande partie entre deux images consécutives de la séquence [C7]. Le mouvement de l'endoscope est estimé non pas en recalant des points 3D, mais en reliant le mouvement 3D au déplacement 2D des structures (textures) dans les images, et en estimant le mouvement 3D permettant de recaler au mieux les images 2D.

Finalement, nous sommes capables de positionner les $8N$ points lasers (où N désigne le nombre d'images de la séquence vidéo) dans un système de coordonnées commun, et de recons-

truire en 3D la surface de l'objet en interpolant les $8N$ points. Notre méthode permet de fournir une cartographie 3D texturée des parois internes de la surface (la vessie) en superposant à la surface 3D reconstruite, les textures contenues dans les images 2D.

Références associées : [R2, C6, C7, C8, C11, CN6].

Deuxième partie

**Algorithmes d'approximation
parcimonieuse pour les problèmes
inverses**

Avant-propos

Dans cette partie, j'ai choisi de présenter plus particulièrement mon activité liée à l'approximation parcimonieuse. J'ai fait ce choix pour plusieurs raisons : d'une part, parce que c'est le thème sur lequel je me suis le plus impliqué ces dernières années. D'autre part, pour privilégier la cohérence entre les chapitres de cette partie. J'ai adopté un découpage en terme de contributions méthodologiques (algorithmiques), appliquées et théoriques (analyse d'algorithmes). Ce découpage reflète assez fidèlement le temps que j'ai consacré à ces trois sujets. Il reflète aussi l'ordre chronologique des travaux. Les chapitres 1, 2 et 3 sont en grande partie basés sur la publications [SOU 11, POL 11, SOU 13b]. Les chapitres 1 et 3 intègrent aussi des travaux publiés ultérieurement [HER 13] ou en cours de finalisation [SOU 13d]. En règle générale, je m'efforce d'adapter ces travaux pour privilégier la pédagogie de l'exposé.

Les travaux liés aux chapitres 1 et 2 ont été menés de concert depuis 2007. Si les algorithmes présentés dans le chapitre 1 sont des algorithmes génériques, ils ont été conçus dans le but de minimiser des critères mixtes ℓ_2 - ℓ_0 comme celui que j'aborde au chapitre 2 pour la segmentation par morceaux de signaux AFM, où le terme ℓ_2 fait intervenir une matrice mal conditionnée. La conception des algorithmes est donc guidée par des problèmes venant de la physique. Comme les algorithmes proposés sont des heuristiques pour minimiser des critères non convexes et généralement multimodaux, il est important de pouvoir fournir des éléments théoriques permettant de prouver, même dans un cadre restrictif, que les algorithmes possèdent de bonnes performances. Ce type de résultat n'est pas une garantie infaillible du bon comportement de l'algorithme sur des données réelles dans des cadres plus généraux (je reviendrai sur ce point dans le chapitre de perspectives), mais ne pas disposer d'arguments théoriques favorables est souvent perçu comme le signe du mauvais comportement d'un algorithme. Le point de vue adopté consiste à comparer des conditions nécessaires et suffisantes garantissant le bon fonctionnement des algorithmes OMP et OLS avec des hypothèses de travail si possible peu restrictives afin de fournir quelques éléments tangibles de discrimination entre les deux. Les algorithmes étudiés étant des algorithmes de reconstruction parcimonieuse, leur comportement est évalué au sens de la capacité de l'algorithme à reconstruire un support inconnu⁵.

Le chapitre 4 dresse finalement quelques perspectives liées aux sujets abordés dans cette partie du manuscrit.

5. ce n'est pas la seule analyse possible, comme nous le verrons.

Liste des notations

NORMES ET OPÉRATIONS MATRICIELLES

$\langle \cdot, \cdot \rangle$	produit scalaire euclidien
$\ \cdot\ _p$	norme ℓ_p , $p > 0$ définie par $\ \mathbf{x}\ _p = (\sum_i x_i ^p)^{1/p}$
$\ \cdot\ _0$	pseudo ⁶ -norme ℓ_0 définie par $\ \mathbf{x}\ _0 \triangleq \text{Card}\{i, x_i \neq 0\}$
\mathbf{I}_p	matrice unité de dimension p
\mathbf{X}^T	transposée de la matrice \mathbf{X}
\mathbf{X}^\dagger	pseudo-inverse de la matrice \mathbf{X}
$\text{spark}(\mathbf{X})$	nombre minimal de colonnes de \mathbf{X} linéairement dépendantes
$\text{Ker}(\mathbf{X})$	noyau de \mathbf{X} . $\text{Ker}(\mathbf{X}) = \{\mathbf{z} : \mathbf{X}\mathbf{z} = \mathbf{0}\}$
$\text{span}(\mathbf{X})$	espace engendré par les colonnes de la matrice \mathbf{X}
$\text{span}(\mathbf{X})^\perp$	espace supplémentaire orthogonal de $\text{span}(\mathbf{X})$

APPROXIMATION PARCIMONIEUSE

$\mathbf{y} \in \mathbb{R}^m$	vecteur des observations
$\mathbf{A} \in \mathbb{R}^{m \times n}$	dictionnaire
$\mathbf{a}_i \in \mathbb{R}^m$	atomes (ou colonnes) du dictionnaire
$\mathbf{x} \in \mathbb{R}^n$	signal parcimonieux
$\mathcal{S}(\mathbf{x}) \subset \{1, \dots, n\}$	support de \mathbf{x}
$\mathbf{r} = \mathbf{y} - \mathbf{A}\mathbf{x}$	résidu de l'approximation
k	degré de parcimonie ($\ \mathbf{x}\ _0 \leq k$)
ε	qualité d'approximation ($\ \mathbf{y} - \mathbf{A}\mathbf{x}\ _2^2 \leq \varepsilon$)
μ	cohérence mutuelle du dictionnaire
(\mathcal{C}_k)	problème ℓ_2 - ℓ_0 contraint $\min_{\mathbf{x}} \ \mathbf{y} - \mathbf{A}\mathbf{x}\ _2^2$ s.c. $\ \mathbf{x}\ _0 \leq k$
$(\mathcal{C}_\varepsilon)$	problème ℓ_2 - ℓ_0 contraint $\min_{\mathbf{x}} \ \mathbf{x}\ _0$ s.c. $\ \mathbf{y} - \mathbf{A}\mathbf{x}\ _2^2 \leq \varepsilon$
(\mathcal{P}_λ)	problème ℓ_2 - ℓ_0 pénalisé $\min_{\mathbf{x}} \{\ \mathbf{y} - \mathbf{A}\mathbf{x}\ _2^2 + \lambda\ \mathbf{x}\ _0\}$

6. Par abus de langage, j'utiliserai parfois la dénomination « norme ℓ_0 » bien que le coût ℓ_0 ne soit pas une norme, car $\|\lambda\mathbf{x}\|_0 = \|\mathbf{x}\|_0 \neq \lambda\|\mathbf{x}\|_0$ pour $|\lambda| \neq 1$ et $\mathbf{x} \neq \mathbf{0}$.

ALGORITHMES GLOUTONS ORTHOGONAUX MONODIRECTIONNELS

$\mathcal{Q} \subset \{1, \dots, n\}$	le support, <i>i.e.</i> , les indices des atomes sélectionnés
$q = \text{Card}[\mathcal{Q}]$	cardinalité, <i>i.e.</i> , nombre d'itérations effectuées pour obtenir \mathcal{Q}
$\ell \notin \mathcal{Q}$	nouvel indice sélectionné à l'itération $q + 1$
$\mathbf{A}_{\mathcal{Q}} \in \mathbb{R}^{m \times q}$	matrice extraite de \mathbf{A} (sélection des colonnes indicées par \mathcal{Q})
$\mathbf{P}_{\mathcal{Q}} = \mathbf{A}_{\mathcal{Q}} \mathbf{A}_{\mathcal{Q}}^{\dagger}$	projection orthogonale sur $\text{span}(\mathbf{A}_{\mathcal{Q}})$
$\mathbf{P}_{\mathcal{Q}}^{\perp} = \mathbf{I}_m - \mathbf{A}_{\mathcal{Q}} \mathbf{A}_{\mathcal{Q}}^{\dagger}$	projection orthogonale sur $\text{span}(\mathbf{A}_{\mathcal{Q}})^{\perp}$
$\mathbf{P}_{\mathcal{Q}} \mathbf{y}$	approximation de \mathbf{y}
$\mathbf{A} \mathbf{x}_{\mathcal{Q}} = \mathbf{A}_{\mathcal{Q}} \mathbf{t}$	approximation de \mathbf{y} (autres expressions) où $\mathbf{t} = \mathbf{A}_{\mathcal{Q}}^{\dagger} \mathbf{y}$ et $\mathbf{x}_{\mathcal{Q}} = \{\mathbf{t}; \mathbf{0}\}$ à une permutation d'indices près
$\mathbf{r}_{\mathcal{Q}} = \mathbf{P}_{\mathcal{Q}}^{\perp} \mathbf{y}$	résidu d'approximation
$\mathcal{E}_{\mathcal{Q}} = \ \mathbf{r}_{\mathcal{Q}}\ _2^2$	erreur quadratique d'approximation

ALGORITHMES GLOUTONS ORTHOGONAUX BIDIRECTIONNELS

$\mathcal{Q} \bullet \ell$	remplacement élémentaire ($\mathcal{Q} \cup \{\ell\}$ ou $\mathcal{Q} \setminus \{\ell\}$)
$\ \mathbf{y} - \mathbf{A} \mathbf{x}\ _2^2$	erreur quadratique d'approximation (terme ℓ_2)
$\mathcal{J}(\mathbf{x}; \lambda) = \ \mathbf{y} - \mathbf{A} \mathbf{x}\ _2^2 + \lambda \ \mathbf{x}\ _0$	critère mixte ℓ_2 - ℓ_0
$\mathcal{J}_{\mathcal{Q}}(\lambda) = \mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}]$	critère mixte ℓ_2 - ℓ_0 fonction du support

ANALYSE D'ALGORITHMES GLOUTONS

$\mathbf{x}^* \in \mathbb{R}^n$	vecteur parcimonieux inconnu
\mathcal{Q}^*	support de \mathbf{x}^*
$k = \text{Card}[\mathcal{Q}^*] = \ \mathbf{x}^*\ _0$	taille du support de \mathbf{x}^*
$\mathcal{Q} \subset \mathcal{Q}^*$	support estimé après $q < k$ itérations
$\tilde{\mathbf{a}}_i = \mathbf{P}_{\mathcal{Q}}^{\perp} \mathbf{a}_i$	atome projeté sur $\text{span}(\mathbf{A}_{\mathcal{Q}})^{\perp}$
$\tilde{\mathbf{b}}_i = \tilde{\mathbf{a}}_i / \ \tilde{\mathbf{a}}_i\ _2$	atome projeté sur $\text{span}(\mathbf{A}_{\mathcal{Q}})^{\perp}$, puis normalisé (vaut $\mathbf{0}$ si $\tilde{\mathbf{a}}_i = \mathbf{0}$)
$\tilde{\mathbf{A}}$	matrice dont les colonnes sont les atomes projetés $\tilde{\mathbf{a}}_i$
$\tilde{\mathbf{B}}$	matrice dont les colonnes sont les atomes projetés $\tilde{\mathbf{b}}_i$
Oxx	désigne indifféremment OMP ou OLS
$\tilde{\mathbf{c}}_i$	désigne indifféremment $\tilde{\mathbf{a}}_i$ (OMP) ou $\tilde{\mathbf{b}}_i$ (OLS)
$\tilde{\mathbf{C}}$	désigne indifféremment $\tilde{\mathbf{A}}$ (OMP) ou $\tilde{\mathbf{B}}$ (OLS)
ERC($\mathbf{A}, \mathcal{Q}^*$)	<i>Exact Recovery Condition</i> pour Oxx à l'itération 0
ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$)	version étendue de l'ERC à l'itération q
BRC($\mathbf{A}, \mathcal{Q}^*$)	<i>Bad Recovery Condition</i>

Chapitre 1

Algorithmes gloutons bidirectionnels pour la minimisation de critères mixtes ℓ_2 - ℓ_0

1.1 Approximation parcimonieuse de signaux

Le problème posé est celui d'approximer un signal observé \mathbf{y} comme une combinaison linéaire d'un faible nombre d'atomes d'un dictionnaire. Formellement, l'approximation s'écrit $\mathbf{y} \approx \mathbf{A}\mathbf{x}$ où \mathbf{x} est un vecteur parcimonieux, *i.e.*, contenant un très faible nombre de composantes non nulles. Dans la littérature de l'approximation parcimonieuse, le dictionnaire $\mathbf{A} \in \mathbb{R}^{m \times n}$ est généralement *sur-dimensionné*, c'est-à-dire que le nombre d'atomes est supérieur voire très supérieur au nombre de données ($m \ll n$) afin de pouvoir approximer au mieux les données avec peu d'atomes. C'est une première différence avec certains problèmes inverses où le nombre de lignes de \mathbf{A} est du même ordre que le nombre de colonnes. Pour fixer les idées, prenons le problème d'école de la déconvolution impulsionnelle $\mathbf{y} = \mathbf{h} \star \mathbf{x} + \mathbf{b}$ où \mathbf{h} est la réponse impulsionnelle et \mathbf{b} est le terme de bruit. Le problème se réécrit $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ où \mathbf{A} est une matrice Toeplitz dont les colonnes sont les versions décalées de \mathbf{h} . Dans ce cas, \mathbf{A} est une matrice approximativement carrée (sur- ou sous-dimensionnée suivant les conditions aux limites). Lorsque \mathbf{h} est un filtre passe-bas, le problème est connu pour être mal conditionné et une façon de régulariser le problème est d'imposer une contrainte de parcimonie sur le signal \mathbf{x} [CHA 01]. Le point de départ de la réflexion menée avec Jérôme Idier est le constat que des algorithmes efficaces pour restaurer des signaux parcimonieux existent depuis les années 80, et ces algorithmes exploitent justement des informations *a priori* sur la parcimonie du signal, à savoir les modèles *Bernoulli-gaussien*. Jérôme Idier m'a donc proposé d'adapter ce modèle au contexte actuel, et d'adapter l'algorithme *Single Most Likely Replacement* (SMLR) qui en découle pour réaliser l'estimation parcimonieuse au sens du maximum *a posteriori* (MAP) [KOR 82, MEN 83, GOU 90, CHA 96].

1.2 Problèmes inverses mal conditionnés

Le caractère mal posé du problème inverse $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ est lié (entre autre) au fait que \mathbf{A} est potentiellement mal conditionnée, car le mauvais conditionnement de \mathbf{A} implique une instabilité de la solution \mathbf{x} par rapport à de très faibles variations de \mathbf{y} . C'est une différence majeure avec le contexte de l'échantillonnage compressé (*compressive sensing*) où la matrice \mathbf{A} , bien que sur-dimensionnée, est relativement bien conditionnée. C'est pour cette raison que

certains algorithmes récents dédiés à l'échantillonnage compressé sont soit inadaptés aux problèmes inverses mal posés, ou ont des performances qui se dégradent sensiblement dans ce contexte [WOH 03, AUS 10, DOS 10].

Plus précisément, le conditionnement matriciel $\kappa(\mathbf{A})$ est défini comme le rapport des valeurs singulières extrêmes de la matrice \mathbf{A} . Les facteurs critiques qui sont traditionnellement mis en avant dans le domaine de l'approximation parcimonieuse ne sont pas le conditionnement de \mathbf{A} , mais celui des matrices extraites de \mathbf{A} obtenues en sélectionnant k colonnes. Pour $k = 2$, ce conditionnement est lié à la cohérence mutuelle, définie comme le produit scalaire maximum entre les colonnes de \mathbf{A} , supposées normalisées :

$$\mu \triangleq \max_{i \neq j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle|. \quad (1.1)$$

Pour $k \geq 2$, on définit la constante d'isométrie (RIC, pour *Restricted Isometry Constant*) d'ordre k comme la plus petite valeur δ_k telle que :

$$\forall \mathbf{x}, \|\mathbf{x}\|_0 \leq k \Rightarrow (1 - \delta_k) \|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1 + \delta_k) \|\mathbf{x}\|_2^2. \quad (1.2)$$

La constante RIC contrôle donc le conditionnement de toutes les sous-matrices $\mathbf{A}_{\mathcal{Q}}$ formées en sélectionnant k colonnes du dictionnaire ($\text{Card}[\mathcal{Q}] = k$) [DAV 09] :

$$\kappa(\mathbf{A}_{\mathcal{Q}}) \leq \sqrt{\frac{1 + \delta_k}{1 - \delta_k}}. \quad (1.3)$$

De plus, la valeur de δ_k est bornée supérieurement en fonction de la cohérence mutuelle : $\delta_k \leq (k - 1)\mu$ [TRO 06, Prop. 21]. En résumé, lorsque la cohérence mutuelle du dictionnaire est suffisamment faible ou que son facteur RIC d'ordre k est inférieur à 1, l'inversion du système $\mathbf{A}_{\mathcal{Q}}\mathbf{t} = \mathbf{y}$ (correspondant à une estimation parcimonieuse du vecteur $\mathbf{x} = \{\mathbf{t}, \mathbf{0}\}$ pour un support \mathcal{Q} de taille k) est numériquement bien conditionnée : la solution obtenue avec des données modifiées $\mathbf{y} + \delta\mathbf{y}$ est proche de celle obtenue avec \mathbf{y} sous réserve que $\delta\mathbf{y}$ est suffisamment faible.

Ce raisonnement suppose implicitement que l'algorithme d'approximation sélectionne le même support \mathcal{Q} avec \mathbf{y} et $\mathbf{y} + \delta\mathbf{y}$ en entrée. De façon générale, la stabilité de l'estimateur, noté $\hat{\mathbf{x}}(\mathbf{y})$, est caractérisée par $\|\hat{\mathbf{x}}(\mathbf{y} + \delta\mathbf{y}) - \hat{\mathbf{x}}(\mathbf{y})\|_2 \leq \kappa \|\delta\mathbf{y}\|_2$ où κ est une constante liée au conditionnement de \mathbf{A} . L'estimation peut donc être potentiellement stable sans que les supports de $\hat{\mathbf{x}}(\mathbf{y} + \delta\mathbf{y})$ et $\hat{\mathbf{x}}(\mathbf{y})$ coïncident. Les études théoriques de stabilité d'algorithmes cherchent généralement à garantir que les supports \mathcal{Q} sont identiques, au moins pour une erreur $\|\delta\mathbf{y}\|_2$ faible (concept de stabilité *locale* [DON 06]). Dans le cas contraire, le facteur qui contrôle la stabilité n'est plus δ_k (où $k = \|\hat{\mathbf{x}}(\mathbf{y})\|_0$) mais $\delta_{k+k'}$ avec $k' = \|\hat{\mathbf{x}}(\mathbf{y} + \delta\mathbf{y})\|_0$ [WOH 03, BAB 10]. En l'absence d'information sur la valeur de $k + k'$, c'est le conditionnement du dictionnaire \mathbf{A} qui contrôle la stabilité.

Dans la suite de ce chapitre et dans les chapitres suivants, mon objectif est de traiter des problèmes où les conditions RIC (impliquant le très bon conditionnement des matrices extraites) ne sont pas nécessairement satisfaites même pour des valeurs très faibles de k . Avant de présenter les algorithmes proposés, je formule le problème d'approximation parcimonieuse comme un problème d'optimisation « mixte » ℓ_2 - ℓ_0 .

1.3 Problème des moindres carrés sous contrainte ℓ_0

L'approximation parcimonieuse est traditionnellement formulée comme le problème d'optimisation sous contrainte :

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.c.} \quad \|\mathbf{x}\|_0 \leq k \quad (\mathcal{C}_k)$$

ou encore, comme

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.c.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \varepsilon. \quad (\mathcal{C}_\varepsilon)$$

Dans la formulation (\mathcal{C}_k) , on impose que le vecteur \mathbf{x} ait au plus k composantes non nulles. Autrement dit, l'approximation $\mathbf{A}\mathbf{x}$ est une combinaison linéaire de k colonnes du dictionnaire. Dans la formulation $(\mathcal{C}_\varepsilon)$, le nombre de colonnes n'est pas connu mais on impose une borne ε sur l'erreur d'approximation, généralement liée à la variance du bruit lorsque celle-ci est connue. Dans le contexte de la régression statistique, ces problèmes sont connus sous le nom de sélection de variables (ou *subset selection*) [MIL 02].

Bien que les coordonnées x_i soient à valeurs réelles, ces deux problèmes sont des problèmes d'optimisation *combinatoire* car \mathbf{x} est parcimonieux. Notons \mathcal{Q} le support de \mathbf{x} . Pour le premier problème, il s'agit de trouver le « meilleur » k -uplet de colonnes de \mathbf{A} , c'est-à-dire le support \mathcal{Q} de taille inférieure ou égale à k conduisant à l'erreur quadratique

$$\mathcal{E}_{\mathcal{Q}} \triangleq \|\mathbf{r}_{\mathcal{Q}}\|_2^2 = \|\mathbf{P}_{\mathcal{Q}}^\perp \mathbf{y}\|_2^2 \quad (1.4)$$

minimale, où $\mathbf{r}_{\mathcal{Q}} = \mathbf{P}_{\mathcal{Q}}^\perp \mathbf{y} = \mathbf{y} - \mathbf{P}_{\mathcal{Q}} \mathbf{y}$ est le résidu de l'approximation, défini comme la projection orthogonale de \mathbf{y} sur l'espace supplémentaire orthogonal de $\text{span}(\mathbf{A}_{\mathcal{Q}})$. La solution du problème (\mathcal{C}_k) s'écrit finalement $\mathbf{x} = \{\mathbf{A}_{\mathcal{Q}}^\dagger \mathbf{y}, \mathbf{0}\}$ à une permutation d'indices près.

Ces problèmes combinatoires sont connus pour être NP-complets [NAT 95] : il n'existe pas d'algorithme optimal permettant de résoudre ces problèmes en un temps polynomial, sauf dans des situations spécifiques (par exemple, la situation triviale où les colonnes du dictionnaire sont orthogonales). En revanche, un grand nombre d'algorithmes sous-optimaux ont été développés [ELA 10]. Parmi les algorithmes sous-optimaux de la littérature, je distingue plusieurs grandes familles (sans vouloir être exhaustif) :

1. Les méthodes qui reposent sur l'approximation continue de la pseudo-norme ℓ_0 , *e.g.*, par la norme ℓ_p ($p \leq 1$) [RAO 03, MOU 10]. Notons que pour obtenir une solution parcimonieuse en résolvant un problème approché, la fonction φ utilisée pour approximer la norme ℓ_0 (approximation de $\|\mathbf{x}\|_0$ par $\sum_i \varphi(x_i)$) doit être non différentiable en 0 [FAN 01]. Parmi les choix possibles, la fonction valeur absolue (méthode dite de *relaxation convexe*) [TRO 06, ZIB 10] est extrêmement populaire : le problème ℓ_2 - ℓ_0 est remplacé par un problème d'optimisation convexe ℓ_2 - ℓ_1 pour lequel un grand nombre d'algorithmes optimaux efficaces sont disponibles [ELA 10]. Plus récemment de nombreux auteurs ont privilégié le choix de fonctions φ non convexes [RAO 03, CAN 08a, MOH 09, MOU 10, WIP 10, GHO 11, RAM 12] et proposé des algorithmes d'optimisation sous-optimaux efficaces.
2. Les algorithmes d'optimisation discrète qui traitent directement les problèmes mixtes ℓ_2 - ℓ_0 originaux (\mathcal{C}_k) et $(\mathcal{C}_\varepsilon)$. Bien que ces problèmes soient réputés plus difficiles à résoudre que leurs versions relaxées (en utilisant une fonction φ convexe ou non) à cause d'un plus grand nombre de minima locaux [CAN 08a, WIP 10], la résolution de problèmes d'optimisation plus simples, même de façon optimale, ne garantit pas que leur solution soit proche de celle des problèmes $(\ell_2 - \ell_0)$. C'est pour cela que nous nous sommes intéressés aux heuristiques pour traiter directement les problèmes mixtes $(\ell_2 - \ell_0)$.

Les algorithmes présentés dans la suite de ce chapitre font donc essentiellement partie de la deuxième catégorie. Des comparaisons seront effectuées avec des algorithmes de la première catégorie à la fois en termes de qualité d'approximation (erreur $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ obtenue à cardinalité $\|\mathbf{x}\|_0$ fixée) et de signal reconstruit (capacité de l'algorithme de retrouver les impulsions $x_i \neq 0$ aux bons endroits).

A l'intérieur de la deuxième catégorie, je distingue deux sous-catégories : les méthodes de type seuillage itératif (CoSaMP, *Subspace Pursuit*, *Gradient Pursuit*, *Hard Thresholding Pursuit*, etc.) [NEE 09, DAI 09, BLU 08, BLU 10, FOU 11] pour lesquelles il n'y a généralement pas de relation d'emboîtement entre les supports obtenus pour deux itérations consécutives, et les algorithmes gloutons qui mettent à jour le support en ajoutant un nouvel indice (voire en retirant un indice, ce qui correspond à dé-sélectionner un atome). Parmi les algorithmes gloutons monodirectionnels (*forward greedy algorithms*) qui ajoutent des atomes un à un, citons par ordre croissant de complexité les algorithmes *Matching Pursuit* (MP) [MAL 93], *Orthogonal Matching Pursuit* (OMP) [PAT 93, DAV 94] et *Orthogonal Least Squares* (OLS) [CHE 89]. Il existe plusieurs variantes d'OMP permettant d'ajouter plusieurs atomes à la fois à chaque itération (tous les atomes les plus corrélés avec le résidu) [HUA 11, DON 12, WAN 12a] ou d'ajouter un atome suivant une règle modifiée [TEM 00, TRO 04, LIU 12, FOU 13a] dans le but de réduire le temps de calcul d'OMP sans perdre en efficacité.

Dans la suite de ce chapitre, je m'intéresse aux algorithmes gloutons bidirectionnels⁷ qui autorisent l'ajout et le retrait d'un atome du support à une itération donnée. Il existe une littérature fournie sur les algorithmes *bidirectionnels* (aussi connus sous le nom de *stepwise regression algorithms*) dans le domaine de l'inférence statistique, proposés pour traiter des problèmes faisant intervenir des dictionnaires *sous-dimensionnés*. Ces algorithmes sont anciens, le premier d'entre eux a été proposé par Efroymsou en 1960 [EFR 60]. D'autres algorithmes de ce type ont été proposés dans les années 1980 [BER 80, BRO 86] avec une structure proche de l'algorithme SMLR, proposé également dans les années 1980 et de façon indépendante pour restaurer des signaux Bernoulli-gaussiens [KOR 82, MEN 83, GOU 90, CHA 96, CHA 01]. Pour une revue de littérature sur les algorithmes bidirectionnels en régression statistique, je renvoie le lecteur au livre [MIL 02, Chap. 3]. Le principe de ces algorithmes est qu'en autorisant à la fois la sélection et la dé-sélection d'atomes, il est possible d'améliorer la performance d'algorithmes monodirectionnels puisque la sélection d'un mauvais atome à une certaine itération de l'algorithme peut être compensée par sa dé-sélection à une itération ultérieure. Au contraire, une mauvaise sélection est irréversible avec un algorithme glouton monodirectionnel.

1.4 Algorithmes de poursuite gloutons orthogonaux

Comme les algorithmes bidirectionnels proposés ont une structure très proche d'OMP et OLS, je commence par rappeler brièvement comment fonctionnent ces algorithmes. Je prêterai peu d'attention aux algorithmes non orthogonaux plus simples qu'OMP, comme MP, car leur performance est limitée pour les problèmes considérés. Je rappelle néanmoins le principe de MP dans le tableau 1.1 à titre de comparaison avec OMP.

Avant de détailler les algorithmes OMP et OLS, notons qu'ils sont connus sous des noms divers dans la littérature : OMP est aussi nommé *Orthogonal Greedy Algorithm* (OGA) [DEV 96, DON 06] et OLS est connu sous les noms de *forward selection* en régression statistique [MIL 02], *the greedy algorithm* [NAT 95], *Order Recursive Matching Pursuit* (ORMP) [COT 99], *Optimized*

7. Au sens premier, le qualificatif « glouton » désigne des algorithmes qui sélectionnent les atomes un à un [TRO 10], donc des algorithmes monodirectionnels. Qualifier les algorithmes bidirectionnels de gloutons est un léger abus de langage [BRU 09] souvent pratiqué [COU 00, HAU 07, ZHA 11a]. J'emploie ici le qualificatif « glouton » car la structure des algorithmes bidirectionnels considérés est très proche de celle d'algorithmes gloutons monodirectionnels : il y a une relation d'emboîtement, dans un sens ou dans l'autre, entre les supports obtenus à des itérations consécutives. Glouton doit donc être compris par opposition aux algorithmes qui mettent à jour *simultanément* un certain nombre d'atomes par itération (*iterative thresholding, subspace pursuit, etc.*) et pour lesquels deux supports consécutifs ne sont pas emboîtés.

Entrées : \mathbf{A} , \mathbf{y} , ε et/ou K
Faire $\mathcal{Q} = \emptyset$, $\mathbf{r}_{\mathcal{Q}} = \mathbf{y}$ et $\mathcal{E}_{\mathcal{Q}} = \ \mathbf{y}\ _2^2$.
Tant que ($\text{Card}[\mathcal{Q}] < K$) et ($\mathcal{E}_{\mathcal{Q}} > \varepsilon$),
[1. Sélection du nouvel atome ℓ]
Faire $\mathcal{Q} = \mathcal{Q} \cup \{\ell\}$ où $\ell \in \arg \max_{i \in \{1, \dots, n\}} \langle \mathbf{r}_{\mathcal{Q}}, \mathbf{a}_i \rangle $.
[2. Mise à jour des amplitudes]
Mettre à jour $\mathbf{r}_{\mathcal{Q}} = \mathbf{r}_{\mathcal{Q}} - \langle \mathbf{r}_{\mathcal{Q}}, \mathbf{a}_{\ell} \rangle \mathbf{a}_{\ell}$.
Fin Tant que
Sortie : support \mathcal{Q}

TABLE 1.1 – Algorithme *Matching Pursuit* (MP) [MAL 93]. Les paramètres d’arrêt ε et/ou K portent respectivement sur la norme ℓ_2 du résidu de l’approximation et sur le nombre d’itérations.

Orthogonal Matching Pursuit (OMP) [REB 02] et *Pure Orthogonal Matching Pursuit* [FOU 11] dans la littérature de l’approximation parcimonieuse. De plus, OMP et OLS sont parfois confondus à cause de leur structure proche. Je détaille maintenant les deux algorithmes en mettant en évidence ce qui les différencie. Le lecteur intéressé par plus de précisions est redirigé vers le rapport technique très éclairant de Blumensath et Davies [BLU 07].

OMP et OLS sont présentés dans les tableaux 1.2 et 1.3. Leur structure est similaire à MP car à chaque itération, deux étapes sont réalisées : la sélection du nouvel atome \mathbf{a}_{ℓ} puis la mise à jour de l’approximation parcimonieuse de \mathbf{y} sous la forme $\mathbf{A}_{\mathcal{Q} \cup \{\ell\}} \mathbf{t}$. Cette mise à jour est commune à OMP et OLS : les amplitudes $\mathbf{t} = \mathbf{A}_{\mathcal{Q} \cup \{\ell\}}^{\dagger} \mathbf{y}$ sont calculées en projetant orthogonalement \mathbf{y} sur le sous-espace $\text{span}(\mathbf{A}_{\mathcal{Q} \cup \{\ell\}})$, et l’erreur d’approximation s’écrit $\mathcal{E}_{\mathcal{Q} \cup \{\ell\}} = \|\mathbf{r}_{\mathcal{Q} \cup \{\ell\}}\|_2^2$ avec $\mathbf{r}_{\mathcal{Q} \cup \{\ell\}} = \mathbf{y} - \mathbf{A}_{\mathcal{Q} \cup \{\ell\}} \mathbf{t}$. En revanche, la règle de sélection du nouvel atome diffère de façon significative. OMP adopte la même règle que MP : l’atome sélectionné \mathbf{a}_{ℓ} est celui dont le produit scalaire avec le résidu est maximal. En revanche, OLS cherche à minimiser l’erreur d’approximation de façon plus précise en résolvant les $n - q$ problèmes de moindres carrés

$$\mathcal{E}_{\mathcal{Q} \cup \{i\}} = \arg \min_{\mathbf{t}} \|\mathbf{y} - \mathbf{A}_{\mathcal{Q} \cup \{i\}} \mathbf{t}\|_2^2 \quad (1.5)$$

pour tout $i \notin \mathcal{Q}$. L’atome sélectionné \mathbf{a}_{ℓ} est finalement celui qui conduit à la plus faible erreur d’approximation $\mathcal{E}_{\mathcal{Q} \cup \{\ell\}}$.

En résumé, $n - q$ inversions du type $\mathbf{t} = \mathbf{A}_{\mathcal{Q} \cup \{i\}}^{\dagger} \mathbf{y}$ sont réalisées à l’itération q d’OLS contre une seule inversion pour OMP, liée à la mise à jour des amplitudes⁸. Cette différence de complexité combinée au constat que pour certains problèmes bien conditionnés, les résultats fournis par OMP et OLS sont de qualité comparable et bien meilleure que MP, expliquent le très fort intérêt de la communauté pour OMP et un relatif désintérêt pour OLS.

8. Les opérations (1.5) coûtent relativement plus cher pour OLS, mais pas $n - q$ fois plus cher. Les implémentations classiques s’appuient sur la procédure d’orthonormalisation de Gram Schmidt pour mettre à jour les erreurs quadratiques $\mathcal{E}_{\mathcal{Q}}$ dans le cas d’OLS [CHE 89, CHE 95] et le résidu $\mathbf{r}_{\mathcal{Q}}$ pour OMP [DAV 94]. Les erreurs d’approximation (1.5) sont évaluées récursivement à partir de la factorisation de Gram-Schmidt courante, et le calcul des amplitudes \mathbf{t} n’est pas nécessaire.

Entrées : \mathbf{A} , \mathbf{y} , ε et/ou K
Faire $\mathcal{Q} = \emptyset$, $\mathbf{r}_{\mathcal{Q}} = \mathbf{y}$ et $\mathcal{E}_{\mathcal{Q}} = \ \mathbf{y}\ _2^2$.
Tant que ($\text{Card}[\mathcal{Q}] < K$) et ($\mathcal{E}_{\mathcal{Q}} > \varepsilon$),
[1. Sélection du nouvel atome ℓ]
Faire $\mathcal{Q} = \mathcal{Q} \cup \{\ell\}$ où $\ell \in \arg \max_{i \notin \mathcal{Q}} \langle \mathbf{r}_{\mathcal{Q}}, \mathbf{a}_i \rangle $.
[2. Mise à jour des amplitudes]
Mettre à jour $\mathbf{r}_{\mathcal{Q}} = \mathbf{P}_{\mathcal{Q}}^{\perp} \mathbf{y}$ et $\mathcal{E}_{\mathcal{Q}} = \ \mathbf{r}_{\mathcal{Q}}\ _2^2$.
Fin Tant que
Sortie : support \mathcal{Q}

TABLE 1.2 – Algorithme *Orthogonal Matching Pursuit* (OMP) [PAT 93, DAV 94]. Les paramètres d'arrêt ε et/ou K portent respectivement sur la norme ℓ_2 du résidu de l'approximation et sur le nombre d'itérations.

Entrées : \mathbf{A} , \mathbf{y} , ε et/ou K
Faire $\mathcal{Q} = \emptyset$ et $\mathcal{E}_{\mathcal{Q}} = \ \mathbf{y}\ _2^2$.
Tant que ($\text{Card}[\mathcal{Q}] < K$) et ($\mathcal{E}_{\mathcal{Q}} > \varepsilon$),
[1. Sélection du nouvel atome ℓ]
Pour tout $i \notin \mathcal{Q}$, calculer l'erreur quadratique $\mathcal{E}_{\mathcal{Q} \cup \{i\}}$.
Faire $\mathcal{Q} = \mathcal{Q} \cup \{\ell\}$ où $\ell \in \arg \min_{i \notin \mathcal{Q}} \mathcal{E}_{\mathcal{Q} \cup \{i\}}$.
[2. Mise à jour des amplitudes]
Mettre à jour $\mathcal{E}_{\mathcal{Q}}$.
Fin Tant que
Sortie : support \mathcal{Q}

TABLE 1.3 – Algorithme *Orthogonal Least Squares* (OLS) [CHE 89, NAT 95]. Les paramètres d'arrêt ε et/ou K portent respectivement sur la norme ℓ_2 du résidu de l'approximation et sur le nombre d'itérations.

1.5 Formulation bayésienne basée sur le modèle Bernoulli-gaussien

Un signal parcimonieux peut être modélisé comme un vecteur aléatoire Bernoulli-gaussien $\mathbf{x} = (\mathbf{q}, \mathbf{r}) \in \{0, 1\}^n \times \mathbb{R}^n$. Chaque échantillon de \mathbf{x} s'écrit $x_i = q_i r_i$ où :

- la variable aléatoire de Bernoulli $q_i \sim \mathcal{B}(\rho)$ code la présence ($q_i = 1$) ou l'absence ($q_i = 0$) de signal en position i , et ρ désigne la probabilité de présence. En d'autres termes, le vecteur $\mathbf{q} = [q_1, \dots, q_n]^T \in \{0, 1\}^n$ représente le support \mathcal{Q} de \mathbf{x} ;
- les amplitudes r_i sont des variables gaussiennes : $r_i | q_i \sim \mathcal{N}(0, q_i \sigma_x^2)$.

Dans le cadre de problèmes inverses $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$, on peut envisager d'estimer \mathbf{x} au sens du maximum *a posteriori* (MAP) ou de la moyenne *a posteriori*, ce qui conduit respectivement à utiliser des algorithmes d'optimisation déterministes [CHA 96] ou de Monte Carlo par chaînes de Markov (MCMC) [CHE 96]. Nous choisissons l'estimation au sens du MAP *joint*, c'est-à-dire qu'on cherche à maximiser la vraisemblance conjointe $\ell(\mathbf{q}, \mathbf{r} | \mathbf{y})$ [MEN 83], une alternative au MAP joint étant le MAP marginal lié à la vraisemblance marginale $\ell(\mathbf{q} | \mathbf{y})$ [CHA 96].

Pour un bruit $\mathbf{b} \sim \mathcal{N}(\mathbf{0}, \sigma_b^2 \mathbf{I}_m)$ blanc, gaussien et indépendant de \mathbf{x} , la règle de Bayes [MEN 83] permet d'exprimer la vraisemblance *a posteriori* (plus précisément, l'opposé de la log-vraisemblance) :

$$-2\sigma_b^2 \log[\ell(\mathbf{q}, \mathbf{r} | \mathbf{y})] = \|\mathbf{y} - \mathbf{A}\mathbf{\Delta}_q \mathbf{r}\|_2^2 + \frac{\sigma_b^2}{\sigma_x^2} \|\mathbf{r}\|_2^2 + \lambda \|\mathbf{q}\|_0 + c \quad (1.6)$$

où $\lambda \triangleq 2\sigma_b^2 \log(1/\rho - 1)$, $\mathbf{\Delta}_q$ est la matrice diagonale de taille n dont les éléments diagonaux sont les q_i (\mathbf{x} s'écrit $\mathbf{x} = \mathbf{\Delta}_q \mathbf{r}$), et c est une constante. On remarque que la log-vraisemblance (1.6) est constituée du terme $\|\mathbf{y} - \mathbf{A}\mathbf{\Delta}_q \mathbf{r}\|_2^2$ mesurant l'erreur quadratique d'approximation $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$, et de deux termes qui pénalisent respectivement les fortes amplitudes ($\|\mathbf{r}\|_2^2 = \|\mathbf{x}\|_2^2$) et les signaux peu parcimonieux ($\|\mathbf{q}\|_0 = \|\mathbf{x}\|_0$). En effectuant le changement de variable $\mathbf{x} = (\mathbf{q}, \mathbf{r})$, la minimisation de (1.6) par rapport à $(\mathbf{q}, \mathbf{r}) \in \{0, 1\}^n \times \mathbb{R}^n$ équivaut à celle de

$$\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \frac{\sigma_b^2}{\sigma_x^2} \|\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0 \quad (1.7)$$

par rapport à $\mathbf{x} \in \mathbb{R}^n$.

Par définition, un signal est parcimonieux s'il comporte un grand nombre de coordonnées x_i égales à 0. Comme cette définition n'impose aucune contrainte sur la valeur des amplitudes non nulles, nous avons choisi d'utiliser un modèle Bernoulli-gaussien *limite* en faisant tendre la variance σ_x^2 vers $+\infty$. Dans ce cas, on montre [SOU 11] que le maximum *a posteriori* est une solution du problème :

$$\min_{\mathbf{x}} \{\mathcal{J}(\mathbf{x}; \lambda) \triangleq \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0\}. \quad (\mathcal{P}_\lambda)$$

L'hyperparamètre λ contrôle le niveau de parcimonie de la solution. Une grande valeur de λ implique une approximation grossière et très parcimonieuse. Inversement, une approximation plus fine mais moins parcimonieuse est obtenue en choisissant λ faible. La minimisation du critère $\mathcal{J}(\mathbf{x}; \lambda)$ défini sur \mathbb{R}^n se ramène à un problème d'optimisation combinatoire :

$$\min_{\mathcal{Q} \subset \{1, \dots, n\}} \{\mathcal{J}_{\mathcal{Q}}(\lambda) \triangleq \mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}]\} \quad (1.8)$$

qui fait apparaître l'erreur quadratique minimale

$$\mathcal{E}_{\mathcal{Q}} = \min_{\mathcal{S}(\mathbf{x}) \subset \mathcal{Q}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad (1.9)$$

pour tout \mathbf{x} dont le support est inclus dans \mathcal{Q} . A partir d'une solution \mathcal{Q} du problème (1.8), les amplitudes \mathbf{x} se déduisent en effectuant la projection orthogonale des données \mathbf{y} sur le sous-espace $\text{span}(\mathbf{A}_{\mathcal{Q}})$ ⁽⁹⁾. Les algorithmes dédiés au problème (1.8) sont donc de nature *exploratoire* : au fur et à mesure des itérations, une succession de supports \mathcal{Q} est testée selon une stratégie de recherche à définir, et $\mathcal{J}_{\mathcal{Q}}(\lambda)$ est évalué pour chaque support. La performance d'un algorithme relève donc du compromis entre le nombre de supports explorés et la capacité de la stratégie de recherche à explorer des supports « intéressants », conduisant à une faible valeur de $\mathcal{J}_{\mathcal{Q}}(\lambda)$.

1.6 Minimisation mixte ℓ_2 - ℓ_0 pour un niveau de parcimonie donné

Le premier algorithme bidirectionnel que nous avons proposé pour minimiser (\mathcal{P}_{λ}) est l'algorithme *Single Best Replacement* (SBR) [SOU 11]. C'est une version revisitée de l'algorithme *Single Most Likely Replacement* (SMLR), dédié à la minimisation de (1.6)-(1.7) et conçu pour restaurer des signaux Bernoulli-gaussien [KOR 82, MEN 83, GOU 90, CHA 96].

1.6.1 Méthode de descente pour l'optimisation combinatoire

SBR est une heuristique (dans le sens où la convergence vers le minimum global n'est pas garantie) permettant d'optimiser le critère pénalisé (\mathcal{P}_{λ}) , ou de façon équivalente, le critère (1.8). C'est un algorithme de descente déterministe dont la structure est simple. Il est basé sur la notion de voisinage entre supports : deux supports sont voisins si l'un est inclus dans l'autre et le plus grand support comporte un élément de plus. Le voisinage d'un support \mathcal{Q} prend la forme :

$$\mathcal{Q} \bullet i \triangleq \begin{cases} \mathcal{Q} \cup \{i\} & \text{si } i \notin \mathcal{Q}, \\ \mathcal{Q} \setminus \{i\} & \text{sinon.} \end{cases} \quad (1.10)$$

et un *remplacement élémentaire* est défini comme l'ajout ($\bullet = \cup$) d'un nouvel indice dans le support ou le retrait ($\bullet = \setminus$) d'un indice du support. Des notions de voisinages étendues sont envisageables comme la substitution d'un indice par un autre indice [CHI 84, MIL 02] ou le remplacement (ajout ou retrait simultané) d'un bloc d'indices [ALL 07]. Les algorithmes proposés dans ce chapitre sont génériques : ils s'étendent naturellement à des voisinages plus complexes que le remplacement élémentaire (1.10).

Contrairement aux problèmes contraints (\mathcal{C}_k) et $(\mathcal{C}_{\varepsilon})$, la minimisation du critère pénalisé (\mathcal{P}_{λ}) permet d'envisager non seulement l'ajout mais aussi le retrait de variables actives dans un algorithme de descente déterministe. Pour fixer les idées, considérons le problème (\mathcal{C}_k) de minimisation quadratique sous contrainte ℓ_0 . Supposons qu'un algorithme glouton ait sélectionné un support \mathcal{Q} . Il est clair qu'un ajout d'atome fait décroître l'erreur quadratique $\mathcal{E}_{\mathcal{Q}}$ et qu'inversement, une dé-sélection d'atome augmente l'erreur d'approximation. Un algorithme de descente conduit donc naturellement à définir une stratégie monodirectionnelle ¹⁰. Ainsi, les algorithmes OMP et OLS peuvent être interprétés comme des algorithmes de descente dédiés au problème (\mathcal{C}_k) .

9. En toute rigueur, le deuxième terme du critère (1.8) devrait s'écrire $\lambda \|\mathbf{A}_{\mathcal{Q}}^{\dagger} \mathbf{y}\|_0$. Choisir la forme plus simple $\lambda \text{Card}[\mathcal{Q}]$, c'est supposer que les amplitudes $\mathbf{A}_{\mathcal{Q}}^{\dagger} \mathbf{y}$ sont toutes non nulles. Cette hypothèse est raisonnable car pour des données bruitées, la probabilité que l'une des amplitudes soit non nulle vaut 0 [SOU 11, Th. 1]. Ce choix permettra aussi de réduire le coût numérique des algorithmes proposés pour minimiser (1.8) en ne calculant jamais explicitement $\mathbf{A}_{\mathcal{Q}}^{\dagger} \mathbf{y}$.

10. De nombreuses heuristiques bidirectionnelles ont été conçues sur la base du critère (\mathcal{C}_k) [MIL 02, Chap. 3] mais ce ne sont pas des algorithmes de descente.

Entrées : \mathbf{A} , \mathbf{y} , λ et le support $\mathcal{Q}_{\text{init}}$ ($\text{Card}[\mathcal{Q}_{\text{init}}] \leq \min(m, n)$)	
Initialisation : Faire $\text{iter} = 1$ et $\mathcal{Q} = \mathcal{Q}_{\text{init}}$.	
Etape 1 :	<p>[1. Sélection du nouvel atome ℓ]</p> <p>Pour $i \in \{1, \dots, n\}$, calculer $\mathcal{J}_{\mathcal{Q} \bullet i}(\lambda)$.</p> <p>Calculer ℓ par (1.11).</p> <p>[2. Test d'arrêt]</p> <p>Si $\mathcal{J}_{\mathcal{Q} \bullet \ell}(\lambda) < \mathcal{J}_{\mathcal{Q}}(\lambda)$,</p> <p style="padding-left: 2em;">Faire $\mathcal{Q} = \mathcal{Q} \bullet \ell$.</p> <p>Sinon,</p> <p style="padding-left: 2em;">Terminer SBR.</p> <p>Fin si.</p> <p>Faire $\text{iter} = \text{iter} + 1$ et retourner à l'étape 1.</p>
Sortie : support $\mathcal{Q} = \text{SBR}(\mathcal{Q}_{\text{init}}; \lambda)$	

TABLE 1.4 – Algorithme *Single Best Replacement* [SOU 11] : extension bidirectionnelle d'OLS. Par défaut, le support initial vaut $\mathcal{Q}_{\text{init}} = \emptyset$. Il n'y a pas de critère d'arrêt : l'algorithme s'arrête en un nombre fini d'itérations.

La situation est différente pour le critère pénalisé $\mathcal{J}(\mathbf{x}; \lambda)$: si une dé-sélection d'atome fait croître l'erreur quadratique, le terme ℓ_0 diminue d'un élément. Il est donc possible de concevoir des algorithmes de descente bidirectionnels pour minimiser le critère \mathcal{J} . L'algorithme SBR présenté ci-dessous est une extension bidirectionnelle d'OLS définie comme un algorithme de descente dédié à (\mathcal{P}_λ) . Un autre algorithme de descente, nommé *Bayesian OMP*, a été proposé en parallèle de notre travail en tant qu'extension bidirectionnelle d'OMP, toujours dans le but de minimiser (\mathcal{P}_λ) [HER 10]. D'autres extensions bidirectionnelles d'OMP [ZHA 11a, JAI 11] et d'OLS [HAU 07, MOG 08] ont été récemment proposées. Elles sont conçues comme des modifications d'algorithmes monodirectionnels en incluant une phase de dé-sélection d'atomes, mais elles ne s'interprètent pas comme des algorithmes de descente pour minimiser le critère pénalisé (\mathcal{P}_λ) .

1.6.2 Adaptation de l'algorithme SMLR : algorithme Single Best Replacement

L'algorithme SBR minimise le critère $\mathcal{J}(\mathbf{x}; \lambda)$ à λ fixé. C'est un algorithme itératif qui effectue un remplacement élémentaire (1.10) par itération, à savoir la sélection d'un nouvel atome de \mathbf{A} ou la dé-sélection d'un atome déjà présent. L'algorithme est résumé dans le tableau 1.4. A chaque itération, le remplacement élémentaire effectué est celui qui permet de faire décroître le plus possible la valeur du critère $\mathcal{J}(\mathbf{x}; \lambda)$:

$$\ell \in \arg \min_{i \in \{1, \dots, n\}} \mathcal{J}_{\mathcal{Q} \bullet i}(\lambda). \quad (1.11)$$

SBR effectue donc n inversions de systèmes linéaires pour calculer les erreurs quadratiques $\mathcal{E}_{\mathcal{Q} \bullet i}$, puis $\mathcal{J}_{\mathcal{Q} \bullet i}(\lambda) = \mathcal{E}_{\mathcal{Q} \bullet i} + \lambda \text{Card}[\mathcal{Q}]$. Ces n inversions se décomposent en $n - q$ inversions pour les

tests d'ajout d'atomes $\mathcal{Q} \cup \{i\}$ et q inversions pour les tests de retrait $\mathcal{Q} \setminus \{i\}$, avec $q = \text{Card}[\mathcal{Q}]$. En comparaison, OLS effectue $n - q$ inversions par itération (tests d'ajout seuls) et OMP une seule inversion lors de la mise à jour du résidu. Effectuer n ou $n - q$ inversions peut paraître important. Cependant, des implémentations efficaces sont possibles pour effectuer ces calculs récursivement et en parallèle [SOU 11].

SBR est une extension de SMLR au cas limite de la restauration d'un processus Bernoulli-gaussien à variance infinie, pour des problèmes où on impose seulement un *a priori* de parcimonie et pas d'*a priori* sur la valeur des amplitudes. La différence principale entre SMLR et SBR est que SMLR réalise l'inversion de matrices de la forme $\mathbf{A}_{\mathcal{Q}}^T \mathbf{A}_{\mathcal{Q}} + \alpha \mathbf{I}_q$ alors que SBR calcule l'inverse de $\mathbf{A}_{\mathcal{Q}}^T \mathbf{A}_{\mathcal{Q}}$. Le terme $\alpha \mathbf{I}_q$ qui apparaît pour SMLR est lié à la régularisation sur les valeurs des amplitudes (*cf.* formulations (1.6)-(1.7)). Cette régularisation permet d'éviter les instabilités lorsque $\mathbf{A}_{\mathcal{Q}}$ est très mal conditionnée au prix du réglage d'un hyperparamètre supplémentaire α . Les instabilités peuvent se produire plus fréquemment en utilisant SBR. De façon à limiter ces problèmes potentiels d'instabilité, nous avons proposé une implémentation qui repose sur la décomposition de Cholesky de la matrice de Gram $\mathbf{A}_{\mathcal{Q}}^T \mathbf{A}_{\mathcal{Q}} = \mathbf{L}_{\mathcal{Q}} \mathbf{L}_{\mathcal{Q}}^T$, la matrice triangulaire $\mathbf{L}_{\mathcal{Q}}$ étant mieux conditionnée que la matrice $\mathbf{A}_{\mathcal{Q}}$ [SOU 11, Sec. IV].

Notons que pour $\lambda = 0$, SBR coïncide avec OLS : le critère $\mathcal{J}_{\mathcal{Q}}(\lambda)$ se simplifie en $\mathcal{E}_{\mathcal{Q}}$, et seul un ajout d'atome permet de faire décroître l'erreur quadratique $\mathcal{E}_{\mathcal{Q}}$. Les tests de retraits ne sont donc pas nécessaires. Une différence notable entre les cas $\lambda > 0$ et $\lambda = 0$ est que dans le premier cas, il n'y a pas de critère d'arrêt ; SBR s'arrête après un nombre fini d'itérations qui dépend de λ , lorsque le critère $\mathcal{J}_{\mathcal{Q}}(\lambda)$ ne peut plus décroître quelque soit le remplacement élémentaire. Pour $\lambda = 0$, on impose un nombre maximal d'itérations ou un seuil sur l'erreur quadratique $\mathcal{E}_{\mathcal{Q}}$ conformément au tableau 1.3.

1.6.3 SBR en tant qu'algorithme de minimisation locale

Le critère pénalisé $\mathcal{J}(\mathbf{x}; \lambda)$ possède généralement un très grand nombre de minima locaux [NIK 13]. C'est un argument classique pour ne pas minimiser ce critère, mais choisir une version approchée de $\mathcal{J}(\mathbf{x}; \lambda)$ avec une pénalisation continue $\varphi(x_i)$ en remplacement de la norme ℓ_0 [CAN 08a, WIP 10] dans l'espoir que le problème continu possède moins de minima locaux que le problème original. Ce raisonnement présente plusieurs inconvénients : d'une part, il suppose implicitement que les minimiseurs (globaux et locaux) du critère approché sont proches de ceux de $\mathcal{J}(\mathbf{x}; \lambda)$. D'autre part, quand la gamme de valeurs non-nulles $x_i \neq 0$ est grande, il est délicat de choisir une bonne approximation $\varphi(x_i)$ de $|x_i|_0$ pour tout x_i . Le choix d'une fonction φ appropriée repose généralement sur l'introduction d'un degré de liberté supplémentaire dont le réglage n'est pas trivial [CAN 08a, MOH 09, NIK 10]. Typiquement, l'algorithme *Iterative Reweighted ℓ_1* (IR ℓ_1 , aussi connu sous le nom plus général de *Difference of Convex functions* ou *DC programming algorithm* [BAC 12, Chap. 7] [TAO 97, HOA 05, HOA 08, GAS 09]) s'interprète comme une méthode d'optimisation pour un critère quadratique pénalisé où la norme ℓ_0 est remplacée par la fonction paramétrique $\varphi(x_i; \eta) = \log(|x_i| + \eta)$ [CAN 08a, WIP 10]. Le paramètre η contrôle le degré de non-convexité de la fonction φ .

Nous avons pris le parti de minimiser directement le critère pénalisé ℓ_2 - ℓ_0 plutôt qu'une version approchée. Si SBR est un algorithme de descente basé sur des mouvements locaux (les remplacements élémentaires $\mathcal{Q} \bullet i$), ce n'est pas un algorithme de minimisation locale associé au critère $\mathcal{J}(\mathbf{x}; \lambda)$ mais un algorithme, bien que déterministe, capable d'explorer de nombreuses vallées associées à différents minimiseurs locaux de $\mathcal{J}(\mathbf{x}; \lambda)$. Dans [SOU 11, Prop. 1], nous avons montré que *chaque* itéré de SBR est presque sûrement un minimiseur local de $\mathcal{J}(\mathbf{x}; \lambda)$ où la localité est à comprendre au sens de la norme ℓ_2 . Ce résultat est lié au fait qu'un itéré de SBR

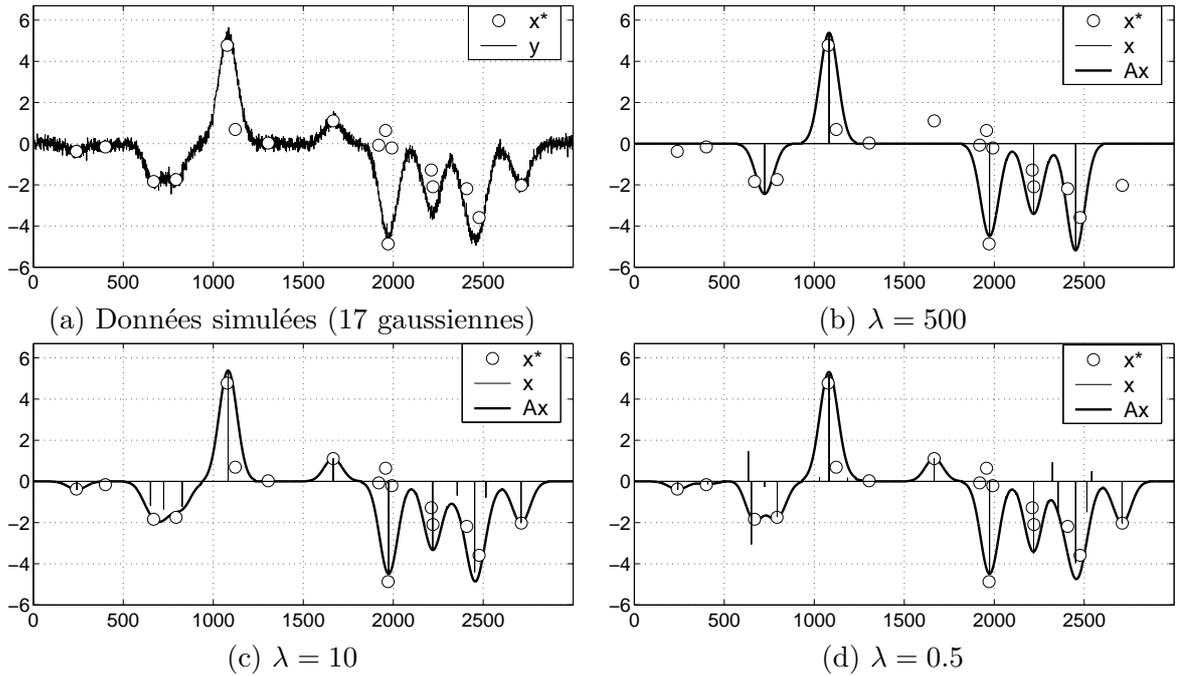


FIGURE 1.1 – Problème de déconvolution impulsionnelle de taille 3000×2700 avec un noyau de convolution \mathbf{h} gaussien (écart-type $\sigma = 50$) [SOU 11]. (a) Données simulées comportant 17 motifs gaussiens, le rapport signal-sur-bruit vaut 20 dB. Les positions exactes des impulsions \mathbf{x}^* portent le label \circ . (b,c,d) Signaux impulsionnels obtenus en sortie de SBR et approximations des données correspondantes avec différents réglages empiriques de λ . Les amplitudes estimées \mathbf{x} sont les signaux impulsionnels avec des barres verticales. Les supports obtenus sont de taille 5, 12 et 18. Les temps de calcul sont toujours inférieurs à 3 secondes (implémentation Matlab).

est obtenu à partir du support \mathcal{Q} à l'itération courante par projection orthogonale, en minimisant l'erreur quadratique $\|\mathbf{y} - \mathbf{A}_{\mathcal{Q}}\mathbf{t}\|_2^2$ par rapport à \mathbf{t} (avec $\mathbf{x} = \{\mathbf{t}, \mathbf{0}\}$ à une permutation d'indices près). En résumé, si SBR est un algorithme de minimisation locale (au sens de la relation de voisinage discrète $\mathcal{Q} \bullet i$) pour minimiser $\mathcal{J}_{\mathcal{Q}}(\lambda)$ par rapport à \mathcal{Q} , SBR est un algorithme (sous-optimal) de minimisation *globale* pour minimiser $\mathcal{J}(\mathbf{x}; \lambda)$ par rapport à \mathbf{x} .

1.6.4 Exemple de la déconvolution impulsionnelle

La figure 1.1 illustre le fonctionnement de SBR et la figure 1.2 donne quelques tendances de comparaisons avec d'autres algorithmes. Ces résultats seront complétés au chapitre 2 par des simulations concernant l'approximation d'un signal par un signal polynomial par morceaux, avec peu de morceaux.

Les résultats de la figure 1.1 concernent le problème de déconvolution impulsionnelle $\mathbf{y} = \mathbf{h} * \mathbf{x}^* + \mathbf{b}$ (résultats extraits de [SOU 11]). Ce problème se réécrit $\mathbf{y} = \mathbf{A}\mathbf{x}^* + \mathbf{b}$ où \mathbf{A} est une matrice de Tœplitz dont les colonnes sont obtenues en décalant le signal \mathbf{h} . La matrice \mathbf{A} est approximativement carrée suivant les conditions aux limites choisies. Sur l'exemple de la figure 1.1, la réponse impulsionnelle \mathbf{h} est un filtre passe-bas (filtre gaussien d'écart-type $\sigma = 50$: un motif gaussien lié à une impulsion $x_i^* \neq 0$ occupe une largeur de taille 301 échantillons) ce qui conduit à une matrice \mathbf{A} de taille 3000×2700 . La figure 1.1(a) montre le signal \mathbf{y} simulé et le signal impulsionnel inconnu \mathbf{x}^* comportant 17 impulsions. On constate que certaines impulsions

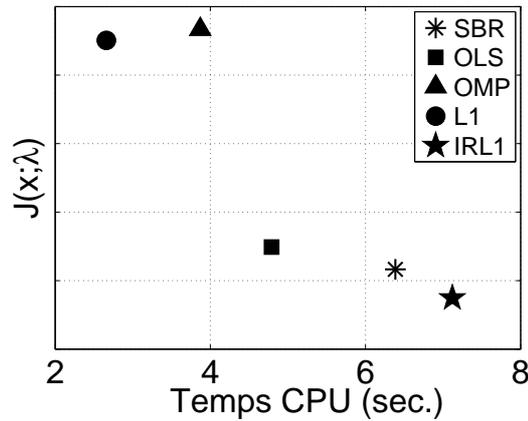


FIGURE 1.2 – Comparaison d’algorithmes d’approximation parcimonieuse en terme de compromis entre la performance (valeur du critère $\mathcal{J}(\mathbf{x}; \lambda)$) et le temps CPU pour le problème de déconvolution de la figure 1.1. SBR($\lambda = 0.5$) fournit un support de cardinalité $k_{\text{sbr}}=18$. Puis, nous exécutons 18 itérations de OLS, OMP, la relaxation convexe (minimisation mixte ℓ_2 - ℓ_1) par l’algorithme d’homotopie [DON 07], et l’algorithme $\text{IR}\ell_1$ (dans la version adaptée au problème (\mathcal{P}_λ) avec le paramètre λ en entrée de l’algorithme) [ZOU 06].

sont suffisamment proches pour que les motifs gaussiens associés se superposent largement, et que d’autres impulsions d’amplitudes plus faibles sont noyées dans le bruit. La figure 1.1(b-d) montre les résultats fournis par SBR avec trois valeurs de λ . Pour les grandes valeurs de λ , seuls les principaux motifs gaussiens sont restaurés. Quand λ décroît, les autres motifs d’amplitudes plus faibles sont reconstruits en même temps que des fausses détections.

Pour ce problème de déconvolution, la matrice \mathbf{A} est mal conditionnée car deux colonnes consécutives de \mathbf{A} sont extrêmement corrélées. Dans ce contexte, les algorithmes de type seuillage itératif (CoSaMP, *Subspace Pursuit*, *Iterative Hard Thresholding*, *Hard Thresholding Pursuit*) conçus pour des problèmes bien conditionnés ont de piètres performances et sont parfois très lents à converger. Le coût par itération de SBR est certes relativement important compte tenu du grand nombre d’inversions de systèmes linéaires, mais le nombre d’itérations est réduit (moins de 25 itérations pour les résultats de la figure 1.1) et de plus, son fonctionnement est simple car il n’y a pas de critère d’arrêt arbitraire. A titre de comparaison, le coût d’une itération de *Iterative Hard Thresholding* (IHT) est très faible mais la convergence d’IHT nécessite au moins 10.000 itérations ce qui conduit à un temps de calcul total supérieur à celui de SBR.

La figure 1.2 dresse un panorama synthétique des performances de SBR, OLS, OMP, des méthodes de relaxation convexe (ℓ_1) et non convexe (algorithme $\text{IR}\ell_1$ associé à la pénalisation $\varphi(x_i; \eta) = \log(|x_i| + \eta)$). Le temps de calcul et la valeur de $\mathcal{J}(\mathbf{x}; \lambda)$ sont respectivement représentés en abscisse et en ordonnée. Cela nous permet d’identifier plusieurs catégories d’algorithmes en fonction de leur position dans le plan 2D : les algorithmes rapides (OMP, régression ℓ_1) se situent dans la région nord-ouest alors que les algorithmes plus efficaces mais plus lents (OLS, SBR et $\text{IR}\ell_1$) se trouvent dans la région sud-est. Les résultats d’algorithmes moins performants qui se situent dans la région nord-est ne sont pas représentés. Cette figure n’est certainement pas un jugement de valeurs dans le but de conclure que tel algorithme est *toujours* meilleur que tel autre. L’objectif est plutôt de regrouper les algorithmes en différentes classes sur la base de leur comportement empirique observé en effectuant de nombreux essais. La figure peut changer de façon significative en fonction du dictionnaire, du niveau de parcimonie λ et de certains pa-

ramètres intervenant dans l'algorithme $\text{IR}\ell_1$. De plus, les performances sont mesurées en terme de valeurs de critères $\mathcal{J}(\mathbf{x}; \lambda)$. Un autre point de vue est de comparer les supports des signaux impulsionnels obtenus en termes de nombre de bonne détection et de fausse alarme [BOU 11b]. D'autres résultats pour le problème de la déconvolution impulsionnelle appliqué au contrôle non destructif par ultrasons sont disponibles dans l'article [SOU 12].

1.6.5 Exemple de la segmentation d'un signal

La segmentation d'un signal est abordée comme son lissage par un signal polynomial par morceaux. Ce problème sera plus amplement repris dans le chapitre 2 (§ 2.3). Le dictionnaire polynomial

$$\mathbf{A} = [\mathbf{A}^0 \mid \dots \mid \mathbf{A}^P] \quad (1.12)$$

est défini comme la concaténation de matrices \mathbf{A}^p dont les colonnes \mathbf{a}_i^p sont les signaux monomiaux $t \mapsto [\max(t - i + 1, 0)]^p$ (signal de Heaviside pour $p = 0$, fonction rampe pour $p = 1$) : cf. fig. 2.7 page 80. En utilisant ce dictionnaire, le produit $\mathbf{A}\mathbf{x}$ génère un signal polynomial par morceaux de degré inférieur ou égal à P . L'approximation parcimonieuse permet d'une part de lisser un signal en l'approchant par un signal polynomial par morceaux et d'autre part, de *segmenter* le signal en détectant ses points de discontinuité (un saut du signal est détecté entre les positions $i - 1$ et i si l'atome \mathbf{a}_i^0 est sélectionné, un changement de pente est lié à l'activation de \mathbf{a}_i^1 , un changement de courbure est lié à l'activation de \mathbf{a}_i^2). Le dictionnaire polynomial est sur-dimensionné, de taille $m \times n$ avec $n = (P + 1)(m - P/2) \geq m$. Contrairement au problème de la déconvolution impulsionnelle où seuls les atomes connexes dans le dictionnaire sont corrélés, tout couple d'atomes \mathbf{a}_i^p et \mathbf{a}_j^q sont corrélés pour le dictionnaire polynomial car les supports de ces signaux se recouvrent. C'est pour cette raison que nous avons préconisé les algorithmes gloutons *bidirectionnels*. Prenons le cas de l'approximation d'un signal \mathbf{y} comportant deux discontinuités aux positions i_1 et i_2 (e.g., un signal constant par morceaux $\mathbf{y}(t)$ avec 3 morceaux $t < i_1$, $t \in [i_1, i_2]$, $t > i_2$). Les algorithmes gloutons commencent par détecter une discontinuité pour une position intermédiaire $i \in [i_1, i_2]$, c'est-à-dire qu'ils sélectionnent un mauvais atome \mathbf{a}_i^p . Les algorithmes monodirectionnels ne peuvent pas retirer ce mauvais atome aux itérations ultérieures : cf. fig. 1.3(b,d).

La figure 1.3 montre une comparaison d'algorithmes en terme de discontinuités détectées (position i et ordre p) pour deux niveaux de parcimonie λ_1 et λ_2 avec $\lambda_1 > \lambda_2$ sur un exemple simple où il s'agit de retrouver les deux positions i des points de changement de pente (abscisses $z_i = -3050$ et -2840). Dans les deux cas, SBR et $\text{IR}\ell_1$ fournissent les meilleurs résultats. SBR est meilleur en terme de localisation précise des positions. Au contraire, les autres algorithmes testés fournissent des résultats moins bons à la fois en terme de localisation et de compromis erreur d'approximation *vs* nombre d'atomes sélectionnés. Notons qu'OLS localise précisément les deux points de changement de pente lorsque cinq itérations sont exécutées (ces points sont la première et la dernière position parmi les cinq) alors qu'OMP échoue même en exécutant davantage d'itérations.

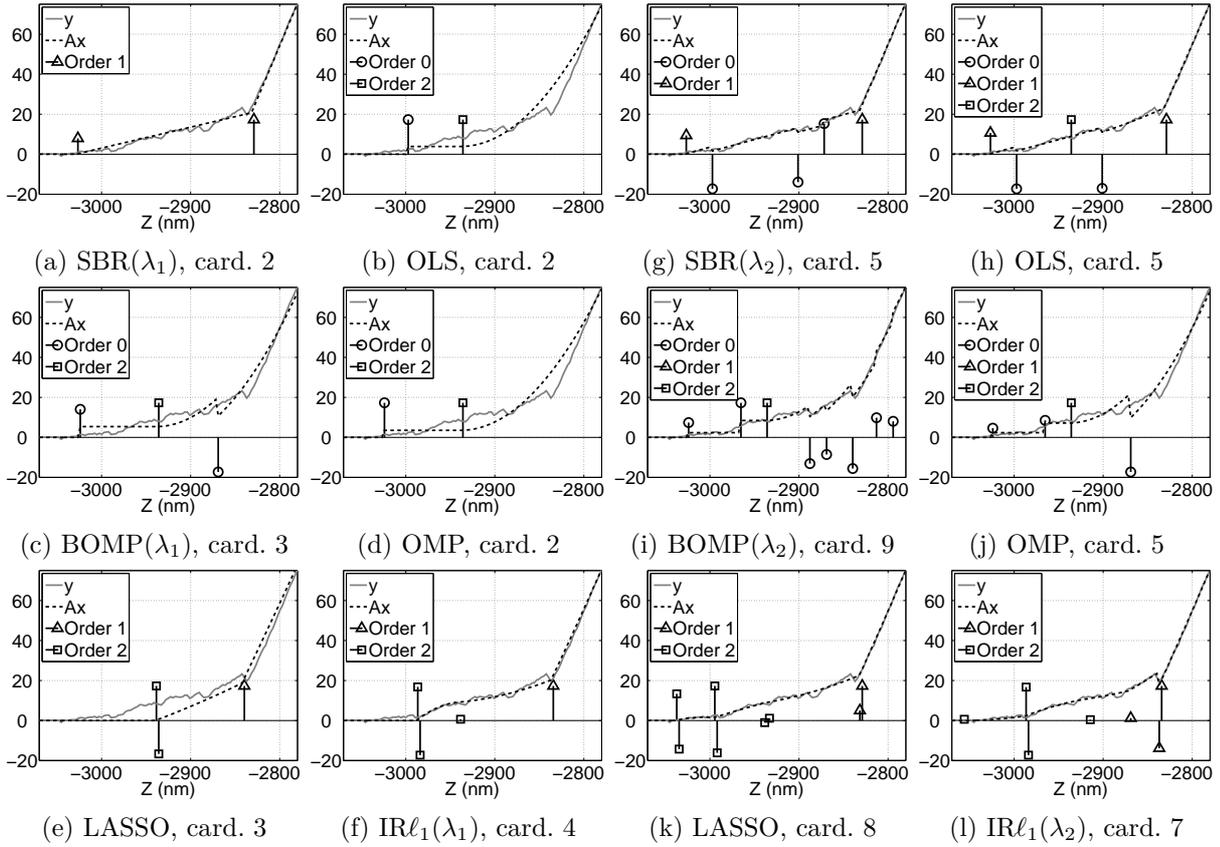


FIGURE 1.3 – Détection de discontinuités d’ordres 0, 1 et 2 (saut du signal, changement de pente et de courbure) à l’aide du dictionnaire polynomial. Les discontinuités trouvées sont représentées par des pics verticaux et un label spécifique représentant leur ordre. (a) Résultat produit par SBR de cardinalité 2 : 4 ajouts et 2 retraits sont effectués ($\lambda_1 = 120$). (b-f) Résultats produits par OLS et OMP après 2 itérations, par BOMP [HER 10], LASSO correspondant à la relaxation ℓ_1 de la norme ℓ_0 , et $\text{IR}\ell_1$ [ZOU 06] avec $\lambda = \lambda_1$. (g-l) Même traitement avec une valeur plus faible de λ ($\lambda_2 = 8.5$). Le support trouvé par SBR est de cardinalité 5 (7 ajouts et 2 retraits) (extrait de [SOU 11]).

1.7 Minimisation mixte ℓ_2 - ℓ_0 pour des niveaux de parcimonie variés

La principale difficulté pour utiliser SBR est le réglage de l’hyperparamètre λ qui est loin d’être trivial. Il est souvent plus simple de régler soit le nombre maximal d’atomes à sélectionner k (en imposant $\|\mathbf{x}\|_0 \leq k$), soit une borne supérieure ε sur l’erreur d’approximation, généralement liée au bruit de mesure. Ces deux solutions sont adoptées dans le chapitre 2 pour le traitement de données réelles en microscopie AFM en utilisant OLS et SBR. La difficulté du réglage de λ conjuguée au fait que pour obtenir des solutions à différents niveaux de parcimonie, il faut exécuter SBR avec plusieurs valeurs de λ (ce qui n’est pas le cas d’OMP ou d’OLS : il suffit simplement d’augmenter le nombre d’itérations et de sauvegarder les résultats intermédiaires) nous a conduit à concevoir des algorithmes étendus fournissant en une exécution, des approximations à des niveaux de parcimonie variés. Les deux algorithmes proposés consistent

donc à minimiser $\mathcal{J}(\mathbf{x}; \lambda)$ par rapport à \mathbf{x} pour une suite de valeurs caractéristiques de λ .

L'algorithme CSBR (pour *Continuation SBR*) effectue des appels récursifs à SBR avec des valeurs décroissantes de λ , calculées de façon adaptative au signal donné \mathbf{y} . CSBR est inspiré par les algorithmes de *continuation* existants (nommés LARS et homotopie) [EFR 04, MAL 05, DON 08] pour résoudre les problèmes d'optimisation quadratiques sous contrainte ℓ_1 comme

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.c.} \quad \|\mathbf{x}\|_1 \leq \nu \quad (1.13)$$

pour un continuum de valeurs de ν , ou de façon équivalente :

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mu \|\mathbf{x}\|_1 \quad (1.14)$$

pour un continuum de valeurs de μ . Ces algorithmes exploitent que le chemin de régularisation est affine par morceaux pour les problèmes mixtes ℓ_2 - ℓ_1 : la solution *optimale* $\mathbf{x}(\nu)$ est une fonction affine par morceaux de ν . Le chemin de régularisation est donc décrit par un nombre fini de solutions pour des valeurs critiques $\nu = \nu_i$ (ou de façon équivalente, pour des valeurs $\mu = \mu_i$) qui définissent les intervalles sur lesquels $\mathbf{x}(\nu)$ est une fonction affine de ν . L'algorithme CSBR généralise ce principe aux problèmes ℓ_2 - ℓ_0 . Il sera détaillé au § 1.7.3.

Nous proposons un deuxième algorithme, ℓ_0 *regularization path track* (ℓ_0 -PT), basé sur une stratégie de recherche globale (au contraire de CSBR qui travaille successivement à des niveaux de parcimonie spécifiques $\lambda = \lambda_j$) pour estimer le chemin de régularisation ℓ_2 - ℓ_0 , caractérisé par la fonction $\lambda \mapsto \min_{\mathbf{x}} \{\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0\}$. Comme nous le verrons plus loin, cette fonction est constant par morceaux et le principe de ℓ_0 -PT est de rechercher les morceaux. Contrairement à CSBR, ℓ_0 -PT n'est pas basé sur des appels à SBR. Il est potentiellement plus efficace au prix d'un coût calculatoire supérieur.

Une adaptation immédiate de ces algorithmes (car ce sont des algorithmes gloutons produisant des supports emboîtés entre deux itérations consécutives) consiste à produire une liste de solutions pour le problème mixte (\mathcal{C}_k) pour des cardinalités consécutives k , ou de façon équivalente pour le problème (\mathcal{C}_ε) avec des valeurs continûment décroissantes de ε . Cette adaptation rend possible l'utilisation de méthodes de sélection d'ordre par différents critères (critère d'Akaike, *Minimum Description Length*, etc.) [WAN 04] puisque des solutions parcimonieuses sont disponibles pour toutes les cardinalités k (§ 1.7.5).

1.7.1 Connexion avec l'optimisation bi-objectif

Le chemin de régularisation associé au problème ℓ_2 - ℓ_0 pénalisé (qui réunit l'ensemble des minimiseurs globaux de $\mathcal{J}(\mathbf{x}; \lambda)$ par rapport à \mathbf{x} pour tous $\lambda > 0$) est *constant* par morceaux [DUA 09c, DUA 10]. Mais comme la norme ℓ_0 est une fonction non convexe, il diffère généralement du chemin de régularisation associé aux problèmes contraints (\mathcal{C}_k) et (\mathcal{C}_ε), *i.e.*, l'ensemble des minimiseurs globaux de ces problèmes pour tout k et pour tout ε [NIK 13]. C'est un résultat classique de l'optimisation multi-objectifs [DAS 07].

Pour être plus précis, les trois formulations (\mathcal{P}_λ), (\mathcal{C}_k) et (\mathcal{C}_ε) sont liées au même problème d'optimisation bi-objectif [MAR 04] : elles ont toutes pour vocation de minimiser à la fois l'erreur d'approximation $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ et la mesure de parcimonie $\|\mathbf{x}\|_0$. Comme nous l'avons déjà vu, ce problème d'optimisation peut être considéré comme un problème discret, celui de l'optimisation conjointe de $\mathcal{E}_{\mathcal{Q}}$ et $\text{Card}[\mathcal{Q}]$ par rapport à \mathcal{Q} , car il y a une correspondance bijective entre les solutions \mathbf{x} et \mathcal{Q} des deux problèmes¹¹. La figure 1.4(a) est une représentation classique du

11. Etablir cette correspondance pose une difficulté technique car les solutions $\mathbf{x}_{\mathcal{Q}}$ qui minimisent l'erreur

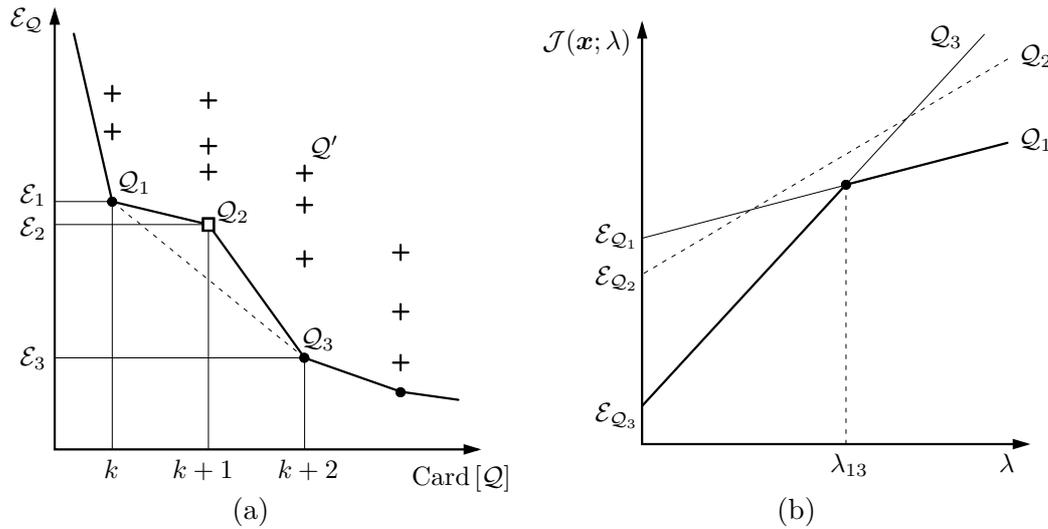


FIGURE 1.4 – Minimisation de critères mixtes ℓ_2 - ℓ_0 en tant que problème d’optimisation bi-objectif. (a) La frontière de Pareto, en trait plein, rassemble les points *non-dominés* : aucun autre point ne peut faire décroître à la fois $\text{Card}[\mathcal{Q}]$ et $\mathcal{E}_{\mathcal{Q}}$. Les points non-dominés sont \mathcal{Q}_1 , \mathcal{Q}_2 et \mathcal{Q}_3 . Les autres points \mathcal{Q}' représentés par une croix sont dominés. (b) Dans le domaine (λ, \mathcal{J}) , chaque support \mathcal{Q} forme une droite $\lambda \mapsto \mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}]$ de pente $\text{Card}[\mathcal{Q}]$. \mathcal{Q}_1 et \mathcal{Q}_3 sont des solutions optimales pour au moins une valeur de λ . La solution \mathcal{Q}_2 localisée dans une zone non convexe sur la frontière de Pareto (a) n’est pas optimale au sens de \mathcal{J} : la droite associée $\lambda \mapsto \mathcal{E}_{\mathcal{Q}_2} + \lambda \text{Card}[\mathcal{Q}_2]$ est strictement au dessus de l’enveloppe concave de l’ensemble des droites (obtenue en considérant leur minimum pour tous les supports \mathcal{Q}) quelquesoit λ .

problème bi-objectif où chaque axe est lié à un objectif particulier [DAS 07] : $\text{Card}[\mathcal{Q}]$ et $\mathcal{E}_{\mathcal{Q}}$ sont représentés horizontalement et verticalement. La figure indique que \mathcal{Q}_1 , \mathcal{Q}_2 et \mathcal{Q}_3 sont des solutions optimales pour le problème contraint (\mathcal{C}_k) pour certaines valeurs de k , ou de façon équivalente, pour $(\mathcal{C}_\varepsilon)$ pour certaines valeurs de l’erreur quadratique ε (avec $\mathcal{E}_{\mathcal{Q}} \leq \varepsilon$) ; ces solutions optimales constituent la frontière de Pareto [MAR 04]. En revanche, seuls \mathcal{Q}_1 et \mathcal{Q}_3 qui sont situés dans des zones convexes de la frontière de Pareto, sont des solutions optimales pour le problème pénalisé (\mathcal{P}_λ) pour certaines valeurs de λ . Ces points sont qualifiés de *supportés* dans la littérature de l’optimisation multi-objectifs [MAR 04].

La stratégie consistant à minimiser la fonction coût $\mathcal{J}_{\mathcal{Q}}(\lambda) = \mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}]$ au lieu de traiter les problèmes contraints (\mathcal{C}_k) ou $(\mathcal{C}_\varepsilon)$ est connue sous le nom de méthode de la somme pondérée (*weighted sum*) : le problème bi-objectif (contraint) est simplifié en un problème mono-objectif dont la fonction coût est une combinaison linéaire des objectifs. Cela permet de simplifier le problème d’optimisation. Mais en contrepartie, en supposant qu’on dispose d’un algorithme optimal pour minimiser $\mathcal{J}_{\mathcal{Q}}(\lambda)$ par rapport à \mathcal{Q} pour tous λ , seules les solutions supportées (\mathcal{Q}_1 et \mathcal{Q}_3 sur la figure 1.4) peuvent être atteintes par la méthode de la somme pondérée [MAR 04].

quadratique à \mathcal{Q} fixé peuvent éventuellement avoir des composantes nulles $x_i = 0$ pour $i \in \mathcal{Q}$. De plus, leur unicité n’est pas garantie si $\mathbf{A}_{\mathcal{Q}}$ est de rang déficient. Je ne m’étends pas davantage sur ce point technique et renvoie le lecteur intéressé au rapport [DUA 09c] et à la référence [NIK 13].

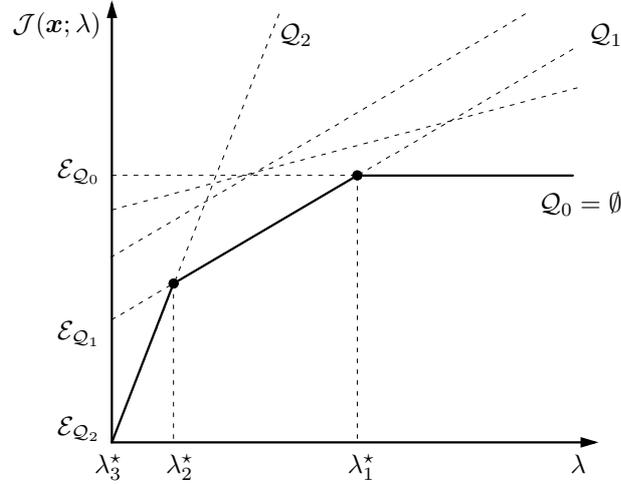


FIGURE 1.5 – Représentation de quelques droites $\lambda \mapsto \mathcal{J}_{\mathcal{Q}}(\lambda) = \mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}]$ pour des supports variés \mathcal{Q} . La « courbe ℓ_0 », en trait continu, est la courbe minimale $\lambda \mapsto \min_{\mathcal{Q}} \{\mathcal{J}_{\mathcal{Q}}(\lambda)\}$. Elle est continue, concave et affine par morceaux avec un nombre fini de morceaux. Elle caractérise le chemin de régularisation ℓ_0 .

1.7.2 Idées sous-jacentes aux algorithmes proposés

Caractérisation du chemin de régularisation ℓ_2 - ℓ_0

Dans toute la suite, la dénomination « chemin de régularisation ℓ_0 » est relative au critère pénalisé $\mathcal{J}(\mathbf{x}; \lambda) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0$.

Les algorithmes proposés ont pour but d'estimer un chemin de régularisation constant par morceaux. Mais il est important de remarquer que ce chemin estimé n'est généralement pas le chemin optimal, le problème d'optimisation ℓ_2 - ℓ_0 étant NP-complet. C'est une différence majeure avec l'algorithme d'homotopie dédié à la minimisation mixte ℓ_2 - ℓ_1 .

Les algorithmes proposés ci-dessous pour caractériser le chemin de régularisation ℓ_2 - ℓ_0 exploitent l'interprétation géométrique très simple de la figure 1.5 :

1. Tout support \mathcal{Q} engendre une droite $\lambda \mapsto \{\mathcal{J}_{\mathcal{Q}}(\lambda) = \mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}]\}$ dans le plan où $\mathcal{E}_{\mathcal{Q}} = \|\mathbf{P}_{\mathcal{Q}}^{\perp} \mathbf{y}\|_2^2$ est le résidu de l'approximation des données \mathbf{y} avec les atomes du dictionnaire indicés par \mathcal{Q} .
2. Le chemin de régularisation ℓ_0 est constant par morceaux [DUA 09c, DUA 10] : il est caractérisé par l'enveloppe concave de cet ensemble de droites. Comme il y a un nombre fini de supports possibles, le nombre de droites $\lambda \mapsto \mathcal{J}_{\mathcal{Q}}(\lambda)$ est fini, donc l'enveloppe concave est un polygone dont le dernier morceau (à l'extrême droite) est horizontal.

Cette enveloppe concave est nommée la « courbe ℓ_0 » : elle définit l'ensemble des valeurs minimales du critère $\mathcal{J}(\mathbf{x}; \lambda)$ pour un continuum de valeurs de λ . Chaque intervalle $(\lambda_{i+1}^*, \lambda_i^*)$ de la courbe ℓ_0 est associé à au moins un support \mathcal{Q}_i qui fournit un¹² minimiseur global de $\mathcal{J}(\mathbf{x}; \lambda)$ pour $\lambda \in (\lambda_{i+1}^*, \lambda_i^*)$. Ce minimiseur global s'écrit $\mathbf{x}_{\mathcal{Q}_i} = \{\mathbf{A}_{\mathcal{Q}_i}^{\dagger} \mathbf{y}; \mathbf{0}\}$. Les sommets du polygone définissent les valeurs critiques $\lambda = \lambda_i^*$ autour desquelles l'ensemble des solutions optimales de $\min_{\mathbf{x}} \mathcal{J}(\mathbf{x}; \lambda)$ change. Sur l'exemple de la figure 1.5, le chemin de régularisation est formé des

12. le minimiseur global de (\mathcal{P}_{λ}) associé à un support optimal \mathcal{Q}_i est unique car $\mathbf{A}_{\mathcal{Q}_i}$ est forcément de rang plein [NIK 13, Th. 4.4].

supports \mathcal{Q}_0 , \mathcal{Q}_1 , et \mathcal{Q}_2 , et des autres supports éventuels dont les droites sont confondues avec les droites $\lambda \mapsto \mathcal{J}_{\mathcal{Q}_i}(\lambda)$ pour $i = 0, 1$ ou 2 (¹³).

Le raisonnement ci-dessus est basé sur les solutions *optimales* du problème d'optimisation (\mathcal{P}_λ). Nous en tirons partie pour définir des heuristiques *sous-optimales* pour minimiser $\mathcal{J}(\mathbf{x}; \lambda)$ pour un continuum de valeurs de λ . Ces heuristiques fournissent également des solutions sous-optimales aux problèmes contraints (\mathcal{C}_k) et (\mathcal{C}_ε) pour des valeurs consécutives de k , et pour un continuum de valeurs de ε .

Principe de continuation

Rappelons que SBR travaille pour un niveau de parcimonie fixé λ . Une stratégie naïve et inélégante pour fournir des solutions à des niveaux de parcimonie variés consiste à exécuter SBR pour de nombreuses valeurs de λ uniformément réparties sur une grille. Cette stratégie est particulièrement inefficace car les exécutions de SBR pour des valeurs parfois éloignées de λ produisent des itérés identiques. Nous avons de plus constaté l'effet inverse : la cardinalité du support en sortie de SBR est susceptible de varier de façon significative pour de très faibles variations de λ . La discrétisation de λ sur une grille uniforme a donc toutes les chances de ne pas détecter ces changements de comportement de SBR.

Nous avons proposé une stratégie de continuation visant à minimiser $\mathcal{J}(\mathbf{x}; \lambda)$ de proche en proche pour des valeurs décroissantes de λ situées sur une grille adaptative [DUA 09a]. De façon très générale, le concept de continuation consiste à traiter un problème difficile en résolvant une suite de problèmes plus simples avec des solutions initiales pertinentes (*warm start initialization*) en faisant varier un hyperparamètre de façon graduelle [WAS 73]. Dans le contexte de l'approximation parcimonieuse, la continuation fait référence à la minimisation mixte ℓ_2 - ℓ_1 : le problème d'optimisation (1.14) est résolu pour des valeurs décroissantes de l'hyperparamètre μ en utilisant la solution courante comme initialisation (*warm start*) pour le problème ℓ_2 - ℓ_1 avec la valeur suivante de l'hyperparamètre μ (¹⁴) [TRO 10]. L'algorithme d'homotopie pour la relaxation convexe $\ell_2 - \ell_1$ [EFR 04, MAL 05, DON 08] tire partie du caractère affine par morceaux du chemin de régularisation $\ell_2 - \ell_1$ (les solutions $\lambda \mapsto \arg \min_{\mathbf{x}} \{\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mu\|\mathbf{x}\|_1\}$ sont une fonction affine par morceaux de λ) et recherche les extrémités des morceaux qui caractérisent le changement du support de la solution. Contrairement à l'algorithme d'homotopie, l'algorithme de continuation proposé pour le problème ℓ_2 - ℓ_0 est sous-optimal : il n'y a pas de garantie que les valeurs de l'hyperparamètre λ trouvées sont liées au chemin de régularisation optimal, ni que la minimisation ℓ_2 - ℓ_0 à λ fixé, en appelant SBR, produise un résultat optimal.

1.7.3 Algorithme Continuation Single Best Replacement

L'algorithme CSBR réalise des appels successifs à SBR avec des valeurs λ_j décroissantes fournies par l'algorithme¹⁵ et avec pour initialisation le support obtenu en sortie du dernier appel à SBR.

13. avec un léger abus de notations, puisque le chemin de régularisation n'est pas défini comme un ensemble de supports mais l'ensemble des solutions $\mathbf{x}_{\mathcal{Q}}$ liées à ces supports.

14. Dans [TRZ 09], le mot « continuation » prend un sens totalement différent. Il fait référence à une méthode de type Non Convexité Graduelle (*Graduated Non Convexity*, GNC) [BLA 87] proche de celle de [MOH 09], où la pseudo-norme ℓ_0 est relaxée par une série de métriques concaves conduisant à résoudre une succession de problèmes d'optimisation continus avec une initialisation de type *warm start*.

15. J'adopte la notation λ_j sans astérisque pour les différencier des valeurs critiques λ_i^* liées au chemin de régularisation optimal (fig. 1.5).

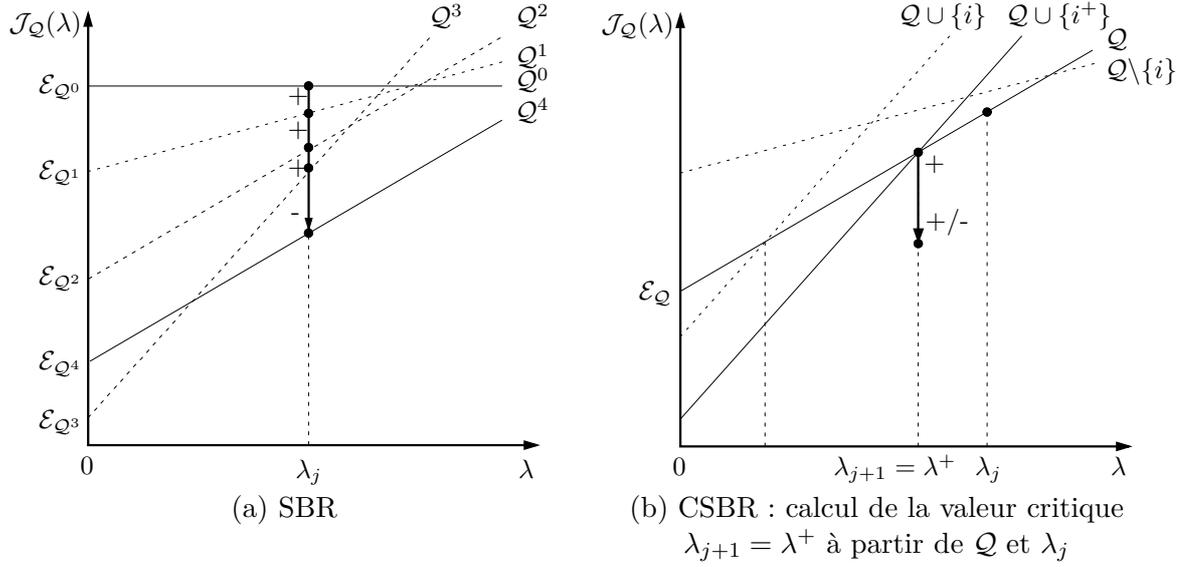


FIGURE 1.6 – Interprétation graphique des algorithmes SBR et CSBR. A chaque support \mathcal{Q} est associée la droite $\lambda \mapsto \mathcal{J}_{\mathcal{Q}}(\lambda)$ de pente $\text{Card}[\mathcal{Q}]$. (a) Algorithme SBR : pour $\lambda = \lambda_j$ fixé, un remplacement élémentaire $\mathcal{Q} \bullet \ell$ est représenté par un déplacement vertical de haut en bas pour relier les droites \mathcal{Q} et $\mathcal{Q} \bullet \ell$. La pente est incrémentée (ajout d’atome) ou décrétementée (retrait) de 1. Le support initial \mathcal{Q}^0 est vide (droite horizontale) ; trois ajouts puis un retrait sont réalisés. (b) Algorithme CSBR : calcul récursif de λ . \mathcal{Q} est le support obtenu en sortie de SBR pour $\lambda = \lambda_j$. Toutes les droites $\mathcal{Q} \bullet i$, en pointillés, se situent donc au dessus de \mathcal{Q} pour $\lambda = \lambda_j$. Un ajout $\mathcal{Q} \cup \{i\}$ produit une droite qui intersecte la droite \mathcal{Q} pour $\lambda < \lambda_j$ (respectivement, un retrait $\mathcal{Q} \setminus \{i\}$ conduit à une intersection pour $\lambda \geq \lambda_j$). La droite $\mathcal{Q} \cup \{i^+\}$ située le plus à droite parmi toutes les droites (parallèles) $\mathcal{Q} \cup \{i\}$ définit la prochaine valeur $\lambda_{j+1} = \lambda^+$ pour laquelle SBR sera exécuté. $\ell = i^+$ est le premier indice à être ajouté dans \mathcal{Q} lors de l’appel de $\text{SBR}(\mathcal{Q}; \lambda_{j+1})$.

Interprétation graphique de l’algorithme SBR

L’algorithme SBR présenté dans la section 1.6 travaille à λ fixé. Avec une représentation graphique du même type que celle de la figure 1.5, les itérations successives de SBR décrivent un mouvement vertical descendant dans le domaine (λ, \mathcal{J}) : à λ fixé, chaque remplacement élémentaire $\mathcal{Q} \bullet \ell$ permet de faire décroître au plus le critère $\mathcal{J}_{\mathcal{Q}}(\lambda) = \mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}]$. Pour un ajout d’atome ($\bullet = \cup$), la droite liée au nouveau support a une pente incrémentée de 1 : $\text{Card}[\mathcal{Q} \cup \{\ell\}] = \text{Card}[\mathcal{Q}] + 1$. Inversement, la pente est décrétementée de 1 pour un retrait : cf. fig. 1.6(a). L’algorithme s’arrête quand les droites liées aux ajouts et aux retraits se situent toutes au dessus de la droite courante pour la valeur de λ considérée ($\forall i, \mathcal{J}_{\mathcal{Q} \bullet i}(\lambda) \geq \mathcal{J}_{\mathcal{Q}}(\lambda)$).

Utilisation récursive de l’algorithme SBR

La figure 1.6(a) illustre l’exécution de SBR pour une valeur spécifique de $\lambda = \lambda_j$. En notant $\mathcal{Q}_{\text{init}}$ le support initial (par défaut, le support vide) et $\mathcal{Q} = \text{SBR}(\mathcal{Q}_{\text{init}}; \lambda_j)$ le support obtenu en sortie de SBR, il est clair que $\text{SBR}(\mathcal{Q}; \lambda_j) = \mathcal{Q}$ car la condition d’arrêt de SBR est satisfaite pour le support \mathcal{Q} . Cette condition s’écrit :

$$\forall i, \mathcal{J}_{\mathcal{Q} \bullet i}(\lambda_j) \geq \mathcal{J}_{\mathcal{Q}}(\lambda_j) \quad (1.15)$$

ou encore

$$\underbrace{\max_{i \notin \mathcal{Q}} \{\mathcal{E}_{\mathcal{Q}} - \mathcal{E}_{\mathcal{Q} \cup \{i\}}\}}_{\lambda^+} \leq \lambda_j \leq \underbrace{\min_{i \in \mathcal{Q}} \{\mathcal{E}_{\mathcal{Q} \setminus \{i\}} - \mathcal{E}_{\mathcal{Q}}\}}_{\lambda^-} \quad (1.16)$$

de par la définition de $\mathcal{J}_{\mathcal{Q}}(\lambda_j) = \mathcal{E}_{\mathcal{Q}} + \lambda_j \text{Card}[\mathcal{Q}]$. L'encadrement (1.16) dissocie explicitement les tests d'arrêt liés à l'ajout et au retrait d'atomes (bornes λ^+ et λ^- , respectivement).

La stratégie de continuation consiste à effectuer un nouvel appel à $\text{SBR}(\mathcal{Q}; \lambda_{j+1})$ pour

$$\lambda_{j+1} = \lambda^+ \triangleq \max_{i \notin \mathcal{Q}} \{\mathcal{E}_{\mathcal{Q}} - \mathcal{E}_{\mathcal{Q} \cup \{i\}}\} \quad (1.17)$$

$$= \lambda_j + \mathcal{J}_{\mathcal{Q}}(\lambda_j) - \min_{i \notin \mathcal{Q}} \mathcal{J}_{\mathcal{Q} \cup \{i\}}(\lambda_j). \quad (1.18)$$

λ^+ est la borne inférieure de (1.16). Pour $\lambda < \lambda^+$, la condition d'arrêt (1.15)-(1.16) est violée, et le critère $\mathcal{J}_{\mathcal{Q}}(\lambda)$ peut décroître en effectuant un remplacement élémentaire (ajout). $\lambda = \lambda^+$ est le cas limite où (1.15) et la première inégalité de (1.16) deviennent des égalités : $\mathcal{J}_{\mathcal{Q}}(\lambda) = \mathcal{J}_{\mathcal{Q} \cup \{i\}}(\lambda)$ pour une valeur de i . Dans ce qui suit et pour lever toute ambiguïté concernant une décroissance nulle du critère, l'exécution de SBR pour $\lambda = \lambda^+$ doit être comprise comme le cas limite du fonctionnement de SBR pour $\lambda \rightarrow \lambda^+$, $\lambda < \lambda^+$.

L'exécution de $\text{SBR}(\mathcal{Q}; \lambda_{j+1})$ réalise donc au moins une itération en sélectionnant un atome

$$i^+ \in \arg \max_{i \notin \mathcal{Q}} \{\mathcal{E}_{\mathcal{Q}} - \mathcal{E}_{\mathcal{Q} \cup \{i\}}\} = \arg \min_{i \notin \mathcal{Q}} \mathcal{J}_{\mathcal{Q} \cup \{i\}}(\lambda_j). \quad (1.19)$$

La valeur du critère $\mathcal{J}_{\mathcal{Q}}(\lambda_{j+1})$ est inchangée lors de cette itération : $\mathcal{J}_{\mathcal{Q} \cup \{i^+\}}(\lambda_{j+1}) = \mathcal{J}_{\mathcal{Q}}(\lambda_{j+1})$. A la deuxième itération de SBR, les tests d'ajout et de retrait d'atomes sont normalement effectués à l'exception du retrait de i^+ pour éviter que l'algorithme ne boucle de façon infinie.

La version adaptée de SBR est résumée dans le tableau 1.5 où les modifications par rapport à la version standard (tableau 1.4) apparaissent en gras : la gestion des deux premières itérations et le calcul de la prochaine valeur λ^+ . SBR possède maintenant trois entrées correspondant au support initial $\mathcal{Q}_{\text{init}}$, à λ_j et au premier atome i_{init} à intégrer $\mathcal{Q}_{\text{init}}$. SBR renvoie comme sorties le support \mathcal{Q} , la prochaine valeur λ^+ et le nouvel atome i^+ à intégrer au support \mathcal{Q} lors de l'appel à SBR pour $\lambda = \lambda^+$. On remarque que la valeur de λ^+ se déduit trivialement des calculs effectués lors de la dernière itération de SBR. En effet, les tests d'insertion conduisent à calculer $\mathcal{J}_{\mathcal{Q} \cup \{i\}}(\lambda_j)$ pour tout $i \notin \mathcal{Q}$. On déduit directement λ^+ par (1.18).

Illustration graphique

Les appels récursifs de SBR et le sens de λ^+ sont illustrés sur la figure 1.6(b). Si $\mathcal{Q} = \text{SBR}(\mathcal{Q}_{\text{init}}; \lambda_j)$, alors :

- pour $\lambda < \lambda_j$, tout retrait d'atome incrémente le critère $\mathcal{J}_{\mathcal{Q}}(\lambda)$. La droite associée à $\mathcal{Q} \setminus \{i\}$, en pointillés, se situe au dessus de la droite associée à \mathcal{Q} pour $\lambda < \lambda_j$;
- pour $\lambda > \lambda^+$, tout ajout d'atome incrémente le critère $\mathcal{J}_{\mathcal{Q}}(\lambda)$. La droite associée à $\mathcal{Q} \cup \{i\}$, en pointillés, se situe au dessus de la droite associée à \mathcal{Q} pour $\lambda > \lambda^+$;
- les droites \mathcal{Q} et $\mathcal{Q} \cup \{i^+\}$ s'intersectent pour $\lambda = \lambda^+$: $\mathcal{J}_{\mathcal{Q}}(\lambda^+) = \mathcal{J}_{\mathcal{Q} \cup \{i^+\}}(\lambda^+)$. Le signe (+) fait référence à l'ajout de i^+ à la première itération de $\text{SBR}(\mathcal{Q}; \lambda^+)$. Les flèches (+/-) représentent les remplacements élémentaires à partir de la deuxième itération de $\text{SBR}(\mathcal{Q}; \lambda^+)$. Ces remplacements font décroître le critère $\mathcal{J}_{\mathcal{Q}}(\lambda^+)$.

Entrées : \mathbf{A} , \mathbf{y} , λ , $\mathcal{Q}_{\text{init}}$ et $i_{\text{init}} \notin \mathcal{Q}_{\text{init}}$

Initialisation : Faire $\mathcal{Q} = \mathcal{Q}_{\text{init}} \cup \{i_{\text{init}}\}$ et **iter** = 2.

Etape 1 : [1. Sélection du nouvel atome ℓ]

 Pour $i \in \{1, \dots, n\}$, calculer $\mathcal{J}_{\mathcal{Q} \bullet i}(\lambda)$.

Si **iter** = 2,

Calculer $\ell \in \arg \min_{i \neq i_{\text{init}}} \mathcal{J}_{\mathcal{Q} \bullet i}(\lambda)$.

Sinon

 Calculer $\ell \in \arg \min_i \mathcal{J}_{\mathcal{Q} \bullet i}(\lambda)$.

Fin Si

 [2. Test d'arrêt]

 Si $\mathcal{J}_{\mathcal{Q} \bullet \ell}(\lambda) < \mathcal{J}_{\mathcal{Q}}(\lambda)$,

 Faire $\mathcal{Q} = \mathcal{Q} \bullet \ell$.

 Sinon,

Calculer λ^+ et i^+ définis par (1.18) et (1.19).

 Terminer SBR.

Fin Si.

 Faire **iter** = **iter** + 1 et retourner à l'étape 1.

Sorties : $[\mathcal{Q}, \lambda^+, i^+] = \text{SBR}(\mathcal{Q}_{\text{init}}; \lambda, i_{\text{init}})$.

TABLE 1.5 – Algorithme *Single Best Replacement* adapté pour faciliter les appels à SBR dans CSBR en évitant les calculs redondants. Les modifications par rapport à la version standard de SBR (tableau 1.4 page 45) apparaissent en gras. SBR prend une entrée supplémentaire i_{init} par rapport à sa version standard. i_{init} est l'indice du premier nouvel atome à intégrer le support $\mathcal{Q}_{\text{init}}$. SBR renvoie deux sorties supplémentaires : la prochaine valeur $\lambda^+ < \lambda$ et l'indice $i_{\text{init}} = i^+$ utilisés par CSBR pour le prochain appel de SBR.

Présentation de l'algorithme CSBR

L'algorithme CSBR est résumé dans le tableau 1.6 où les appels successifs

$$\mathcal{Q}_j = \text{SBR}(\mathcal{Q}_{j-1}; \lambda_j) \tag{1.20}$$

produisent des supports \mathcal{Q}_j pour des valeurs décroissantes de λ_j , avec $\lambda_0 = +\infty$ et $\mathcal{Q}_0 = \emptyset$. Conformément à la figure 1.6(b), \mathcal{Q}_j est la solution sous-optimale du problème (\mathcal{P}_λ) pour tout $\lambda \in (\lambda_{j+1}, \lambda_j]$, et la liste des supports \mathcal{Q}_j obtenus couvre un continuum de valeurs de $\lambda \in (\lambda_{J+1}, +\infty)$ quand J itérations de CSBR ont été exécutées. A la première itération, $\mathcal{Q}_0 = \emptyset$. (1.17) et (1.19) se simplifient en :

$$i^+ \in \arg \max_{i \in \{1, \dots, n\}} \frac{|\langle \mathbf{y}, \mathbf{a}_i \rangle|}{\|\mathbf{a}_i\|_2} \quad \text{et} \quad \lambda_1 = \lambda^+ = \frac{\langle \mathbf{y}, \mathbf{a}_{i^+} \rangle^2}{\|\mathbf{a}_{i^+}\|_2^2}. \tag{1.21}$$

On impose les deux conditions d'arrêt habituelles qui reposent sur la cardinalité maximale K et le seuil ε sur l'erreur quadratique d'approximation : CSBR s'arrête lorsque \mathcal{Q}_j est de

Entrées : \mathbf{A} , \mathbf{y} , ε et/ou K	
Initialisation : $\lambda_0 = +\infty$, $\mathcal{Q}_0 = \emptyset$.	
	Faire $j = 1$. Calculer λ_1 et i^+ par (1.21).
Etape 1 :	Appeler $[\mathcal{Q}_j, \lambda_{j+1}, i^+] = \text{SBR}(\mathcal{Q}_{j-1}; \lambda_j, i^+)$.
	Si $(\lambda_{j+1} \leq 0)$ ou $(\text{Card}[\mathcal{Q}_j] \geq K)$ ou $(\mathcal{E}_{\mathcal{Q}_j} \leq \varepsilon)$,
	Terminer CSBR.
	Fin Si
	Faire $j = j + 1$ et revenir à l'étape 1.
Sorties :	<ul style="list-style-type: none"> • Chemin de régularisation estimé : liste des valeurs critiques λ_j, listes des supports \mathcal{Q}_j liés aux intervalles $(\lambda_{j+1}, \lambda_j]$ et des erreurs quadratiques $\mathcal{E}_{\mathcal{Q}_j}$. • Meilleurs itérés de SBR $\mathcal{Q}_k^{\text{CSBR}}$ et erreurs associées $\mathcal{E}_k^{\text{CSBR}}$ ($k = 0, \dots, K$).

TABLE 1.6 – Algorithme *Continuation Single Best Replacement*. Les paramètres d'arrêt ε et/ou K portent respectivement sur la norme ℓ_2 du résidu de l'approximation et sur la cardinalité. CSBR appelle de façon répétée la version (très légèrement) adaptée de SBR décrite par le tableau 1.5.

cardinalité supérieure à K ou lorsque l'erreur quadratique $\mathcal{E}_{\mathcal{Q}_j}$ est inférieure à ε . CSBR s'arrête également si $\lambda_{j+1} \leq 0$, c'est-à-dire si tout le chemin de régularisation $\lambda \in (0, +\infty)$ a été parcouru. Ces différents tests portent sur les itérés \mathcal{Q}_j de CSBR mais pas sur les itérés intermédiaires de SBR. SBR n'est donc jamais arrêté avant sa convergence.

Problèmes contraints (\mathcal{C}_k) et (\mathcal{C}_ε)

CSBR est conçu pour traiter le problème (\mathcal{P}_λ) pour un continuum de valeurs de λ (pour $\lambda \in (\lambda_{j+1}, \lambda_j]$, la solution de (\mathcal{P}_λ) est la même que pour $\lambda = \lambda_j$). Il peut être trivialement adapté pour fournir des approximations parcimonieuses pour toutes les cardinalités consécutives $k = 1, \dots, K$. Pour cela, il suffit de mémoriser les meilleurs itérés de SBR au fil des itérations dans un tableau $\mathcal{Q}_k^{\text{CSBR}}$: $\mathcal{Q}_k^{\text{CSBR}}$ est défini comme l'itéré de cardinalité k dont l'erreur quadratique d'approximation ($\mathcal{E}_k^{\text{CSBR}}$) est minimale parmi tous les supports de cardinalité k explorés durant les différents appels à SBR. $\mathcal{Q}_k^{\text{CSBR}}$ and $\mathcal{E}_k^{\text{CSBR}}$ sont donc mis à jour à chaque itération de SBR et fournis en sortie de CSBR. CSBR retourne de plus la liste des valeurs critiques λ_j décrivant le chemin de régularisation estimé et les solutions \mathcal{Q}_j obtenues en sortie de SBR pour ces niveaux de parcimonie λ_j .

1.7.4 Algorithme de reconstruction du chemin de régularisation ℓ_0

Nous proposons un algorithme glouton bidirectionnel, nommé ℓ_0 *regularization path tracking* (ℓ_0 -PT), permettant de rechercher le chemin de régularisation ℓ_2 - ℓ_0 (article [SOU 13d] en cours de finalisation). Nous avons vu que le chemin de régularisation optimal est caractérisé par une courbe ℓ_0 polygonale et concave (fig. 1.5). Le principe de l'algorithme ℓ_0 -PT est de construire et raffiner de façon graduelle une courbe ℓ_0 polygonale et concave en mettant à jour la liste $\{\mathcal{Q}_0, \dots, \mathcal{Q}_J\}$ des supports qui constituent le chemin de régularisation et la liste des valeurs

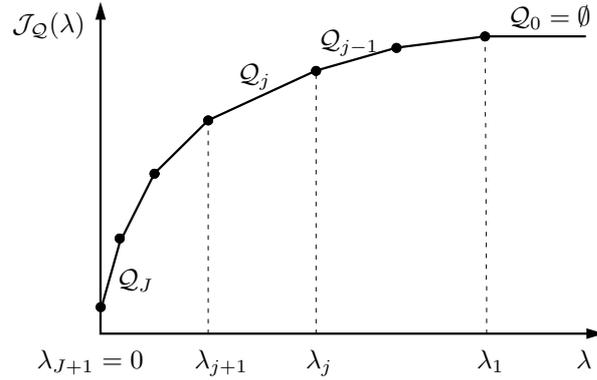


FIGURE 1.7 – Paramétrisation du chemin de régularisation ℓ_2 - ℓ_0 estimé par les algorithmes CSBR et ℓ_0 -PT. \mathcal{Q}_j sont les supports formant le chemin de régularisation ℓ_0 et λ_j sont les valeurs critiques de λ autour desquelles le support de la solution change, classées par ordre décroissant. La courbe ℓ_0 estimée est décrite par les sommets du polygone de coordonnées $(\lambda_j, \mathcal{J}_{\mathcal{Q}_j}(\lambda_j))$ où $\mathcal{J}_{\mathcal{Q}_j}(\lambda_j) = \mathcal{J}_{\mathcal{Q}_{j-1}}(\lambda_j)$. L'arête liée à l'intervalle $(\lambda_{j+1}, \lambda_j)$ coïncide avec la droite $\lambda \mapsto \mathcal{J}_{\mathcal{Q}_j}(\lambda)$.

critiques $\lambda_0 \triangleq +\infty > \lambda_1 > \dots > \lambda_J > \lambda_{J+1} \triangleq 0$. Ces notations sont illustrées sur la figure 1.7 : le support \mathcal{Q}_j est la solution sous-optimale du problème (\mathcal{P}_λ) pour $\lambda \in (\lambda_{j+1}, \lambda_j)$.

Pour l'algorithme CSBR, j correspond à la fois au numéro de l'itération courante et à l'indice du support \mathcal{Q}_j car CSBR construit le chemin de régularisation de proche en proche en considérant chaque intervalle contigu $(\lambda_{j+1}, \lambda_j)$ en allant des plus grandes jusqu'aux plus faibles valeurs de λ_j . A la j -ème itération de CSBR, le chemin de régularisation est partiellement construit pour $\lambda \in (\lambda_{j+1}, +\infty)$. Cette partie du chemin de régularisation reste figée au cours des itérations ultérieures qui explorent le domaine $\lambda \in (0, \lambda_{j+1}]$. La différence essentielle avec ℓ_0 -PT est qu'à chaque itération de ℓ_0 -PT, le chemin de régularisation ℓ_0 est déjà construit sur l'intégralité du domaine $\lambda \in (0, +\infty)$. *Toutes* les valeurs critiques λ_j déjà estimées sont cette fois susceptibles de changer au cours des itérations suivantes de façon à faire « décroître » le plus possible la courbe ℓ_0 estimée : à chaque itération, la courbe ℓ_0 mise à jour se trouve en dessous de la courbe ℓ_0 à l'itération précédente.

Concrètement, une itération de ℓ_0 -PT est liée à un support $\mathcal{Q} = \mathcal{Q}_j$ du chemin de régularisation courant. L'inspection de ce support consiste à tester tous les remplacements élémentaires $\mathcal{Q} \bullet i$ comme pour SBR et CSBR, en calculant les erreurs quadratiques $\mathcal{E}_{\mathcal{Q} \bullet i}$. On recherche les deux remplacements $\mathcal{Q} \cup \{i^+\}$ et $\mathcal{Q} \setminus \{i^-\}$ conduisant aux plus faibles erreurs quadratiques possibles en ajoutant et en retirant un élément dans/de \mathcal{Q} :

$$\begin{cases} i^+ \in \arg \max_{i \notin \mathcal{Q}} \{\mathcal{E}_{\mathcal{Q}} - \mathcal{E}_{\mathcal{Q} \cup \{i\}}\} \\ i^- \in \arg \min_{i \in \mathcal{Q}} \{\mathcal{E}_{\mathcal{Q} \setminus \{i\}} - \mathcal{E}_{\mathcal{Q}}\}. \end{cases} \quad (1.22)$$

Les nouveaux supports $\mathcal{Q}_j \cup \{i^+\}$ et $\mathcal{Q}_j \setminus \{i^-\}$ sont intégrés dans le chemin de régularisation si les droites associées $\lambda \mapsto \mathcal{J}_{\mathcal{Q} \cup \{i^+\}}(\lambda)$ et $\lambda \mapsto \mathcal{J}_{\mathcal{Q} \setminus \{i^-\}}(\lambda)$ intersectent la courbe ℓ_0 , permettant ainsi de produire une courbe ℓ_0 située en dessous de la courbe courante. Lorsque la courbe ℓ_0 est mise à jour, les nouvelles valeurs λ_j se déduisent des calculs d'intersection entre les droites liées aux nouveaux supports et la courbe ℓ_0 précédente. Les différents cas de figure sont résumés par la figure 1.8.

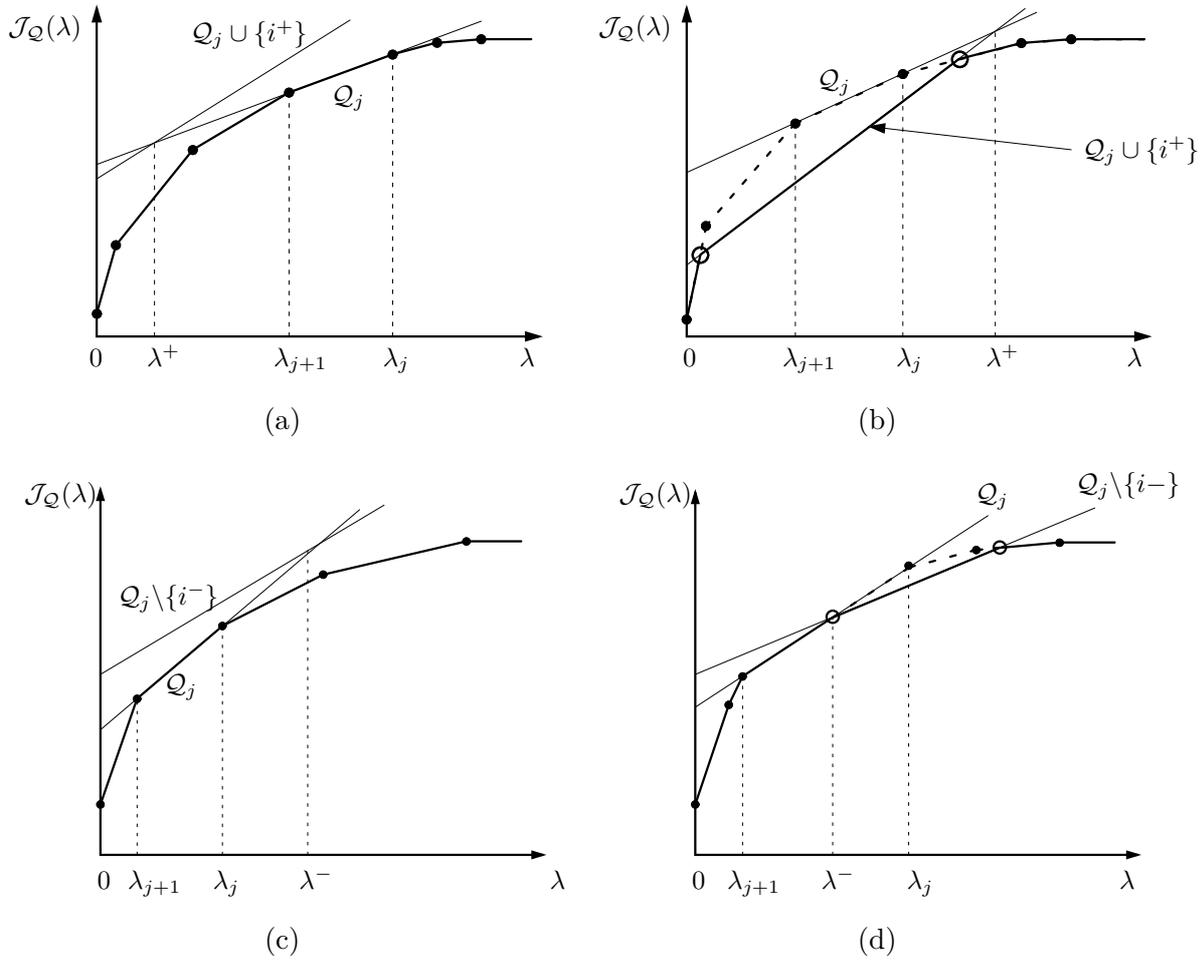


FIGURE 1.8 – Algorithme ℓ_0 -PT : tests de remplacements élémentaires. Les droites associées à \mathcal{Q}_j et $\mathcal{Q}_j \cup \{i^+\}$ s'intersectent pour $\lambda^+ = \mathcal{E}_{\mathcal{Q}_j} - \mathcal{E}_{\mathcal{Q}_j \cup \{i^+\}}$. De façon similaire, \mathcal{Q}_j et $\mathcal{Q}_j \setminus \{i^-\}$ s'intersectent pour $\lambda^- = \mathcal{E}_{\mathcal{Q}_j \setminus \{i^-\}} - \mathcal{E}_{\mathcal{Q}_j}$. (a,b) Tentative d'insertion de $\mathcal{Q}_j \cup \{i^+\}$ dans le chemin de régularisation ℓ_2 - ℓ_0 pour améliorer la courbe ℓ_0 . Quand $\lambda^+ \leq \lambda_{j+1}$, la nouvelle droite $\mathcal{J}_{\mathcal{Q}_j \cup \{i^+\}}(\lambda)$ se trouve au dessus de la courbe ℓ_0 et aucune amélioration n'est possible (a). Quand $\lambda^+ > \lambda_{j+1}$, $\mathcal{J}_{\mathcal{Q}_j \cup \{i^+\}}(\lambda)$ intersecte la courbe ℓ_0 . La courbe ℓ_0 est mise à jour en incluant $\mathcal{Q}_j \cup \{i^+\}$ comme nouvelle arête (b). L'ancienne et la nouvelle courbes ℓ_0 sont représentées en traits pointillés et en trait plein. (c,d) Tentative d'insertion du support $\mathcal{Q}_j \setminus \{i^-\}$. Quand $\lambda^- \geq \lambda_j$, aucune amélioration n'est possible (c). Quand $\lambda^- < \lambda_j$, la nouvelle droite $\mathcal{J}_{\mathcal{Q}_j \setminus \{i^-\}}(\lambda)$ intersecte la courbe ℓ_0 qui est mise à jour (d).

Précisons à présent la règle de sélection du support \mathcal{Q}_j à explorer. Les supports $\mathcal{Q}_0, \dots, \mathcal{Q}_J$ qui constituent le chemin de régularisation ont un statut « déjà exploré » ou « non encore exploré ». Pour un support \mathcal{Q}_j déjà exploré, tous les remplacements élémentaires associés $\mathcal{Q}_j \bullet i$ ont été testés selon la règle (1.22). A chaque itération, ℓ_0 -PT sélectionne le support non encore exploré \mathcal{Q}_j dont la valeur associée λ_j est la plus grande de telle sorte que $\mathcal{Q}_0, \dots, \mathcal{Q}_{j-1}$ ont tous déjà été explorés.

L'algorithme ℓ_0 -PT est résumé dans le tableau 1.7. A titre de comparaison, notons que pour CSBR, la construction progressive du chemin de régularisation repose sur le calcul de la borne

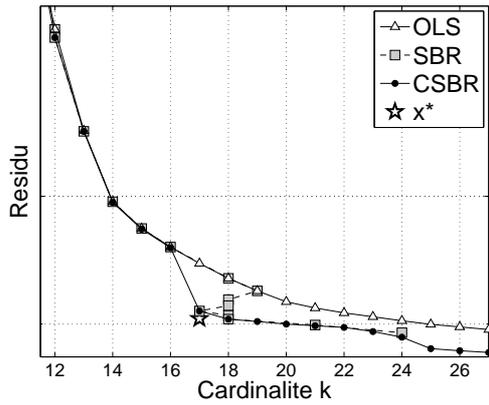
Entrées : \mathbf{A} , \mathbf{y} , ε et/ou K
Etape 0 : Initialisation : $J = 0$, $\mathcal{Q}_0 = \emptyset$, $\lambda_0 = +\infty$, $\lambda_1 = 0$. Faire exploré (0) = 0 et $j = 0$.
Etape 1 : [Exploration de \mathcal{Q}_j] Calculer le meilleur ajout i^+ selon (1.22) et faire $\lambda^+ = \mathcal{E}_{\mathcal{Q}_j} - \mathcal{E}_{\mathcal{Q}_j \cup \{i^+\}}$. Si $\text{Card}[\mathcal{Q}_j] > 2$, Calculer le meilleur retrait i^- par (1.22) et faire $\lambda^- = \mathcal{E}_{\mathcal{Q}_j \setminus \{i^-\}} - \mathcal{E}_{\mathcal{Q}_j}$. Fin Si. Faire exploré (j) = 1. Si ($\lambda^+ > \lambda_{j+1}$), Ajouter $\mathcal{Q}_j \cup \{i^+\}$ au chemin de régularisation avec le statut exploré à 0. Mettre à jour le chemin de régularisation et la courbe ℓ_0 . Fin Si If ($\lambda^- < \lambda_j$), Tester si $\mathcal{J}_{\mathcal{Q}_j \setminus \{i^-\}}(\lambda)$ intersecte la courbe ℓ_0 . S'il y a des intersections, Ajouter $\mathcal{Q}_j \setminus \{i^-\}$ au chemin de régularisation avec le statut exploré à 0. Mettre à jour le chemin de régularisation et la courbe ℓ_0 . Fin Si Fin Si [Recherche de la première arête non-explorée de la courbe ℓ_0] Sélectionner le plus faible indice j tel que exploré (j) = 0. Si ($\text{Card}[\mathcal{Q}_j] \geq K$) ou ($\mathcal{E}_{\mathcal{Q}_j} < \varepsilon$), Terminer ℓ_0 -PT. Sinon, Retourner à l'étape 1. Fin Si
Sorties : • Chemin de régularisation ℓ_0 estimé $\{\mathcal{Q}_j, j = 0, \dots, J\}$. • Liste des valeurs critiques λ_j et des erreurs d'approximation $\mathcal{E}_{\mathcal{Q}_j}$. • Meilleurs itérés $\mathcal{Q}_k^{\text{PT}}$ et erreurs quadratiques associées $\mathcal{E}_k^{\text{PT}}$ ($k = 0, \dots, K$)

TABLE 1.7 – Algorithme ℓ_0 -PT. Les paramètres d'arrêt ε et/ou K portent respectivement sur la norme ℓ_2 du résidu de l'approximation et sur la cardinalité. En sortie, on renvoie en plus du chemin de régularisation (liste des supports \mathcal{Q}_j et des valeurs critiques λ_j), les meilleurs supports $\mathcal{Q}_k^{\text{PT}}$ pour toute cardinalité $k = 0, \dots, K$, où $\mathcal{Q}_k^{\text{PT}}$ est le support exploré de cardinalité k de plus faible erreur quadratique, notée $\mathcal{E}_k^{\text{PT}}$.

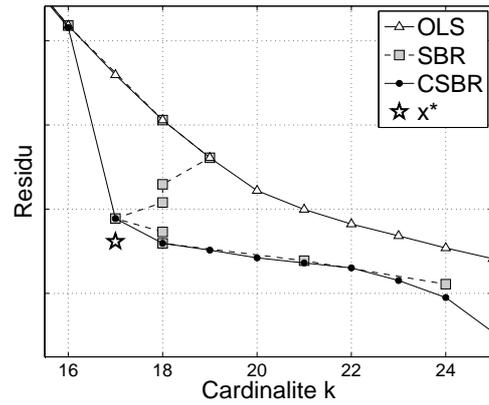
inférieure λ^+ de (1.16). Le fonctionnement de ℓ_0 -PT exploite à la fois les bornes inférieure (λ^+) et supérieure (λ^-) de (1.16). Cela permet de raffiner le chemin de régularisation non seulement pour des niveaux de parcimonie λ inférieurs au niveau courant (solutions moins parcimonieuses) mais aussi d'améliorer les solutions très parcimonieuses déjà disponibles (valeurs supérieures de λ). ℓ_0 -PT est potentiellement plus efficace que CSBR mais sa complexité est plus importante en terme de nombre d'itérations (pour K et/ou ε donnés) et de structure algorithmique.

1.7.5 Evaluation numérique des algorithmes proposés

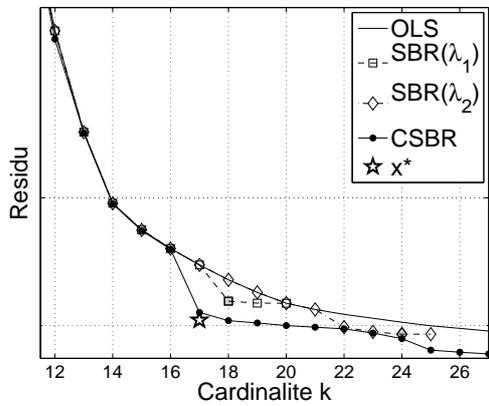
Tout au long de cette section, les données traitées sont celles de la figure 1.1(a) pour le problème de déconvolution impulsionnelle $\mathbf{y} = \mathbf{h} * \mathbf{x}^* + \mathbf{b}$ avec un filtre gaussien.



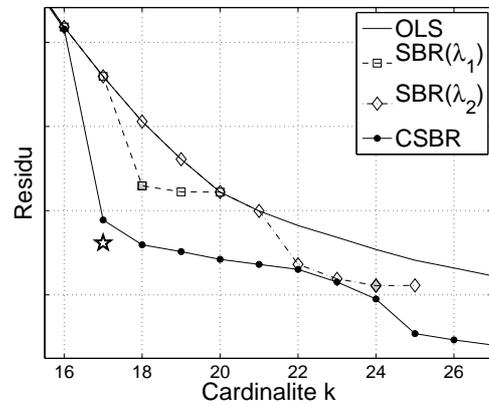
(a) OLS vs sorties de SBR(λ) vs CSBR



(b) Zoom de (a)



(c) CSBR vs itérés de SBR(λ)



(d) Zoom de (c)

FIGURE 1.9 – Comparaison des algorithmes OLS, SBR et CSBR pour le problème de déconvolution impulsionnelle $\mathbf{y} = \mathbf{h} * \mathbf{x}^* + \mathbf{b}$ présenté dans la figure 1.1(a) page 47. Les résultats sont représentés dans le plan $(\text{Card}[\mathcal{Q}], \mathcal{E}_{\mathcal{Q}})$. L'étoile représente la solution inconnue : $\|\mathbf{x}^*\|_0 = 17$ et $\|\mathbf{h} * \mathbf{x}^*\|_2^2 = \|\mathbf{b}\|_2^2$. (a,b) OLS et CSBR sont exécutés une fois avec les conditions d'arrêt $K = 25$ et $\varepsilon = 0$. SBR est exécuté pour toutes les valeurs λ_j obtenues en sortie de CSBR (chemin de régularisation estimé). On représente les supports \mathcal{Q}_k , $k = 1, \dots, K$ fournis par OLS et CSBR, et les supports \mathcal{Q}_j obtenus en sortie de SBR. (c,d) Même exécution d'OLS et de CSBR. SBR est exécuté pour deux valeurs de λ seulement, notées λ_1 et λ_2 avec $\lambda_1 > \lambda_2$. A chaque fois, tous les itérés de SBR sont représentés (carrés/losanges non coloriés). Les supports obtenus en sortie de SBR($\emptyset; \lambda_1$) et SBR($\emptyset; \lambda_2$) ont pour cardinalités respectives 18 et 24 (coloriage en gris).

Comparaison CSBR vs SBR

Pour illustrer l'intérêt de l'approche par continuation, commençons par comparer les performances de SBR et CSBR. La figure 1.9(a,b) compare les supports obtenus pour chaque cardinalité k avec OLS et CSBR, et pour les sorties \mathcal{Q}_j de SBR exécuté pour des valeurs $\lambda = \lambda_j$ décroissantes. Les résultats sont représentés en terme de compromis erreur d'approximation vs cardinalité dans le plan $(\text{Card}[\mathcal{Q}], \mathcal{E}_{\mathcal{Q}})$. On constate que la courbe liée à CSBR se situe en dessous de celles obtenues avec OLS et SBR (ces trois courbes sont confondues pour $\text{Card}[\mathcal{Q}] \leq 11$). De plus, la courbe de SBR présente des irrégularités qui traduisent une grande sensibilité à des

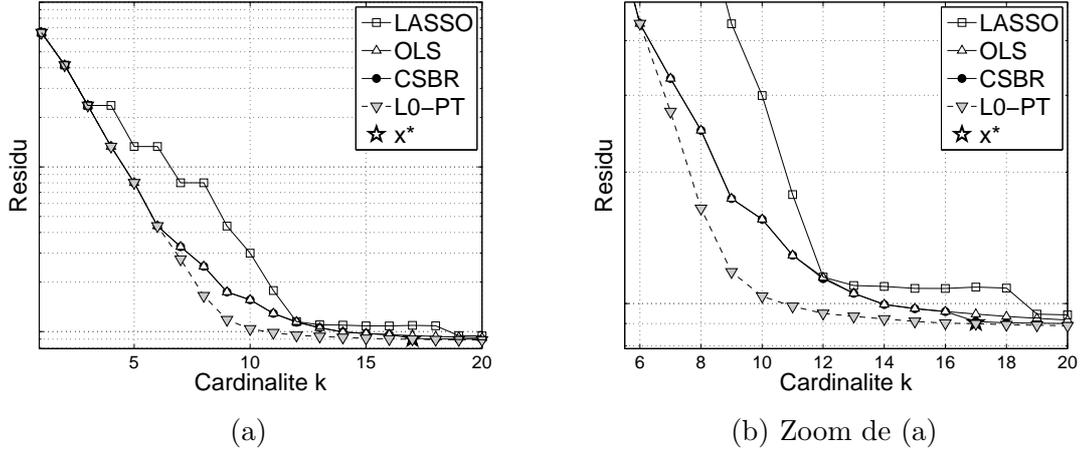


FIGURE 1.10 – Comparaison de OLS, CSBR, et ℓ_0 -PT pour le problème de continuation $\ell_2\text{-}\ell_0$ et de l'algorithme d'homotopie pour le problème de continuation $\ell_2\text{-}\ell_1$, aussi appelé LASSO. La comparaison est effectuée pour le problème de déconvolution impulsionnelle $\mathbf{y} = \mathbf{h} * \mathbf{x}^* + \mathbf{b}$ présenté dans la figure 1.1(a).

petites variations de λ pour des valeurs critiques de λ . En particulier, la cardinalité du support obtenu avec SBR ne croît pas systématiquement lorsque λ décroît : elle vaut successivement 19, puis 18, 17, 18 et 21 (figure 1.9(b)).

L'intérêt réel de CSBR est qu'il fournit *en une seule exécution* et sans réglage délicat de paramètres (hormis les critères d'arrêt habituels ε et/ou K) un ensemble de solutions $(\mathcal{Q}_k^{\text{CSBR}}, \mathcal{E}_k^{\text{CSBR}})$ pour chaque cardinalité k alors que SBR retourne un seul support associé à une valeur de λ . On peut bien sûr envisager de sauvegarder les meilleurs supports obtenus en tant qu'itérés de $\text{SBR}(\emptyset; \lambda)$ à λ fixé. La figure 1.1(c,d) montre que cette stratégie est inefficace : l'exécution de SBR pour deux valeurs $\lambda_1 > \lambda_2$ fournit des meilleurs itérés dont l'erreur d'approximation est parfois largement supérieure à celle des solutions de CSBR de même cardinalité. D'où l'intérêt de CSBR.

Comparaison des algorithmes de continuation

On compare quatre stratégies qui reconstruisent des supports pour toutes les cardinalités k : les algorithmes OLS, CSBR et ℓ_0 -PT pour la continuation $\ell_2\text{-}\ell_0$ et l'algorithme d'homotopie qui résout le problème de continuation $\ell_2\text{-}\ell_1$. Contrairement à la norme ℓ_0 , la régularisation ℓ_1 impose une pénalisation sur les amplitudes $|x_i|$ en plus d'une contrainte de parcimonie (*cf.* formulation (1.14)). Les erreurs quadratiques $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ des solutions obtenues par homotopie ne coïncident donc pas avec la meilleure approximation possible pour le support $\mathcal{Q} = \mathcal{S}(\mathbf{x})$, pour laquelle l'erreur d'approximation vaut $\mathcal{E}_{\mathcal{Q}} = \|\mathbf{P}_{\mathcal{Q}}^\perp \mathbf{y}\|_2^2$. Pour rendre pertinente la comparaison avec les algorithmes dédiés à la régularisation ℓ_0 , un post-traitement (débiaisage) s'impose. Il consiste à calculer la projection orthogonale $\mathbf{P}_{\mathcal{Q}}^\perp \mathbf{y}$ puis l'erreur $\mathcal{E}_{\mathcal{Q}} = \|\mathbf{P}_{\mathcal{Q}}^\perp \mathbf{y}\|_2^2$ pour tout support \mathcal{Q} obtenu par l'algorithme d'homotopie.

La figure 1.10(a) montre que même avec ce débiaisage, les solutions fournies par l'algorithme d'homotopie sont bien moins précises que celles d'OLS, CSBR et surtout ℓ_0 -PT. ℓ_0 -PT améliore sensiblement les performances de CSBR au prix d'un coût de calcul accru : avec les paramètres d'arrêt $K = 34$ et $\varepsilon = 0$, 34 itérations d'OLS ont été réalisées contre 58 itérations (nombre de remplacements élémentaires depuis le support vide) pour CSBR et 113 itérations

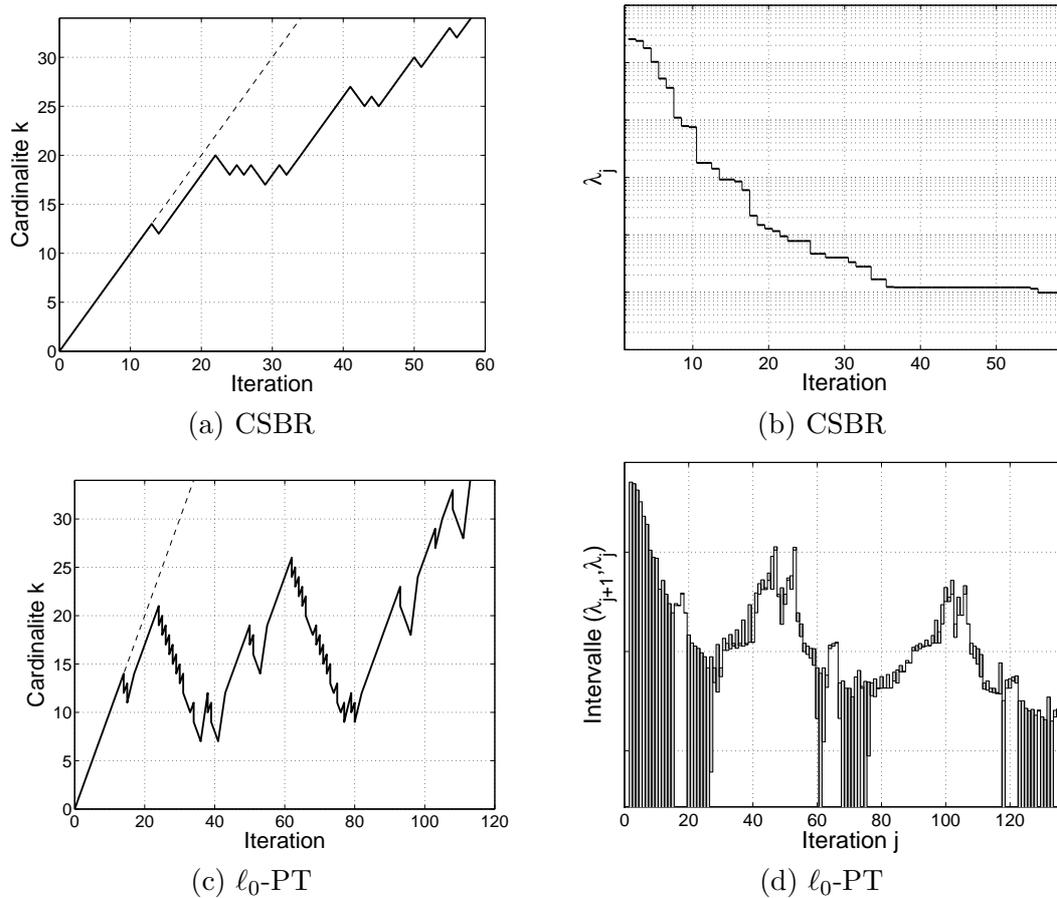


FIGURE 1.11 – Fonctionnements empiriques de CSBR et ℓ_0 -PT pour le problème test de la déconvolution impulsionnelle. (a) Cardinalité du support courant après chaque remplacement élémentaire durant l’exécution de CSBR. (b) Liste des valeurs λ_j explorées durant l’exécution de CSBR. Pour chaque valeur de λ_j , l’exécution de SBR produit au moins un remplacement élémentaire. Le nombre de remplacements élémentaires correspond à la taille des paliers horizontaux sur la figure. (c) Cardinalité des supports intégrés dans le chemin de régularisation au cours des itérations de ℓ_0 -PT. A chaque itération, 0, 1 ou 2 supports sont intégrés, d’où certains paliers verticaux dans la figure lorsque deux supports sont simultanément intégrés. (d) Représentation du « degré de parcimonie » $(\lambda_{j+1}, \lambda_j)$ lié au support \mathcal{Q}_j manipulé lors de l’itération courante de ℓ_0 -PT (*cf.* fig. 1.8 pour les notations λ_j et \mathcal{Q}_j). La partie grisée matérialise l’intervalle $(\lambda_{j+1}, \lambda_j)$. Lorsqu’elle atteint le bas de l’image, la borne inférieure vaut $\lambda_{j+1} = 0$.

pour ℓ_0 -PT. Plus précisément, en ce qui concerne ℓ_0 -PT, 8% des itérations effectuées ont été infructueuses : aucun remplacement élémentaire n’améliore le chemin de régularisation courant qui reste inchangé. Pour 62% des itérations, un seul nouveau support est intégré au chemin de régularisation courant. Enfin, deux nouveaux supports y sont intégrés pour 30% des itérations. En considérant que le coût par itération de ℓ_0 -PT est quasi-identique à celui de CSBR¹⁶, le prix à payer pour la meilleure performance de ℓ_0 -PT est donc un temps de calcul double.

La figure 1.11 fournit un éclairage intéressant sur les fonctionnements de ℓ_0 -PT et CSBR. Au cours des 25 premières itérations, ℓ_0 -PT effectue majoritairement des ajouts d’atomes à la

16. à la mise à jour de la courbe ℓ_0 près.

manière d'OLS. Les supports explorés sont donc de cardinalité croissante (figure 1.11(c)) et le niveau de parcimonie λ décroît (sous-figure (d)). Entre les itérations 25 et 40, l'algorithme raffine des solutions très parcimonieuses obtenues précédemment pour $k = 20, 19, \dots, 7$ en effectuant une série de dé-sélection d'atomes. Ces solutions très parcimonieuses sont à nouveau améliorées autour de l'itération 80. En comparaison, CSBR explore des supports de cardinalité « globalement » croissante (même si des dé-sélections sont effectuées) et les solutions très parcimonieuses ne sont pas améliorées au cours des itérations, car CSBR travaille avec des valeurs décroissantes de λ (figure 1.11(a,b)).

Pour tous les résultats présentés dans les figures 1.9, 1.10 et 1.11, je me suis borné à faire une comparaison en termes d'erreur quadratique pour des raisons de simplicité. Des résultats plus complets intégreraient les performances de localisation des impulsions du signal \mathbf{x}^* en termes de nombre de bonnes détections et de fausses alarmes. Ces tests de localisation sont cependant délicats pour deux raisons. D'abord, il doivent être répétés pour chaque cardinalité possible et chaque impulsion. De plus, une fausse alarme n'est pas toujours représentative du comportement réel d'un algorithme : une erreur de localisation est comptée comme une fausse alarme si la position d'une impulsion n'est pas correctement trouvée. Une fausse alarme peut donc se produire quelquesoit la distance entre la « vraie » impulsion et son estimation. Il est souvent plus pertinent d'utiliser des tests de bonne localisation non-binaires qui tiennent compte de la distance entre les vraies impulsions et leur estimation erronée [ROS 01, CAR 13].

Estimation de l'ordre du modèle

Les algorithmes de continuation proposés sont tout naturellement compatibles avec les méthodes classiques de sélection d'ordre de modèle [WAN 04]. En effet, CSBR et ℓ_0 -PT fournissent des solutions (sous-optimales) $\mathcal{Q}_k^{\text{CSBR}}$ et $\mathcal{Q}_k^{\text{PT}}$ à tous les ordres $k = 0, \dots, K$. Parmi le grand nombre de méthodes disponibles pour estimer k , nous avons testé deux catégories de critères :

1. ceux qui reposent sur la minimisation d'un critère composite de type :

$$\arg \min_k \{m \log(\mathcal{E}_k) + \alpha k\} \quad (1.23)$$

où m est le nombre de mesures (la taille de \mathbf{y}) et la constante α vaut respectivement 2, $\log m$ et $2 \log \log m$ pour les critères d'Akaike (AIC) [AKA 74], du *Minimum Description Length* (MDL) [RIS 83] et de Hannan et Quinn [HAN 79].

2. les critères de type validation croisée [GOL 79] dont le principe est de partitionner les données \mathbf{y} en deux sous-ensembles, *e.g.*, en mettant de côté une observation y_i ce qui conduit à former un signal observé réduit $\mathbf{y}^{[i]}$ de taille $m-1$. La validation croisée consiste à tester si l'estimation $\mathbf{x}(\mathbf{y}^{[i]}, k)$ de \mathbf{x} à partir des données réduites en utilisant le modèle d'ordre k conduit à une bonne prédiction de l'observation y_i [WAH 77]. Formellement, cela conduit à résoudre le problème :

$$\arg \min_k \left\{ \frac{1}{m} \sum_i \left(y_i - \hat{y}_i^{[i]}(k) \right)^2 \right\} \quad (1.24)$$

où $\hat{y}_i^{[i]}(k)$ est la i -ème entrée de $\mathbf{Ax}(\mathbf{y}^{[i]}, k)$.

Le cas particulier des algorithmes d'approximation parcimonieuse conduit à des expressions simplifiées du critère de validation croisée et de sa version généralisée. Le lecteur intéressé par plus de précisions est redirigé vers les livres [MIL 02, Chap. 5] et [WAN 04].

Pour le problème de déconvolution impulsionnelle étudié, la tendance générale est que les critères d'Akaike et de type validation croisée surestiment très largement le nombre d'impulsions attendues (elles estiment entre 45 et 50 impulsions pour CSBR et ℓ_0 -PT alors que le signal inconnu \mathbf{x}^* n'en comporte que 17), de même que leurs versions généralisées. Le critère MDL fournit les résultats les plus réalistes (25 et 20 impulsions respectivement pour CSBR et ℓ_0 -PT). De plus, il sous-estime le nombre d'impulsions pour des niveaux de bruit plus importants : par exemple, 8 impulsions sont trouvés à la fois pour CSBR et ℓ_0 -PT pour un rapport signal-sur-bruit de 0 dB. Ce comportement est souhaitable car à fort niveau de bruit, les petites impulsions sont noyées dans le bruit ce qui justifie que les algorithmes de déconvolution impulsionnelle ne peuvent les détecter.

1.8 Conclusion

Nous nous sommes intéressés à la parcimonie pour des problèmes inverses induisant des dictionnaires dont les (ou certaines) colonnes sont très corrélées [SOU 11, BOU 11b], et des dictionnaires de grande taille pour permettre une bonne approximation du signal à reconstruire avec peu d'atomes [BOU 11a, SAH 12b]. Le choix d'un algorithme est un compromis entre sa performance et son coût calculatoire. Les algorithmes proposés dans ce chapitre en tant qu'extension d'OLS sont relativement coûteux numériquement mais ils font partie des plus performants (avec l'algorithme $\text{IR}\ell_1$). Une explication est que les algorithmes gloutons bidirectionnels basés OLS ont la capacité de « s'échapper » de minimiseurs locaux du critère $\mathcal{J}(\mathbf{x}; \lambda) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$ pour un niveau de parcimonie λ fixé. En fait, chaque itéré de l'algorithme SBR est un minimiseur local du critère \mathcal{J} . Cette caractéristique est remarquable car d'autres algorithmes concurrents n'ont pas cette capacité à s'échapper d'un minimiseur local de \mathcal{J} . Dans la littérature, cet argument est souvent mis en avant pour proposer des algorithmes d'optimisation qui ne sont pas dédiés au critère \mathcal{J} mais à un critère approximant \mathcal{J} et comportant moins de minimiseurs locaux. L'efficacité des algorithmes bidirectionnels basés OLS pour minimiser \mathcal{J} repose sur leur très faible nombre d'itérations avant convergence malgré un coût par itération important. En particulier, SBR possède une structure très simple et aucun réglage *ad hoc* n'est nécessaire pour l'arrêt de l'algorithme. Son efficacité est reconnue par d'autres chercheurs de la communauté (voir [SEL 13] pour une comparaison de 15 algorithmes où SBR fait partie des tout meilleurs) et les algorithmes dérivés de SBR (CSBR et ℓ_0 -PT) permettent encore d'améliorer ses performances tout en fournissant des solutions pour un continuum de niveaux de parcimonie λ . Ces algorithmes de continuation permettent en outre de mettre très facilement en œuvre les méthodes classiques de sélection d'ordre, comme le *Minimum Description Length*, pour estimer le niveau de parcimonie de façon adaptative. Les implémentations rapides des algorithmes proposés sont possibles y compris pour des problèmes de grande taille, et en particulier pour des dictionnaires structurés (*e.g.*, pour les problèmes convolutifs) comme nous le verrons au chapitre 2 (§ 2.3.5). Dans le chapitre 4, je dresserai quelques perspectives visant à proposer des algorithmes plus rapides que SBR et potentiellement plus performants pour des problèmes inverses spécifiques.

Chapitre 2

Analyse d'images force-volume en microscopie AFM

Ce projet a pour cadre une collaboration entre le CRAN et le LCPME. Il a pour objectif le développement d'outils originaux de traitement des images force-volume en microscopie de force atomique qui intègrent le savoir-faire des laboratoires, à savoir la spectroscopie moléculaire, et plus particulièrement la microscopie de force atomique pour le LCPME (G. Francius), le traitement du signal et de l'image (C. Soussen, D. Brie) pour le CRAN. Ce chapitre est une adaptation du chapitre d'ouvrage [SOU 13a].

La microscopie de force atomique (AFM pour *Atomic Force Microscopy*) est une technologie récente permettant de caractériser les propriétés physico-chimiques d'objets et en particulier d'échantillons biologiques à l'échelle nanométrique. Nous nous sommes focalisés sur la modalité dite *force-volume* qui consiste à mesurer des courbes de force sur une grille spatiale, chaque courbe de force représentant la force d'interaction entre une pointe AFM et la surface du nano-objet en fonction de leur distance. Une courbe de force est composée de deux signaux relatifs à l'approche et au retrait de la pointe. Les méthodes disponibles pour interpréter ces données tridimensionnelles sont essentiellement empiriques et requièrent de nombreux réglages manuels. Nous proposons une approche automatique pour reconstruire un ensemble d'images 2D représentant chacune une propriété physico-chimique du nano-objet (topographie, élasticité, *etc.*). Ce traitement repose sur les modèles physiques disponibles dans les phases d'approche et de retrait. Pour le type d'échantillons étudiés (bactéries), nous utilisons les modèles d'interaction électrostatique et mécanique dans la phase d'approche et le modèle *Freely Jointed Chain* (FJC) qui décrit l'étirement de bio-polymères dans la phase de retrait.

L'algorithme consiste à lisser par morceaux chaque courbe de force de façon à déterminer les régions d'intérêt dans lesquelles les modèles physiques s'appliquent. Le lissage par morceaux est reformulé comme un problème d'approximation parcimonieuse utilisant un dictionnaire polynomial [SOU 11]. Les paramètres d'intérêt sont finalement estimés par ajustement des modèles physiques par moindres carrés dans chaque région d'intérêt. Une cartographie (image 2D) de chacun des paramètres physiques est obtenue en répétant l'analyse des courbes de force pour chaque pixel. La méthode est appliquée, d'une part, à l'analyse des propriétés électrostatiques et mécaniques d'un mutant de la bactérie *Escherichia coli* et, d'autre part, à l'analyse d'exopolymères produits par la bactérie *Pseudomonas fluorescens* [POL 11].

Je commence par présenter le type de données que l'on peut acquérir en imagerie force-volume et les paramètres physiques que l'on cherche à extraire de ces données. Ces paramètres sont étroitement liés aux modèles physiques décrivant l'interaction de deux nano-objets avant et

après contact. La deuxième partie de ce chapitre décrit les problématiques de traitement. Ils sont basés sur la notion de parcimonie, au sens où les courbes de force contiennent quelques points caractéristiques (sauts, changements de pente ou de courbure) qu'il s'agit d'estimer très précisément. Les algorithmes d'approximation parcimonieuse choisis sont des méthodes de recherche de points caractéristiques sur une grille discrète. Ce sont des applications des algorithmes génériques OLS et SBR présentés dans le chapitre 1 pour l'approximation parcimonieuse. L'adaptation de ces algorithmes au cadre du lissage par morceaux de signaux est simple sur le principe, mais pose néanmoins des difficultés techniques liées à la structure du dictionnaire considéré. La troisième partie du chapitre rappelle très brièvement le principe de ces algorithmes et décrit les adaptations techniques nécessaires. Enfin, la dernière partie du chapitre illustre la méthode par quelques résultats types obtenus pour des données réelles.

2.1 Introduction à la microscopie de force atomique

2.1.1 Caractérisation de cellules biologiques

Le développement des techniques de microscopie à champ proche et en particulier de la microscopie AFM [BIN 86, BIN 88] permet de déterminer *in situ* des propriétés physico-chimiques locales (électriques, magnétiques, vibrationnelles, mécaniques) [WIE 94]. Ces techniques fournissent des mesures spectroscopiques à l'échelle nanométrique indépendamment de la nature des échantillons (biologique, organique, minérale) et des images tridimensionnelles, dites *images force-volume* en microscopie AFM [BUT 05].

La caractérisation des propriétés physico-chimiques des objets biologiques, en particulier les micro-organismes (bactéries, champignons, *etc.*) et les cellules animales, est un enjeu majeur dans des domaines aussi divers que la biologie, la microbiologie, l'industrie pharmaceutique et la médecine clinique. Par exemple, la détermination de la charge électrostatique et de l'élasticité des bactéries est une question fondamentale pour comprendre les mécanismes d'adhésion de bactéries et les processus d'infection. La microscopie AFM est un outil puissant car les mesures peuvent être réalisées en milieu aqueux ou dans des conditions physiologiques avec une résolution spatiale sub-nanométrique. La spectroscopie de force AFM rend possible l'observation et la quantification des propriétés mécaniques d'échantillons biologiques déformables et la mesure des interactions intramoléculaires entre bio-molécules. Etant donné que certaines tumeurs cancéreuses et cellules souches sont régulées par des propriétés mécaniques, des travaux en nanomédecine ont montré que certaines maladies peuvent être diagnostiquées par microscopie AFM [KAS 08, CRO 09]. Une revue de littérature sur le potentiel de l'AFM en cancérologie est disponible dans l'annexe A.

2.1.2 Fonctionnement d'un microscope AFM

Le fonctionnement d'un microscope AFM est basé sur la détection des forces inter-atomiques (capillaires, électrostatiques, Van der Waals, frictions) s'exerçant entre une pointe associée à un levier de constante de raideur fixe et la surface d'un échantillon [BUT 05] (*cf.* fig. 2.1). On distingue généralement deux modes d'acquisition de données.

Images isoforce et isodistance

Les données résultent du balayage par la pointe de la surface de l'échantillon selon deux modes opératoires :

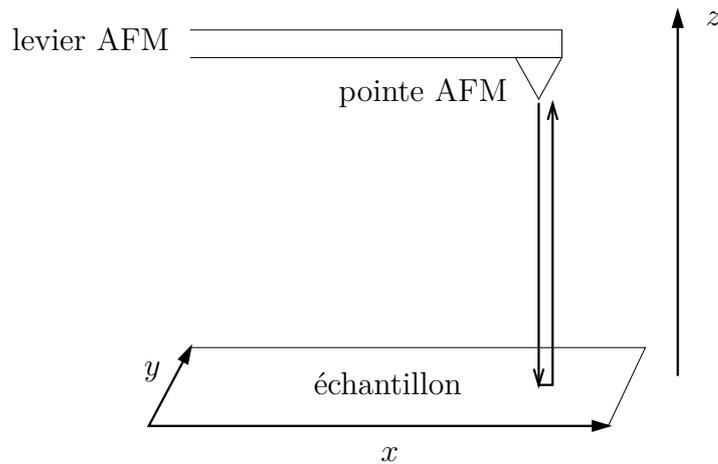


FIGURE 2.1 – Spectroscopie de force. En un point de la surface de l'échantillon, la pointe nanométrique s'approche jusqu'au contact déformable (l'échantillon est mécaniquement comprimé par l'avancée de la pointe) puis non déformable (à un certain stade, la pointe ne peut plus déformer davantage l'échantillon). Puis la pointe se retire progressivement et perd finalement le contact. Dans chaque phase, un signal $f(z)$ est mesuré. Ces signaux sont appelés courbes d'approche et de retrait.

- le mode contact : un système d'asservissement permet de maintenir constante la force exercée sur le levier supportant la pointe au cours du balayage. Cette force est proportionnelle à la déflexion du levier qui est la grandeur mesurée. Ce mode permet d'obtenir directement la topographie, c'est-à-dire le relief des surfaces ;
- le mode intermittent (ou vibrant) : une variation des forces d'interaction se traduit par une variation de la fréquence de vibration du levier qui est la grandeur mesurée. Par inversion de la relation liant la force à la fréquence, on obtient une image de la force. Dans ce cas, le système d'asservissement permet de maintenir constante soit la distance entre la pointe et l'échantillon, soit la force agissant sur le levier. Ce mode permet donc d'obtenir des images isodistance ou isoforce.

Spectroscopie de force

Contrairement au mode d'acquisition précédent, la spectroscopie de force est une analyse *ponctuelle* de l'échantillon, obtenue en enregistrant la déflexion du levier en fonction de la distance z de séparation entre la pointe AFM et la surface de l'échantillon (fig. 2.1). Une courbe de force $f(z)$ montre donc l'évolution de la force en fonction de la distance z en un point de l'échantillon.

La forme générale d'une courbe de force est schématisée sur la figure 2.2. L'intensité de la force est calculée à partir de la mesure de la déflexion du levier. Elle est représentée en fonction du déplacement de la pointe z , où les plus grandes valeurs de la distance z correspondent aux positions de la pointe les plus éloignées de l'échantillon. Dans une courbe de force, on distingue deux signaux correspondant à l'approche et au retrait de la pointe (en trait continu et en traits pointillés respectivement). Dans ce qui suit, nous décrivons des zones particulières sur ces courbes pendant les phases d'approche (zones A-C) et de retrait (zones D-E).

- *Zone A.* Aucune interaction n'est observée lorsque la pointe est placée à une distance suffisamment grande de l'échantillon ;

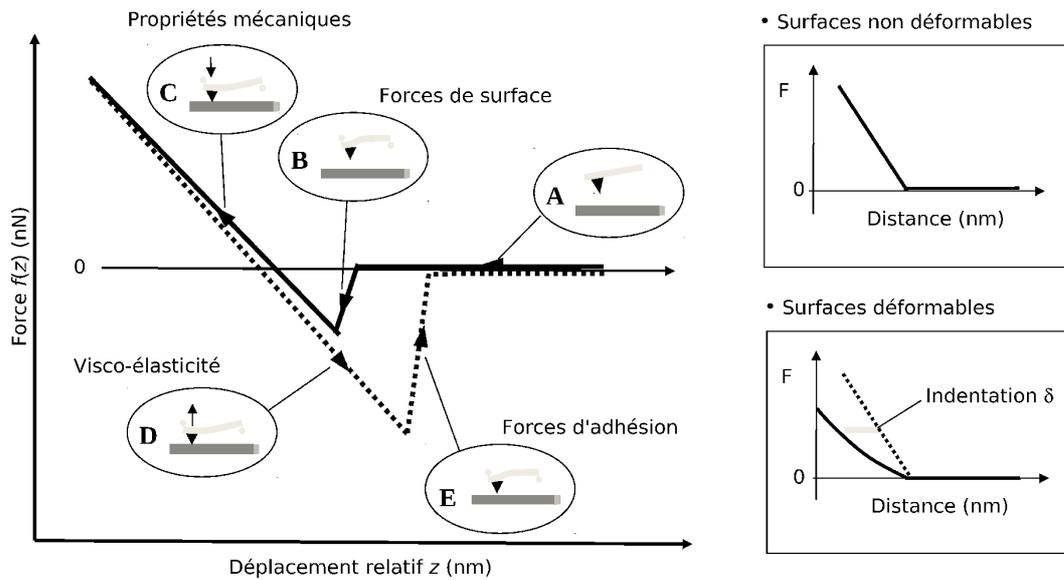


FIGURE 2.2 – Forme générale d'une courbe de force : courbes d'approche (en trait continu) et de retrait (en pointillés). Les valeurs négatives (positives) représentent des forces d'attraction (répulsion). L'indentation $\delta(z)$ est la profondeur de déformation de l'échantillon compressé par la pointe. Pour un échantillon non déformable, $\delta = 0$. Pour un échantillon déformable, δ augmente progressivement jusqu'à atteindre une valeur limite correspondant à la fin du régime déformable et au début du régime non déformable (figure adaptée de [GAB 07]).

- Zone B. Elle correspond aux forces de surface (électrostatiques, Van der Waals). Ces interactions sont négatives (attraction entre la pointe et la surface) ou positives (répulsion). Dans cette zone, la pointe n'est pas en contact avec l'échantillon ;
- Entre les zones B et C se situe le point de contact entre la pointe et l'échantillon ;
- Zone C : interactions mécaniques du levier et/ou de l'échantillon. Pour un échantillon non déformable, le comportement observé est essentiellement dû à la déformation linéaire du levier. En revanche, pour un échantillon déformable, des processus de compression et/ou d'indentation de l'échantillon conduisent à des comportements linéaires ou non linéaires ;
- Zone D. La présence d'une hystérèse entre les courbes d'approche et de retrait traduit les propriétés viscoélastiques de l'échantillon. Pour les surfaces non déformables, cette hystérèse est nulle ;
- Zone E. Au cours du retrait, les courbes peuvent présenter une force de rappel importante en fonction de l'aire et du temps de contact mais aussi et surtout de l'énergie de surface échantillon-pointe. Pour des surfaces de micro-organismes, ce domaine présente de nombreuses ruptures qu'il convient d'analyser.

La figure 2.3(c,d) présente des courbes de force mesurées pour des bactéries *Escherichia coli* (approche) et *Pseudomonas fluorescens* (retrait) en milieu aqueux. *Escherichia coli* est une bactérie présente dans le corps humain dont certaines formes pathogènes sont responsables d'infections intestinales sévères. *Pseudomonas fluorescens* est une bactérie présente dans les réseaux de distribution d'eau. Elle est responsable de la formation de biofilms en produisant et excréant des exopolymères. Ces biofilms sont à l'origine de certaines maladies nosocomiales. Pour la courbe d'approche (c), le contact mécanique est déformable puis non déformable (parties convexe et affine sur la droite) alors que pour la courbe de retrait (d), on observe de nombreux

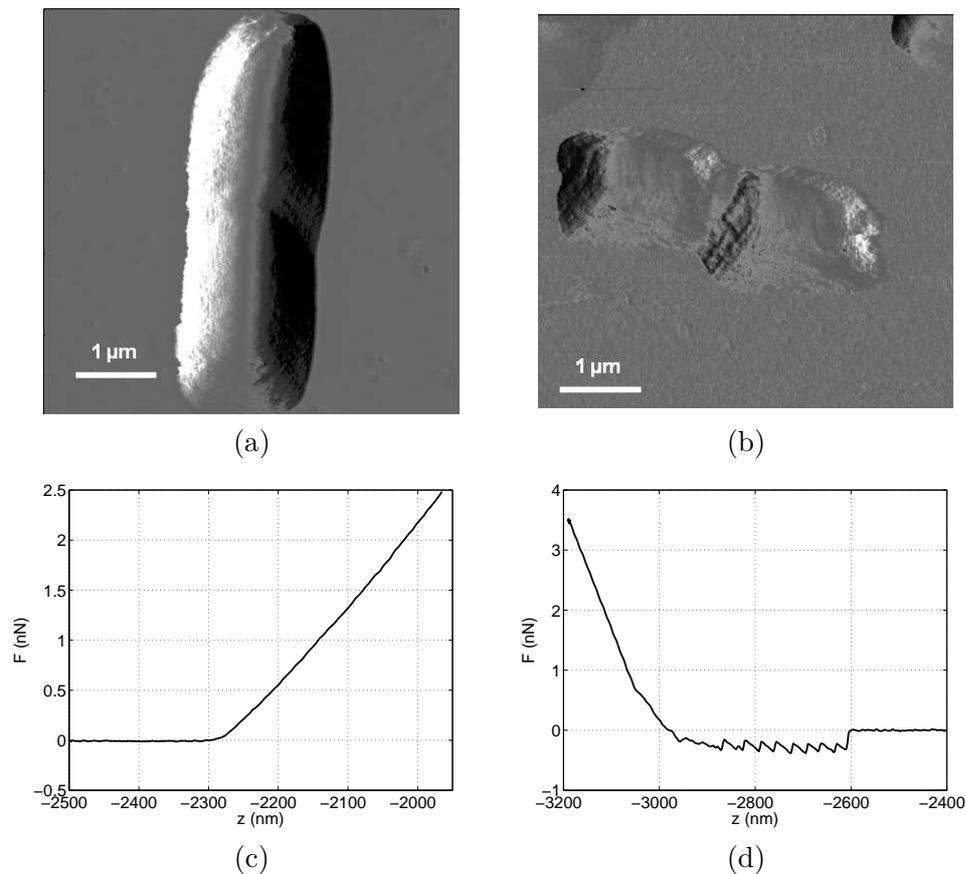


FIGURE 2.3 – Images d’un mutant de la bactérie *Escherichia coli* (a) et d’une bactérie *Pseudomonas fluorescens* (b). (c) Courbe d’approche relative à la bactérie *Escherichia coli*. L’axe z est inversé par rapport aux figures 2.1 et 2.2. (d) Courbe de retrait correspondant au déroulement d’un exo-polymère produit et excrété par la bactérie *Pseudomonas fluorescens*.

décrochages liés au déroulement progressif des exo-polymères situés sur la surface de la bactérie. Chaque zone décroissante est une zone de contact entre la pointe et le polymère dans laquelle il existe des modèles physiques décrivant l’étirement du polymère.

Imagerie force-volume

Reproduire l’analyse précédente en balayant la surface de l’échantillon conduit à former une image *force-volume* $f(x, y, z)$. Cette image résulte de la collection des courbes de force $f(z)$ sur une grille (x, y) représentant la surface de l’échantillon (fig. 2.4). La résolution en z est inférieure au nanomètre et la résolution latérale (x, y) est de l’ordre du nanomètre¹⁷. Un traitement typique (souvent réalisé empiriquement par les physiciens) consiste à estimer le point de contact pointe-échantillon pour une courbe de force donnée. Une image 2D représentant la topographie de l’échantillon est obtenue en répétant ce traitement pour chaque pixel (x, y) , mais cette analyse reste partielle car les données ont un contenu plus riche que la seule information de topographie.

17. Les images force-volume traitées dans ce chapitre contiennent 32×32 pixels. Notons l’essor de la technologie d’imagerie AFM ultra-rapide (*FastScan*) depuis 2011. Cette technologie permet d’enregistrer des images force-volume dites *PeakForce Mapping*, qui contiennent jusqu’à cent fois plus de courbes, acquises en un temps dix fois plus rapide pour une résolution et précision des mesures qui augmente dans la même proportion.

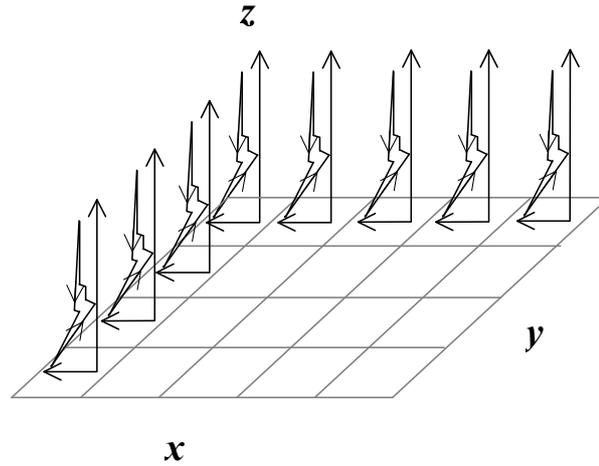


FIGURE 2.4 – Imagerie force-volume : acquisition d'un ensemble de courbes de force sur une grille (x, y) représentant la surface de l'échantillon.

Notre objectif est donc de concevoir des méthodes de traitement robustes pour caractériser la topographie et les autres propriétés physico-chimiques de l'échantillon.

2.1.3 Modèles physiques par morceaux

Dans la suite de ce chapitre, je me focalise sur la spectroscopie de force et l'imagerie force-volume. La description schématisée de la figure 2.2 est une modélisation par morceaux des courbes d'approche et de retrait, où un morceau correspond à une région pour laquelle un type d'interaction se produit. J'introduis maintenant les modèles physiques retenus pour chacun des morceaux. Ces modèles sont illustrés graphiquement par la figure 2.5. Le but de cette section n'est pas de dresser un panorama exhaustif des modèles existants mais de présenter quelques modèles physiques adaptés aux échantillons biologiques étudiés dans la section 2.4. Pour un panorama plus complet, voir les références [BUT 05, JAN 00, LIN 07a].

Modèles pour la phase d'approche

Pour la phase d'approche, les modèles physiques mettent en jeu les propriétés mécaniques d'échantillons déformables, à savoir les propriétés viscoélastiques de cellules vivantes comme l'élasticité de cellules humaines et de bactéries, et d'autres paramètres physico-chimiques comme la densité de charge et la pression de turgescence de bactéries recouvertes par des protéines spécifiques. Pour la bactérie *Escherichia coli*, on observe successivement une interaction électrostatique avant contact puis des interactions mécaniques après contact. On utilise un modèle exponentiel pour l'interaction électrostatique [VEL 02] :

$$f_{\text{ELEC}}(z; \kappa^{-1}, A, Z_1) = A \exp(\kappa(z - Z_1)) \quad (2.1)$$

pour $z \leq Z_1$, où Z_1 est le « point de contact virtuel » et κ^{-1} la longueur de Debye, la distance au delà de laquelle une séparation significative d'objets chargés peut avoir lieu. Notons que lorsque la pointe et l'échantillon ont une charge de même signe, le point de contact n'est pas bien défini car l'interaction est répulsive. Une définition alternative du point de contact est le premier échantillon du signal d'approche où l'interaction électrostatique est détectable dans les mesures (le point Z_0 de la figure 2.5(a)).

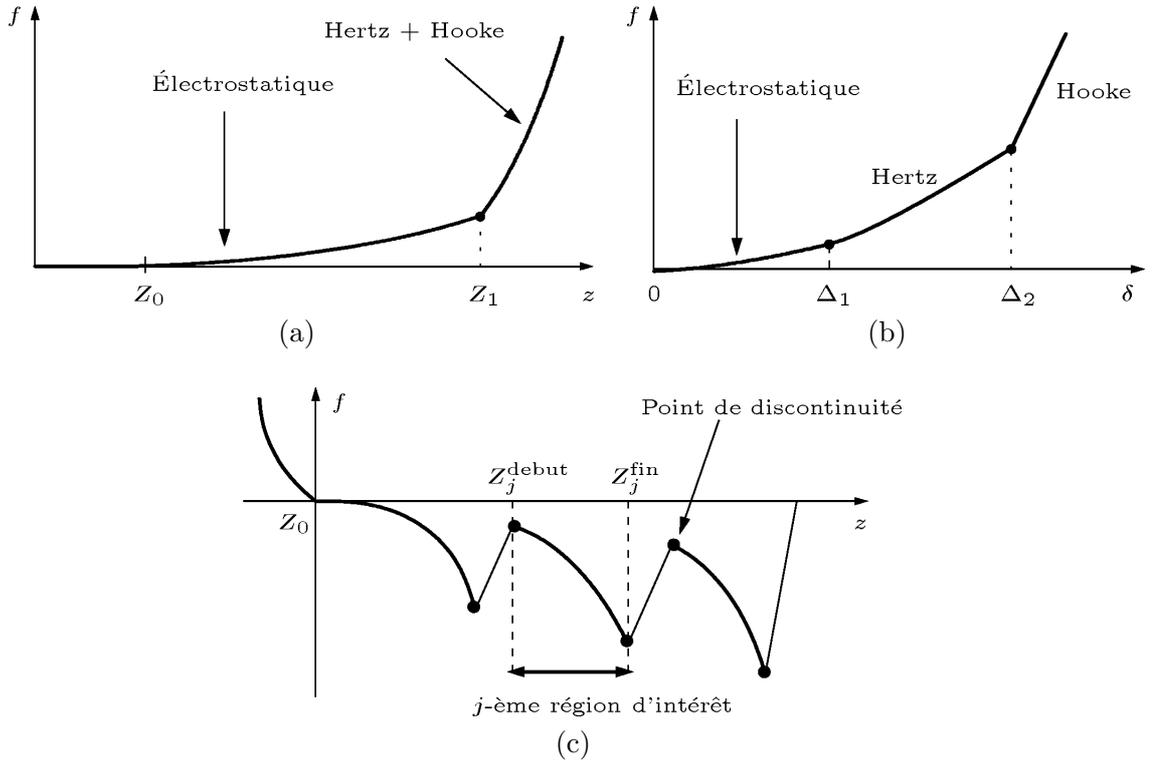


FIGURE 2.5 – Modèles par morceaux pour les courbes d’approche et de retrait (adapté de [POL 11]). Pour l’approche, on définit trois régions d’intérêt dans les deux domaines $z \mapsto f(z)$ (a) et $\delta \mapsto f(\delta)$ (b), où $\delta(z)$ est l’indentation définie par (2.2). Les points de transition $z = Z_0$ et $z = Z_1$ indiquent le début et la fin de l’interaction électrostatique. Z_1 indique aussi le début du contact déformable (loi de Hertz). Dans le domaine z , il est souvent difficile de voir la transition Z_2 entre les contacts déformable et non déformable, c’est pourquoi elle n’est pas représentée ici. Dans le domaine δ , l’indentation Δ_2 correspondant à Z_2 marque la transition entre les contacts déformable et non déformable (loi de Hooke). (c) Pour une courbe de retrait relative à des décrochages successifs de monomères, les régions d’intérêt sont les intervalles $[Z_j^{\text{début}}, Z_j^{\text{fin}}]$ dans lesquels le signal est décroissant, et les zones de décrochage sont les parties croissantes de la courbe. On utilise le modèle FJC dans chaque région d’intérêt.

Au contraire du modèle électrostatique qui dépend directement de la distance pointe-échantillon z , les interactions mécaniques après contact dépendent de l’indentation de l’échantillon, c’est-à-dire de la profondeur de déformation de sa surface externe, définie par

$$\delta(z) = z - Z_0 - \frac{f(z) - f(Z_0)}{k_c} \quad (2.2)$$

où k_c est la constante de raideur du levier sur lequel repose la pointe. L’indentation dépend non seulement de z mais aussi de la force mesurée $f(z)$. Dans l’intervalle $[\Delta_1, \Delta_2]$ définissant le contact déformable (où $\Delta_1 = \delta(Z_1)$; cf. fig. 2.5(a,b)), le modèle de Hertz s’écrit

$$f_{\text{HERTZ}}(\delta; E, \Delta_2) \propto E\delta^2, \quad (2.3)$$

où E est le module d’Young caractéristique de l’élasticité de l’objet et \propto indique la proportionnalité. Notons que $f_{\text{HERTZ}}(\delta; E, \Delta_2)$ dépend de Z_0 car δ dépend implicitement de Z_0 . Le paramètre

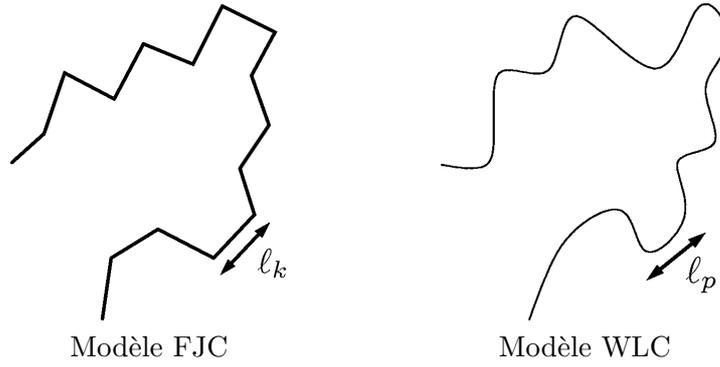


FIGURE 2.6 – Modèles *Freely-Jointed Chain* et *Worm-Like Chain* pour les polymères, vus respectivement comme une chaîne de segments rigides ou flexibles.

k_c est une constante fournie par le constructeur de la pointe AFM. Le modèle (2.3) est spécifique à une pointe de géométrie conique ou pyramidale. Il existe des modèles plus complexes pour une pointe sphérique ou plate [LIN 07a]. Pour le contact non déformable, on utilise le modèle de Hooke. C'est une interaction affine [SEN 05] du type :

$$f_{\text{HOOKE}}(\delta; a, k_{\text{cell}}, \Delta_2) = a + k_{\text{cell}}\delta \quad (2.4)$$

où $\delta > \Delta_2$. La pente k_{cell} est la constante d'élasticité de la bactérie. Elle est étroitement liée à sa pression de turgescence [YAO 02, POL 11].

Modèles pour la phase de retrait

Pendant le retrait de la pointe AFM, les bio-macromolécules situées sur la surface de l'échantillon biologique sont progressivement étirées et détachées de la pointe jusqu'à la perte de contact définitive, de façon similaire au détachement d'une bande auto-agrippante velcro. Les étirements successifs correspondent sur la figure 2.5(c) aux intervalles $[Z_j^{\text{début}}, Z_j^{\text{fin}}]$ sur lesquels la courbe est décroissante (voir aussi la courbe expérimentale de la figure 2.3(d)). Ces intervalles sont séparés par des parties croissantes liées aux décrochages successifs de monomères de la pointe. Les modèles les plus utilisés dans la littérature sont appelés *Freely Jointed Chain* (FJC) et *Worm-Like Chain* (WLC) [JAN 00]. Pour le modèle FJC, les monomères sont modélisés par des segments rigides connectés deux à deux par des jonctions flexibles. Pour le modèle WLC, chaque élément de la chaîne n'est plus un segment rigide mais flexible (fig. 2.6). Ce dernier modèle est bien adapté à la modélisation des constituants des protéines. Il existe aussi des extensions inélastiques nommées FJC+ et WLC+, qui mettent en jeu des paramètres physiques supplémentaires [JAN 00].

Dans ce chapitre, le modèle FJC est retenu car il est bien adapté aux exopolymères étudiés qui sont des polysaccharides. Contrairement au modèle WLC, le modèle FJC n'est pas défini par une fonction $z \mapsto f(z)$ mais par la fonction réciproque pour les échantillons du j -ème intervalle $z \in [Z_j^{\text{début}}, Z_j^{\text{fin}}]$ (fig. 2.5(c)) :

$$z_{\text{FJC}}(f; L_C, \ell_k) = -L_C \left(\coth \left(\frac{f\ell_k}{k_B T} \right) - \left(\frac{f\ell_k}{k_B T} \right)^{-1} \right) \quad (2.5)$$

où k_B , T et ℓ_k sont la constante de Boltzmann, la température, et la longueur de Kuhn caractérisant la longueur moyenne d'un segment de la chaîne, et L_C représente la longueur totale du contour. Le nombre de monomères est défini par le rapport L_C/ℓ_k .

2.2 Traitement de données en spectroscopie AFM

2.2.1 Objectifs et méthodologie de traitement

Nous retenons qu'une courbe de force est décrite par un ensemble de modèles paramétriques par morceaux. Pour pouvoir en extraire des informations physiques (topographie, mesure d'élasticité de la bactérie, nombres de monomères, *etc.*), il est essentiel de détecter les régions d'intérêt (les morceaux) où les modèles paramétriques s'appliquent. C'est la principale difficulté du traitement. Une fois les régions d'intérêt détectées, les paramètres physiques sont estimés au sens des moindres carrés dans chaque région. Le traitement d'une courbe de force se résume finalement en trois étapes :

1. recherche des points de discontinuité dans la courbe (segmentation) et lissage par morceaux de la courbe ;
2. détermination à partir des résultats du lissage, des régions d'intérêt dans lesquelles les modèles paramétriques s'appliquent. A ce stade, on obtient la topographie (Z_1) de l'échantillon pour la phase d'approche et les intervalles $[Z_j^{\text{début}}, Z_j^{\text{fin}}]$ qui décrivent le déroulement successif des polymères durant la phase de retrait ;
3. ajustement de chaque modèle paramétrique dans la région d'intérêt où il s'applique. Pour la phase d'approche, cela conduit à estimer la longueur de Debye pour l'interaction électrostatique, le module d'Young caractéristique de l'élasticité, la pression de turgescence et la constante d'élasticité de la bactérie dans la zone de contact mécanique. Pour le retrait, l'analyse des protéines adhésives (polymères) recouvrant le corps de la bactérie conduit à estimer le nombre de monomères, la longueur totale du polymère et la longueur moyenne des monomères.

Cette méthodologie est commune aux courbes d'approche et de retrait. La segmentation des morceaux, qui est le point le plus critique, sera amplement développée au paragraphe 2.2.2. Pour les deux derniers points et notamment pour la détection des régions d'intérêt, nous proposons des solutions aussi automatiques que possible mais il est parfois nécessaire d'introduire quelques paramètres de réglage empiriques. Nous résumons à présent ces deux derniers points.

Détection des régions d'intérêt

La figure 2.5 illustre que pour les courbes d'approche comme pour les courbes de retrait, les extrémités de chaque région d'intérêt sont des *points de discontinuité* où la courbe de force possède un saut et/ou un changement de pente et/ou de courbure. Pour les courbes d'approche, le point de contact pointe-échantillon est défini comme le point de discontinuité ($z = Z_1$) de la dérivée $f'(z)$ et la zone de contact est matérialisée par l'intervalle $z \geq Z_1$ (fig. 2.5(a)). Pour les courbes de retrait, les régions d'intérêt sont les intervalles pour lesquels le signal est décroissant (fig. 2.5(c)).

Ajustement de modèles paramétriques

Une fois les régions d'intérêt détectées, il ne reste plus qu'à ajuster chaque modèle paramétrique dans la région d'intérêt correspondante. Les paramètres physiques sont estimés au sens des moindres carrés en minimisant un critère du type :

$$\sum_{z \in \mathcal{R}_j} [f(z) - f_j(z; \boldsymbol{\theta}_j)]^2 \quad (2.6)$$

où \mathcal{R}_j désigne la j -ème région d'intérêt et $f_j(z; \boldsymbol{\theta}_j)$ est le modèle paramétrique défini pour cette région. $\boldsymbol{\theta}_j$ regroupe les paramètres de forme de la courbe de force. Par exemple, $\boldsymbol{\theta}_j$ vaut $\{A, \kappa^{-1}\}$ pour le modèle électrostatique (2.1) et $\{\Delta_2, E, a, k_{\text{cell}}\}$ pour les modèles Hertz (2.3) et Hooke (2.4) réunis. Pour la phase de retrait, le modèle FJC est du type $f \mapsto z_j(f; \boldsymbol{\theta}_j)$. Le critère des moindres carrés porte cette fois sur les erreurs quadratiques $(z - z_j(f; \boldsymbol{\theta}_j))^2$ dans le domaine z au lieu du domaine f . Bien que le nombre de paramètres à estimer dépasse rarement cinq, l'optimisation du critère (2.6) nécessite de prendre des précautions car le critère est susceptible de présenter des minima locaux ou des vallées plates.

2.2.2 Segmentation d'une courbe de force par approximation parcimonieuse

Ce traitement a pour but de reconstruire un signal $g(z)$ lisse par morceaux à partir d'une courbe de force expérimentale $f(z)$ où g est une version débruitée du signal f . La notion de parcimonie est liée au faible nombre de morceaux ou de façon équivalente, au faible nombre de points de discontinuité correspondant à des sauts du signal et/ou des changements de pente et/ou de courbure. Rechercher ces points nous conduira à résoudre un problème d'approximation parcimonieuse. Ce traitement s'apparente au lissage par splines polynomiales [DIE 95] à la différence que les points de contrôle ne sont pas choisis arbitrairement mais de façon adaptative [FRI 91, SMI 96]. Il s'apparente également à l'estimation d'un signal polynomial par morceaux à partir de données échantillonnées [VET 02]. A la différence de l'approche présentée dans [VET 02], notre approche repose sur la minimisation d'un critère mixte composé d'un terme d'adéquation aux données quadratique et d'un terme de régularisation permettant la détection conjointe des discontinuités d'ordres 0 et/ou 1, ..., P , où une discontinuité d'ordre p est définie comme un saut dans la dérivée p -ème du signal.

Enfin, notons quelques rapprochements avec le domaine de la modélisation et de la compression d'images, où un grand nombre de travaux visent à approximer une image en utilisant des représentations géométriques *adaptatives* [PEY 11]. La difficulté spécifique au traitement d'images est de modéliser finement les contours des objets. Parmi les nombreuses représentations possibles, la segmentation par arbre quaternaire (*prune-join quadtree*) [SHU 05] permet de modéliser des images polynomiales par morceaux avec des contours réguliers (images dites *cartoon*). Une autre possibilité est d'utiliser une triangulation adaptative permettant de segmenter l'image en cellules élémentaires (ou éléments finis). Une approximation polynomiale est alors réalisée sur chaque cellule [DEM 06, BOU 09].

Détection des sauts dans un signal

Dans un premier temps, on considère l'ordre 0. Pour reconstruire un signal g constant par morceaux, on propose de minimiser le critère :

$$\mathcal{J}(g; \lambda) = \sum_z [f(z) - g(z)]^2 + \lambda \sum_z \phi(g'(z)). \quad (2.7)$$

$\phi : \mathbb{R} \rightarrow \mathbb{R}_+$ est une fonction paire et croissante sur \mathbb{R}_+ , et le terme de pénalisation $\sum_z \phi(g'(z))$ a pour rôle de contraindre le signal $g(z)$ à être constant par morceaux avec peu de morceaux¹⁸. Le paramètre λ règle le compromis entre le degré de parcimonie (le nombre de morceaux) et la qualité de l'approximation de f par g . Pour une grande valeur de λ , l'erreur d'approximation

¹⁸. J'adopte une notation continue mais $g'(z)$ désigne en fait la dérivée discrète de g au sens des différences finies.

est grande mais le nombre de morceaux est faible. Des approximations de meilleure qualité sont obtenues pour des faibles valeurs de λ au prix d'un plus grand nombre de morceaux.

Pour contraindre la dérivée première du signal g à être parcimonieuse, c'est-à-dire non nulle pour un nombre limité de valeurs de z seulement, il faut que ϕ soit non différentiable en 0 [FAN 01]. C'est par exemple le cas de la pénalisation ℓ_1 $\phi(t) = |t|$. Ce choix présente néanmoins l'inconvénient de pénaliser davantage les fortes valeurs de $|g'(z)|$ et rend délicate la restauration des sauts de grande amplitude dans le signal f . Choisir une fonction ϕ non convexe permet de pallier ce problème. Nous choisissons d'utiliser le coût ℓ_0 :

$$\phi(t) = |t|_0 = \begin{cases} 0 & \text{si } t = 0, \\ 1 & \text{si } t \neq 0. \end{cases} \quad (2.8)$$

Le critère \mathcal{J} est un critère mixte « $\ell_2 - \ell_0$ » car il est composé d'un terme quadratique et d'un terme lié à la pseudo-norme ℓ_0 qui compte le nombre de morceaux :

$$\mathcal{J}(g; \lambda) = \sum_z [f(z) - g(z)]^2 + \lambda \sum_z |g'(z)|_0. \quad (2.9)$$

Cette formulation est aussi qualifiée d'*approche analyse* pour la parcimonie dans la littérature [ELA 07, VAI 13, GIR 14]. Elle diffère de l'*approche synthèse* présentée au chapitre 1 (§ 1.3) dans la mesure où le terme de régularisation porte non pas sur le signal g lui-même, mais sur une version filtrée de g (par le filtre dérivateur). Nous nous ramènerons par la suite (§ 2.3) à une formulation de type synthèse par le biais du changement de variables $x(z) = g'(z)$ ⁽¹⁹⁾. Dans ce cas, le terme quadratique du nouveau critère dépendant de x fait intervenir un opérateur d'intégration numérique.

Notons que comme $\phi(t) = |t|_0$ est fortement non convexe, le critère \mathcal{J} n'est pas convexe et présente généralement des minima locaux. Il existe un algorithme optimal efficace spécifique au critère (2.9) [BLA 89] mais pour minimiser les critères introduits ci-dessous en tant qu'extensions de (2.9), il faudra recourir à des algorithmes sous-optimaux.

Détection conjointe de discontinuités à des ordres différents

Pour détecter des points de discontinuité à plusieurs ordres simultanément dans un signal (un point de discontinuité est soit un saut du signal, soit un changement de pente ou de courbure), on remplace le critère (2.9) par :

$$\mathcal{J}(g_0, \dots, g_P; \lambda) = \sum_z \left[f(z) - \sum_{p=0}^P g_p(z) \right]^2 + \lambda \sum_{p=0}^P \sum_z |g_p^{(p+1)}(z)|_0. \quad (2.10)$$

Le signal lissé par morceaux s'exprime maintenant sous la forme $g = \sum_p g_p$, où chaque signal g_p est un polynôme de degré p par morceaux (g_0 est constant par morceaux, g_1 est affine par morceaux, *etc.*). Les points de discontinuité d'ordre p sont les points en lesquels la dérivée d'ordre $p + 1$ de g_p est non nulle. Par exemple, pour $P = 2$, le signal lissé par morceaux s'exprime par $g = g_0 + g_1 + g_2$ où g'_0 , $g_1^{(2)}$ et $g_2^{(3)}$ sont parcimonieux. g n'est rien d'autre qu'une fonction quadratique par morceaux.

19. Ce changement de variables est possible dans le cas particulier du critère (2.9), mais en toute généralité, on ne peut systématiquement remplacer une formulation de type analyse en une formulation de type synthèse.

Sélection de variables scalaires et vectorielles

Nous avons envisagé deux stratégies alternatives pour la recherche des points de discontinuité, appelées sélection de variables scalaires et vectorielles :

1. Pour la sélection de variables scalaires, les discontinuités détectées à des ordres différents ne sont généralement pas en une même position z . C'est ce qui se produit lorsqu'on minimise (2.10). Une « variable scalaire » est définie comme un couple (z, p) qui décrit la position z d'une discontinuité et son ordre p ;
2. Pour reconstruire un signal polynomial par morceaux avec un nombre minimal de morceaux, il est avantageux d'imposer la détection de discontinuités à *tous les ordres* $p = 0, \dots, P$ en une même position z , en comptant 1 (soit une *position* de discontinuité) au lieu de $P + 1$ ($P + 1$ variables). Ainsi, une discontinuité à la position z est associée à une « variable vectorielle » qui regroupe les amplitudes $\{g_p^{(p+1)}(z), p = 0, \dots, P\}$ des discontinuités relatives à la position z , pour un coût unitaire. Pour cela, on propose de remplacer le terme de régularisation du critère (2.10) par $\lambda \|g\|_{\infty,0}$, où

$$\|g\|_{\infty,0} \triangleq \sum_z \left| \max_{p=0:P} |g_p^{(p+1)}(z)| \right|_0 \quad (2.11)$$

Ce terme est une pseudo-norme mixte $\ell_{\infty,0}$ [BAC 12, Sec. 1.3], définie comme la composition des normes ℓ_{∞} dans la dimension p (à z fixé, on calcule le maximum des valeurs $g_p^{(p+1)}(z)$, $p = 0, \dots, P$) et de la norme ℓ_0 dans la dimension z , qui compte le nombre de valeurs maximales non nulles. L'activation des discontinuités à tous les ordres en une même position z est favorisée puisque la pénalisation vaut λ (et pas plus) dès lors que $g_p^{(p+1)}(z) \neq 0$ pour au moins une valeur de p .

2.3 Approximation parcimonieuse pour la segmentation

Plutôt que de développer un algorithme spécifique à la minimisation de (2.10) ou de l'extension pénalisée (2.11), nous nous ramenons au cadre générique du chapitre 1 : l'approximation d'un signal à partir d'un nombre limité d'éléments d'un dictionnaire incluant les structures recherchées. La minimisation du critère (2.10) et de son extension « vectorielle » se ramène en effet à ce cadre, où le dictionnaire est une matrice de Tœplitz ou de Tœplitz par blocs qui représente un ou des opérateurs d'intégration numérique.

2.3.1 Reformulation de la détection conjointe de discontinuités

Nous montrons à présent que le problème (2.10) peut être mis sous la forme :

$$\min_{\|\mathbf{x}\|_0 \leq k} \|\mathbf{f} - \mathbf{A}\mathbf{x}\|_2^2 \quad (2.12)$$

ou

$$\min_{\mathbf{x} \in \mathbb{R}^n} \{\mathcal{J}(\mathbf{x}; \lambda) = \|\mathbf{f} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0\}. \quad (2.13)$$

où le vecteur $\mathbf{f} = [f(z_1), \dots, f(z_m)]^T \in \mathbb{R}^m$ regroupe les données ($f(z)$ pour tout z) et le dictionnaire \mathbf{A} est matrice de taille $m \times n$. Le problème de sélection de variables (scalaires) dans un dictionnaire consiste à trouver une représentation de \mathbf{f} à partir d'un faible nombre de colonnes de \mathbf{A} , c'est-à-dire de trouver une approximation $\mathbf{f} \approx \mathbf{A}\mathbf{x}$ telle que $\mathbf{x} \in \mathbb{R}^n$ soit un

vecteur parcimonieux (*cf.* chapitre 1). L'entier k ou le réel λ contrôle le degré de parcimonie. La formulation (2.12) impose explicitement que k colonnes au plus sont sélectionnées alors que (2.13) se rapproche de la formulation (2.9)-(2.10) introduite pour le lissage par morceaux. Le cas de la variante vectorielle utilisant la pénalisation (2.11) au lieu du coût ℓ_0 sera traité au paragraphe 2.3.4.

Pour minimiser les critères (2.9) et (2.10), on ne peut pas directement utiliser les algorithmes classiques et efficaces d'approximation parcimonieuse de type synthèse, car le terme de régularisation ne prend pas la forme $\lambda\|\mathbf{g}\|_0$: c'est une combinaison de termes de type $\lambda\|\mathbf{D}^p\mathbf{g}\|_0$ où \mathbf{D}^p est un opérateur de dérivation discrète d'ordre p . Pour se ramener au cadre classique, nous avons recours à un changement de variable du type $\mathbf{x} = \mathbf{D}^p\mathbf{g} \iff \mathbf{g} = \mathbf{A}^p\mathbf{x}$ où \mathbf{A}^p est un opérateur d'intégration numérique. L'effet de ce changement de variable est que le premier terme $\|\mathbf{f} - \mathbf{g}\|_2^2$ dans (2.9) et (2.10) devient $\|\mathbf{f} - \mathbf{A}^p\mathbf{x}\|_2^2$.

2.3.2 Construction du dictionnaire

La détection de discontinuités est abordée comme un problème de sélection de variables, où une variable (i, p) est associée à la fois à la position i et à l'ordre p d'une discontinuité (la position $z = z_i$ d'un point de discontinuité est décrite par l'indice $i \in \{1, \dots, m\}$ de l'échantillon correspondant). Le dictionnaire²⁰ est construit de façon à ce que tous les types de discontinuité recherchés puissent être présents pour une position donnée. Ainsi, pour les discontinuités d'ordre p , nous construisons la matrice de Toeplitz \mathbf{A}^p dont les colonnes résultent de l'échantillonnage sur une grille régulière du monome $z \mapsto z^p$ en envisageant tous les points de discontinuité possibles :

$$\mathbf{A}^p = \begin{bmatrix} 1^p & 0 & \dots & 0 \\ 2^p & 1^p & & \vdots \\ 3^p & 2^p & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \vdots & \vdots & \dots & 1^p \\ m^p & (m-1)^p & \dots & 2^p \end{bmatrix}. \quad (2.14)$$

\mathbf{A}^p est un opérateur d'intégration numérique qui consiste à intégrer p fois un signal impulsionnel en position i , avec la condition aux limites que le signal résultant est nul avant la position i . La figure 2.7 illustre cette définition dans les cas $p = 0, 1$ et 2 . Les colonnes des matrices $\mathbf{A}^0, \mathbf{A}^1$ et \mathbf{A}^2 sont respectivement les fonctions de Heaviside discrètes, les fonctions rampes, les fonctions quadratiques positives.

Le dictionnaire permettant la détection conjointe des discontinuités d'ordres $p = 0, \dots, P$ est finalement obtenu en concaténant l'ensemble des dictionnaires :

$$\mathbf{A} = [\mathbf{A}^0 | \mathbf{A}^1 | \dots | \mathbf{A}^P] \quad (2.15)$$

et le lissage d'un signal polynomial par morceaux (2.10) se réécrit finalement comme l'approximation de \mathbf{f} par $\mathbf{Ax} = \sum_{p=0}^P \mathbf{A}^p\mathbf{x}^p$ où \mathbf{x}^p désignent les blocs qui constituent le vecteur \mathbf{x} . Les éléments x_i^p représentent ici l'amplitude des discontinuités (analogues à $g_p^{(p+1)}(z)$ dans (2.10)).

20. Ce problème a été rapidement introduit dans le chapitre 1 (§ 1.6.5) en tant qu'exemple. Je rappelle ici cette définition et je l'illustre davantage pour favoriser la clarté du chapitre et pour qu'il soit lisible indépendamment du chapitre 1.

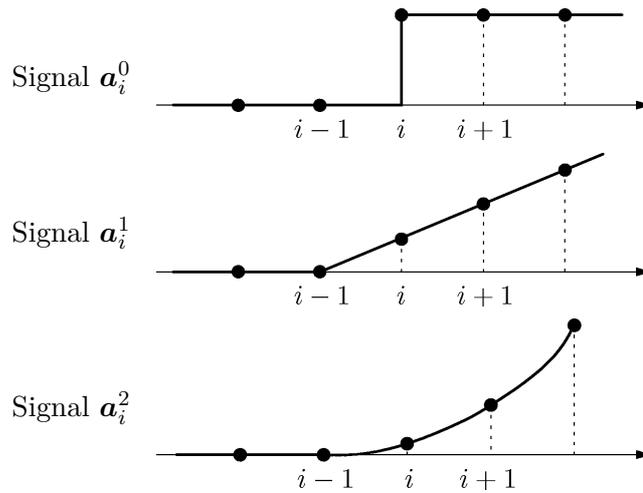


FIGURE 2.7 – Signaux \mathbf{a}_i^p liés à l'activation d'une discontinuité d'ordre p pour la position i . \mathbf{a}_i^0 est la fonction de Heaviside discrète, \mathbf{a}_i^1 est la fonction rampe, et \mathbf{a}_i^2 est la fonction quadratique positive. Chaque signal vaut 1 pour la position i et son support est égal à $\{i, \dots, m\}$.

2.3.3 Sélection de variables scalaires

La sélection de variables scalaires est formulée par les problèmes (2.12) et (2.13). Pour traiter ce problème, l'étude comparative du chapitre 1 (*cf.* fig. 1.3 page 50) montre que les algorithmes OLS et SBR font partie des plus performants parmi les algorithmes testés.

OLS (*cf.* tableau 1.3 page 42) fournit d'abord des approximations grossières $\mathbf{Ax}(k)$ avec peu de points de discontinuité (faibles valeurs de k), puis des approximations plus fines avec davantage de points de discontinuité pour k croissant. Le critère d'arrêt retenu pour OLS est l'erreur quadratique d'approximation : on considère que les principales discontinuités ont été trouvées lorsque l'erreur d'approximation $\|\mathbf{f} - \mathbf{Ax}\|_2^2$ est inférieure à un seuil. La valeur du seuil est réglée automatiquement en fonction de la variance du bruit de mesure, elle-même estimée à partir des données expérimentales (dans la zone plate des courbes de force, où la force est théoriquement nulle).

SBR fait partie des algorithmes bidirectionnels (*forward-backward*) : *cf.* tableau 1.4 page 45. L'intérêt de ce type d'algorithme par rapport aux algorithmes monodirectionnels comme OLS est qu'une variable préalablement activée peut être désactivée dès lors que l'erreur quadratique n'augmente pas significativement [MIL 02, chapitre 3]. Contrairement à OLS, SBR s'arrête en un nombre fini d'itérations. Aucun réglage n'est donc nécessaire pour l'arrêt de SBR. La difficulté pratique est le réglage de λ que nous réalisons de façon empirique. A parcimonie égale, SBR fournit des approximations plus précises qu'OLS pour un temps de calcul (lié au nombre de remplacements élémentaires effectués) plus important [SOU 11].

Des exemples de résultats liés à OLS et SBR pour la sélection de variables scalaires sont disponibles au § 1.6.5 du chapitre 1. Dans la suite de ce chapitre, la sélection de variables vectorielles est privilégiée.

2.3.4 Sélection de variables vectorielles

Dans l'approche scalaire, le terme $\|\mathbf{x}\|_0$ compte le nombre de discontinuités pour chaque ordre et pour chaque position. La détection de discontinuités à des ordres différents et *pour une même position* (par exemple, on impose de détecter simultanément un saut et un changement de

dérivée dans le signal pour la même position i) consiste à sélectionner (activer) plusieurs colonnes de la matrice \mathbf{A} simultanément. Pour formuler mathématiquement la sélection de variables vectorielles, on introduit la notation vectorielle $\mathbf{x}_i \in \mathbb{R}^{P+1}$ pour regrouper toutes les coordonnées de \mathbf{x} liées à une même position i et on réordonne le vecteur \mathbf{x} comme la concaténation des vecteurs $\mathbf{x}_i : \mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$. Le décompte du nombre de positions i sélectionnées s'écrit

$$\|\mathbf{x}\|_{\infty,0} \triangleq \sum_{i=1}^m \|\mathbf{x}_i\|_{\infty}|_0$$

où $\|\mathbf{x}_i\|_{\infty}|_0$ vaut 1 si $\mathbf{x}_i \neq \mathbf{0}$, et 0 sinon. La sélection de variables vectorielles est finalement formulée comme la minimisation des critères

$$\min_{\|\mathbf{x}\|_{\infty,0} \leq k} \|\mathbf{f} - \mathbf{A}\mathbf{x}\|_2^2 \quad (2.16)$$

ou

$$\min_{\mathbf{x}} \{\|\mathbf{f} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_{\infty,0}\}, \quad (2.17)$$

où k contrôle le nombre maximal de points de discontinuité autorisés. Il s'agit d'autoriser la sélection d'au plus k variables vectorielles, soit $k(P+1)$ variables scalaires.

Les algorithmes utilisés pour minimiser les critères (2.16) et (2.17) sont de simples adaptations d'OLS et SBR. A chaque itération, il s'agit maintenant de tester tous les ajouts et retraits possibles d'une variable vectorielle, c'est-à-dire l'ajout simultané de $P+1$ discontinuités scalaires en une position i ($\mathbf{x}_i \neq \mathbf{0}$) ou, si la position i est active, le retrait simultané des discontinuités présentes ($\mathbf{x}_i = \mathbf{0}$).

2.3.5 Difficultés techniques et aspects implémentation

Une partie importante du travail lié au traitement de données réelles a consisté à développer des implémentations rapides pour traiter des signaux comportant environ 3000 échantillons ce qui conduit à un dictionnaire polynomial de taille 3000×9000 (pour $P=2$). L'implémentation générique des algorithmes gloutons orthogonaux repose sur le stockage mémoire de la matrice $\mathbf{G} = \mathbf{A}^T \mathbf{A}$ de façon à pouvoir extraire la matrice de Gram $\mathbf{G}_{\mathcal{Q}} \triangleq \mathbf{A}_{\mathcal{Q}}^T \mathbf{A}_{\mathcal{Q}}$ de \mathbf{G} ; le facteur de Cholesky $\mathbf{L}_{\mathcal{Q}}$ (avec $\mathbf{G}_{\mathcal{Q}} = \mathbf{L}_{\mathcal{Q}} \mathbf{L}_{\mathcal{Q}}^T$) est systématiquement mis à jour quand \mathcal{Q} subit une modification élémentaire.

Pour un dictionnaire convolutif lié à une réponse impulsionnelle *finie* [SOU 12], il n'est pas nécessaire de stocker \mathbf{G} car elle possède une structure de Toeplitz et ses éléments se déduisent directement de l'auto-corrélation de la réponse impulsionnelle [MAI 09]. Cela permet de traiter des problèmes de grande taille car les seules quantités stockées au cours de l'algorithme sont des matrices et des vecteurs de taille $k = \text{Card}[\mathcal{Q}]$ (typiquement, la matrice $\mathbf{L}_{\mathcal{Q}}$).

Le dictionnaire polynomial est une concaténation de dictionnaires convolutifs liés à des filtres intégrateurs dont la réponse impulsionnelle est *infinie*. Les matrices \mathbf{A}^p possèdent toujours une structure de Toeplitz mais pas les matrices de Gram $\mathbf{A}^{pT} \mathbf{A}^p$. Néanmoins, ces matrices possèdent une structure assez simple (de même que les matrices $\mathbf{A}^{pT} \mathbf{A}^q$) qui permet, au prix de quelques adaptations techniques, de déduire le calcul rapide des blocs de $\mathbf{G} = \mathbf{A}^T \mathbf{A}$ de l'auto-corrélation des filtres intégrateurs liés sans avoir à stocker la matrice \mathbf{G} .

La réelle difficulté technique que pose le dictionnaire polynomial provient des dépendances linéaires entre les atomes du dictionnaire. Plus précisément, le nombre minimal de colonnes de \mathbf{A} linéairement dépendantes vaut $\text{spark}(\mathbf{A}) = 3$ car la différence entre deux rampes discrètes à des instants consécutifs i et $i+1$ produit la fonction de Heaviside discrète à l'instant i :

$$\mathbf{a}_i^1 - \mathbf{a}_{i+1}^1 = \mathbf{a}_i^0. \quad (2.18)$$

Les implémentations traditionnelles des algorithmes gloutons orthogonaux reposent sur l'hypothèse que \mathbf{A}_Q est de rang plein. Elles exploitent l'inverse de la matrice de Gram pour calculer les amplitudes non nulles $\mathbf{t} = \mathbf{A}_Q^\dagger \mathbf{y} = \mathbf{G}_Q^{-1} \mathbf{A}_Q^T \mathbf{y}$ et l'erreur quadratique d'approximation :

$$\mathcal{E}_Q = \|\mathbf{y}\|_2^2 - \|\mathbf{L}_Q^{-1} \mathbf{A}_Q^T \mathbf{y}\|_2^2. \quad (2.19)$$

Cette implémentation n'est plus valide lorsque \mathbf{A}_Q est de rang déficient.

Pour le dictionnaire polynomial, la déficience de rang de \mathbf{A}_Q est susceptible de se produire pour des supports Q de très faible cardinalité. Les alternatives que nous avons envisagé (et testé) sont :

- d'interdire l'ajout d'un atome conduisant à une matrice \mathbf{A}_Q de rang déficient ; c'est le choix retenu dans [SOU 11]. Cette solution est implémentable en pratique dans la mesure où on dispose d'une condition sur Q facile à tester numériquement²¹ qui garantit que \mathbf{A}_Q est de rang plein. Pour le dictionnaire polynomial, nous avons établi une condition suffisante (mais non nécessaire) portant directement sur Q [SOU 11, Prop. 2]. Pour fixer les idées, dans le cas $P = 1$, on peut interdire la sélection conjointe de deux atomes à la même position (\mathbf{a}_i^0 et \mathbf{a}_i^1) pour éviter que les trois atomes \mathbf{a}_i^1 , \mathbf{a}_{i+1}^1 et \mathbf{a}_i^0 soient sélectionnés conjointement. Cette condition est néanmoins non nécessaire au sens où \mathbf{A}_Q peut être de rang plein en sélectionnant les deux atomes \mathbf{a}_i^0 et \mathbf{a}_i^1 .
- une stratégie plus fine qui consiste à anticiper l'obtention d'une matrice de rang déficient en envisageant des substitutions : pour le dictionnaire polynomial, compte tenu de (2.18), si deux atomes \mathbf{a}_i^1 et \mathbf{a}_{i+1}^1 sont sélectionnés, on teste des remplacements visant à supprimer les deux atomes et à ajouter le nouvel atome \mathbf{a}_i^0 .

La première stratégie est retenue. C'est une simple adaptation des algorithmes gloutons alors que la seconde stratégie modifie la structure du voisinage (le voisinage d'un support Q ne se restreint plus aux seuls remplacements élémentaires $Q \cup \{i\}$ et $Q \setminus \{i\}$ compte tenu du mouvement supplémentaire de substitution).

2.4 Traitement de données réelles

Les courbes de force et les images force-volume ont été acquises par un microscope MFP3D-BIO (Asylum Res. Tech., Atomic Force F&E GmbH, Mannheim). Les pointes AFM, de forme conique, ont été construites par Veeco (MLCT-AUNM, Veeco Instruments SAS, Dourdan). Avant de procéder à l'acquisition de données, la géométrie de la pointe est contrôlée par un système de vision 3D (TGT1, NT-MTD Company, Moscou). Chaque courbe de force est acquise en adoptant une vitesse de $1 \mu\text{m.s}^{-1}$ durant l'approche et le retrait de la pointe. L'échantillon étudié dans la phase d'approche est un mutant K-12 de la bactérie *Escherichia coli* (E2152) fourni par l'Institut Pasteur de Paris. Les données traitées dans la phase de retrait sont relatives à la bactérie *Pseudomonas fluorescens* (fig. 2.3). Cette bactérie, présente dans des réseaux de distribution d'eau, produit et excrète des exopolysaccharides. Ces sucres complexes sont à l'origine de la formation de biofilms, agglomérats de bactéries fixées sur les surfaces des réseaux de distribution d'eau et qui sont à l'origine de nombreux cas de maladies nosocomiales.

2.4.1 Segmentation d'une courbe de retrait : comparaison de stratégies

La figure 2.8 présente quelques résultats comparatifs des deux méthodes de sélection de variables pour une courbe de retrait (une comparaison plus poussée est disponible dans [DUA 09b]).

21. Une condition plus simple et moins approximative que le seuillage des valeurs singulières de \mathbf{A}_Q .

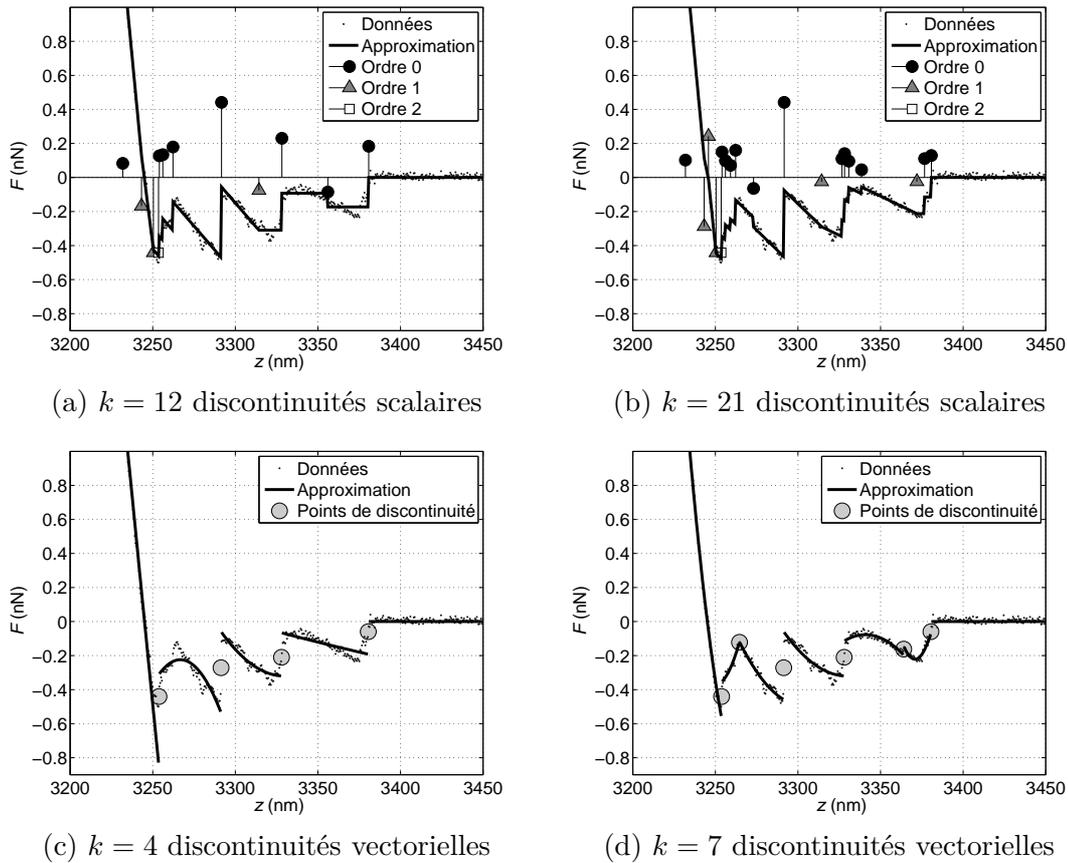


FIGURE 2.8 – Lissage par morceaux d’une courbe de force par sélection de variables scalaires et vectorielles en utilisant l’algorithme SBR. Le signal lissé est quadratique par morceaux, c’est-à-dire qu’on recherche des discontinuités aux ordres $p = 0, 1$ et 2 . La courbe de force expérimentale $f(z)$ est affichée en pointillés et son approximation est en trait plein. (a-b) Sélection de variables scalaires obtenue avec deux degrés de parcimonie différents. (c-d) Sélection de variables vectorielles pour deux niveaux de parcimonie. Les points de discontinuités sont représentés par un seul cercle car des discontinuités scalaires aux ordres $0, 1$, et 2 sont présentes simultanément.

Les discontinuités d’ordres $0, 1$ et 2 sont simultanément recherchées dans le signal, ce qui revient à l’approcher par un polynôme de degré 2 par morceaux. L’algorithme utilisé est SBR, et le paramètre λ est fixé empiriquement en fonction de la qualité d’approximation désirée. L’approche par sélection de variables scalaires ne détecte généralement pas des discontinuités à plusieurs ordres et pour des positions identiques (fig. 2.8(a,b)). Il y a autant de variables scalaires sélectionnées que de points de discontinuités (12 et 21 variables respectivement, c’est-à-dire 12 et 21 positions). L’approche vectorielle conduit à seulement 4 et 7 points de discontinuité suivant le choix du paramètre λ , ce qui revient (aussi) à sélectionner 12 et 21 variables scalaires dans le dictionnaire. Les intervalles délimités par les points de discontinuité permettent d’identifier de façon plus fiable les régions décroissantes de la courbe de force dans lesquelles le modèle FJC s’applique.

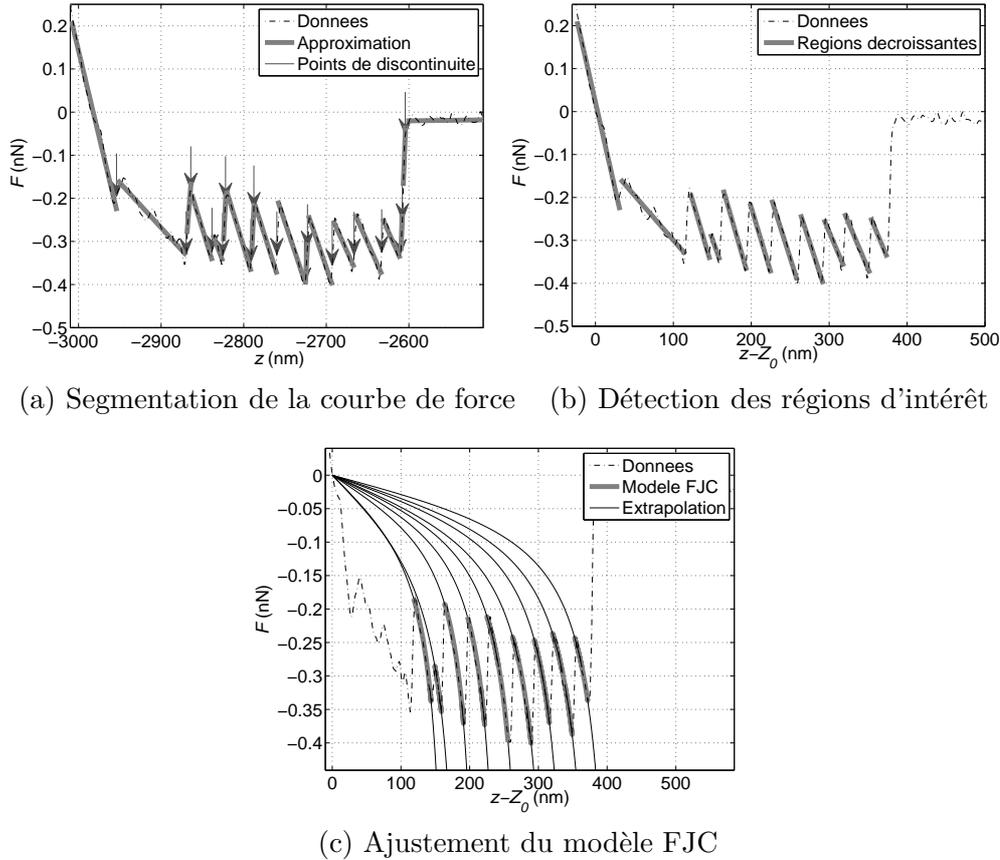


FIGURE 2.9 – Traitement complet de la courbe de retrait présentée sur la figure 2.3(d) : segmentation par l’algorithme OLS vectoriel (a) ; détection des régions décroissantes du signal lissé (b) ; et ajustement du modèle FJC aux données expérimentales dans chaque région décroissante (c). Les lignes en gras représentent l’ajustement du modèle FJC dans chaque région, et les lignes en trait fin indiquent l’extrapolation du modèle en dehors des régions d’intérêt (notamment vers le point de contact $z = Z_0$).

2.4.2 Traitement complet d’une courbe de retrait

Pour le traitement complet et systématique d’une courbe de retrait, nous appliquons l’algorithme OLS couplé à la sélection de variables vectorielles. Bien que SBR fournisse des résultats légèrement supérieurs à OLS, il est en pratique plus aisé de fixer le degré de parcimonie k c’est-à-dire le nombre d’itérations d’OLS que le paramètre λ pour SBR. k est réglé automatiquement de telle sorte que le résidu $\|\mathbf{f} - \mathbf{A}\mathbf{x}(k)\|_2^2$ de l’approximation des données avec k variables soit inférieur à un seuil ε fixé, lié à la variance empirique du bruit de mesure que l’on calcule dans la partie plate du signal expérimental (zone à l’extrême droite, où la force est théoriquement nulle). Les résultats de la segmentation, de la détection de régions d’intérêt (régions décroissantes) et de l’ajustement de la courbe de retrait par le modèle FJC sont présentés sur la figure 2.9 pour la courbe de retrait de la figure 2.3(d). Pour la segmentation, on utilise un lissage affine par morceaux du signal. On constate que la plupart des régions décroissantes sont finement détectées et que l’ajustement du modèle FJC est très précis. L’algorithme a par ailleurs été testé sur un très grand nombre de signaux réels. En termes quantitatifs, il fournit des résultats co-

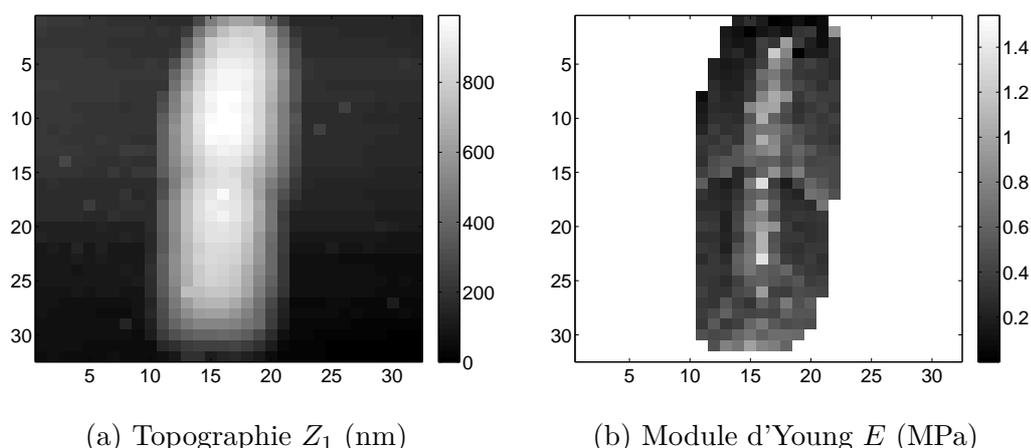


FIGURE 2.10 – Traitement complet d’une image force-volume dans la phase d’approche. Pour un échantillon mutant K-12 de la bactérie *Escherichia coli* (E2152), on représente les deux images (a) de la topographie, c’est-à-dire du relief de la surface, et (b) du module d’Young caractérisant l’élasticité.

hérents avec les valeurs connues dans la littérature pour les paramètres physiques des bactéries étudiées [POL 11].

2.4.3 Traitement d’une image force-volume dans la phase d’approche

On illustre le traitement d’une image force-volume dans la phase d’approche. Pour chaque courbe de force de l’image, on applique l’algorithme OLS vectoriel et on estime sept paramètres par la procédure décrite au paragraphe 2.2.1, à savoir la topographie Z_1 , l’indentation Δ_2 qui caractérise la transition entre les modes d’interaction déformable et non déformable, le préfacteur A et la longueur de Debye κ^{-1} pour l’interaction électrostatique, le module d’Young E , la constante d’élasticité k_{cell} de la bactérie et sa pression de turgescence pour l’interaction mécanique. On reconstruit finalement sept images 2D des paramètres physiques. Sur la figure 2.10 sont représentées l’image de la topographie Z_1 et celle du module d’Young E . Des résultats plus complets, y compris pour la phase de retrait, sont disponibles dans [POL 11]. Dans la phase de retrait, la visualisation des résultats est moins aisée car le nombre de régions d’intérêt (et donc de paramètres physiques estimés L_c et ℓ_k) diffère d’une courbe de force à l’autre.

2.5 Conclusion

Le problème inverse traité dans ce chapitre est la détection conjointe des discontinuités à différents ordres dans un signal et son approximation par un signal polynomial par morceaux. Par le choix d’un dictionnaire adapté, nous nous sommes ramenés à un problème d’approximation parcimonieuse en utilisant une régularisation de type ℓ_0 . Traiter ce problème, c’est résoudre un problème d’optimisation discrète pour rechercher la position et l’ordre des points de discontinuité. Les algorithmes OLS et SBR fournissent des résultats très précis pour la segmentation, en un temps de calcul certes plus important que d’autres algorithmes d’approximation parcimonieuse plus simples comme OMP et les algorithmes qui reposent sur la relaxation de la pseudo-norme ℓ_0 par la norme ℓ_1 [SOU 11]. Cependant, l’exploitation de la structure du dictionnaire permet une mise en œuvre rapide et adaptée à des données de grande taille (quelques

milliers de points par courbe).

L'approche scalaire permet de détecter simultanément les discontinuités à différents ordres dans un signal sans imposer leur détection à la même position (un saut et un changement de pente ne se produisent pas nécessairement de façon simultanée). Comme il est parfois souhaitable de favoriser cette détection simultanée, nous avons développé une variante vectorielle dans laquelle on impose l'activation de discontinuités à tous les ordres possibles et à une position donnée pour un coût unitaire. Cette variante se rapproche du concept de parcimonie de groupe [KOW 09, KOW 10, BAC 12] où on définit une relation de connexité entre les colonnes du dictionnaire pour favoriser l'activation simultanée de colonnes connexes. Nous avons montré la pertinence de cette approche pour segmenter très finement les courbes de force expérimentales avec un nombre limité de points de discontinuité.

L'ajustement des modèles physiques dans chaque intervalle segmenté est effectué au sens des moindres carrés. C'est un problème d'optimisation continu classique dont les variables sont les paramètres des modèles physiques. Comme le nombre de paramètres est faible, l'application de méthodes d'optimisation locale est rapide bien qu'il faille prendre des précautions car certains modèles sur-paramétrés induisent des critères avec des vallées plates ou des minima locaux (voir le rapport [SOU 07, fig. 8] pour une illustration graphique de ce type de critères).

Par opposition au traitement indépendant d'un ensemble de courbes de force, une perspective consiste à envisager un traitement conjoint. Pour cela, on fera l'hypothèse qu'une courbe de force résulte du mélange de signaux sources où une source est liée à un type d'interaction élémentaire et les mélanges sont des combinaisons linéaires avec retard des signaux sources. Ce modèle a été validé sur des données expérimentales dans la phase d'approche [SOU 08b]. Le traitement conjoint de plusieurs courbes de force reposera finalement sur une procédure de séparation de sources retardées, où chaque source est décrite par un signal polynomial par morceaux et les retards sont liés à la topographie du nano-objet.

Chapitre 3

Analyse en k itérations des algorithmes OMP et OLS

3.1 Motivations

Le point de départ de cette analyse est le travail décrit au chapitre 1 concernant les aspects algorithmiques de l'approximation parcimonieuse, formulée comme un problème de minimisation quadratique sous contrainte ℓ_0 :

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_2^2 \quad \text{s.c.} \quad \|\mathbf{x}\|_0 \leq k \quad (3.1)$$

ou²² comme

$$\min_{\mathbf{x}} \{\|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda \|\mathbf{x}\|_0\}. \quad (3.2)$$

Les algorithmes d'optimisation discrète décrits au chapitre 1 sont des heuristiques sous-optimales ayant pour but de retrouver le « meilleur » support. Les problèmes ci-dessus sont en fait transformés en des problèmes d'optimisation combinatoire :

$$\min_{\mathcal{Q}} \mathcal{E}_{\mathcal{Q}} \quad \text{s.c.} \quad \text{Card}[\mathcal{Q}] \leq k \quad (3.3)$$

$$\min_{\mathcal{Q}} \{\mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}]\} \quad (3.4)$$

car le support $\mathcal{Q} \subset \{1, \dots, n\}$ peut être codé comme un vecteur de $\{0, 1\}^n$ où n est le nombre d'atomes du dictionnaire (en codant 1 si un atome de \mathbf{A} est présent dans \mathcal{Q} , et 0 sinon). La transformation des problèmes (3.1)-(3.2) en (3.3)-(3.4) repose sur l'expression $\mathcal{E}_{\mathcal{Q}} \triangleq \|\mathbf{r}_{\mathcal{Q}}\|_2^2$ de l'erreur quadratique minimale $\|\mathbf{y} - \mathbf{Ax}\|_2^2$, où $\mathbf{r}_{\mathcal{Q}} \triangleq \mathbf{P}_{\mathcal{Q}}^\perp \mathbf{y} = \mathbf{y} - \mathbf{P}_{\mathcal{Q}} \mathbf{y}$ est le résidu de l'approximation. L'approximation des données $\mathbf{P}_{\mathcal{Q}} \mathbf{y}$ est définie comme la projection orthogonale de \mathbf{y} sur $\text{span}(\mathbf{A}_{\mathcal{Q}})$, et $\mathbf{r}_{\mathcal{Q}}$ est la projection orthogonale de \mathbf{y} sur l'espace supplémentaire orthogonal de $\text{span}(\mathbf{A}_{\mathcal{Q}})$. Les algorithmes gloutons *orthogonaux* OMP, OLS et SBR effectuent cette projection orthogonale : à chaque itération, un nouveau support \mathcal{Q} est exploré et les données \mathbf{y} sont projetées orthogonalement sur $\text{span}(\mathbf{A}_{\mathcal{Q}})$. La solution de (3.1)-(3.2) s'écrit $\mathbf{x} = \{\mathbf{A}_{\mathcal{Q}}^\dagger \mathbf{y}, \mathbf{0}\}$ à une permutation d'indices près.

L'un des objectifs de mon séjour de délégation CNRS à l'IRCCyN (année 2010-2011) était d'examiner les aspects théoriques liés aux algorithmes gloutons bidirectionnels (SBR notamment). Comme les algorithmes gloutons étudiés (OMP, OLS, SBR, CSBR, ℓ_0 -PT) sont tous des

22. de façon non équivalente : cf. § 1.7.1.

heuristiques de recherche, la convergence vers la solution optimale des problèmes (3.3) et (3.4) n'est pas garantie. Il s'agit donc d'étudier si dans certaines situations, les algorithmes fournissent des garanties d'optimalité.

Les algorithmes gloutons proposés au chapitre 1 ont été définis comme des extensions bidirectionnelles d'OLS. Nous nous sommes donc naturellement posé la question des conditions liées à OLS. Nous avons fait le constat que contrairement à OMP, il n'existait pas de résultat de reconstruction exacte théorique pour OLS (ce constat était partagé par d'autres chercheurs de la communauté [DAV 12]). Avec Rémi Gribonval et Cédric Herzet de l'INRIA Rennes, nous nous sommes donc consacrés à l'établissement de tels résultats pour OLS avec comme point de départ les résultats relatifs à la reconstruction exacte de supports par OMP [TRO 04, DAV 10]. L'analyse des algorithmes bidirectionnels (comme SBR et les extensions bidirectionnelles d'OMP) est reportée à de futurs travaux (section 4.2 du chapitre 4).

3.2 Analyse en k itérations

3.2.1 Définition

Plaçons-nous dans le cadre non bruité où les données \mathbf{y} sont des combinaisons linéaires de k atomes indicés par \mathcal{Q}^* (avec $k \triangleq \text{Card}[\mathcal{Q}^*]$) :

$$\mathbf{y} = \mathbf{A}\mathbf{x}^* = \mathbf{A}_{\mathcal{Q}^*}\mathbf{t} \quad (3.5)$$

où \mathbf{t} désigne la restriction du vecteur \mathbf{x}^* au support \mathcal{Q}^* (\mathbf{x}^* se décompose en $\mathbf{x}^* = \{\mathbf{t}, \mathbf{0}\}$ en réordonnant ses composantes) et $\|\mathbf{t}\|_0 \leq k$. On s'intéresse à la reconstruction exacte du support de \mathbf{x}^* à partir de \mathbf{y} en effectuant au plus k itérations d'OMP ou d'OLS. Quand \mathcal{Q}^* est le support de \mathbf{x}^* , *i.e.*, quand toutes les amplitudes t_i sont non nulles, la reconstruction exacte est obtenue en exactement k itérations lorsque les k atomes indicés par \mathcal{Q}^* sont successivement choisis. Notons que comme OMP et OLS sont des algorithmes *orthogonaux*, le même atome ne peut pas être sélectionné plusieurs fois, et la sélection de tous les atomes de \mathcal{Q}^* implique l'annulation du résidu : $\mathbf{P}_{\mathcal{Q}^*}^\perp \mathbf{y} = \mathbf{0}$. C'est une différence essentielle avec *Matching Pursuit*, où un même atome peut être sélectionné plusieurs fois et l'annulation du résidu ne se produit plus en un nombre fini d'itérations. Dans ce cas, la reconstruction exacte prend un sens différent : il s'agit de garantir la sélection d'un atome de \mathcal{Q}^* à chaque itération $k \geq 0$ et l'annulation asymptotique du résidu à une vitesse exponentielle [GRI 06].

Dans la littérature, l'analyse des algorithmes gloutons orthogonaux en au plus k itérations est nommée *k-step analysis* [DON 08] : « *any vector with at most k nonzeros can be recovered from the related noise-free observation in at most k iterations* ». C'est ce type d'analyse que je vais adopter. Par abus de langage, la terminologie « analyse en k itérations » couvre aussi le cas pathologique où le support de \mathbf{x}^* est strictement inclus dans \mathcal{Q}^* , et moins de k itérations sont nécessaires à sa reconstruction. Il existe une autre approche, suivie par Foucart [FOU 13a], pour analyser les algorithmes gloutons orthogonaux et en particulier OLS. Son principe est de minorer la décroissance du résidu $\mathcal{E}_{\mathcal{Q}} - \mathcal{E}_{\mathcal{Q} \cup \{\ell\}}$. Cette approche permet de façon élégante d'analyser OMP, OLS et la forme faible d'OMP dans un cadre unifié (sous hypothèse RIP). Contrairement à l'approche présentée ici, ce n'est pas une analyse en k itérations car la reconstruction du support \mathcal{Q}^* n'est pas garantie après k itérations, mais en un nombre d'itérations proportionnel à k .

3.2.2 Extension bruitée

Le modèle (3.5) n'est pas réaliste car en pratique, les données sont bruitées et bien souvent, le signal \mathbf{x}^* inconnu n'est pas strictement parcimonieux : \mathbf{x}^* comporte quelques variables x_i^* de grande amplitude et beaucoup de variables non nulles de très faible amplitude (parcimonie au sens faible). Dans la littérature, les conditions de reconstruction exacte pour le cas bruité et/ou pour un signal \mathbf{x}^* parcimonieux au sens faible sont souvent définies comme une extension des conditions dédiées au problème (3.5). En notant $\hat{\mathbf{x}}$ le signal parcimonieux estimé et $\mathbf{x}_{\mathcal{Q}^*}^*$ la version seuillée de \mathbf{x}^* en ne gardant que les plus grandes amplitudes indicées par \mathcal{Q}^* , l'erreur quadratique $\|\hat{\mathbf{x}} - \mathbf{x}^*\|_2$ ou $\|\hat{\mathbf{x}} - \mathbf{x}_{\mathcal{Q}^*}^*\|_2$ est majorée par une expression qui fait intervenir la variance du bruit et le terme $\mathbf{A}(\mathbf{x} - \mathbf{x}_{\mathcal{Q}^*}^*)$ lié aux faibles amplitudes [CAI 11, BEN 10, FOU 13a]. L'analyse proposée ci-dessous peut également être étendue au cas bruité ou pour un signal faiblement parcimonieux.

3.2.3 Hypothèses de travail

Aucune hypothèse particulière n'est émise sur le dictionnaire \mathbf{A} comme une faible cohérence mutuelle [DON 06, BEN 10, CAI 11] ou comme la condition RIP (*Restricted Isometry Property*) [DAV 10, LIU 12, MO 12, FOU 13a, DIN 13]. Ces cas sont abondamment traités dans la littérature, et l'objectif principal est ici d'analyser des problèmes pour lesquels les garanties habituelles de reconstruction exacte ne sont pas satisfaites sans imposer d'hypothèse forte sur la structure du dictionnaire.

Pour des raisons qui seront éclaircies par la suite (§ 3.5), nous serons amenés à faire l'hypothèse que les $q < k$ premières itérations d'OMP/OLS ont toutes sélectionné un atome de \mathcal{Q}^* , et à définir des conditions *partielles* de reconstruction de \mathcal{Q}^* en k itérations. Ce sont des conditions plus faibles que les conditions suffisantes à l'itération $q = 0$. D'autres travaux de la littérature récente de l'échantillonnage compressé émettent cette hypothèse de connaissance partielle du support, où la partie connue du support peut éventuellement contenir des erreurs. Je renvoie le lecteur aux références [KAR 12] pour une analyse partielle d'OMP sous hypothèse RIP, à [VAS 10] pour le problème de relaxation convexe ℓ_2 - ℓ_1 et à [FRI 12] pour un problème de relaxation non convexe. L'hypothèse d'une connaissance partielle accessible du support se justifie lorsqu'on doit traiter des données séquentielles $\mathbf{y}(t)$, et que les atomes liés au signal $\mathbf{y}(t)$ à un instant t sont approximativement les mêmes que les atomes trouvés à l'instant précédent $t - 1$. Ces différents travaux ont établi des conditions suffisantes pour la reconstruction exacte beaucoup plus faibles que les conditions standard (sans connaissance *a priori* du support) en faisant l'hypothèse que les erreurs sur la partie connue du support sont limitées.

Dans notre travail [SOU 13b], la motivation de l'analyse partielle est de pouvoir discriminer les performances d'OMP et d'OLS en termes de garantie de reconstruction exacte : nous verrons que ces conditions sont identiques à l'itération $q = 0$ (sans connaissance partielle) et qu'elles diffèrent pour $q \geq 1$.

3.3 Différence entre OMP et OLS : considérations géométriques

Les algorithmes OMP et OLS ont été définis au chapitre 1 à la page 42. J'ai adopté un point de vue de type « optimisation » qui consiste à présenter OLS (et OMP) comme une heuristique pour minimiser l'erreur quadratique $\mathcal{E}_{\mathcal{Q}}$: à chaque itération d'OLS, l'atome sélectionné est celui qui conduit à la plus forte décroissance de l'erreur d'approximation $\mathcal{E}_{\mathcal{Q} \cup \{i\}}$ [CHE 89, MIL 02, FOU 13a]. Un autre point de vue est d'interpréter OLS en termes de projections orthogonales [NAT 95, COT 99, REB 02] ce qui rend la structure d'OLS très proche de celle d'OMP.

Le vecteur \mathbf{y} et les atomes du dictionnaire sont projetés sur le sous-espace $\text{span}(\mathbf{A}_{\mathcal{Q}})^\perp$ orthogonal aux atomes déjà sélectionnés. Pour OMP, le nouvel atome sélectionné est celui dont la projection est la plus corrélée avec le résidu $\mathbf{r}_{\mathcal{Q}}$. Pour OLS, le nouvel atome est celui dont la projection *normalisée* est la plus corrélée avec $\mathbf{r}_{\mathcal{Q}}$.

Formellement, on introduit la notation

$$\tilde{\mathbf{a}}_i \triangleq \mathbf{P}_{\mathcal{Q}}^\perp \mathbf{a}_i \quad (3.6)$$

pour les atomes projetés sur $\text{span}(\mathbf{A}_{\mathcal{Q}})^\perp$ où par simplicité, la dépendance par rapport à \mathcal{Q} est omise. Notons que $\tilde{\mathbf{a}}_i = \mathbf{0}$ pour $i \in \mathcal{Q}$. De plus, on définit les projections normalisées par :

$$\tilde{\mathbf{b}}_i = \begin{cases} \tilde{\mathbf{a}}_i / \|\tilde{\mathbf{a}}_i\|_2 & \text{si } \tilde{\mathbf{a}}_i \neq \mathbf{0}, \\ \mathbf{0} & \text{sinon.} \end{cases} \quad (3.7)$$

Avec ces notations, les règles de sélection d'atomes pour OMP et OLS s'écrivent :

$$\ell^{\text{OMP}} \in \arg \max_{i \notin \mathcal{Q}} |\langle \mathbf{r}_{\mathcal{Q}}, \tilde{\mathbf{a}}_i \rangle| \quad (3.8)$$

$$\ell^{\text{OLS}} \in \arg \max_{i \notin \mathcal{Q}} |\langle \mathbf{r}_{\mathcal{Q}}, \tilde{\mathbf{b}}_i \rangle| \quad (3.9)$$

(ce résultat est trivial pour OMP, je renvoie le lecteur à [REB 02] pour OLS). Autrement dit, OMP maximise le produit scalaire entre $\tilde{\mathbf{a}}_i$ et $\mathbf{r}_{\mathcal{Q}}$ alors qu'OLS minimise l'angle entre les deux vecteurs. Cette différence est illustrée sur la figure 3.1 (voir également l'illustration graphique du rapport technique de Blumensath et Davies [BLU 07]).

Compte tenu de la symétrie entre les expressions (3.8) et (3.9), on adoptera la notation unifiée :

$$\tilde{\mathbf{c}}_i \triangleq \begin{cases} \tilde{\mathbf{a}}_i & \text{pour OMP,} \\ \tilde{\mathbf{b}}_i & \text{pour OLS.} \end{cases} \quad (3.10)$$

Les algorithmes OMP et OLS sont finalement réécrits de façon générique (algorithme noté Oxx) dans le tableau 3.1 en utilisant :

$$\ell^{\text{Oxx}} \in \arg \max_{i \notin \mathcal{Q}} |\langle \mathbf{r}_{\mathcal{Q}}, \tilde{\mathbf{c}}_i \rangle|. \quad (3.11)$$

Comme on s'intéresse à une analyse en k itérations, Oxx est arrêté si le résidu est nul ou lorsque k itérations ont été effectuées. Obtenir un résidu $\mathbf{r}_{\mathcal{Q}}$ nul en au plus k itérations n'implique pas nécessairement que la représentation $\mathbf{y} = \mathbf{A}_{\mathcal{Q}^*} \mathbf{t}$ a été trouvée. Le vecteur \mathbf{y} peut en effet posséder des représentations multiples $\mathbf{y} = \mathbf{A}_{\mathcal{Q}^*} \mathbf{t} = \mathbf{A}_{\mathcal{Q}'^*} \mathbf{t}'$. L'hypothèse $\text{spark}(\mathbf{A}) > 2k$ est une condition nécessaire et suffisante pour garantir l'unicité de toute représentation avec au plus k atomes [DON 03]. Ainsi, lorsque $\text{spark}(\mathbf{A}) > 2k$, la sélection d'un mauvais atome ($i \notin \mathcal{Q}^*$) par Oxx rend impossible la reconstruction d'un résidu nul en k itérations [TRO 04].

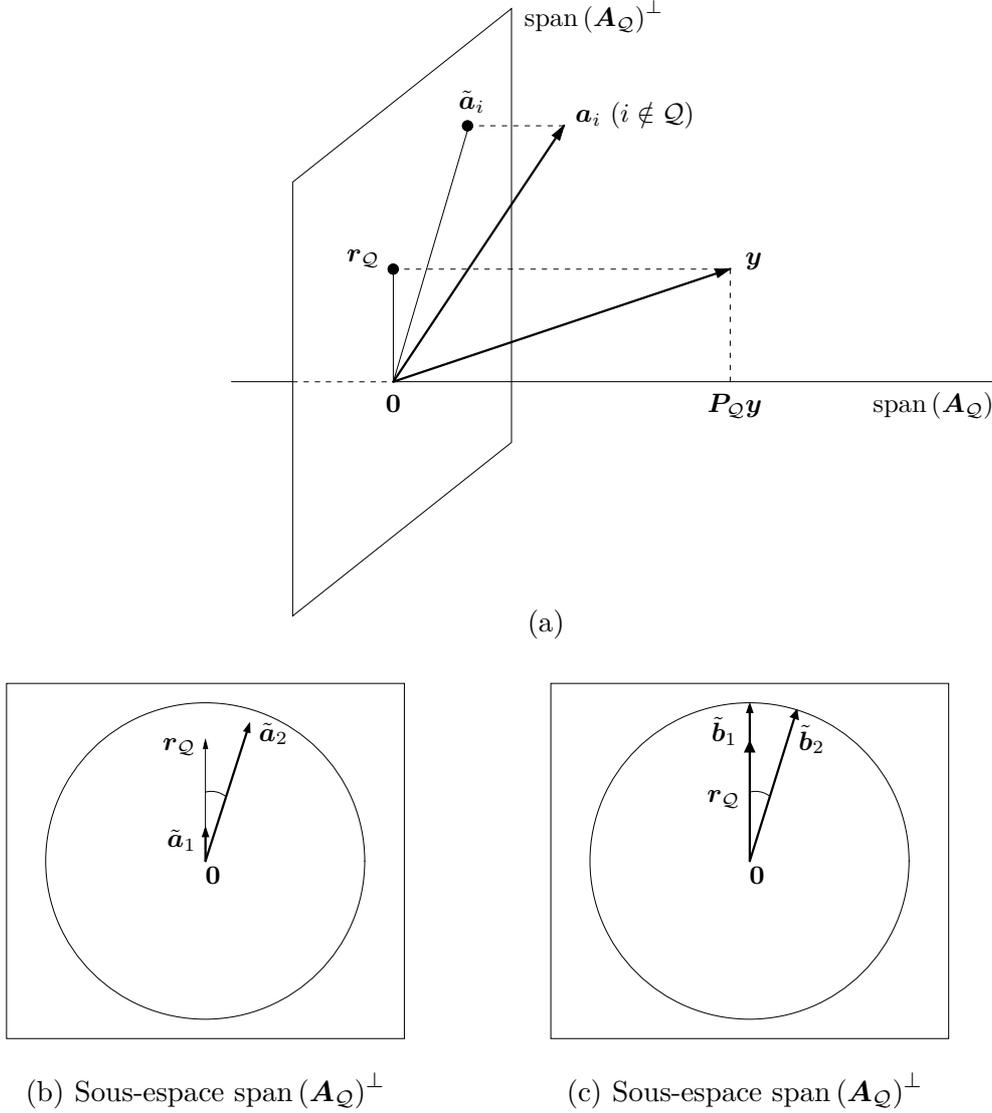


FIGURE 3.1 – Illustration graphique des algorithmes OMP et OLS. (a) La décomposition orthogonale des données \mathbf{y} sur les sous-espaces $\text{span}(\mathbf{A}_Q)$ et $\text{span}(\mathbf{A}_Q)^\perp$ produit l'approximation $P_Q \mathbf{y}$ des données et le résidu d'approximation $\mathbf{r}_Q = P_Q^\perp \mathbf{y}$. Les atomes \mathbf{a}_i du dictionnaire se projettent en $\tilde{\mathbf{a}}_i = P_Q^\perp \mathbf{a}_i$ avec $\tilde{\mathbf{a}}_i = \mathbf{0}$ pour $i \in Q$. (b) Vue du plan $\text{span}(\mathbf{A}_Q)^\perp$ et de la boule unité ($\|\tilde{\mathbf{a}}_i\|_2 \leq \|\mathbf{a}_i\|_2 = 1$). OMP sélectionne l'atome projeté dont le produit scalaire avec le résidu est maximal (sélection de l'atome \mathbf{a}_2). OLS sélectionne l'atome projeté dont l'angle avec le résidu est minimal (atome \mathbf{a}_1). (c) Visualisation des atomes projetés normalisés $\tilde{\mathbf{b}}_i$. OLS sélectionne l'atome qui maximise $|\langle \mathbf{r}_Q, \tilde{\mathbf{b}}_i \rangle|$ (c'est-à-dire \mathbf{a}_1 sur cette figure). Pour un dictionnaire quasi-orthogonal (hypothèse de faible cohérence mutuelle ou d'isométrie restreinte), la situation (b,c) ne peut se produire car les vecteurs $\tilde{\mathbf{a}}_i$ sont tous de norme quasiment unitaire. Dans ce cas, $\tilde{\mathbf{b}}_i$ et $\tilde{\mathbf{a}}_i$ sont quasiment confondus et il n'y a pas de différence notable entre les comportements d'OMP et d'OLS [FOU 13a].

Entrées : \mathbf{A} , \mathbf{y} , k et le support $\mathcal{Q}_{\text{init}}$.
Faire $\mathcal{Q} = \mathcal{Q}_{\text{init}}$. Calculer $\mathbf{r}_{\mathcal{Q}} = \mathbf{P}_{\mathcal{Q}}^{\perp} \mathbf{y}$ et $\tilde{\mathbf{c}}_i$ par (3.10) pour $i \notin \mathcal{Q}$.
Tant que ($\text{Card}[\mathcal{Q}] < k$) et ($\mathbf{r}_{\mathcal{Q}} \neq \mathbf{0}$),
[1. Sélection du nouvel atome ℓ]
Faire $\mathcal{Q} = \mathcal{Q} \cup \{\ell\}$ où $\ell \in \arg \max_{i \notin \mathcal{Q}} \langle \mathbf{r}_{\mathcal{Q}}, \tilde{\mathbf{c}}_i \rangle $.
[2. Mise à jour du résidu et des atomes projetés]
Calculer $\mathbf{r}_{\mathcal{Q}} = \mathbf{P}_{\mathcal{Q}}^{\perp} \mathbf{y}$, et $\tilde{\mathbf{c}}_i$ par (3.10) pour $i \notin \mathcal{Q}$.
Fin Tant que
Sortie : support \mathcal{Q}

TABLE 3.1 – Présentation unifiée des algorithmes OMP et OLS (algorithme générique noté Oxx) pour un dictionnaire supposé normalisé ($\|\mathbf{a}_i\|_2 = 1$). Par défaut, $\mathcal{Q}_{\text{init}} = \emptyset$. Dans ce cas, la première ligne devient $\mathbf{r}_{\mathcal{Q}} = \mathbf{y}$ et $\tilde{\mathbf{c}}_i = \mathbf{a}_i$. L'algorithme est arrêté lorsque le résidu d'approximation est nul ou quand k itérations ont été exécutées.

3.4 Analyse au pire cas

Nous abordons la reconstruction exacte du support \mathcal{Q}^* sous l'angle d'une analyse au pire cas. Les conditions proposées sont valables pour tous les vecteurs $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$. Elles dépendent donc du dictionnaire \mathbf{A} et du support inconnu \mathcal{Q}^* . Je reviendrai dans la conclusion de ce chapitre et dans le chapitre 4 sur les limitations de l'analyse au pire cas.

Nous parlerons de *succès* d'Oxx avec un vecteur $\mathbf{y} = \mathbf{A}_{\mathcal{Q}^*} \mathbf{t}$ en entrée lorsqu'aucun mauvais atome \mathbf{a}_j , $j \notin \mathcal{Q}^*$ n'est sélectionné. Compte tenu des conditions d'arrêt choisies pour Oxx, le résidu s'annule alors après au plus k itérations et le support obtenu en sortie d'Oxx vérifie $\mathcal{Q} \subset \mathcal{Q}^*$. Sous réserve que $\mathbf{A}_{\mathcal{Q}^*}$ soit de rang plein et que les amplitudes t_i sont toutes non nulles, \mathcal{Q} coïncide avec \mathcal{Q}^* . Au contraire, un *échec* signifie la sélection d'un mauvais atome au cours d'une itération de l'algorithme. On considère en particulier qu'un échec se produit lorsque la règle de sélection d'Oxx produit des solutions multiples comprenant un mauvais atome :

$$\max_{i \in \mathcal{Q}^* \setminus \mathcal{Q}} |\langle \mathbf{r}_{\mathcal{Q}}, \tilde{\mathbf{c}}_i \rangle| = \max_{j \notin \mathcal{Q}^*} |\langle \mathbf{r}_{\mathcal{Q}}, \tilde{\mathbf{c}}_j \rangle|. \quad (3.12)$$

Pour OMP, la condition nécessaire et suffisante pour la reconstruction du support \mathcal{Q}^* est la condition ERC (*Exact Recovery condition*) établie par Tropp.

Théorème 1. [TRO 04, Ths. 3.1 et 3.10]. *Si $\mathbf{A}_{\mathcal{Q}^*}$ est de rang plein et*

$$\max_{j \notin \mathcal{Q}^*} \|\mathbf{A}_{\mathcal{Q}^*}^{\dagger} \mathbf{a}_j\|_1 < 1, \quad \text{ERC}(\mathbf{A}, \mathcal{Q}^*)$$

alors un succès se produit avec OMP pour toute entrée $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^})$. De plus, lorsque $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ n'est pas satisfaite, il existe $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ pour laquelle un mauvais atome est sélectionné à la première itération d'OMP. Quand $\text{spark}(\mathbf{A}) > 2k$, OMP ne peut donc pas retrouver les k atomes qui constituent la représentation (unique) de \mathbf{y} .*

Ce résultat est plus fort que d'autres conditions établies sous l'hypothèse d'isométrie restreinte (*Restricted Isometry Property*) [DAV 10, HUA 11, LIU 12, MO 12, WAN 12b]. En effet,

ces résultats sont également des conditions suffisantes de reconstruction exacte au pire cas. Comme la condition ERC est aussi une condition nécessaire, ERC est forcément vérifiée lorsque ces conditions le sont.

Nous nous sommes posé les questions suivantes :

1. Peut-on étendre le résultat du théorème 1 à OLS ?
2. Y a-t-il une implication entre les conditions de reconstruction exacte au pire cas liées à OMP et OLS ?

La réponse est positive pour la première question et négative pour la deuxième car les conditions au pire cas coïncident pour OMP et OLS (§ 3.5). C'est ce qui nous a poussé à élaborer une analyse plus poussée des deux algorithmes sous l'hypothèse de connaissance partielle du support \mathcal{Q}^* (§ 3.6).

3.5 Reconstruction exacte par OLS

Nous avons établi un résultat analogue au théorème 1 pour OLS.

Théorème 2. [SOU 13b, Th. 2] *Si $\mathbf{A}_{\mathcal{Q}^*}$ est de rang plein et $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ est satisfaite, alors un succès se produit avec OLS pour toute entrée $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$. De plus, lorsque $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ n'est pas satisfaite, il existe $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ pour laquelle un mauvais atome est sélectionné à la première itération d'OLS. Quand $\text{spark}(\mathbf{A}) > 2k$, OLS ne peut donc pas retrouver les k atomes qui constituent la représentation (unique) de \mathbf{y} .*

La condition nécessaire est un résultat immédiat car la toute première itération d'OLS coïncide avec celle d'OMP et $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ est une condition nécessaire à la première itération pour OMP : quand $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ n'est pas vérifiée, OLS échoue avec toute donnée \mathbf{y} pour laquelle OMP échoue à la première itération. La preuve de la condition suffisante du théorème 2 est en revanche très technique. Elle repose sur l'analyse de reconstruction exacte de la section 3.6 en initialisant OLS par un support partiellement connu $\mathcal{Q} \subset \mathcal{Q}^*$ et sur une relation récursive entre les conditions de reconstruction obtenues pour des supports \mathcal{Q} et \mathcal{Q}' emboîtés ($\mathcal{Q} \subset \mathcal{Q}' \subset \mathcal{Q}^*$).

Comme $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ est une condition nécessaire à la première itération pour OMP comme pour OLS, aucune discrimination des deux algorithmes n'est possible à la première itération. Mais que se passe-t-il aux itérations suivantes, lorsque $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ n'est pas satisfaite mais néanmoins, les $q < k$ premières itérations de Oxx ont sélectionné des atomes de $\mathcal{Q} \subset \mathcal{Q}^*$? C'est la question (outre la preuve du théorème 2) qui a motivé l'analyse partielle décrite ci-dessous.

3.6 Reconstruction exacte par OMP et OLS à partir d'une connaissance partielle du support

Le mot partiel a parfois un sens différent dans la littérature : dans [DON 06, VAI 13], il désigne la reconstruction d'un support \mathcal{Q} inclus dans \mathcal{Q}^* . Ici, \mathcal{Q} est une information connue *a priori*. Il s'agit d'établir des conditions qui garantissent la reconstruction exacte de \mathcal{Q}^* pour tout $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ en supposant, au choix, que :

1. les $q < k$ premières itérations de OMP/OLS ont réussi, et $k - q$ itérations restent à effectuer.
2. Oxx est initialisé avec le support \mathcal{Q} (au lieu du support vide) et on s'intéresse à la reconstruction de $\mathcal{Q}^* \setminus \mathcal{Q}$ en $k - q$ itérations.

Les conditions obtenues sont nommées ERC-OMP et ERC-OLS. Comme l'analyse d'OMP et OLS est parallèle, on notera de façon générique ERC-Oxx.

3.6.1 Conditions pour un support initial \mathcal{Q} fixé

Les conditions de reconstruction exacte pour OMP et OLS sont définies pour $\mathcal{Q} \subset \mathcal{Q}^*$ par :

$$\max_{j \notin \mathcal{Q}^*} \|\tilde{\mathbf{C}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{c}}_j\|_1 < 1 \quad \text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$$

où la matrice $\tilde{\mathbf{C}}_{\mathcal{Q}^* \setminus \mathcal{Q}} = \{\tilde{\mathbf{c}}_i, i \in \mathcal{Q}^* \setminus \mathcal{Q}\}$ de taille $m \times (k - q)$ rassemble les atomes projetés $\tilde{\mathbf{a}}_i$ pour OMP et les atomes $\tilde{\mathbf{b}}_i$ projetés et normalisés pour OLS. Cette définition est une extension de la condition de Tropp ERC($\mathbf{A}, \mathcal{Q}^*$) correspondant au cas $\mathcal{Q} = \emptyset$.

La reconstruction exacte par OMP et OLS après q itérations réussies signifie qu'à partir de l'itération $q + 1$, des atomes de $\mathcal{Q}^* \setminus \mathcal{Q}$ sont sélectionnés jusqu'à annulation du résidu. Le résultat prend la forme générique suivante.

Théorème 3. [SOU 13b, Ths. 3 et 4] *Supposons que $\mathbf{A}_{\mathcal{Q}^*}$ est de rang plein. Si Oxx avec $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ en entrée sélectionne $\mathcal{Q} \subset \mathcal{Q}^*$ et si ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) est vérifiée, alors un succès se produit en au plus k itérations.*

Considérons à présent le deuxième point de vue où Oxx est initialisé par le support \mathcal{Q} au lieu du support vide. Le théorème 3 devient :

Théorème 4. [HER 13, Th. 1] *Supposons que $\mathbf{A}_{\mathcal{Q}^*}$ est de rang plein. Si ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) est vérifiée, alors un succès se produit en au plus $k - q$ itérations en initialisant Oxx avec le support \mathcal{Q} quelquesoit $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$. Inversement, si ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) n'est pas vérifiée, alors il existe $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ pour lequel Oxx avec \mathcal{Q} pour initialisation sélectionne un mauvais atome $j \notin \mathcal{Q}^*$ à la première itération.*

Le lecteur aura probablement remarqué que dans le théorème 4, ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) est une condition *nécessaire* et suffisante de reconstruction exacte alors que dans le théorème 3, la même condition n'est que suffisante. Il y a là une subtilité car pour établir une condition nécessaire à la q -ème itération dans le cadre du théorème 3, il faut prouver l'existence d'un vecteur \mathbf{y} pour lequel \mathcal{Q} est atteint en q itérations (puis, l'itération $q + 1$ échoue). Pour OLS, tout support \mathcal{Q} est effectivement atteignable (cf. § 3.9). On peut alors en déduire que ERC-OLS($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) est une condition nécessaire.

Théorème 5. [SOU 13b, Th. 4] *Soit $\mathcal{Q} \subset \mathcal{Q}^*$ un support de cardinalité q . Supposons que $\mathbf{A}_{\mathcal{Q}^*}$ est de rang plein et que $\text{spark}(\mathbf{A}) \geq q + 2$. Si ERC-OLS($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) n'est pas vérifiée, il existe $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ pour lequel OLS sélectionne \mathcal{Q} durant les q premières itérations, puis un mauvais atome $j \notin \mathcal{Q}^*$ à l'itération $q + 1$.*

En revanche, certains supports peuvent ne pas être atteints par OMP (cf. § 3.8) et on ne peut rien conclure sur le caractère nécessaire de ERC-OMP($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) dans le cadre du théorème 3 sans émettre d'hypothèse supplémentaire [SOU 13b, Th. 5].

3.6.2 Conditions pour une itération q fixée

La condition ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) dédiée à un support \mathcal{Q} particulier (connu) est étendue en une condition qui dépend de \mathcal{Q}^* et de l'itération q . La nouvelle condition ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, q$) est définie comme l'intersection des conditions ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) pour tous $\mathcal{Q} \subset \mathcal{Q}^*$ de cardinalité q :

$$\max_{\substack{\mathcal{Q} \subset \mathcal{Q}^* \\ \text{Card}[\mathcal{Q}] = q}} \max_{j \notin \mathcal{Q}^*} \|\tilde{\mathbf{C}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{c}}_j\|_1 < 1 \quad \text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, q)$$

La condition suffisante du théorème 3 se généralise ainsi :

Corollaire 1. [SOU 13b, Cor. 1] *Supposons que $\mathbf{A}_{\mathcal{Q}^*}$ est de rang plein. Si Oxx avec $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ en entrée sélectionne des atomes de \mathcal{Q}^* au cours des q premières itérations et si $\text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, q)$ est satisfaite, alors un succès d'Oxx se produit en au plus k itérations.*

Les théorèmes 4 et 5 se généralisent de la même façon : $\text{ERC-OMP}(\mathbf{A}, \mathcal{Q}^*, q)$ est une condition suffisante de reconstruction exacte par OMP et $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, q)$ est nécessaire et suffisante pour OLS.

3.6.3 Cas particulier $q = 0$

Comme $\text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ coïncident avec $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ pour $\mathcal{Q} = \emptyset$, les conditions nécessaires et suffisantes des théorèmes 1 (OMP) et 2 (OLS) apparaissent comme le cas particulier du théorème 4 à l'itération 0.

Techniquement, le résultat clé permettant de prouver que $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ est une condition suffisante de reconstruction exacte pour OLS relève d'un raisonnement d'emboîtement : $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ implique $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}')$ pour $\mathcal{Q} \subset \mathcal{Q}' \subset \mathcal{Q}^*$. Plus précisément, la condition $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ a été définie en suivant le raisonnement de Tropp [TRO 04, Th. 3.1] de façon à garantir la sélection d'un bon atome $\ell \in \mathcal{Q}^*$ à l'itération $q + 1$. Pour pouvoir conclure sur la sélection des atomes de \mathcal{Q}^* aux itérations suivantes $q + 2, \dots, k$, il s'agit de prouver que $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}')$ est vérifiée pour $\text{Card}[\mathcal{Q}'] = q + 2, \dots, k$. La propriété d'emboîtement des conditions $\text{ERC-OMP}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ est triviale pour OMP mais elle requiert une analyse sophistiquée pour $\text{ERC-OLS}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ à cause des normalisations d'atomes $\tilde{\mathbf{b}}_i = \tilde{\mathbf{a}}_i / \|\tilde{\mathbf{a}}_i\|_2$ qui interviennent dans la règle de sélection (3.11).

3.7 Evaluation empirique des conditions partielles de reconstruction

Les conditions ERC-Oxx sont hautement techniques, et (reconnaissons le!) difficilement exploitables en pratique car leur évaluation requiert de nombreux calculs de projection orthogonale et des inversions de systèmes linéaires. Dans [HER 13], nous avons établi des conditions suffisantes beaucoup plus simples à exploiter, portant sur la cohérence mutuelle du dictionnaire :

Théorème 6. [HER 13, Th. 3] *Si Oxx avec $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ en entrée sélectionne des atomes de \mathcal{Q}^* au cours des q premières itérations et si*

$$\mu < \frac{1}{2k - q - 1}, \tag{3.13}$$

alors un succès se produit pour Oxx en au plus k itérations. Réciproquement, il existe un dictionnaire \mathbf{A} et une combinaison linéaire de k atomes de \mathbf{A} telle que (i) $\mu = 1/(2k - q - 1)$ et (ii) Oxx sélectionne des atomes de \mathcal{Q}^ pendant les q premières itérations puis un mauvais atome à l'itération $q + 1$.*

Ce résultat étend la condition bien connue à l'itération $q = 0$ [TRO 04, CAI 10] qui s'écrit $\mu < 1/(2k - 1)$. La condition (3.13), bien que très simple à évaluer, est une condition suffisante (beaucoup) plus forte que les conditions $\text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ ou $\text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, q)$. De plus, elle ne permet pas de discriminer OMP et OLS.

Nous avons proposé une évaluation empirique des conditions $\text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ dans le but d'établir cette discrimination [SOU 13b, Sec. IV]. Nos premières conclusions montrent qu'**aucun des deux algorithmes n'est uniformément meilleur que l'autre à l'itération $q \geq 1$** :

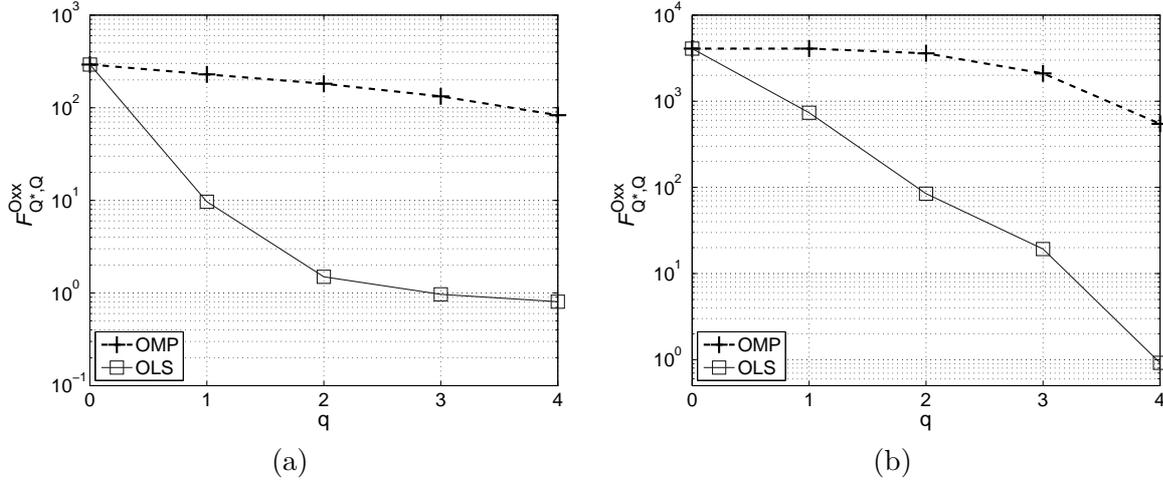


FIGURE 3.2 – Courbes représentant $F_{Q^*, Q}^{Oxx} \triangleq \max_{j \notin Q^*} \|\tilde{C}_{Q^* \setminus Q}^\dagger \tilde{c}_j\|_1$ en fonction de $q = \text{Card}[Q]$ pour un problème de déconvolution avec une réponse impulsionnelle gaussienne. Q^* est fixé et le support partiel $Q \subset Q^*$ est incrémenté progressivement. Q^* et Q sont respectivement formés des $k = 5$ et des $q < k$ premiers échantillons. (a) Réponse impulsionnelle gaussienne de largeur $\sigma = 50$, dictionnaire de taille 3000×2710 . (b) $\sigma = 10$, dictionnaire de taille 1000×4940 .

- il n'y a pas d'implication automatique, ni dans un sens ni dans l'autre, entre les conditions ERC-OMP et ERC-OLS : il existe des contre-exemples (*e.g.*, des triplets (\mathbf{A}, Q^*, Q)) pour lesquels ERC-OMP(\mathbf{A}, Q^*, Q) est satisfait et ERC-OLS(\mathbf{A}, Q^*, Q) ne l'est pas, et *vice versa*.
- il existe des triplets (\mathbf{A}, Q^*, q) pour lesquels ERC-OMP(\mathbf{A}, Q^*, q) est satisfait et ERC-OLS(\mathbf{A}, Q^*, q) ne l'est pas, et *vice versa*.

Pour fixer les idées, si ERC-OMP(\mathbf{A}, Q^*, Q) est vérifiée et ERC-OLS(\mathbf{A}, Q^*, Q) ne l'est pas, il existe un vecteur $\mathbf{y} \in \text{span}(\mathbf{A}_{Q^*})$ pour lequel OLS, initialisé par Q , sélectionne un mauvais atome à la première itération (théorème 4). Toujours d'après le théorème 4, OMP reconstruit le support de la représentation de \mathbf{y} car ERC-OMP(\mathbf{A}, Q^*, Q) est satisfaite par hypothèse. Le même raisonnement est valable en inversant les rôles d'OMP et OLS.

En résumé, une discrimination uniforme systématique (pour tout dictionnaire \mathbf{A}) d'OMP et OLS n'est pas possible. Une étude plus poussée basée sur l'évaluation numérique des conditions ERC-Oxx(\mathbf{A}, Q^*, Q) fournit **des tendances claires de discrimination pour différents types de problèmes avec des dictionnaires corrélés**. Ces simulations consistent à établir des « diagrammes de transition de phase » [DON 08, MAL 10, DON 12] en : (1) tirant aléatoirement le dictionnaire \mathbf{A} et le support Q^* ; (2) et en incrémentant le support $Q \subset Q^*$ d'un élément à la fois jusqu'à ce que la condition ERC-Oxx(\mathbf{A}, Q^*, Q) soit satisfaite. Les conclusions de cette étude sont que :

- pour les dictionnaires peu corrélés, il y a très peu de différence entre les conditions liées à OMP et à OLS : la reconstruction garantie est atteinte pour une valeur moyenne identique de $q = \text{Card}[Q]$.
- en revanche, **pour les dictionnaires très corrélés, on observe une nette différence en faveur d'OLS** : la reconstruction garantie de Q^* est obtenue pour une valeur plus faible de $q = \text{Card}[Q]$ pour OLS. De plus, ERC-OMP(\mathbf{A}, Q^*, Q) n'est parfois même pas satisfaite pour $q = k - 1$.

La figure 3.2 illustre cette tendance pour le problème de déconvolution impulsionnelle $\mathbf{y} = \mathbf{h} \star \mathbf{x}^*$

avec un filtre gaussien \mathbf{h} . La quantité représentée en ordonnée est le facteur

$$F_{\mathcal{Q}^*, \mathcal{Q}}^{\text{Oxx}} \triangleq \max_{j \notin \mathcal{Q}^*} \|\tilde{\mathbf{C}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{c}}_j\|_1 \quad (3.14)$$

qui intervient dans la condition ERC-Oxx(\mathbf{A} , \mathcal{Q}^* , \mathcal{Q}). Rappelons que ERC-Oxx(\mathbf{A} , \mathcal{Q}^* , \mathcal{Q}) est satisfaite si ce facteur est inférieur à 1. Les valeurs de $F_{\mathcal{Q}^*, \mathcal{Q}}^{\text{OMP}}$ sont très largement supérieures à 1 même pour $q = k - 1$ alors que celles de $F_{\mathcal{Q}^*, \mathcal{Q}}^{\text{OLS}}$ sont plus proches de 1, toujours inférieures à 1 pour $q = k - 1$ et parfois inférieures à 1 pour $q < k - 1$.

La déconvolution impulsionnelle par un filtre passe-bas est un problème réputé difficile à analyser. D'autres auteurs ont étudié ce problème en évaluant la condition de Tropp ERC(\mathbf{A} , \mathcal{Q}^*) à l'itération $q = 0$, et la condition ERC au sens faible²³ [GRI 08a, Lemme 3] :

$$\max_{j \notin \mathcal{Q}^*} \|\mathbf{A}_{\mathcal{Q}^*}^T \mathbf{a}_j\|_1 + \max_{i \in \mathcal{Q}^*} \|\mathbf{A}_{\mathcal{Q}^* \setminus \{i\}}^T \mathbf{a}_i\|_1 < 1. \quad \text{WERC}(\mathbf{A}, \mathcal{Q}^*)$$

WERC(\mathbf{A} , \mathcal{Q}^*) est une condition suffisante plus forte que ERC(\mathbf{A} , \mathcal{Q}^*), garantissant la reconstruction du support \mathcal{Q}^* à l'itération 0. Elle est plus simple à évaluer car elle ne fait intervenir que des produits scalaires entre atomes du dictionnaire, et aucune inversion de système. Différentes études [DOS 05b, DOS 05a, DEN 09, LOR 09, LOR 11] ont montré que les conditions ERC ou WERC peuvent être vérifiées pour des problèmes de déconvolution impulsionnelle avec un filtre passe-bas sous réserve que les impulsions sont suffisamment espacées (autrement dit, qu'il y ait un faible recouvrement entre les motifs gaussiens résultant de la convolution $\mathbf{h} \star \mathbf{x}$). Notre analyse vise à explorer le cas général où on n'impose pas de distance minimale entre les impulsions, en supposant que les impulsions principales ont été correctement retrouvées dans les premières itérations de l'algorithme. Les simulations de la figure 3.2 concernent le cas extrême où les impulsions sont positionnées en des échantillons consécutifs. Les garanties de reconstruction exacte du support ne sont jamais atteintes pour OMP et atteintes à partir d'un nombre élevé d'itérations (pour q proche de k) pour OLS. Ce résultat (les conditions suffisantes sont peu souvent vérifiées) reflète la difficulté du problème, mais il montre néanmoins que pour des problèmes réputés difficiles, il est possible de dépasser les analyses traditionnelles qui font des hypothèses très restrictives sur les dictionnaires et de fournir des tendances sur le comportement et la discrimination des performances des algorithmes dans des situations plus difficiles.

3.8 Garantie de non-reconstruction pour OMP

Considérons le cas particulier $q = k - 1$. \mathcal{Q}^* s'écrit $\mathcal{Q}^* = \mathcal{Q} \cup \{i\}$. Pour les exemples de la figure 3.2, $F_{\mathcal{Q}^*, \mathcal{Q}}^{\text{OMP}}$ est toujours supérieur à 1, c'est-à-dire que la condition ERC-OMP(\mathbf{A} , \mathcal{Q}^* , \mathcal{Q}) n'est pas satisfaite : si \mathcal{Q} est atteint après $k - 1$ itérations, alors la sélection du dernier atome $i = \mathcal{Q}^* \setminus \mathcal{Q}$ est impossible. Au contraire, ERC-OLS(\mathbf{A} , \mathcal{Q}^* , \mathcal{Q}) est toujours satisfaite ce qui garantit la sélection du dernier atome par OLS.

Les deux théorèmes ci-dessous généralisent ce constat empirique. Ils constituent une **discrimination d'OMP et OLS à la dernière itération**.

Théorème 7. [SOU 13b, Th. 6] *Si $[\mathbf{A}_{\mathcal{Q}^*}, \mathbf{a}_j]$ est de rang plein pour tout $j \notin \mathcal{Q}^*$, alors ERC-OLS(\mathbf{A} , \mathcal{Q}^* , $k - 1$) est vraie : si les $k - 1$ premières itérations d'OLS sélectionnent des atomes de \mathcal{Q}^* , alors le dernier atome de \mathcal{Q}^* est retrouvé à la k -ème itération.*

²³. Condition *Weak ERC*, notée WERC(\mathbf{A} , \mathcal{Q}^*). Cette condition est également appelée « condition ERC de Neumann » dans [LOR 11].

Théorème 8. [SOU 13b, Th. 7] Si $\mathbf{A}_{\mathcal{Q}^*}$ est de rang plein et

$$\min_{\substack{\mathcal{Q} \subset \mathcal{Q}^* \\ \text{Card}[\mathcal{Q}] = k-1}} \left[\max_{j \notin \mathcal{Q}^*} \|\tilde{\mathbf{A}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{a}}_j\|_1 \right] \geq 1, \quad (\text{BRC-OMP}(\mathbf{A}, \mathcal{Q}^*))$$

alors \mathcal{Q}^* ne peut être reconstruit par OMP pour aucun vecteur $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$.

Le théorème 7 se comprend de façon très intuitive en considérant la définition d'OLS du point de vue de l'optimisation de l'erreur quadratique (tableau 1.3 page 42) : OLS sélectionne le nouvel atome conduisant à l'erreur quadratique la plus faible. Il est clair qu'à la k -ème itération, l'erreur quadratique obtenue avec le dernier atome $i = \mathcal{Q}^* \setminus \mathcal{Q}$ est nulle : $\mathcal{E}_{\mathcal{Q} \cup \{i\}} = \mathcal{E}_{\mathcal{Q}^*} = 0$ car $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$.

La condition de non-reconstruction (*BRC* pour *Bad Recovery Condition*) du théorème 8 est définie comme l'intersection des négations logiques des conditions ERC-OMP($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) pour tout $\mathcal{Q} \subset \mathcal{Q}^*$ de cardinalité $k-1$. C'est un résultat beaucoup plus fort qu'une condition nécessaire du type :

« il existe un vecteur $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ pour lequel \mathcal{Q}^* n'est pas reconstruit »

puisque BRC-OMP($\mathbf{A}, \mathcal{Q}^*$) signifie que :

« pour tout $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$, \mathcal{Q}^* n'est pas reconstruit ».

Plus précisément, un mauvais atome est sélectionné soit durant les $k-1$ premières itérations, soit à la k -ème itération si jamais les $k-1$ premières ont réussi, puisque ERC-OMP($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) est vérifié pour tout \mathcal{Q} de cardinalité $k-1$.

Nous avons montré dans [SOU 13b] que pour des dictionnaires corrélés (*e.g.*, pour le problème de la déconvolution impulsionnelle), la condition BRC-OMP($\mathbf{A}, \mathcal{Q}^*$) est très fréquemment satisfaite. Pour fixer les idées, voici un exemple à quatre atomes dans \mathbb{R}^3 où la non-reconstruction de \mathcal{Q}^* est assurée pour tout $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$. Pour cet exemple, \mathcal{Q}^* ne peut être atteint en deux itérations à partir d'aucun vecteur $\mathbf{y} \in \mathbb{R}^3$.

Exemple 1. [SOU 13b, Exemple 1] Soit le dictionnaire

$$\mathbf{A} = \begin{bmatrix} \cos \theta_1 & \cos \theta_1 & 0 & 0 \\ -\sin \theta_1 & \sin \theta_1 & \cos \theta_2 & \cos \theta_2 \\ 0 & 0 & \sin \theta_2 & -\sin \theta_2 \end{bmatrix}$$

avec $\mathcal{Q}^* = \{1, 2\}$. La valeur de $\theta_2 \in (0, \pi/2)$ est arbitrairement fixée. Quand $\theta_1 \neq 0$ est proche de 0, alors BRC-OMP($\mathbf{A}, \mathcal{Q}^*$) est vérifiée. De plus, \mathcal{Q}^* ne peut être atteint après deux itérations d'OMP pour aucun $\mathbf{y} \in \mathbb{R}^3$.

Démonstration. Je fournis une preuve géométrique qui s'appuie sur la figure 3.3 (voir [SOU 13b] pour une preuve formelle). Supposons qu'un atome de \mathcal{Q}^* a été sélectionné à la première itération (par exemple \mathbf{a}_1). Il s'agit de montrer que l'autre atome de \mathcal{Q}^* (\mathbf{a}_2) ne peut être sélectionné à la seconde itération. La règle de sélection du deuxième atome (3.8) fait intervenir les atomes projetés $\tilde{\mathbf{a}}_i = \mathbf{P}_{\{1\}}^\perp \mathbf{a}_i$ et le résidu $\mathbf{r}_{\{1\}} = \mathbf{P}_{\{1\}}^\perp \mathbf{y}$ situés dans le plan $\text{span}(\mathbf{a}_1)^\perp$. Quand θ_1 est faible, \mathbf{a}_1 et \mathbf{a}_2 sont (quasiment) colinéaires, donc $\tilde{\mathbf{a}}_2$ est de très faible norme. Au contraire, \mathbf{a}_3 et \mathbf{a}_4 qui sont (quasiment) orthogonaux à \mathbf{a}_1 , se projettent en des vecteurs de norme (quasiment) unité. Les vecteurs \mathbf{r} qui vérifient $|\langle \mathbf{r}, \tilde{\mathbf{a}}_2 \rangle| \geq |\langle \mathbf{r}, \tilde{\mathbf{a}}_3 \rangle|$ décrivent le cône de la figure 3.3 orienté dans

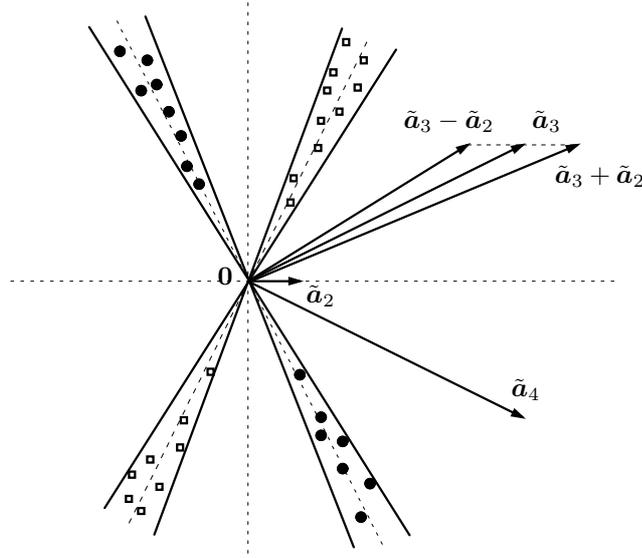


FIGURE 3.3 – Exemple 1 : dessin du plan $\text{span}(\mathbf{a}_1)^\perp$. La notation « $\tilde{\cdot}$ » se réfère implicitement au support $\mathcal{Q} = \{1\}$. Quand θ_1 est proche de 0, $\tilde{\mathbf{a}}_2$ est de très faible norme, et $\tilde{\mathbf{a}}_3$ et $\tilde{\mathbf{a}}_4$ sont des vecteurs de normé unité. Les inégalités $|\langle \mathbf{r}, \tilde{\mathbf{a}}_2 \rangle| \geq |\langle \mathbf{r}, \tilde{\mathbf{a}}_3 \rangle|$ et $|\langle \mathbf{r}, \tilde{\mathbf{a}}_2 \rangle| \geq |\langle \mathbf{r}, \tilde{\mathbf{a}}_4 \rangle|$ sont les équations des cônes dont les directions centrales sont respectivement orthogonales à $\tilde{\mathbf{a}}_3$ et $\tilde{\mathbf{a}}_4$ (lignes en pointillés). Comme ces deux cônes s'intersectent en $\mathbf{r} = \mathbf{0}$, OMP ne peut pas sélectionner \mathbf{a}_2 à la deuxième itération après avoir sélectionné \mathbf{a}_1 .

les directions sud-est/nord-ouest (résidus \mathbf{r} matérialisés par des disques). De même, le deuxième cône orienté dans les directions sud-ouest/nord-est caractérise l'inégalité $|\langle \mathbf{r}, \tilde{\mathbf{a}}_2 \rangle| \geq |\langle \mathbf{r}, \tilde{\mathbf{a}}_4 \rangle|$. Comme ces deux cônes s'intersectent en $\mathbf{r} = \mathbf{0}$, on déduit :

$$\forall \mathbf{r} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}, |\langle \mathbf{r}, \tilde{\mathbf{a}}_2 \rangle| < \max(|\langle \mathbf{r}, \tilde{\mathbf{a}}_3 \rangle|, |\langle \mathbf{r}, \tilde{\mathbf{a}}_4 \rangle|).$$

Donc \mathbf{a}_2 ne peut être sélectionné à la deuxième itération. \square

3.9 Transition de phase pour la non-reconstruction

La vision d'OMP et OLS comme algorithmes exploratoires est liée au problème d'optimisation (3.3) : les algorithmes gloutons explorent une séquence de supports \mathcal{Q} de taille croissante. La dénomination d'algorithmes exploratoires peut sembler abusive dans le cas d'algorithmes gloutons monodirectionnels puisque dans tous les cas, k itérations sont effectuées (selon les critères d'arrêt dans ce chapitre). On ne peut donc pas prétendre qu'un algorithme glouton monodirectionnel explore plus qu'un autre. En revanche, certains algorithmes explorent *mieux*.

L'exemple 1 ci-dessus et le théorème 8 montrent qu'OMP a des capacités exploratoires moindres qu'OLS car certains supports ne peuvent pas être atteints par OMP :

- \mathcal{Q}^* ne peut être exactement reconstruit en k itérations à partir d'aucune combinaison linéaire $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}^*})$ (théorème 8) ;
- \mathcal{Q}^* ne peut être *atteint* en k itérations à partir d'un vecteur \mathbf{y} quelconque (exemple 1).

Au contraire, OLS possède des propriétés favorables pour atteindre tout support.

Lemme 1. [SOU 13b, Lemme 3] *Tout support \mathcal{Q} de cardinalité $\text{Card}[\mathcal{Q}] \leq \text{spark}(\mathbf{A}) - 2$ peut être atteint par OLS à partir d'un vecteur $\mathbf{y} \in \text{span}(\mathbf{A}_{\mathcal{Q}})$.*

De façon un peu surprenante, le phénomène de « non-reconstruction » peut également se produire avec *basis pursuit* [CHE 01] dont le principe est de trouver la représentation parcimonieuse de norme ℓ_1 minimale :

$$\arg \min_{\{\mathbf{x}: \mathbf{y}=\mathbf{A}\mathbf{x}\}} \|\mathbf{x}\|_1. \quad (3.15)$$

L'analyse de *basis pursuit* est plus directe que celle d'OMP. En effet, la capacité de *basis pursuit* à reconstruire un support \mathcal{Q}^* ne dépend que du noyau de \mathbf{A} et du signe des amplitudes non nulles $\text{sgn}(\mathbf{t}) \in \{-1, 1\}^k$ (dans la décomposition $\mathbf{y} = \mathbf{A}_{\mathcal{Q}^*}\mathbf{t}$) [FUC 04, DOS 10] alors que le comportement d'OMP dépend aussi de la valeur des amplitudes t_i . Nous avons établi une condition nécessaire et suffisante de non-reconstruction par *basis pursuit* en tant que conséquence directe de la condition NSP (*Null-Space Property*) [DON 01],[GRI 03, Lemme 1] qui contrôle la reconstruction exacte par *basis pursuit*. La condition de non-reconstruction prend la forme [SOU 13b, Prop. 2] :

$$\forall \boldsymbol{\varepsilon} \in \{-1, 1\}^k, \exists \mathbf{x} \in \text{Ker}(\mathbf{A}) : \sum_{i \notin \mathcal{Q}^*} |x_i| < \sum_{i \in \mathcal{Q}^*} \varepsilon_i x_i. \quad \text{BRC-BP}(\mathbf{A}, \mathcal{Q}^*)$$

Cette condition s'interprète comme l'intersection (pour tous les orthants possibles, paramétrés par le vecteur signe $\boldsymbol{\varepsilon} \triangleq \text{sgn}(\mathbf{t})$) de la négation de la condition de Fuchs [FUC 04], qui est une condition de reconstruction exacte pour un vecteur signe $\boldsymbol{\varepsilon}$ donné.

Nous avons analysé plus en détails les conditions de non-reconstruction par OMP et *basis pursuit* dans la communication [SOU 13c]. Cette étude consiste, pour un algorithme donné, à classer l'ensemble des supports de $\{1, \dots, n\}$ en trois catégories :

1. Les supports garantis d'être uniformément (pour tout \mathbf{y}) atteints par l'algorithme (la condition ERC / NSP est vérifiée) ;
2. Les supports garantis de ne pas être atteints pour un moins un vecteur \mathbf{y} (la condition ERC / NSP n'est pas vérifiée) ;
3. Les supports garantis de ne pas être atteints pour tout \mathbf{y} (la condition de non-reconstruction est vérifiée).

Cela nous permet d'étendre le concept traditionnel de *diagramme de transition de phase*. La transition de phase fait généralement référence à la « frontière » entre les deux premières catégories [MAL 10, DOS 10]. Elle caractérise donc la reconstruction exacte. Ici, nous cherchons à caractériser une transition de phase pour la non-reconstruction, délimitant les supports de la deuxième et de la troisième catégorie. Cette idée est illustrée sur la figure 3.4.

3.10 Conclusions

La première contribution de ce travail (théorème 2) est une *première* analyse d'OLS en k itérations qui montre que la condition ERC de Tropp est non seulement une condition nécessaire et suffisante pour la reconstruction exacte de support avec OMP, mais aussi avec OLS. La technique de preuve de ce résultat repose sur une étude de reconstruction exacte par OLS en supposant que les q premières itérations ont toutes sélectionné des atomes du support inconnu \mathcal{Q}^* . Nous avons défini des conditions partielles de reconstruction exacte à la q -ème itération d'OLS, qui garantissent la sélection d'un atome de \mathcal{Q}^* à l'itération $q + 1$ en supposant que q atomes de \mathcal{Q}^* ont déjà été sélectionnés. Une analyse minutieuse permet d'établir une relation d'emboîtement entre les conditions partielles aux itérations q et $q + 1$, et ainsi de prouver la

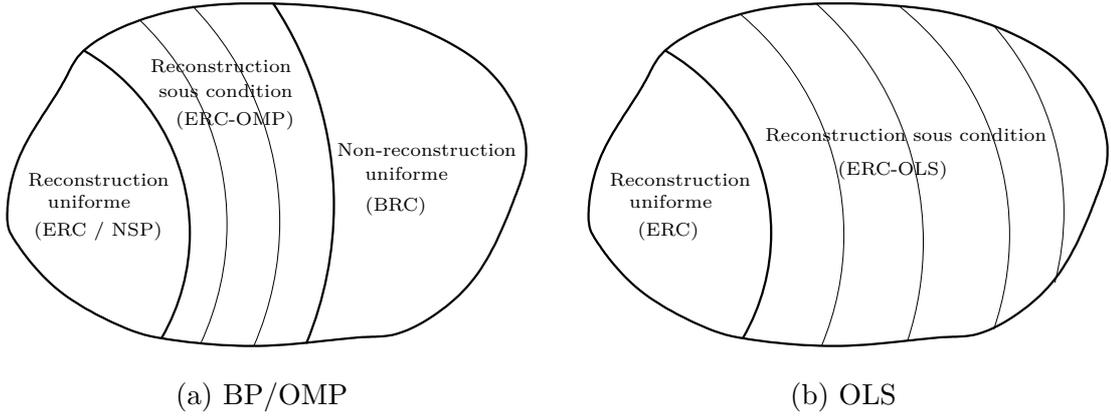


FIGURE 3.4 – Transitions de phase pour OMP, *basis pursuit* et OLS. (a) Pour OMP et BP, on distingue deux transitions de phase (en trait gras) pour la reconstruction uniforme et la non-reconstruction uniforme. Elles délimitent trois catégories de supports : les supports qui peuvent être uniformément reconstruits, ceux qui sont garantis de ne pas être reconstruits, et une catégorie intermédiaire dans laquelle une reconstruction partielle (*i.e.*, sous condition) est possible. Pour OMP, les supports vérifiant les conditions partielles ERC-OMP($\mathbf{A}, \mathcal{Q}^*, q$) pour $q \geq 1$ (mais pas la condition ERC) apparaissent dans la catégorie intermédiaire. Les frontières fines délimitent les supports vérifiant les conditions partielles pour q croissant. (b) Pour OLS, la non-reconstruction uniforme ne peut se produire (lemme 1). Il n’y a donc qu’une seule transition de phase correspondant à la reconstruction exacte.

sélection de $k - q$ atomes de \mathcal{Q}^* jusqu’à l’annulation du résidu, c’est-à-dire la reconstruction exacte du support \mathcal{Q}^* (théorème 2).

La seconde contribution est une étude parallèle d’OMP et OLS qui exploite les conditions partielles de reconstruction lorsque les premières d’itérations ont réussi. Dans ce cas, aucun des deux algorithmes n’est uniformément meilleur que l’autre, mais plusieurs éléments empiriques et analytiques permettent de distinguer les deux algorithmes. L’étude empirique montre qu’OLS se comporte mieux qu’OMP lorsque le dictionnaire est très corrélé. L’étude analytique met en évidence les limites d’OMP pour atteindre certains supports : il existe des dictionnaires pour lesquels OMP est incapable de reconstruire certains supports. Au contraire, OLS peut atteindre n’importe quel support.

Le meilleur comportement d’OLS pour des dictionnaires corrélés est un résultat cohérent avec l’analyse empirique du chapitre 1, mais également avec d’autres études empiriques qui soulignent qu’OLS est plus performant qu’OMP au prix d’un coût calculatoire supérieur [REB 02, DYM 11, CHA 12, DAV 12]. Contrairement aux analyses théoriques basées sur l’hypothèse d’isométrie restreinte pour analyser OMP [DAV 10, FOU 11, LIU 12, MO 12] et OLS [FOU 13a], notre analyse ne requiert pas d’hypothèse spécifique sur le dictionnaire, ce qui nous permet d’établir une (ébauche de) discrimination entre OMP et OLS. Cette discrimination n’est pas possible avec l’hypothèse d’isométrie restreinte (voir les explications fournies avec la figure 3.1 page 91).

Malgré ces quelques tendances en faveur d’OLS, il n’est pas aisé de mieux discriminer OMP et OLS car l’analyse en k itérations est une analyse *au pire cas* : les conditions proposées garantissent la reconstruction du support *pour toute* combinaison linéaire \mathbf{y} d’atomes indicés par le support. Elles sont en réalité très pessimistes au regard de la performance réelle des algorithmes. D’autres auteurs ont souligné le pessimisme excessif des conditions au pire cas et la nécessité de recourir à des conditions en moyenne pour mieux refléter le comportement

réel des algorithmes [DON 06, Sec. VI] [SCH 07, GRI 08b]. Les études basées sur des arguments probabilistes sont moins générales car restreintes à certains modèles de dictionnaires aléatoires [TRO 07, FLE 09, WAI 09]. Pour obtenir des résultats généraux et moins restrictifs, une alternative intéressante (ou complémentaire) est d'effectuer une analyse *différée*, en plus de k itérations. Cette analyse autorise la sélection de « mauvais » atomes $i \notin \mathcal{Q}^*$, et la notion de bonne reconstruction d'un support est définie au sens large comme la reconstruction en plus de k itérations d'un support englobant \mathcal{Q}^* [PLU 07, FOU 11, ZHA 11b].

L'objectif initial de mon projet était de fournir une analyse d'algorithmes gloutons bidirectionnels orthogonaux. Il est pour l'instant relégué au chapitre de perspectives. Cette analyse me semble néanmoins envisageable en s'inspirant de l'analyse partielle d'OMP/OLS présentée dans ce chapitre, et des techniques d'analyse différée de ces algorithmes [PLU 07, FOU 11, ZHA 11b]. En effet, les analyses différées autorisent la sélection de mauvais atomes $\mathcal{Q} \not\subset \mathcal{Q}^*$ au cours des itérations de l'algorithme et conduisent à des garanties de reconstruction de $\mathcal{Q}^* \cup \mathcal{Q}$. Ce principe se rapproche du fonctionnement des algorithmes bidirectionnels, où une reconstruction exacte est possible malgré de mauvaises sélections d'atomes à un certain stade de l'exécution de l'algorithme. Pour obtenir des conditions de reconstruction exacte de \mathcal{Q}^* , il s'agira de garantir que les mauvais atomes sélectionnés seront effectivement désélectionnés aux itérations ultérieures. Cette perspective est détaillée au chapitre 4 (§ 4.2.5).

Chapitre 4

Quelques perspectives à court et moyen terme

Je présente dans ce chapitre les principaux projets que j'envisage de mener au cours des prochaines années en terme de développement méthodologique, algorithmique et appliqué. Les projets mis en avant dans les sections 4.1, 4.2 et 4.3 sont respectivement liés aux trois premiers chapitres de ce manuscrit. Je m'efforcerai de souligner l'originalité de l'approche proposée et les aspects novateurs par rapport aux travaux passés et à la littérature actuelle.

4.1 Développement d'algorithmes d'approximation parcimonieuse

Collaborateurs

- Sébastien Bourguignon, Jérôme Idier (IRCCyN) ;
- Vincent Mazet (ICube).

Contexte

Les idées de développement de nouveaux algorithmes d'approximation parcimonieuse sont nées d'échanges avec Jérôme Idier depuis 2010, puis avec Sébastien Bourguignon, Maître de conférences à l'IRCCyN depuis 2011 qui s'intéresse à l'application d'algorithmes gloutons (OMP, OLS) et leur implémentation dans le contexte de la restauration d'images hyperspectrales, à la parcimonie structurée et à la conception d'algorithmes bidirectionnels plus complexes que SBR pour des applications en contrôle non destructif.

Vincent Mazet s'intéresse à des problèmes de décomposition paramétriques de signaux spectroscopiques (ou spectres) en motifs élémentaires : dans ces problèmes, on cherche typiquement à décomposer un spectre comme une combinaison linéaire de motifs gaussiens dont le centre, la largeur et l'amplitude positive sont inconnues. La spécificité des problèmes abordés au laboratoire ICube est qu'ils ne mettent pas en jeu un seul spectre mais une séquence (par exemple temporelle) de spectres, l'objectif étant de décomposer chaque spectre et d'effectuer une classification des motifs pour suivre temporellement la trajectoire de chaque motif. Dans le cadre du projet exploratoire PEPS-CNRS SpectroDec (2012-2013), nous (Vincent Mazet, El-Hadi Djermoune et moi pour le CRAN) avons montré la pertinence de l'utilisation d'algorithmes parcimonieux pour ces problèmes. Nous avons identifié de nouveaux problèmes pour lesquels des développements méthodologiques sont envisagés.

Avec les partenaires mentionnés ci-dessus, nous envisageons au cours des prochaines années de déposer un projet de type ANR afin de développer, entre autres, ces thématiques.

Je détaille à présent chacune des thématiques, à savoir :

- les méthodes de parcimonie intégrant des contraintes de positivité ;
- la conception d’algorithmes bidirectionnels plus efficaces que SBR pour l’optimisation de critères mixtes « ℓ_2 - ℓ_0 » ;
- les méthodes de « super-résolution » où on cherche à reconstruire un signal mieux résolu que le signal observé ;
- l’approximation conjointe d’une séquence de spectres en motifs élémentaires.

4.1.1 Parcimonie et contrainte de positivité

Ces dernières années, la communauté du traitement du signal et de l’image a connu un engouement significatif pour les problèmes de restauration, reconstruction et séparation de signaux sous contrainte de positivité. A titre d’exemple, la communauté française a organisé une journée intitulée « Contraintes de non-négativité en traitement du signal et des images » en janvier 2011 dans le cadre du GDR ISIS. Imposer une contrainte de positivité est tout à fait naturel dans un grand nombre de problèmes (restauration d’images en microscopie optique et en astronomie [BER 09], reconstruction tomographique [KAK 87], séparation de sources pour l’imagerie hyperspectrale [MOU 08], identification de systèmes [CHE 11]) où l’on sait que le signal ou l’image à reconstruire est nécessairement à valeurs positives. Intégrer une contrainte de positivité est aussi un moyen de régulariser des problèmes inverses mal posés en restreignant leurs solutions admissibles, et en rendant leurs solutions plus stables [BAR 08].

Dans le cadre du projet PEPS SpectroDec, nous avons identifié le besoin de méthodes de décomposition parcimonieuse positive de signaux pour différentes applications en spectroscopie moléculaire ou en radio-astronomie [MAZ 13] (voir également la thèse [GUO 11] pour la spectroscopie de résonance magnétique). De façon générale, un signal spectroscopique représente la distribution de particules physiques ou d’ondes électromagnétiques en fonction de leur énergie, de leur longueur d’onde, *etc.* C’est un signal à valeurs positives, constitué d’un ensemble de composantes spectrales, appelées raies, dont les caractéristiques (*e.g.*, la position, l’amplitude et la largeur de motifs gaussiens) informent sur l’objet analysé. La décomposition parcimonieuse ressemble au problème de la déconvolution impulsioneuse abordée au chapitre 1 de ce manuscrit (§ 1.6.4) à la différence que la largeur des raies est variable. Dans ces deux problèmes, il n’est pas trivial de séparer des raies proches car les signaux qui en résultent s’entremêlent²⁴. Nous avons constaté que l’introduction de la contrainte de positivité permet d’améliorer la séparation des motifs.

Le projet est de concevoir des algorithmes parcimonieux gloutons permettant de contraindre les amplitudes des atomes sélectionnés à être positives. Formellement, cela conduit à résoudre un problème d’optimisation ℓ_2 - ℓ_0 sous contrainte de positivité :

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.c.} \quad \begin{cases} \|\mathbf{x}\|_0 \leq k \\ \mathbf{x} \geq \mathbf{0} \end{cases} \quad (4.1)$$

ou

$$\min_{\mathbf{x} \geq \mathbf{0}} \{\mathcal{J}(\mathbf{x}; \lambda) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0\}. \quad (4.2)$$

24. Par exemple, la somme de deux gaussiennes à des positions proches produit un signal qui peut ressembler à une gaussienne plus large ; en déconvolution impulsioneuse, les algorithmes d’approximation ont toutes les chances de trouver une gaussienne à un mauvais emplacement, entre les deux vraies gaussiennes.

Ces problèmes me semblent loin d'être triviaux.

A ce stade, il me semble important de préciser que certaines approches pour résoudre des problèmes inverses sous contrainte de positivité sont inadaptées à la décomposition parcimonieuse positive. Les problèmes de la séparation de sources ou de la restauration d'images peuvent être classiquement formulés comme la minimisation sous contrainte de positivité d'un critère des moindres carrés régularisé, convexe ou non [BIO 12]. La contrainte de positivité est souvent gérée en choisissant un algorithme de type point intérieur [CHO 10] qui impose la positivité au sens strict [NOC 06] : aucun des pixels de l'image restaurée ne vaut exactement 0. Inversement, les méthodes de type lagrangien augmenté n'imposent la contrainte de positivité qu'asymptotiquement [GIO 05, AFO 11, HEN 13]. Dans ce cas, les pixels de l'image restaurée pour lesquels la contrainte est active sont non nuls et peuvent prendre des valeurs négatives. La différence fondamentale avec l'approximation parcimonieuse est que dans ce domaine, on souhaite qu'un grand nombre de coefficients x_i soient exactement nuls.

Dans la littérature, il est classique de remplacer le problème d'optimisation (4.1) ou (4.2) par le problème relaxé :

$$\min_{\mathbf{x} \geq 0} \{ \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 \} \quad (4.3)$$

en utilisant la régularisation ℓ_1 [ZIB 10]. La contrainte de positivité peut alors être gérée naturellement par différents algorithmes classiques comme les algorithmes de seuillage itératif [DAU 04, COM 07, FIG 07, BEC 09, DUP 10, PUS 10] ou l'algorithme *Alternating Direction Method of Multipliers* (ADMM) [AFO 10, BOY 11]. Le problème est en effet un problème de programmation quadratique exprimé sous forme standard :

$$\min_{\mathbf{x} \geq 0} \left\{ \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \sum_i x_i \right\} \quad (4.4)$$

pour lequel des algorithmes efficaces existent.

Comme nous l'avons vu au chapitre 1, utiliser des pénalisations non convexes (relaxation non convexe de la pseudo-norme ℓ_0) peut permettre une amélioration substantielle des résultats de l'approximation parcimonieuse par rapport à la relaxation ℓ_1 à la fois en termes d'erreur d'approximation et de sélection d'atomes. Dans [MAZ 13], nous avons appliqué l'algorithme *Iterative Reweighted ℓ_1* ($\text{IR}\ell_1$) pour minimiser le critère :

$$\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \sum_i \varphi(x_i), \quad \text{où} \quad \varphi(t) = \log(|t| + \eta) \quad (4.5)$$

avec $\eta > 0$ et faible (φ est une fonction paire, concave sur \mathbb{R}^+ et non différentiable en 0). $\text{IR}\ell_1$ est un algorithme sous-optimal efficace dédié à la minimisation ce type de critère [ZOU 06, CAN 08a, WIP 10]. Comme nous l'avons vu au chapitre 1, $\text{IR}\ell_1$ conduit à des solutions de très bonne qualité même lorsque les atomes du dictionnaire sont corrélés les uns aux autres [SOU 11]. De plus, la structure de l'algorithme $\text{IR}\ell_1$ permet tout naturellement de gérer la contrainte de positivité [WIP 10] car son principe est de résoudre une suite de problèmes d'optimisation ℓ_2 - ℓ_1 du type (4.3) pour lesquels des algorithmes standard sont applicables.

J'aimerais ces prochaines années **proposer de nouvelles solutions basées sur l'adaptation d'algorithmes gloutons (notamment d'OLS) pour traiter le problème** (4.2). Les algorithmes gloutons présentent en effet une structure beaucoup plus simple et directe que celle d' $\text{IR}\ell_1$. $\text{IR}\ell_1$ est davantage une stratégie (de type majorisation-minimisation) pour optimiser un critère non convexe qu'un algorithme précis. Cette stratégie repose sur la majoration de la

fonction φ par sa tangente autour du point courant, menant à l'optimisation de critères mixtes ℓ_2 - ℓ_1 . Notons que ces problèmes ℓ_2 - ℓ_1 auxiliaires peuvent être traitées par divers algorithmes au choix de l'utilisateur. La structure emboîtée de $\text{IR}\ell_1$ peut potentiellement poser des problèmes de propagation d'erreur : une erreur numérique dans la résolution du problème ℓ_2 - ℓ_1 se répercute sur le support estimé par $\text{IR}\ell_1$ à l'itération courante, et donc sur le comportement futur de l'algorithme²⁵. Ce problème motive une réflexion autour de la conception d'algorithmes non emboîtés potentiellement plus simples.

Parmi les algorithmes gloutons monodirectionnels, seul le plus simple d'entre eux, *Matching Pursuit* (MP), se prête sans aucune difficulté à une extension positive car aucune projection orthogonale n'est effectuée pour le calcul des amplitudes. A chaque itération, toutes les amplitudes gardent les mêmes valeurs qu'à l'itération précédente à l'exception de celle du nouvel atome x_ℓ . x_ℓ est calculé en minimisant un critère 1D correspondant à la norme quadratique du résidu $\mathbf{r}_{\mathcal{Q} \cup \{\ell\}} = \mathbf{r}_{\mathcal{Q}} - \mathbf{a}_\ell x_\ell$ (cf. tableau 1.1 page 41). Pour obtenir une version positive de MP, il suffit d'intégrer la contrainte de positivité lors de la phase de sélection du nouvel atome :

$$\ell \in \arg \min_i \left\{ \min_{x_i \geq 0} \|\mathbf{r}_{\mathcal{Q}} - \mathbf{a}_i x_i\|_2^2 \right\}. \quad (4.6)$$

On peut facilement vérifier que seuls les ajouts d'indices $\mathcal{Q} \cup \{i\}$ pour lesquels $\langle \mathbf{r}_{\mathcal{Q}}, \mathbf{a}_i \rangle > 0$ permettent de faire décroître la norme du résidu (en supposant les atomes normalisés). Dans ce cas, la décroissance maximale est obtenue pour $x_i = \langle \mathbf{r}_{\mathcal{Q}}, \mathbf{a}_i \rangle$, et $\|\mathbf{r}_{\mathcal{Q} \cup \{i\}}\|_2^2 = \|\mathbf{r}_{\mathcal{Q}}\|_2^2 - \langle \mathbf{r}_{\mathcal{Q}}, \mathbf{a}_i \rangle^2$. La règle de sélection d'atome (4.6) se réécrit :

$$\ell \in \arg \max_{\{i: \langle \mathbf{r}_{\mathcal{Q}}, \mathbf{a}_i \rangle > 0\}} \langle \mathbf{r}_{\mathcal{Q}}, \mathbf{a}_i \rangle. \quad (4.7)$$

La version positive de *Matching Pursuit* est aussi connue sous le nom de CLEAN en restauration d'images astronomiques [THI 10].

Contrairement à MP, les algorithmes de poursuite *orthogonaux* procèdent à des projections orthogonales, et les amplitudes des atomes $\mathbf{A}_{\mathcal{Q} \cup \{\ell\}}^\dagger \mathbf{y}$ sont *toutes* mises à jour. Ces algorithmes ne se prêtent pas aussi facilement à une extension positive. Pour OMP, adopter la règle de sélection de MP (4.7) ne garantit pas que $\mathbf{A}_{\mathcal{Q} \cup \{\ell\}}^\dagger \mathbf{y}$ est un vecteur positif. Bruckstein *et al.* [BRU 08] ont proposé une version positive d'OMP avec (4.7) comme règle de sélection, où la mise à jour des amplitudes est formulée comme un problème de moindres carrés contraints :

$$\mathbf{x}_{\mathcal{Q} \cup \{\ell\}} = \{\hat{\mathbf{t}}, \mathbf{0}\}, \text{ où } \hat{\mathbf{t}} = \arg \min_{\mathbf{t} \geq 0} \|\mathbf{y} - \mathbf{A}_{\mathcal{Q} \cup \{\ell\}} \mathbf{t}\|_2^2. \quad (4.8)$$

Notons qu'en adoptant cette règle, il est fort possible que certains atomes de $\mathcal{Q} \cup \{\ell\}$ soient pondérés par $t_i = 0$. Cette règle ouvre des perspectives de rapprochement avec des versions bidirectionnelles d'OMP qui effectuent également des retraits d'atomes [HER 10, ZHA 11a].

Pour OLS, la règle de sélection à la q -ème itération conduit à résoudre $n - q$ problèmes des moindres carrés. Un tel algorithme peut facilement s'écrire sur le papier avec des contraintes de positivité : les $n - q$ problèmes prennent la forme (4.8) où $\mathcal{Q} \cup \{\ell\}$ est remplacé par les supports $\mathcal{Q} \cup \{i\}$ testés pour tout $i \notin \mathcal{Q}$. Cependant, concevoir une implémentation compétitive en temps de calcul n'est pas une tâche triviale. Les implémentations typiques de la version non contrainte (standard) d'OLS résolvent les problèmes des moindres carrés $\min_{\mathbf{t}} \|\mathbf{y} - \mathbf{A}_{\mathcal{Q}} \mathbf{t}\|_2^2$ de

25. J'ai effectivement rencontré ce problème en pratique en utilisant l'implémentation `SolveLasso` de la bibliothèque `SparseLab` [DON 07] pour résoudre les problèmes de minimisation ℓ_2 - ℓ_1 . `SolveLasso` met en jeu différents paramètres de réglage pour éviter les erreurs numériques, et une variation infime de ces paramètres engendre parfois une modification très importante des itérés d' $\text{IR}\ell_1$.

façon récursive lorsque la taille de \mathcal{Q} est incrémentée d'un élément [CHE 89]. Elles exploitent que certaines quantités peuvent être mises en mémoire une seule fois pour tester les $n - q$ supports candidats $\mathcal{Q} \cup \{i\}$, $i \notin \mathcal{Q}$ [SOU 11]. Concevoir des implémentations récursives efficaces dans le cas contraint est un problème ouvert à ma connaissance, et loin d'être trivial. En résumé, il y a une réflexion à mener pour développer des variantes positives des algorithmes gloutons orthogonaux, et en particulier d'OLS.

4.1.2 Algorithmes plus efficaces que SBR pour la minimisation $\ell_2\text{-}\ell_0$

Comme nous l'avons vu au chapitre 1, le choix d'un algorithme d'approximation parcimonieuse relève d'un compromis entre la performance désirée et le temps de calcul que l'on souhaite consacrer. De ce point de vue, SBR fait partie des algorithmes les plus coûteux mais les plus efficaces (voir [SEL 13] par exemple). Le coût calculatoire de SBR est essentiellement lié aux n inversions de systèmes linéaires (où n est le nombre d'atomes du dictionnaire) effectuées à chaque itération pour tester les n remplacements élémentaires. Dans la version standard de SBR, le mouvement élémentaire correspond à un seul remplacement, *i.e.*, à l'ajout ou au retrait d'un atome (noté $\mathcal{Q} \bullet i$ où $i \in \{1, \dots, n\}$ et \mathcal{Q} est le support courant). Le principe de SBR reste bien sûr valable en travaillant avec un voisinage plus large (par exemple : faire deux remplacements d'atomes simultanément) mais cela rend explosive la complexité de l'algorithme. Pour éviter une explosion numérique, on peut envisager de ne pas faire tous les tests de remplacement, mais seulement certains tests qui ont le plus de chance d'être concluants. Isoler ces tests potentiellement concluants nécessite d'exploiter la structure du dictionnaire. C'est le choix effectué dans la thèse d'E. Carcreff [CAR 14] pour un dictionnaire convolutif en tenant compte de la forme (oscillante) de la réponse impulsionnelle pour proposer des remplacements intelligents. Des nouveaux algorithmes bidirectionnels y sont proposés, exploitant la nature convolutive du problème, en tant qu'extension de l'algorithme SBR.

Avec Sébastien Bourguignon et Jérôme Idier, nous envisageons de concevoir des versions adaptées et améliorées de SBR pour des problèmes spécifiques dans la lignée de la thèse d'E. Carcreff. Nous allons explorer plusieurs pistes pour concevoir des heuristiques au moins aussi efficaces en un coût calculatoire réduit. L'idée sous-jacente est d'une part, de **ne plus effectuer la totalité des n inversions par itération** pour réduire le coût de calcul, et d'autre part, d'**autoriser des remplacements élémentaires plus complexes** que l'ajout ou le retrait d'un seul élément ($\mathcal{Q} \bullet i$) pour améliorer les performances. Pour l'exemple de la déconvolution à réponse impulsionnelle finie, on peut envisager de privilégier la recherche d'impulsions dans une zone locale où des impulsions ont déjà été détectées et ainsi, travailler avec un sous-problème (sous-dictionnaire) de taille réduite et diminuer le coût calculatoire.

Une étude préliminaire consistera à tester les extensions monodirectionnelles d'OMP et d'OLS récemment proposées par Chatterjee *et al.* [CHA 12] dans le but de sélectionner plus efficacement les atomes pour des problèmes faisant intervenir des dictionnaires corrélés. Cette sélection intelligente s'appuie d'une part, sur une recherche restreinte (les atomes candidats ne sont pas tous les atomes du dictionnaire mais une pré-sélection de certains atomes), et d'autre part, sur des tests d'insertion plus poussés (plus coûteux) où l'effet de la sélection d'un atome sur les itérations suivantes de l'algorithme est anticipé, ce qui conduit à une règle de sélection améliorée.

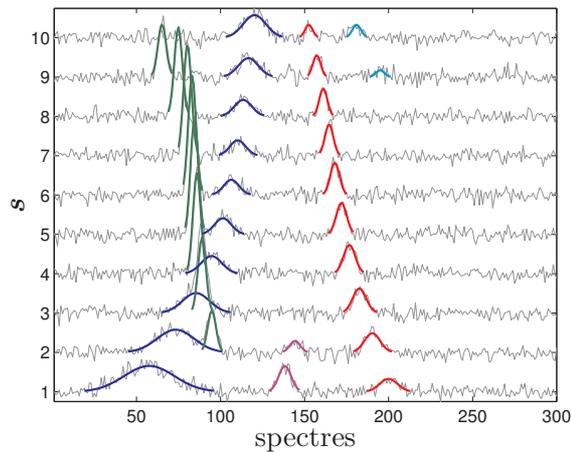


FIGURE 4.1 – Séquence simulée de 10 spectres de photo-électrons. Chaque spectre est décomposé comme un mélange de motifs gaussiens dont les positions, amplitudes et largeurs sont inconnues. Problème traité par Vincent Mazet au laboratoire ICube ayant fait l’objet d’une collaboration avec le CRAN dans le cadre du projet PEPS CNRS SpectroDec [MAZ 13].

4.1.3 Algorithmes hybrides pour la décomposition en motifs paramétriques

Ces dernières années, je me suis intéressé à des problèmes de *décomposition en motifs élémentaires*. La déconvolution impulsionnelle en est un exemple simple où les motifs élémentaires ont une forme parfaitement connue, celle de la réponse impulsionnelle. Il s’agit essentiellement de détecter la position des impulsions, et de déduire leurs amplitudes [SOU 11, BOU 11b, SOU 12]. Au CRAN et au laboratoire ICube, nous avons abordé d’autres problèmes de décomposition en motifs élémentaires *paramétriques* pour des problèmes de spectroscopie (spectroscopie RMN, spectroscopie de photo-électrons) : cette fois, les motifs ont non seulement une position et une amplitude, mais aussi une forme inconnue. Il faut donc estimer les paramètres de forme de chacun des motifs [SAH 12a, MAZ 11] : voir la figure 4.1 pour une illustration où les motifs sont des formes gaussiennes de largeur inconnue.

Pour ce type de problèmes, la solution que nous avons retenue jusqu’à présent consiste à discrétiser les positions [STO 12] et les paramètres de forme (la largeur de la gaussienne) sur une grille de façon à construire un très grand dictionnaire contenant tous les motifs potentiellement présents dans le signal [AUS 10]. Les atomes sont donc « multidimensionnels » puisqu’ils sont décrits par une valeur pour la position du motif et une ou plusieurs valeurs discrètes pour les paramètres de forme. Le dictionnaire devient très grand lorsqu’il y a plusieurs paramètres de forme [SAH 12a] ou pour des applications de type « super-résolution » où le signal inconnu \mathbf{x}^* est discrétisé plus finement (mieux résolu) que le signal observé $\mathbf{y} = \mathbf{A}\mathbf{x}^*$ [CAR 13]. Dans ce type de situation, le dictionnaire est d’autant moins bien conditionné que l’on souhaite reconstruire un signal bien résolu.

Dans la thèse de S. Sahnoun au CRAN [SAH 12a], nous avons proposé des stratégies multi-grilles pour réduire le coût calculatoire d’algorithmes gloutons (OMP, OLS, SBR) en discrétisant le problème sur des grilles de plus en plus fines, et en excluant progressivement certaines zones de la grille en fonction du signal parcimonieux trouvé. A chaque niveau de résolution, un atome est associé à un point de la grille multidimensionnelle obtenue en discrétisant les paramètres continus. En procédant de la sorte, il est possible de travailler à des niveaux de discrétisation très fins avec des dictionnaires de taille limitée en nombre d’atomes, et d’obtenir des résultats compéti-

tifs [SAH 12b]. L'inconvénient de cette approche est que la décision d'exclure certaines zones de la grille et de ne pas les raffiner a forcément un impact (non maîtrisé) sur les approximations parcimonieuses aux échelles suivantes.

Une alternative à l'approche qui discrétise tous les paramètres continus pour former un très grand dictionnaire « paramétrique » est de **discrétiser uniquement certains paramètres** (typiquement, la position des motifs à rechercher dans un signal) **et de garder les autres paramètres continus** (les paramètres de forme du motif). Le modèle s'écrit :

$$\mathbf{y} = \sum_{i=1}^n x_i \mathbf{a}(c_i; \boldsymbol{\theta}_i) \quad (4.9)$$

où le signal $\mathbf{x} = [x_1, \dots, x_n]^T$ est toujours parcimonieux, les positions (ou centres des motifs) c_i sont discrètes, et les atomes du dictionnaire $\mathbf{a}(c_i; \boldsymbol{\theta}_i)$ dépendent de paramètres de forme $\boldsymbol{\theta}_i$ continus. L'approximation parcimonieuse consisterait à coupler une approche de type discrète (recherche de la position c_i des motifs par algorithme glouton) avec une optimisation locale continue pour mettre à jour les paramètres de forme $\boldsymbol{\theta}_i$. Le travail de thèse de F. Dublanche au L2S à la fin des années 1990 [DUB 97] est une source d'inspiration qui nous a conduit à imaginer cette approche hybride continu/discret. Dans cette thèse, les impulsions du signal étaient traitées en discret à une certaine étape de l'algorithme pour introduire des nouveaux pics, et en continu à une autre étape pour raffiner la position des pics sur la grille continue.

Une étape d'optimisation continue ne peut que fournir des résultats plus fins que nos résultats actuels en terme de valeur des paramètres de forme. De plus, cela permettra d'éviter de sélectionner deux atomes proches (relatifs à deux positions discrètes adjacentes sur la grille, qui encadrent une position continue) et de favoriser la sélection d'un seul « atome continu ».

4.1.4 Décomposition conjointe de signaux en motifs

Dans un grand nombre de domaines de la physique, une collection de signaux (spectres) \mathbf{y}_s , $s = 1, \dots, S$ sont mesurés sur une grille spatiale $s = (x, y)$ ou temporelle $s = t$, ce qui conduit à la formation d'images multispectrales ou hyperspectrales pour une grille spatiale. On souhaite souvent décomposer conjointement ces signaux sur un même dictionnaire en imposant des hypothèses de continuité : deux spectres spatialement ou temporellement voisins ont des décompositions proches, suivant une notion de voisinage prédéfinie entre les atomes du dictionnaire.

Parmi les applications que j'ai abordées dans ce manuscrit, une procédure de décomposition conjointe serait utile pour améliorer le traitement des images force-volume en microscopie AFM (chapitre 2). Actuellement, chaque courbe de force \mathbf{y}_s est décomposée indépendamment des autres. La figure 4.1 illustre un problème analogue dans le contexte de la spectroscopie de photo-électrons [MAZ 11] où l'on cherche à décomposer une séquence temporelle de spectres \mathbf{y}_s en motifs élémentaires dans le but de faire la classification et le suivi temporel des motifs. Avec Vincent Mazet [MAZ 13], nous avons d'abord proposé une solution séquentielle visant à (i) décomposer chaque spectre indépendamment dans un dictionnaire paramétrique ($\mathbf{y}_s \approx \mathbf{A}\mathbf{x}_s$); puis (ii) appliquer un algorithme de classification pour mettre en correspondance les motifs obtenus pour tous les spectres \mathbf{y}_s .

Puis, nous avons ébauché des solutions pour décomposer conjointement la séquence de spectres en imposant que les motifs présents dans deux spectres consécutifs se trouvent à des positions proches et ont des formes ressemblantes. Une approche possible est de minimiser un

critère pénalisé du type :

$$\sum_{s=1}^S \|\mathbf{y} - \mathbf{A}\mathbf{x}_s\|_2^2 + \mu \sum_{s=1}^{S-1} \|\mathbf{A}\mathbf{x}_s - \mathbf{A}\mathbf{x}_{s+1}\|_2^2 + \lambda \sum_{s=1}^S \|\mathbf{x}_s\|_0 \quad (4.10)$$

conjointement par rapport à $\{\mathbf{x}_1, \dots, \mathbf{x}_S\}$ où le terme du milieu est une régularisation qui contraint les approximations consécutives $\mathbf{A}\mathbf{x}_s$ et $\mathbf{A}\mathbf{x}_{s+1}$ à être proches. Il impose indirectement une proximité entre les signaux impulsionnels \mathbf{x}_s et \mathbf{x}_{s+1} . L'intérêt de cette régularisation est que le critère (4.10) est un critère ℓ_2 - ℓ_0 pénalisé dont la minimisation peut être directement mise en œuvre en utilisant l'algorithme SBR. Cette approche ne peut être concluante que lorsque les motifs évoluent lentement d'un spectre à l'autre, *i.e.*, il y a suffisamment de recouvrement entre un même motif présent dans deux spectres consécutifs s et $s + 1$. Dans le cas contraire, nous avons envisagé des solutions plus complexes qui prennent en compte des distances « discrètes » explicites entre les signaux impulsionnels \mathbf{x}_s et \mathbf{x}_{s+1} comme la distance de Hausdorff [MAZ 13]. Il sera certainement intéressant de comparer cette approche avec les techniques de parcimonie structurée [BAC 12] et de parcimonie sociale [KOW 13] qui reposent sur l'utilisation d'une norme mixte pour favoriser l'activation d'atomes connexes pour deux spectres consécutifs.

4.2 Analyse d'algorithmes d'approximation parcimonieuse

Collaborateurs

- Cédric Herzet, Rémi Gribonval (INRIA Rennes).
- Jérôme Idier (IRCCyN).

4.2.1 Bilan et positionnement par rapport à la littérature

L'approche développée dans le chapitre 3 est une analyse des algorithmes OMP et OLS sans émettre d'hypothèse sur la structure du dictionnaire ou le modèle statistique des amplitudes du signal parcimonieux. Les analyses traditionnelles reposent sur des hypothèses fortes sur la structure du dictionnaire comme l'hypothèse d'isométrie restreinte (RIP). Cette condition présente en outre l'inconvénient d'être très délicate à tester en pratique sauf pour certains modèles de dictionnaires aléatoires [CAN 08b, BAN 13]. La littérature récente souligne que l'hypothèse RIP mène à des résultats de reconstruction exacte (garantie de reconstruction d'un support, stabilité) trop pessimistes. La tendance est donc de s'en affranchir pour établir des conditions suffisantes (et si possible nécessaires) plus faibles [XU 11, CAN 11, GRA 11, HAL 13].

L'idée de relaxer progressivement des conditions théoriques restreintes rejoint finalement la démarche (inverse) que j'ai adoptée dans ce manuscrit, qui consiste à :

- adopter un point de vue pragmatique et choisir des algorithmes empiriquement efficaces pour les problèmes inverses faisant intervenir des dictionnaires mal conditionnés (chapitre 1) ;
- chercher à les analyser, certes dans un contexte simplifié (signal fortement parcimonieux, modèle sans bruit ou avec peu de bruit) dans l'espoir de discriminer les algorithmes pour les problèmes inverses mal posés.

Les travaux très récents relatifs à l'approche *analyse* pour la parcimonie [ELA 07, NAM 13, VAI 13, GIR 14] ouvrent des perspectives intéressantes pour l'analyse théorique de problèmes inverses. L'approche analyse est basée sur la minimisation de critères du type

$$\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda \|\mathbf{D}\mathbf{x}\|_0 \quad (4.11)$$

où \mathbf{D} joue le rôle d'un opérateur de régularisation et \mathbf{H} celui de la matrice d'observation. Elle se distingue de l'approche *synthèse* que j'ai adoptée dans ce manuscrit :

$$\|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda\|\mathbf{x}\|_0 \quad (4.12)$$

où le dictionnaire $\mathbf{A} = \mathbf{HW}$ est le produit de la matrice d'observation \mathbf{H} et d'une matrice \mathbf{W} qui synthétise le signal \mathbf{Wx} à restaurer. La forme (4.12) est moins générale mais plus confortable car elle rend possible l'utilisation d'algorithmes standard d'optimisation. C'est pourquoi je l'ai privilégiée. Le problème d'optimisation (4.11) peut être transformé sous la forme (4.12) *via* le changement de variable $\mathbf{x} \leftarrow \mathbf{Dx}$ sous réserve que \mathbf{D} soit de rang colonne plein. Je renvoie le lecteur au chapitre 2 pour l'exemple du dictionnaire lié au lissage d'un signal par un polynôme par morceaux.

La minimisation de critères régularisés du type (4.11) est le formalisme habituel pour les problèmes inverses : la solution de (4.11) correspond à l'estimation au sens du maximum *a posteriori* et \mathbf{D} joue (typiquement) le rôle d'un opérateur de différences finies pour restaurer des signaux ou images à bords francs. Le développement de nouveaux algorithmes et concepts (*co-sparse analysis*) [NAM 13, PEL 13, VAI 13, GIR 14] peut potentiellement améliorer et étendre les analyses actuelles pour les problèmes inverses mal posés en découplant les hypothèses liées au dictionnaire \mathbf{A} en des hypothèses séparées liées à la structure de la matrice d'observation \mathbf{H} et à celle de \mathbf{D} .

4.2.2 Discrimination OMP *vs* OLS

L'objectif du travail présenté au chapitre 3 était de comprendre si une analyse théorique des algorithmes OMP et OLS permet de distinguer leur comportement et de valider le constat empirique (chapitre 1) qu'OLS présente de meilleures garanties, à la fois en termes d'erreur d'approximation et de position des atomes sélectionnés, pour les problèmes inverses mal conditionnés. Le bilan de cette étude est qu'une analyse au pire cas en k itérations ne permet pas cette distinction : aucun algorithme n'est uniformément meilleur que l'autre (pour tout dictionnaire et pour toute combinaison linéaire de k atomes du dictionnaire), y compris en supposant que les $q < k - 1$ premières itérations d'OMP/OLS ont toutes sélectionné de bons atomes. Le cas extrême $q = k - 1$ permet néanmoins cette distinction puisque OLS sélectionne le dernier bon atome de façon certaine.

Contrairement à l'analyse au pire cas (quelquesoit le dictionnaire), notre étude empirique²⁶ [SOU 13b] confirme très clairement que pour des dictionnaires corrélés ou très mal conditionnés, OLS fournit de meilleures garanties de reconstruction uniforme. Très souvent, OMP ne fournit aucune garantie pour des dictionnaires mal conditionnés, même pour $q = k - 1$. Cette garantie de bon comportement d'OLS est remarquable car la plupart des conditions traditionnelles sont inapplicables dans le contexte des dictionnaires mal conditionnés. Pour fixer les idées, la condition classique $\mu < 1/(2k - 1)$ portant sur la cohérence mutuelle du dictionnaire garantit la reconstruction de toute combinaison linéaire de k atomes pour seulement $k \leq 3$ pour des dictionnaires peu corrélés (dictionnaires aléatoires gaussiens de taille 300×300 où μ vaut en moyenne 0,25). La condition nécessaire et suffisante $\text{ERC}(\mathbf{A}, \mathcal{Q}^*)$ à l'itération 0 est à peine moins restrictive ; pour l'exemple précédent, elle fournit des garanties jusqu'à $k = 11$. Dans [SOU 13b], nous avons montré qu'en travaillant à la q -ème itération ($q > 0$), on peut obtenir des garanties

26. Le mot « empirique » prend ici un sens particulier : les conditions $\text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ de reconstruction uniforme du support \mathcal{Q}^* à partir d'un support partiel $\mathcal{Q} \subset \mathcal{Q}^*$, sont empiriquement calculées pour des dictionnaires \mathbf{A} divers, mais les algorithmes Oxx ne sont jamais réellement exécutés.

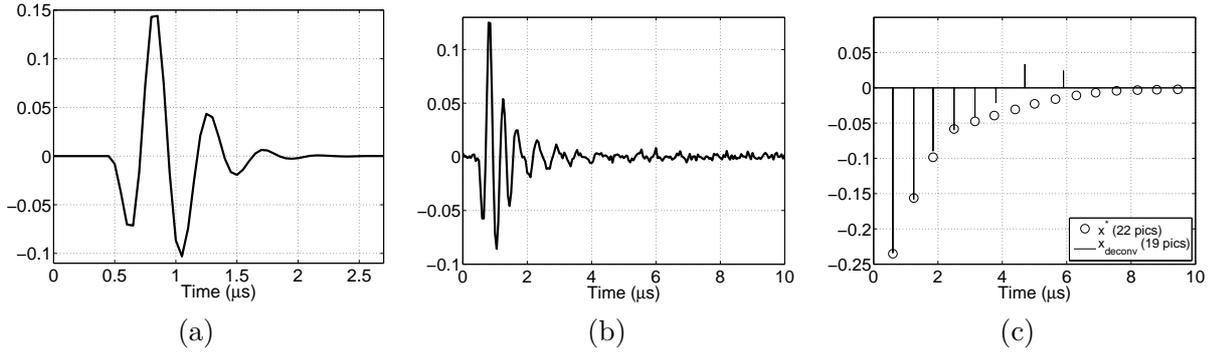


FIGURE 4.2 – Déconvolution de données acoustiques simulées (figure extraite de [SOU 12]). (a) Ondelette acoustique réaliste. (b) Données simulées $\mathbf{y} = \mathbf{h} * \mathbf{x}^* + \mathbf{b}$ correspondant à la propagation de l’ondelette dans un bloc homogène d’aluminium. Le signal impulsionnel \mathbf{x} recherché modélise les échos acoustiques résultant des réflexions multiples de l’ondelette incidente dans le matériau. (c) L’algorithme de déconvolution (SBR) réalise une estimation sans erreur des pics de plus grande amplitude. La localisation des derniers échos (pics de plus faibles amplitudes) est incertaine.

beaucoup plus intéressantes avec à la fois OMP et OLS pour des dictionnaires peu corrélés, mais aussi pour des dictionnaires corrélés et mal conditionnés dans le cas d’OLS.

L’hypothèse de connaissance partielle du support est fréquente dans la littérature récente [VAS 10, FRI 12, KAR 12]. Elle est émise pour des problèmes dynamiques où une séquence de signaux \mathbf{y}_s doit être décomposée sur le même dictionnaire et la décomposition de deux signaux successifs \mathbf{y}_s et \mathbf{y}_{s+1} implique des atomes communs (ou proches). Dans [VAS 10, FRI 12, KAR 12, HER 13], le support partiellement connu peut contenir des faux atomes alors que dans notre article initial [SOU 13b], nous avons supposé que des vrais atomes $\mathcal{Q} \subset \mathcal{Q}^*$ ont été sélectionnés aux q premières itérations. Cette hypothèse se justifie typiquement dans des problèmes de déconvolution impulsionnelle en contrôle non destructif [SOU 12] :

$$\mathbf{y} = \mathbf{h} * \mathbf{x}^* + \mathbf{b} \quad (4.13)$$

où les atomes de la décomposition parcimonieuse sont liés aux échos (retard) de l’onde incidente \mathbf{h} suite à ses réflexions et réfractions dans le milieu, et les amplitudes \mathbf{x}^* représentent son atténuation (*cf.* fig. 4.2). Les atomes liés aux plus grandes impulsions (\mathcal{Q}) s’obtiennent sans ambiguïté soit en maximisant l’inter-corrélation entre le signal mesuré \mathbf{y} et l’onde incidente \mathbf{h} , soit en exécutant les premières itérations d’Oxx. La difficulté réelle du problème est de séparer les échos entremêlés résultant de deux temps d’écho proches ou de trouver des échos de faibles amplitudes (*cf.* fig. 4.2(c)). C’est ce que nous cherchons à comprendre en analysant les algorithmes à partir de la q -ème itération.

La première perspective envisagée est de fournir des conditions analytiques permettant de **mieux discriminer les algorithmes OMP et OLS sous l’hypothèse de connaissance partielle du support pour des problèmes inverses mal conditionnés**, alors que nos comparaisons actuelles restent empiriques. Il s’agit de borner les facteurs intervenant dans les conditions ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) :

$$\max_{j \notin \mathcal{Q}^*} \|\tilde{\mathbf{C}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{c}}_j\|_1 < 1$$

en fonction du conditionnement de la matrice \mathbf{A} . L'objectif est de mettre en évidence que pour des dictionnaires mal conditionnés, le facteur $\|\tilde{\mathbf{B}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{b}}_j\|_1$ caractérisant OLS est inférieur à celui d'OMP $\|\tilde{\mathbf{A}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{a}}_j\|_1$, et que $\|\tilde{\mathbf{B}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{b}}_j\|_1$ décroît plus rapidement que $\|\tilde{\mathbf{A}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{a}}_j\|_1$ quand \mathcal{Q} croît. Ce type de majoration est particulièrement technique pour le facteur lié à OLS car contrairement à OMP, il n'existe pas de relation récursive explicite reliant $\|\tilde{\mathbf{B}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{b}}_j\|_1$ à $\|\tilde{\mathbf{B}}_{\mathcal{Q}^* \setminus \mathcal{Q}'}^\dagger \tilde{\mathbf{b}}_j\|_1$ pour $\mathcal{Q} \subset \mathcal{Q}'$ [SOU 13b].

La deuxième perspective est d'**adapter nos conditions définies pour un support partiel exact** $\mathcal{Q} \subset \mathcal{Q}^*$ [SOU 13b] **au cas d'un support partiel erroné** $\mathcal{Q} \not\subset \mathcal{Q}^*$. Dans ce cas, la reconstruction des représentations indicées par \mathcal{Q}^* n'est plus possible en k itérations, mais en $k + k'$ itérations avec $k' = \text{Card}[\mathcal{Q} \setminus \mathcal{Q}^*]$. Par exemple, nous avons montré [HER 13, Th. 3] que la condition $\mu < 1/(2k + k' - \text{Card}[\mathcal{Q} \cap \mathcal{Q}^*] - 1)$ basée cohérence mutuelle permet une reconstruction du support $\mathcal{Q} \cup \mathcal{Q}^*$ en $k + k'$ itérations lorsqu'Oxx est initialisé avec le support \mathcal{Q} . Différentes techniques d'analyse utilisant ce concept de reconstruction différée (*delayed recovery* [PLU 07]) ont été proposées récemment pour OMP [FOU 11, ZHA 11b] et OLS [FOU 13a]. Les conditions proposées garantissent l'annulation du résidu après au plus $31k$ itérations sous l'hypothèse RIP. En d'autres termes, le nombre de mauvaises sélections k' est majoré par $30k$. Soulignons que [FOU 13a] ne discrimine pas OMP et OLS tout comme notre condition basée cohérence mutuelle [HER 13].

Une extension de notre technique d'analyse parallèle des deux algorithmes [SOU 13b] au cadre de la reconstruction différée me semble envisageable, mais relier les conditions ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) obtenues pour $\mathcal{Q} \subset \mathcal{Q}^*$ aux conditions adaptées ERC-Oxx($\mathbf{A}, \mathcal{Q}^* \cup \mathcal{Q}, \mathcal{Q}$) pour $\mathcal{Q} \not\subset \mathcal{Q}^*$ n'est pas acquis. De façon générale, il n'y a pas de relation d'implication entre les conditions de type ERC relatives à des supports emboîtés : ERC($\mathbf{A}, \mathcal{Q}^*$) n'implique pas ERC($\mathbf{A}, \mathcal{Q}^* \cup \mathcal{Q}$) [PLU 07, Th. 10].

4.2.3 Forme faible des conditions ERC-Oxx

A l'itération $q = 0$, la condition ERC($\mathbf{A}, \mathcal{Q}^*$) s'écrit :

$$\max_{j \notin \mathcal{Q}^*} \|\mathbf{A}_{\mathcal{Q}^*}^\dagger \mathbf{a}_j\|_1 < 1. \quad \text{ERC}(\mathbf{A}, \mathcal{Q}^*)$$

Son évaluation requiert $n - k$ inversions de systèmes linéaires $\mathbf{A}_{\mathcal{Q}^*}^\dagger \mathbf{a}_j$. La condition ERC au sens faible (*Weak ERC*) est définie par [GRI 08a, Lemme 3] :

$$\max_{j \notin \mathcal{Q}^*} \|\mathbf{A}_{\mathcal{Q}^*}^T \mathbf{a}_j\|_1 + \max_{i \in \mathcal{Q}^*} \|\mathbf{A}_{\mathcal{Q}^* \setminus \{i\}}^T \mathbf{a}_i\|_1 < 1. \quad \text{WERC}(\mathbf{A}, \mathcal{Q}^*)$$

Cette condition plus simple ne fait intervenir que des produits scalaires entre atomes du dictionnaire. Son évaluation est immédiate à partir de la matrice de Gram $\mathbf{A}^T \mathbf{A}$. WERC est une condition suffisante plus forte que la condition ERC, mais beaucoup plus faible que la condition $\mu < 1/(2k - 1)$ portant sur la cohérence mutuelle [TRO 04].

A l'itération $q > 0$, la condition :

$$\max_{j \notin \mathcal{Q}^*} \|\tilde{\mathbf{C}}_{\mathcal{Q}^* \setminus \mathcal{Q}}^\dagger \tilde{\mathbf{c}}_j\|_1 < 1 \quad \text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$$

est encore plus complexe à évaluer que ERC($\mathbf{A}, \mathcal{Q}^*$) puisqu'il faut projeter orthogonalement les atomes du dictionnaire sur $\text{span}(\mathbf{A}_{\mathcal{Q}})^\perp$ en plus des inversions de systèmes linéaires. L'introduction des conditions ERC-Oxx($\mathbf{A}, \mathcal{Q}^*, \mathcal{Q}$) est essentiellement d'ordre « motivationnel » (pour montrer qu'il est possible de relaxer la condition ERC à la q -ème itération). Dans le cas où \mathcal{Q} est

a priori inconnu, l'utilisation concrète de ces conditions nécessite d'évaluer $\text{ERC-Oxx}(\mathbf{A}, \mathcal{Q}^*, \mathcal{Q})$ pour tout $\mathcal{Q} \subset \mathcal{Q}^*$ de cardinalité q donnée. Il me semble intéressant d'**étendre les conditions ERC-Oxx en des conditions faibles** WERC-Oxx dédiées à la q -ème itération d'Oxx, qui seraient plus facilement manipulables dans la mesure où WERC-Oxx ne ferait intervenir que des produits scalaires entre atomes. Cette extension ne peut conduire qu'à des conditions suffisantes (et non nécessaires) de reconstruction à partir partiellement connus, elle ne sera donc d'aucune aide pour discriminer OMP et OLS. Nous avons établi un premier résultat simple utilisant les produits scalaires entre atomes sous la forme d'une condition de type cohérence mutuelle [HER 13] : la condition $\mu < 1/(2k - 1)$ à l'itération 0 est relaxée en $\mu < 1/(2k - q - 1)$ (théorème 6 page 95).

4.2.4 Analyse approfondie des conditions de non-reconstruction

La condition de non-reconstruction (*BRC* pour *Bad Recovery Condition*) pour OMP est une condition suffisante de non-reconstruction du support \mathcal{Q}^* à partir de *toute* combinaison linéaire de $k = \text{Card}[\mathcal{Q}^*]$ atomes avec des pondérations non nulles (*cf.* § 3.8) [SOU 13b]. L'idée sous-jacente est de compléter les conditions existantes qui cherchent à caractériser :

1. les supports garantis d'être reconstruits (condition suffisante de reconstruction exacte) ;
2. ceux pour lesquels il n'y a pas de garantie (la condition nécessaire de reconstruction exacte n'est pas vérifiée : *il existe* une combinaison linéaire d'atomes indicés par \mathcal{Q}^* qui n'est pas reconstruite) ;

par une troisième catégorie de supports :

3. les supports non atteignables quelquesoit la combinaison d'atomes.

Imposer des hypothèses comme la connaissance partielle du support \mathcal{Q}^* pour relaxer les conditions de bonne reconstruction n'a de sens que pour les supports de la catégorie 2 qui ne font pas partie de la catégorie 3. L'analyse approfondie de la condition de non-reconstruction a justement pour but de comprendre si les supports de la catégorie 3 sont marginaux (parmi tous les supports possibles) ou si au contraire, la condition BRC-OMP couvre un vaste ensemble de supports.

Dans la communication [SOU 13c], nous avons montré que pour le problème de déconvolution impulsionnelle par un filtre passe-bas [DOS 05a, LOR 09], la condition BRC-OMP est fréquemment satisfaite dès lors que deux impulsions sont proches, et la condition de non-reconstruction par *basis pursuit* l'est parfois. Une prolongation de ce résultat consistera à établir des liens entre les supports non reconstruits par OMP et par *basis pursuit*, et à étendre la notion de non-reconstruction garantie à un cadre plus large qu'OMP et *basis pursuit*. Par exemple, il serait intéressant de comprendre si le phénomène de non-reconstruction garantie peut également se produire avec la relaxation ℓ_p de la pseudo-norme ℓ_0 ou s'il se restreint à la relaxation convexe.

4.2.5 Analyse d'algorithmes bidirectionnels

L'analyse partielle des algorithmes OMP et OLS est une analyse *sans bruit* en k itérations pour la reconstruction d'un support de cardinalité k . C'est un point de départ vers une analyse étendue des algorithmes bidirectionnels *Bayesian* OMP [HER 10] et SBR [SOU 11] conçus pour minimiser le critère pénalisé $\ell_2\text{-}\ell_0$: $\mathcal{J}(\mathbf{x}; \lambda) = \|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda\|\mathbf{x}\|_0$.

Dans la littérature, la minimisation de critères pénalisés est souvent vue comme une façon naturelle de traiter des problèmes inverses bruités :

$$\mathbf{y} = \mathbf{Ax}^* + \mathbf{b} \tag{4.14}$$

car le résidu \mathbf{b} correspondant à la solution inconnue \mathbf{x}^* est non nul et la solution $\hat{\mathbf{x}}$ obtenue en minimisant $\mathcal{J}(\mathbf{x}; \lambda)$ produit également un résidu non nul pour $\lambda > 0$; voir [FUC 04, FUC 05, DON 06, TRO 06] pour l'analyse de problèmes bruités par minimisation de critères pénalisés ℓ_2 - ℓ_1 .

En proposant d'analyser *Bayesian* OMP et SBR, on souhaite à la fois traiter le cas bruité mais aussi confirmer que sans bruit, ces algorithmes peuvent présenter des meilleures garanties de bonne reconstruction en comparaison avec les algorithmes monodirectionnels OMP et OLS de par la possibilité de suppression d'atomes [JAI 11]. Le cas sans bruit correspond au cas limite $\lambda \rightarrow 0$ avec $\lambda \neq 0$ de façon à autoriser la suppression d'atomes. Pour SBR, notons que les mauvais atomes présents seront nécessairement dé-sélectionnés pour des données non bruitées si le support \mathcal{Q} à l'itération courante vérifie $\mathcal{Q} \supset \mathcal{Q}^*$. Dans ce cas, le retrait de tout mauvais atome $i \in \mathcal{Q} \setminus \mathcal{Q}^*$ fait décroître le critère $\mathcal{J}_{\mathcal{Q}}(\lambda)$ étant donné que $\mathcal{E}_{\mathcal{Q}} = \mathcal{E}_{\mathcal{Q} \setminus \{i\}} = 0$ et $\text{Card}[\mathcal{Q} \setminus \{i\}] < \text{Card}[\mathcal{Q}]$. En d'autres termes, l'obtention d'un support $\mathcal{Q} \supset \mathcal{Q}^*$ garantit la reconstruction exacte de \mathcal{Q}^* .

Plusieurs contributions récentes ont proposé d'analyser des algorithmes bidirectionnels basés OMP sous hypothèse RIP [ZHA 11a, JAI 11]. OMP *with replacement* (OMPR) [JAI 11] est inspiré des algorithmes de seuillage itératif : le support garde toujours la même cardinalité au fil des itérations. Chaque itération effectue l'ajout d'un atome selon une règle de sélection proche de celle d'OMP, puis le retrait d'un atome par seuillage. Pour *Adaptive Forward-Backward Greedy Algorithm* (FoBa) [ZHA 11a], le nombre de retraits par itération dépend de leur impact sur l'augmentation de l'erreur quadratique d'approximation, ce qui rapproche FoBa des algorithmes *Bayesian* OMP et SBR. Jain *et al.* [JAI 11] montrent que l'introduction de retraits permet d'obtenir la condition la moins restrictive en terme de facteur RIP parmi un grand nombre de méthodes dans une analyse au pire cas (pour toute combinaison linéaire de k atomes) : cette condition s'écrit $\delta_{2k} < 0,499$. Les techniques d'analyse de [ZHA 11a, JAI 11] s'appuient toutes les deux sur l'hypothèse RIP. Une question intéressante est d'étendre ces résultats à un cadre plus général.

La technique envisagée pour analyser les algorithmes bidirectionnels repose sur la minimisation du critère $\mathcal{J}(\mathbf{x}; \lambda) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$ ou de façon équivalente, sur le problème d'optimisation combinatoire :

$$\min_{\mathcal{Q}} \{ \mathcal{J}_{\mathcal{Q}}(\lambda) \triangleq \mathcal{E}_{\mathcal{Q}} + \lambda \text{Card}[\mathcal{Q}] \} \quad (4.15)$$

qui porte sur tous les supports $\mathcal{Q} \subset \{1, \dots, n\}$. Pour des données bruitées (4.14) associées à un vecteur \mathbf{x}^* parcimonieux, le support inconnu $\mathcal{Q}^* = \mathcal{S}(\mathbf{x}^*)$ n'est pas nécessairement un minimiseur global de (4.15). Il s'agit alors de **fournir des garanties permettant de relier les supports obtenus $\hat{\mathcal{Q}}$ en sortie de *Bayesian* OMP et SBR à \mathcal{Q}^* et $\mathbf{a}(\mathbf{x})$ minimiseur global de (4.15)** voire de garantir sous certaines conditions que $\hat{\mathcal{Q}}$ est un minimiseur global. Remarquons que le support $\hat{\mathcal{Q}}$ obtenu en sortie de SBR est un minimum local de $\lambda \mapsto \mathcal{J}_{\mathcal{Q}}(\lambda)$ de par la définition de SBR : tous les remplacements élémentaires $\hat{\mathcal{Q}} \bullet i$ font croître le critère $\mathcal{J}_{\mathcal{Q}}(\lambda)$ (*cf.* tableau 1.4 page 45). Ce n'est pas le cas de *Bayesian* OMP qui adopte une règle plus simple dans la phase de sélection d'atomes, de la même façon qu'OMP adopte une règle plus simple qu'OLS en évitant de résoudre des problèmes de moindres carrés [BLU 07] (*cf.* section 3.3). Pour *Bayesian* OMP, une question supplémentaire est donc de garantir que le support $\hat{\mathcal{Q}}$ obtenu est bien un minimum local de $\lambda \mapsto \mathcal{J}_{\mathcal{Q}}(\lambda)$.

4.3 Analyse de l'invasivité de cellules cancéreuses par microscopie AFM

4.3.1 Bilan et contexte

Collaborateurs

- Claire Barbieux, David Brie, Stéphanie Grandemange, Rémi Pannequin (CRAN).
- Grégory Francius (LCPME).
- Angéline Razafitianamaharavo (LIEC).

Financement

La mise en place de ce projet est financée pour une année (2013-2014) par le programme « Projets Exploratoires Premier Soutien » (PEPS Mirabelle) conjoint au CNRS et à la Région Lorraine : projet BIOFORCE (*Analyse d'images PeakForce Mapping de tissus biologiques comme outil de diagnostic et de caractérisation de cellules tumorales et métastatiques*) dont je suis le porteur. A l'issue de cette mise en place, nous envisageons de déposer un plus gros projet de type ANR ou Plan Cancer.

Département SBS du CRAN

Ce projet est la poursuite de la collaboration entre le CRAN et le LCPME. Son objectif est d'étendre l'approche proposée dans [POL 11] à l'analyse de cellules cancéreuses. La nouveauté significative est que l'on étudie le potentiel de la microscopie AFM comme une aide à la décision thérapeutique pour le cancer. Le projet intègre de nouvelles compétences en biologie cellulaire (S. Grandemange, C. Barbieux). Le LIEC (nouvellement créé en 2013, ex Laboratoire Environnement et Minéralurgie - LEM) est un partenaire privilégié du LCPME concernant l'acquisition de données AFM puisqu'il dispose depuis 2011 d'un microscope AFM de dernière génération doté de la technologie ultra-rapide (*Dimension FastScan*, Bruker AXS), qui permet l'acquisition d'images nettement plus résolues et précises que les images force-volume acquises jusqu'ici avec le microscope Asylum du LCPME.

Ce projet s'inscrit pleinement dans la politique scientifique du département SBS (Santé – Biologie – Signal) du CRAN, créé en 2013, qui vise à développer des approches innovantes pour l'analyse et la compréhension du cancer en favorisant l'interaction de chercheurs en biologie, neurologie et traitement du signal autour de projets d'ingénierie pour la santé, et en particulier de projets orientés cancer. Stéphanie Grandemange (rattachée au CRAN depuis 2013) est spécialiste de la modification génétique de cellules cancéreuses *via* des protéines spécifiques permettant de moduler (augmenter ou diminuer) leur invasivité [MCE 12, ENN 11, ENN 13].

Le projet comprend également un ingénieur de recherche (R. Pannequin) travaillant sur la valorisation d'un logiciel de traitement automatique d'images force-volume en microscopie AFM. Ce logiciel doit être adapté à la dernière génération de microscopes AFM, dotée du mode *PeakForce Mapping*. Cette technologie d'imagerie AFM ultra-rapide est apparue en 2011 ; elle permet d'acquérir pour une même image, cent fois plus de courbes (256×256 pixels contre 32×32 pixels pour une image force-volume standard) en un temps dix fois plus rapide pour une résolution et précision des mesures qui augmente dans la même proportion.

Etat actuel du projet sur la microscopie AFM

La collaboration entre le CRAN (C. Soussen, D. Brie) et le LCPME (F. Gaboriaud puis G. Francius) sur le projet de la microscopie AFM a débuté en 2007. Elle a conduit au développement d'un **outil automatique de traitement** numérique des images force-volume utilisant les modèles physiques de l'interaction (électrostatique et mécanique) entre deux nano-objets (*cf.* chapitre 2). Contrairement à beaucoup d'approches *ad hoc* de la littérature, cette méthode nécessite très peu de réglages empiriques et permet de reconstruire des images 2D représentant la **cartographie de paramètres physico-chimiques** à partir de données expérimentales.

Jusqu'à présent, nous nous sommes concentrés sur l'analyse d'échantillons bactériens présents dans l'environnement. Le projet futur concerne l'utilisation de la microscopie AFM pour mieux comprendre les phénomènes physiques complexes se produisant en **biologie cellulaire**. Plus précisément, nous proposons d'analyser des cellules biologiques cancéreuses dans le but de déterminer leur pouvoir invasif : **la cellule cancéreuse imagée par le microscope est-elle métastatique ou non ?**

4.3.2 Analyse de cellules cancéreuses

Le diagnostic clinique systématique du comportement métastatique de cellules cancéreuses reste une perspective à long terme. Cependant, les techniques spectroscopiques et microscopiques à l'échelle infra-cellulaire ont récemment fourni des éléments prometteurs pour **caractériser qualitativement l'invasivité de cellules cancéreuses**. Une caractérisation quantitative précise reste un problème ouvert. L'objectif du projet est de **concevoir des méthodes avancées de traitement du signal** pour les données acquises en microscopie AFM. Il s'agira d'estimer des propriétés adhésives qui sont censées différer pour les cellules cancéreuses et de les relier aux propriétés biologiques.

Les cellules tumorales mammaires étudiées peuvent être génétiquement modifiées de façon à moduler leur invasivité et leur faculté migratoire (c'est-à-dire leur comportement métastatique)²⁷. En effectuant ces modifications avec plusieurs gènes dont l'influence sur les cellules diffère, il est possible de moduler le comportement métastatique, *i.e.*, créer des cellules modifiées hautement métastatiques, ou au contraire diminuer le caractère métastatique des cellules. Pour pouvoir caractériser précisément la faculté migratoire de différentes lignées de cellules, des données AFM seront acquises à partir de plusieurs lignées de cellules cancéreuses mammaires génétiquement modifiées. L'analyse d'une cellule sera effectuée par la modalité *PeakForce Mapping* de la microscopie AFM. Ces données apporteront un éclairage sur les propriétés mécaniques et biologiques de cellules dans le but de distinguer les cellules métastatiques des cellules non métastatiques.

Les résultats attendus sont :

1. la validation que l'on peut effectivement distinguer des cellules (génétiquement modifiées) cancéreuses métastatiques des cellules cancéreuses non métastatiques provenant d'une même lignée cellulaire par microscopie AFM ;
2. comprendre la capacité de changement invasif pour d'autres types de modifications génétiques (pour d'autres protéines) ;
3. la conception de nouvelles méthodes de traitement du signal pour l'analyse d'images force-volume et au format *PeakForce Mapping*.

27. Plus précisément, les biologistes du CRAN étudient des modifications génétiques de façon à contrôler le taux de la protéine DDB2 (*Damage DNA Binding 2*) dans les cellules, car cette protéine influence très fortement l'invasivité des cellules cancéreuses [ENN 13].

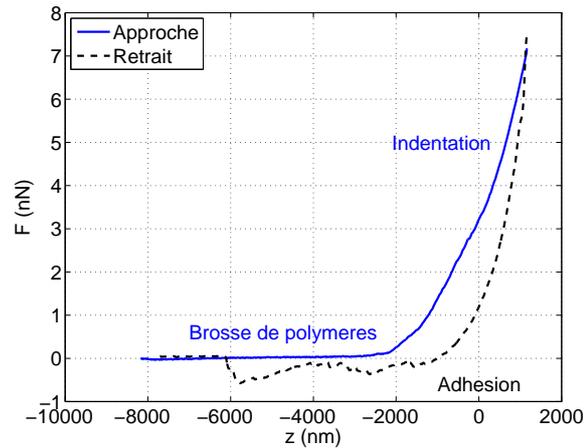


FIGURE 4.3 – Modèle de courbe de force pour une cellule biologique.

Différentes études ont démontré que les cellules métastatiques sont moins adhésives que les cellules non métastatiques [CRO 07, CRO 08, WAT 12, CRO 11] (ce qui leur permet de migrer plus facilement d'un organe à un autre) et que leur surface est moins rigide que les cellules non métastatiques, leur conférant un caractère hautement invasif [XU 12]²⁸. L'analyse des courbes de forces mesurées par la modalité *PeakForce Mapping* doit permettre de reconstruire des paramètres physiques liés à la rigidité de la cellule pour la phase d'approche, et à l'énergie d'adhésion dans la phase de retrait. Notre projet est donc de concevoir ces outils d'analyse quantitative qui devraient nous conduire à des résultats sensiblement différents pour des cellules cancéreuses ou non.

4.3.3 Traitement du signal

Besoin de nouveaux outils de traitement du signal non supervisé

En biologie, l'interprétation quantitative des données repose souvent sur des outils de traitement soit très simples, soit plus complexes mais empiriques. L'analyse d'images force-volume que nous proposons a pour but de fournir une cartographie, *i.e.*, de reconstruire la distribution spatiale des paramètres physiques pertinents pour caractériser un échantillon biologique.

Dans [POL 11], nous avons proposé une méthode non-supervisée pour segmenter des courbes de force, où les régions d'intérêt sont délimitées par des points de changement (saut du signal, changement de pente, changement de courbure) qui sont estimés de façon adaptative. Cette méthode a été appliquée avec succès aux données acquises sur des bactéries (*cf.* chapitre 2). Cette approche repose sur l'hypothèse que les différentes interactions (par exemple l'interaction avant contact, puis l'interaction mécanique déformable et non déformable) correspondent à des régions disjointes de la courbe de force délimitées par des points de discontinuité. Pour les cellules biologiques, cette hypothèse n'est pas vérifiée pour les données expérimentales que nous avons déjà acquises : les courbes de force sont relativement lisses. Un exemple est donné sur la figure 4.3. On distingue deux zones sur la courbe d'approche :

- une première zone, avant contact, liée à l'interaction de la pointe AFM avec la brosse de polymères recouvrant la cellule ;

²⁸. Le lecteur intéressé est renvoyé à l'annexe A pour une revue de littérature, effectuée par S. Grandemange, concernant la microscopie AFM comme outil de diagnostic de tumeurs.

- une zone, après contact, qui résulte de l'interaction mécanique de la pointe avec la cellule (déformation/indentation de la cellule) ;
- enfin, tout le signal dans la phase de retrait est lié à l'adhésion de la pointe avec la bactérie.

Pour résumer, l'allure des courbes de force (et les modèles d'interaction que l'on utilisera) diffère des courbes observées dans le cas de bactéries. Pour la phase d'approche, on observe deux régions d'interaction (avant contact, et contact déformable) alors que pour les bactéries, on observait une troisième interaction liée au contact non déformable. Cette troisième interaction n'est plus présente car une cellule est un objet mou. Pour les courbes de retrait, on n'observe plus des décrochages successifs comme pour l'analyse de polymères. Si les outils de traitement disponibles sont assez facilement adaptables pour les courbes d'approche [POL 11], ce n'est pas le cas des courbes de retrait et de nouvelles solutions doivent être proposées.

Traitements envisagés basés sur les modèles physiques

Le traitement des courbes d'approche sera effectué en appliquant (moyennant quelques adaptations) la méthode déjà disponible au laboratoire [POL 11]. Cela devrait permettre de reconstruire des images 2D représentant chacune un paramètre physique (la topographie, le module d'Young caractéristique de l'élasticité de l'objet, *etc.*) et de vérifier que les paramètres de rigidité sont effectivement très différents pour les cellules métastatiques et non métastatiques.

Pour les courbes de retrait, la conception de nouveaux outils de traitement nécessitera au préalable une étude bibliographique concernant les modèles physiques adéquats (modèles adhésifs d'interactions entre cellules). Un grand nombre de modèles sont envisageables. Parmi les plus simples, citons les modèles JKR (Johnson, Kendall, Roberts) et DMT (Derjagin, Muller, Toropov) [BUT 05]. Il s'agira de recenser les modèles disponibles [BUT 05, JAN 00, LIN 07a] et de les valider avec des données réelles déjà acquises. Une étude préliminaire des modèles de contacts adhésifs de types JKR et Sirghi [LIN 07b, SIR 08] a déjà été effectuée au CRAN en 2012 [LAH 12].

Par la suite, le traitement de courbes de retrait en utilisant ces modèles devrait permettre la reconstruction de cartes spatiales 2D représentant l'énergie d'adhésion et la viscosité. En particulier, nous espérons discriminer de façon significative les cellules tumorales métastatiques des cellules tumorales non métastatiques en comparant leur énergie d'adhésion.

Autres développements méthodologiques envisagés pour le traitement de données

Nous envisageons de développer des nouvelles approches pour le traitement d'images force-volume visant à étudier les effets de la convolution due à la géométrie de la pointe AFM et de les retirer des données acquises. En effet, lors de l'acquisition des courbes de force sur une grille très résolue spatialement, la pointe n'est pas en contact avec un seul pixel mais également avec ses voisins (fig. 4.4), induisant un phénomène de convolution [SOU 08a]. Retirer l'effet de la convolution sur les données est un problème ouvert car tous les outils de traitement actuels considèrent chaque pixel indépendamment des autres. Ce problème n'est pas un problème classique de déconvolution d'image car le modèle convolutif est tridimensionnel.

Une autre perspective intéressante est d'améliorer le traitement effectué dans [POL 11], où une courbe de force est modélisée par des morceaux disjoints (*e.g.*, interactions électrostatique, mécanique avec déformation de l'objet, puis mécanique sans déformation de l'objet). Un modèle plus réaliste est basé sur le chevauchement de morceaux. On considère alors que les observations peuvent résulter d'un mélange de différents types d'interactions (*e.g.*, les modèles électrostatique et mécanique sont actifs simultanément autour du point de contact). Nous avons étudié la

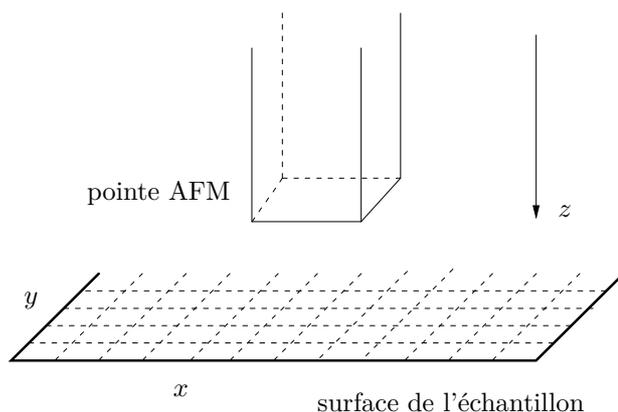


FIGURE 4.4 – Prise en compte de la géométrie de la pointe AFM. Une géométrie de forme parallélépipédique induit une convolution 2D dans le domaine (x, y) mais des formes plus réalistes (conique, parabolique) induisent une convolution 3D dans le domaine (x, y, z) .

validation de ce modèle dans la thèse de Junbo Duan et proposé des solutions préliminaires pour la séparation de sources [DUA 10]. Nous pensons qu'une telle procédure permettra d'améliorer les estimations quantitatives de grandeurs physiques.

4.3.4 Liens avec d'autres projets pour le diagnostic du cancer

Je suis le porteur du sujet de la microscopie AFM. Ce sujet n'est cependant pas un sujet isolé dans l'équipe : il fait partie d'un projet plus large dont le but est d'associer différentes techniques (spectroscopie et imagerie Raman, cytométrie en flux, microscopie AFM) pour observer différents paramètres physiques et biologiques à différentes échelles (les autres chercheurs impliqués sont David Brie et Sebastian Miron pour le CRAN et Manuel Dossot pour le LCPME). Les autres modalités de spectroscopie moléculaire utilisées (en dehors de la microscopie AFM) permettent de caractériser les cellules cancéreuses à l'échelle macroscopique (à l'échelle d'une population de cellules) ou à l'échelle d'une cellule avec une résolution spatiale limitée, alors que la microscopie AFM permet une caractérisation d'une cellule à l'échelle nanométrique, *i.e.*, de comprendre les interactions mécaniques se produisant à l'intérieur d'une cellule. Par cette approche multi-échelle, nous espérons comprendre comment les processus physico-chimiques induisent des effets biologiques et caractériser précisément leur signification statistique : un phénomène physique affecte-t-il la population entière ou seulement une partie de la population de cellules ? Répondre à cette question repose fortement sur le développement de méthodes originales de traitement du signal et des images pour estimer quantitativement les propriétés vibrationnelles, mécaniques et biologiques à partir de données acquises sur différents types de cellules cancéreuses.

Annexe A

Intérêt de la microscopie AFM pour le diagnostic de tumeurs

Cette revue de littérature sur le potentiel de l'AFM en cancérologie a été effectuée par Stéphanie Grandemange, collègue biologiste du CRAN.

Au cours de la dernière décennie, la microscopie AFM a été intensément utilisée pour la recherche biologique et en particulier pour la recherche sur le cancer [LEK 09, CRO 09, KAS 08, SEN 10]. Cette technologie permet, en balayant directement la surface de cellules sur leur support naturel, d'obtenir une information topographique mais aussi de mesurer des forces d'adhésion de cellules. Durant la genèse de tumeurs, les cellules sont sujettes à plusieurs modifications et diffèrent des cellules normales en termes de taille, morphologie, d'adhésion à des matrices extra-cellulaires, d'organisation de cyto squelette et d'interaction cellulaire. L'AFM est capable de détecter la plupart de ces changements. En cancérologie, l'AFM est principalement utilisée pour des recherches exploratoires sur la morphologie de cellules cancéreuses [KON 09], la rigidité des cellules [KUZ 07], ou pour explorer les changements en temps-réel de l'ultrastructure de la membrane cellulaire induits par les interactions avec des drogues [WAN 09]. De plus, une étude récente montre que la rigidité cellulaire, déterminée par mesures AFM, pourrait être un biomarqueur potentiel pour les cellules cancéreuses ovariennes [XU 12]. En effet, **la rigidité des cellules** discrimine non seulement les cellules cancéreuses ovariennes des cellules bénignes mais de plus, **discrimine les cellules cancéreuses invasives des cellules non invasives**. Dans la lignée de cette étude, il a été démontré que les biopsies de cancer mammaire présentent un profil particulier de rigidité en comparaison avec les tissus sains, comme l'attestent les mesures en AFM [PLO 12]. Les auteurs concluent que la microscopie AFM pourrait être un outil intéressant pour le diagnostic clinique.

Une thèse récente [MCE 12, WU 10] a mis en évidence l'intérêt de plusieurs modalités de spectroscopie pour le diagnostic du caractère métastatique (ou non métastatique) de tumeurs cancéreuses. Pour cela, la thèse propose de trouver des propriétés caractéristiques d'objets biologiques et de comprendre comment ces propriétés interagissent avec l'environnement biologique des cellules. Cette thèse montre que d'autres équipes s'intéressent au même sujet. Cependant, notre approche est originale car nous proposons des méthodes de traitement du signal non supervisées pour caractériser les propriétés invasives de cellules cancéreuses (ce qui n'est pas fait dans [MCE 12, WU 10]).

Index

- Algorithme ℓ_0 -PT, 58–61
- Algorithme *Bayesian* OMP, 45, 115
- Algorithme *Iterative Hard Thresholding*, 40, 48
- Algorithme CSBR, 51, 54
- Algorithme d’homotopie, 51, 54
- Algorithme glouton bidirectionnel, 24, 40, 114
- Algorithme glouton monodirectionnel, 24, 40
- Algorithme glouton orthogonal, 40, 87
- Algorithme $IR\ell_1$, 48, 105
- Algorithme MP, 41, 106
- Algorithme OLS, 40, 42, 80, 92, 99, 111
- Algorithme OMP, 40, 42, 92, 111
- Algorithme OMP positif, 106
- Algorithme SBR, 45, 80, 107, 114
- Algorithme SMLR, 24, 44
- Analyse au pire cas, 26, 92, 101
- Analyse d’OMP/OLS en k itérations, 26, 88
- Analyse différée, 102, 113
- Analyse partielle, 89, 93, 112
- Approche bayésienne, 43
- Approches analyse *vs* synthèse, 77, 110
- Approximation parcimonieuse, 23, 37, 78

- Bactérie, 18, 71, 85
- Basis poursuit, 100
- Bio-capteur bactérien, 20

- Calibration d’un système Procam, 28
- Cancer, 116, 121
- Cellule métastatique, 117, 121
- Chemin de régularisation, 51, 53
- Cohérence mutuelle, 27, 38, 95
- Compressive sensing, 37
- Condition *Weak* ERC, 97, 113
- Condition BRC, 98, 114
- Condition d’isométrie restreinte (RIP), 38, 91
- Condition ERC, 26, 92
- Condition ERC-Oxx, 27, 93, 113
- Conditionnement, 23, 38, 113
- Continuation, 51, 54

- Contrôle non destructif, 15, 25, 112
- Courbe de force, 18, 69
- Critère mixte ℓ_2 - ℓ_0 , 23, 38, 77, 87
- Critère mixte ℓ_2 - ℓ_1 , 51, 53, 105

- Déconvolution d’image force-volume, 119
- Déconvolution d’image hyperspectrale, 21
- Déconvolution impulsionnelle, 47, 96
- Dictionnaire, 23, 37, 78
- Dictionnaire convolutif, 48, 97
- Dictionnaire corrélé, 27, 48, 96, 98, 111
- Dictionnaire de grande taille, 25, 109
- Dictionnaire paramétrique, 25, 109
- Dictionnaire polynomial, 25, 49, 79
- Dictionnaire structuré, 107
- Dictionnaire Tœplitz, 47, 78

- Endoscopie, 27
- Estimation de mouvement 3D, 29

- Fluorescence, 20
- Frontière de Pareto, 52

- Image 3D parcimonieuse, 16
- Imagerie force-volume, 17, 71
- Imagerie hyperspectrale, 21
- Indentation, 73, 85

- Lissage par morceaux, 25, 76, 83
- Longueur de Debye, 72
- Longueur de Kuhn, 74
- Lumière structurée, 28

- Maximum *a posteriori*, 24, 43
- Microscopie AFM, 17, 68, 116, 121
- Microscopie confocale, 20
- Modèle électrostatique, 72
- Modèle Bernoulli-gaussien, 24, 43
- Modèle de Hertz, 73
- Modèle de Hooke, 74
- Modèle FJC, 74
- Modèle polynomial par morceaux, 25, 49, 79

Modèle WLC, 74
Modèles physiques en microscopie AFM, 72, 119
Modification génétique, 117
Mosaïquage d'images, 28
Motifs paramétriques, 25, 108

Norme mixte, 78

Optimisation bi-objectif, 51
Optimisation combinatoire, 39, 43, 87
Optimisation locale, 44, 46, 109

Pénalisation ℓ_0 , 24, 43
Parcimonie de groupe, 86
Polymère, 70, 74
Positivité, 104
Problèmes inverses, 16, 18, 22, 37, 43, 111
Projection orthogonale, 26, 44, 89

Recalage d'images, 27
Reconstruction exacte de support, 25, 87
Reconstruction exacte par OLS, 26, 93
Reconstruction surfacique, 29
Reconstruction tomographique, 15
Relaxation convexe, 39, 105
Relaxation non-convexe, 39
Remplacement élémentaire, 44, 55, 59, 82, 107

Sélection de variables, 23, 39, 78–80
Sélection de variables vectorielles, 81
Séparation d'image hyperspectrale, 21
Segmentation, 18, 49, 76, 84
Spectroscopie de fluorescence, 20
Spectroscopie de force, 69
Stabilité, 38, 46
Super-résolution, 108

Transition de phase, 96, 99–101

Vision active, 28

Bibliographie

- [AFO 10] AFONSO M. V., BIOUCAS-DIAS J. M., FIGUEIREDO M. A. T., « Fast image recovery using variable splitting and constrained optimization », *IEEE Trans. Image Process.*, vol. 19, n° 9, pp. 2345–2356, sep. 2010.
- [AFO 11] AFONSO M. V., BIOUCAS-DIAS J. M., FIGUEIREDO M. A. T., « An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems », *IEEE Trans. Image Process.*, vol. 20, n° 3, pp. 681–695, mars 2011.
- [AKA 74] AKAIKE H., « A new look at the statistical model identification », *IEEE Trans. Automat. Contr.*, vol. AC-19, n° 6, pp. 716–723, déc. 1974.
- [ALL 07] ALLAIN M., IDIER J., « Efficient binary reconstruction for non-destructive evaluation using gammagraphy », *Inverse Problems*, vol. 23, n° 4, pp. 1371–1393, août 2007.
- [AUS 10] AUSTIN C. D., MOSES R. L., ASH J. N., ERTIN E., « On the relation between sparse reconstruction and parameter estimation with model order selection », *IEEE J. Sel. Top. Signal Process.*, vol. 4, n° 3, pp. 298–309, juin 2010.
- [BAB 10] BABAIE-ZADEH M., JUTTEN C., « On the stable recovery of the sparsest overcomplete representations in presence of noise », *IEEE Trans. Signal Process.*, vol. 58, n° 10, pp. 5396–5400, oct. 2010.
- [BAC 12] BACH F., JENATTON R., MAIRAL J., OBOZINSKI G., « Optimization with sparsity-inducing penalties », *Foundations and Trends in Machine Learning*, vol. 4, n° 1, pp. 1–106, jan. 2012.
- [BAN 13] BANDEIRA A. S., DOBRIBAN E., MIXON D. G., SAWIN W. F., « Certifying the restricted isometry property is hard », *IEEE Trans. Inf. Theory*, vol. 59, n° 6, pp. 3448–3450, juin 2013.
- [BAR 08] BARDSLEY J., MERIKOSKI J., VIO R., « The stabilizing properties of nonnegativity constraints in least-squares image reconstruction », *Int. J. of Pure and Applied Mathematics*, vol. 43, n° 1, pp. 95–109, 2008.
- [BEC 09] BECK A., TEOULLE M., « A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems », *SIAM J. Imaging Sci.*, vol. 2, n° 1, pp. 183–202, mars 2009.
- [BEN 10] BEN-HAIM Z., ELДАР Y. C., ELAD M., « Coherence-based performance guarantees for estimating a sparse vector under random noise », *IEEE Trans. Signal Process.*, vol. 58, n° 10, pp. 5030–5043, oct. 2010.
- [BER 80] BERK K. N., « Forward and backward stepping in variable selection », *J. Statist. Comput. Simul.*, vol. 10, n° 3-4, pp. 177–185, avr. 1980.
- [BER 09] BERTERO M., BOCCACCI P., DESIDERÀ G., VICIDOMINI G., « Image deblurring with Poisson data : from cells to galaxies », *Inverse Problems*, vol. 12, pp. 1–26, déc. 2009.
- [BIG 04] BIGGS D. S. C., « Clearing up deconvolution », *Biophotonics International*, vol. 11, n° 2, pp. 32–37, fév. 2004.

- [BIN 86] BINNIG G., QUATE C. F., GERBER C., « Atomic force microscope », *Physical Review Letters*, vol. 56, n° 9, pp. 930–933, mars 1986.
- [BIN 88] BINNIG G., « Atomic force microscope and method for imaging surfaces with atomic resolution », *US Patent*, vol. 4, n° 724, pp. 1-318, fév. 1988.
- [BIO 12] BIOCAS-DIAS J. M., PLAZA A. J., DOBIGEON N., PARENTE M., DU Q., GADER P. D., CHANUSSOT J., « Hyperspectral unmixing overview : Geometrical, statistical, and sparse regression-based approaches », *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, n° 2, pp. 354-379, avr. 2012.
- [BLA 87] BLAKE A., ZISSERMAN A., *Visual Reconstruction*, The MIT Press, Cambridge, MA, USA, jan. 1987.
- [BLA 89] BLAKE A., « Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction », *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-11, n° 1, pp. 2–12, jan. 1989.
- [BLU 07] BLUMENSATH T., DAVIES M. E., On the difference between Orthogonal Matching Pursuit and Orthogonal Least Squares, Rapport technique, University of Edinburgh, mars 2007.
- [BLU 08] BLUMENSATH T., DAVIES M. E., « Iterative thresholding for sparse approximations », *J. Fourier Anal. Appl.*, vol. 14, n° 5, pp. 629-654, déc. 2008.
- [BLU 10] BLUMENSATH T., DAVIES M. E., « Normalized iterative hard thresholding : Guaranteed stability and performance », *IEEE J. Sel. Top. Signal Process.*, vol. 4, n° 2, pp. 298-309, avr. 2010.
- [BOU 09] BOUGLEUX S., PEYRÉ G., COHEN L. D., « Image compression with geodesic anisotropic triangulations », *ICCV*, Kyoto, Japon, pp. 2343-2348, oct. 2009.
- [BOU 11a] BOURGUIGNON S., MARY D., SLEZAK E., « Restoration of astrophysical spectra with sparsity constraints : Models and algorithms », *IEEE J. Sel. Top. Signal Process.*, vol. 5, n° 5, pp. 1002-1013, sep. 2011.
- [BOU 11b] BOURGUIGNON S., SOUSSEN C., CARFANTAN H., IDIER J., « Sparse deconvolution : Comparison of statistical and deterministic approaches », *IEEE Workshop Stat. Sig. Proc.*, Nice, pp. 317-320, juin 2011.
- [BOY 11] BOYD S., PARIKH N., CHU E., PELEATO B., ECKSTEIN J., « Distributed optimization and statistical learning *via* the alternating direction method of multipliers », *Foundations and Trends in Machine Learning*, vol. 3, n° 1, pp. 1–122, Now Publishers Inc., juil. 2011.
- [BRE 96] BRETTE S., IDIER J., « Optimized single site update algorithms for image deblurring », *Proc. IEEE ICIP*, vol. 3, Lausanne, Suisse, pp. 65-68, sep. 1996.
- [BRO 86] BROERSEN P. M. T., « Subset regression with stepwise directed search », *J. R. Statist. Soc. C*, vol. 35, n° 2, pp. 168-177, 1986.
- [BRU 08] BRUCKSTEIN A. M., ELAD M., ZIBULEVSKY M., « On the uniqueness of nonnegative sparse solutions to underdetermined systems of equation », *IEEE Trans. Inf. Theory*, vol. 54, n° 11, pp. 4813-4820, nov. 2008.

- [BRU 09] BRUCKSTEIN A. M., DONOHO D. L., ELAD M., « From sparse solutions of systems of equations to sparse modeling of signals and images », *SIAM Rev.*, vol. 51, n° 1, pp. 34-81, fév. 2009.
- [BUT 05] BUTT H.-J., CAPPELLA B., KAPPL M., « Force measurements with the atomic force microscope : Technique, interpretation and applications », *Surf. Sci. Rep.*, vol. 59, n° 1-6, pp. 1-152, oct. 2005.
- [CAI 10] CAI T. T., XU G., « Stable recovery of sparse signals using an oracle inequality », *IEEE Trans. Inf. Theory*, vol. 56, n° 7, pp. 3516-3522, juil. 2010.
- [CAI 11] CAI T. T., WANG L., « Orthogonal matching pursuit for sparse signal recovery with noise », *IEEE Trans. Inf. Theory*, vol. 57, n° 7, pp. 4680-4688, juil. 2011.
- [CAN 08a] CANDÈS E. J., WAKIN M. B., BOYD S. P., « Enhancing sparsity by reweighted ℓ_1 minimization », *J. Fourier Anal. Appl.*, vol. 14, n° 5-6, pp. 877-905, déc. 2008.
- [CAN 08b] CANDÈS E. J., WAKIN M. B., « An introduction to compressive sampling », *IEEE Trans. Signal Process. Mag.*, vol. 25, n° 2, pp. 21-30, mars 2008.
- [CAN 11] CANDÈS E. J., PLAN Y., « A probabilistic and RIPless theory of compressed sensing », *IEEE Trans. Inf. Theory*, vol. 57, n° 11, pp. 7235-7254, nov. 2011.
- [CAR 13] CARCREFF E., BOURGUIGNON S., IDIER J., SIMON L., « Resolution enhancement of ultrasonic signals by up-sampled sparse deconvolution », *Proc. IEEE ICASSP*, Vancouver, Canada, pp. 6511-6515, mai 2013.
- [CAR 14] CARCREFF E., Déconvolution adaptative de signaux ultrasonores pour le contrôle de matériaux complexes, thèse de doctorat, IRCCyN, Nantes, en cours, 2011-2014.
- [CHA 96] CHAMPAGNAT F., GOUSSARD Y., IDIER J., « Unsupervised deconvolution of sparse spike trains using stochastic approximation », *IEEE Trans. Signal Process.*, vol. 44, n° 12, pp. 2988-2998, déc. 1996.
- [CHA 01] CHAMPAGNAT F., GOUSSARD Y., GAUTIER S., IDIER J., « Déconvolution impulsionnelle », IDIER J., Ed., *Approche bayésienne pour les problèmes inverses*, Paris, Traité IC2, Série traitement du signal et de l'image, Hermès, pp. 115-138, nov. 2001.
- [CHA 03] CHAN M., LIN W., ZHOU C., QU J. Y., « Miniaturized three-dimensional endoscopic imaging system based on active stereovision », *Applied Optics*, vol. 42, n° 10, pp. 1888-1898, avr. 2003.
- [CHA 12] CHATTERJEE S., SUNDMAN D., VEHKAPERÄ M., SKOGLUND M., « Projection-based and look-ahead strategies for atom selection », *IEEE Trans. Signal Process.*, vol. 60, n° 2, pp. 634-647, fév. 2012.
- [CHE 89] CHEN S., BILLINGS S. A., LUO W., « Orthogonal least squares methods and their application to non-linear system identification », *Int. J. Control*, vol. 50, n° 5, pp. 1873-1896, nov. 1989.
- [CHE 95] CHEN S., WIGGER J., « Fast orthogonal least squares algorithm for efficient subset model selection », *IEEE Trans. Signal Process.*, vol. 43, n° 7, pp. 1713-1715, juil. 1995.

- [CHE 96] CHENG Q., CHEN R., LI T.-H., « Simultaneous wavelet estimation and deconvolution of reflection seismic signals », *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 377–384, mars 1996.
- [CHE 01] CHEN S. S., DONOHO D. L., SAUNDERS M. A., « Atomic decomposition by basis pursuit », *SIAM Rev.*, vol. 43, n° 1, pp. 129–159, 2001.
- [CHE 11] CHEN J., RICHARD C., BERMUDEZ J. C. M., HONEINE P., « Nonnegative least-mean-square algorithm », *IEEE Trans. Signal Process.*, vol. 59, n° 11, pp. 5225–5235, nov. 2011.
- [CHI 84] CHI C. Y., MENDEL J. M., « Improved maximum-likelihood detection and estimation of Bernoulli-Gaussian processes », *IEEE Trans. Inf. Theory*, vol. 30, pp. 429–435, mars 1984.
- [CHO 10] CHOZENOUX E., Recherche de pas par majoration-minoration. Application à la résolution de problèmes inverses, thèse de doctorat, Écode Centrale de Nantes, IRCCyN, déc. 2010.
- [COM 07] COMBETTES P. L., PESQUET J.-C., « A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery », *IEEE J. Sel. Top. Signal Process.*, vol. 1, n° 4, pp. 564–574, déc. 2007.
- [COT 99] COTTER S. F., ADLER J., RAO B. D., KREUTZ-DELGADO K., « Forward sequential algorithms for best basis selection », *IEE Proc. Vision, Image and Signal Processing*, vol. 146, n° 5, pp. 235–244, oct. 1999.
- [COU 00] COUVREUR C., BRESLER Y., « On the optimality of the backward greedy algorithm for the subset selection problem », *SIAM J. Matrix Anal. Appl.*, vol. 21, n° 3, pp. 797–808, fév. 2000.
- [CRO 07] CROSS S. E., JIN Y. S., RAO J., GIMZEWSKI J. K., « Nanomechanical analysis of cells from cancer patients », *Nature nanotechnology*, vol. 2, n° 12, pp. 780–783, déc. 2007.
- [CRO 08] CROSS S. E., JIN Y. S., TONDRE J., GIMZEWSKI J. K., « AFM-based analysis of human metastatic cancer cells », *Nanotechnology*, vol. 19, n° 38, pp. 384003.1–384003.8, sep. 2008.
- [CRO 09] CROSS S. E., JIN Y.-S., RAO J., GIMZEWSKI J. K., « Applicability of AFM in cancer detection », *Nature nanotechnology*, vol. 4, n° 2, pp. 72–73, fév. 2009.
- [CRO 11] CROSS S. E., JIN Y. S., LU Q. Y., RAO J. Y., GIMZEWSKI J. K., « Green tea extract selectively targets nanomechanics of live metastatic cancer cells », *Nanotechnology*, vol. 22, n° 21, page 215101, mai 2011.
- [DAI 09] DAI W., MILENKOVIC O., « Subspace pursuit for compressive sensing signal reconstruction », *IEEE Trans. Inf. Theory*, vol. 55, n° 5, pp. 2230–2249, mai 2009.
- [DAS 07] DAS I., DENNIS J. E., « A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems », *Structural optimization*, vol. 14, n° 1, pp. 63–69, août 2007.
- [DAU 04] DAUBECHIES I., DEFRISE M., DE MOL C., « An iterative thresholding algorithm for linear inverse problems with a sparsity constraint », *Comm. Pure Appl. Math.*, vol. 57, n° 11, pp. 1413–1457, nov. 2004.

-
- [DAV 94] DAVIS G., MALLAT S., ZHANG Z., « Adaptive time-frequency decompositions with matching pursuits », *Opt. Eng.*, vol. 33, n° 7, pp. 2183-2191, juil. 1994.
- [DAV 09] DAVIES M. E., GRIBONVAL R., « Restricted isometry constants where ℓ^p sparse recovery can fail for $0 \ll p \leq 1$ », *IEEE Trans. Inf. Theory*, vol. 55, n° 5, pp. 2203-2214, mai 2009.
- [DAV 10] DAVENPORT M. A., WAKIN M. B., « Analysis of orthogonal matching pursuit using the restricted isometry property », *IEEE Trans. Inf. Theory*, vol. 56, n° 9, pp. 4395-4401, sep. 2010.
- [DAV 12] DAVIES M. E., ELДАР Y. C., « Rank awareness in joint sparse recovery », *IEEE Trans. Inf. Theory*, vol. 58, n° 2, pp. 1135-1146, fév. 2012.
- [DEM 06] DEMARET L., DYN N., ISKE A., « Image compression by linear splines over adaptive triangulations », *Signal Process.*, vol. 86, n° 7, pp. 1604-1616, juil. 2006.
- [DEN 09] DENIS L., LORENZ D. A., TREDE D., « Greedy solution of ill-posed problems : Error bounds and exact inversion », *Inverse Problems*, vol. 25, n° 11, pp. 115017-115030, nov. 2009.
- [DEV 96] DEVORE R. A., TEMLYAKOV V. N., « Some remarks on greedy algorithms », *Adv. Comput. Math.*, vol. 5, pp. 173-187, 1996.
- [DIE 95] DIERCKX P., *Curve and surface fitting with splines*, Monographs on Numerical Analysis, Oxford University Press, Inc., New York, 2ème édition, mai 1995.
- [DIN 90] DINTEN J.-M., Tomographie à partir d'un nombre limité de projections : régularisation par champs markoviens, thèse de doctorat, Université de Paris-Sud, jan. 1990.
- [DIN 13] DING J., CHEN L., GU Y., « Perturbation analysis of orthogonal matching pursuit », *IEEE Trans. Signal Process.*, vol. 61, n° 2, pp. 398-410, jan. 2013.
- [DIP 05] DIPANDA A., WOO S., « Towards a real-time 3D shape reconstruction using a structured light system », *Pattern Recogn.*, vol. 38, n° 10, pp. 1632-1650, oct. 2005.
- [DON 01] DONOHO D. L., HUO X., « Uncertainty principles and ideal atomic decomposition », *IEEE Trans. Inf. Theory*, vol. 47, n° 7, pp. 2845-2862, nov. 2001.
- [DON 03] DONOHO D. L., ELAD M., « Optimally sparse representation in general (non-orthogonal) dictionaries via ℓ^1 minimization », *Proc. Natl. Acad. Sci. USA*, vol. 100, n° 5, pp. 2197-2202, mars 2003.
- [DON 06] DONOHO D. L., ELAD M., TEMLYAKOV V. N., « Stable recovery of sparse over-complete representations in the presence of noise », *IEEE Trans. Inf. Theory*, vol. 52, n° 1, pp. 6-18, jan. 2006.
- [DON 07] DONOHO D. L., STODDEN V., TSAIG Y., About SparseLab, Rapport technique, Stanford University, mars 2007.
- [DON 08] DONOHO D. L., TSAIG Y., « Fast solution of ℓ_1 -norm minimization problems when the solution may be sparse », *IEEE Trans. Inf. Theory*, vol. 54, n° 11, pp. 4789-4812, nov. 2008.

- [DON 12] DONOHO D. L., TSAIG Y., DRORI I., STARCK J.-L., « Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit », *IEEE Trans. Inf. Theory*, vol. 58, n° 2, pp. 1094-1121, fév. 2012.
- [DOS 05a] DOSSAL C., MALLAT S., « Sparse spike deconvolution with minimum scale », *Signal Processing with Adaptive Sparse Structured Representations (SPARS workshop)*, Rennes, pp. 1-4, avr. 2005.
- [DOS 05b] DOSSAL C., Estimation de fonctions géométriques et déconvolution, thèse de doctorat, École Polytechnique, Palaiseau, déc. 2005.
- [DOS 10] DOSSAL C., PEYRÉ G., FADILI J., « A numerical exploration of compressed sampling recovery », *Linear Alg. Appl.*, vol. 432, n° 7, pp. 1663-1679, mars 2010.
- [DUA 09a] DUAN J., SOUSSEN C., BRIE D., IDIER J., « A continuation approach to estimate a solution path of mixed L2-L0 minimization problems », *Signal Processing with Adaptive Sparse Structured Representations (SPARS workshop)*, Saint-Malo, pp. 1-6, avr. 2009.
- [DUA 09b] DUAN J., SOUSSEN C., BRIE D., IDIER J., « Détection conjointe de discontinuités d'ordres différents dans un signal par minimisation de critère L2-L0 », *Actes 22^e coll. GRETSI*, Dijon, sep. 2009.
- [DUA 09c] DUAN J., SOUSSEN C., BRIE D., IDIER J., On the properties of the solution path of the constrained and penalized L2-L0 problems, Rapport technique, Centre de Recherche en Automatique de Nancy, fév. 2009.
- [DUA 10] DUAN J., Restauration et séparation de signaux polynômiaux par morceaux. Application à la microscopie de force atomique, thèse de doctorat, Université Henri Poincaré, CRAN, Nancy, nov. 2010.
- [DUB 97] DUBLANCHET F., IDIER J., DUVAUT P., « Direction-of-arrival and frequency estimation using Poisson-Gaussian modeling », *Proc. IEEE ICASSP*, Munich, Allemagne, pp. 3501-3504, avr. 1997.
- [DUP 09] DUPÉ F.-X., FADILI J. M., STARCK J.-L., « A proximal iteration for deconvolving Poisson noisy images using sparse representations », *IEEE Trans. Image Process.*, vol. 18, n° 2, pp. 310-321, fév. 2009.
- [DUP 10] DUPÉ F.-X., Restauration et reconnaissance des formes dans des images avec bruit de Poisson : application à l'analyse des neurones en microscopie de fluorescence, thèse de doctorat, Université de Caen/Basse-Normandie, Caen, jan. 2010.
- [DYM 11] DYMARSKI P., MOREAU N., RICHARD G., « Greedy sparse decompositions : a comparative study », *Eurasip Journal on Advances in Signal Processing*, vol. 2011, n° 34, pp. 1-16, août 2011.
- [EFR 60] EFROYMSON M. A., « Multiple regression analysis », RALSTON A., WILF H. S., Eds., *Mathematical Methods for Digital Computers*, vol. 1, pp. 191-203, Wiley, New York, 1960.
- [EFR 04] EFRON B., HASTIE T., JOHNSTONE I., TIBSHIRANI R., « Least angle regression », *Ann. Statist.*, vol. 32, n° 2, pp. 407-499, avr. 2004.

-
- [ELA 07] ELAD M., MILANFAR P., RUBINSTEIN R., « Analysis versus synthesis in signal priors », *Inverse Problems*, vol. 23, n° 3, pp. 947–968, juin 2007.
- [ELA 10] ELAD M., *Sparse and redundant representations : From theory to applications in signal and image processing*, Springer, New York, août 2010.
- [ENN 11] ENNEN M., MINIG V., GRANDEMANGE S., TOUCHE N., MERLIN J.-L., BESANCENOT V., BRUNNER E., DOMENJOUR L., BECUWE P., « Regulation of the high basal expression of the manganese superoxide dismutase gene in aggressive breast cancer cells », *Free Radic. Biol. Med.*, vol. 15, n° 50, pp. 1771-1779, juin 2011.
- [ENN 13] ENNEN M., KLOTZ R., TOUCHE N., PINEL S., BARBIEUX C., BESANCENOT V., BRUNNER E., THIEBAUT D., JUNG A. C., LEDRAPPIER S., DOMENJOUR L., ABECASSIS J., PLÉNAT F., GRANDEMANGE S., BECUWE P., « DDB2 : a novel regulator of NF- κ B and breast tumor invasion », *Cancer Research*, vol. 73, n° 16, pp. 5040-5052, août 2013.
- [FAN 01] FAN J., LI R., « Variable selection via nonconcave penalized likelihood and its oracle properties », *J. Acoust. Soc. Amer.*, vol. 96, n° 456, pp. 1348–1360, déc. 2001.
- [FIG 07] FIGUEIREDO M. A. T., NOWAK R. D., WRIGHT S. J., « Gradient projection for sparse reconstruction : Application to compressed sensing and other inverse problems », *IEEE J. Sel. Top. Signal Process.*, vol. 1, n° 4, pp. 586-597, déc. 2007.
- [FLE 09] FLETCHER A. K., RANGAN S., « Orthogonal matching pursuit from noisy measurements : A new analysis », BENGIO Y., SCHUURMANS D., LAFFERTY J., WILLIAMS C. K. I., CULOTTA A., Eds., *Neural Information Processing Systems*, vol. 22, MIT Press, pp. 540–548, 2009.
- [FOU 11] FOU CART S., « Hard thresholding pursuit : An algorithm for compressive sensing », *SIAM J. Num. Anal.*, vol. 49, n° 6, pp. 2543-2563, 2011.
- [FOU 13a] FOU CART S., « Stability and robustness of weak orthogonal matching pursuits », BILYK D., DE CARLI L., PETUKHOV A., STOKOLOS A. M., WICK B. D., Eds., *Recent advances in harmonic analysis and applications*, vol. 25, Springer proceedings in Mathematics & Statistics, pp. 395-405, 2013.
- [FOU 13b] FOU CART S., RAUHUT H., *A mathematical introduction to compressive sensing*, Applied and Numerical Harmonic Analysis, Birkhäuser, Basel, juin 2013.
- [FRI 91] FRIEDMAN J. H., « Multivariate Adaptive Regression Splines », *Ann. Statist.*, vol. 19, n° 1, pp. 1–67, mars 1991.
- [FRI 12] FRIEDLANDER M. P., MANSOUR H., SAAB R., YILMAZ O., « Recovering compressively sampled signals using partial support information », *IEEE Trans. Inf. Theory*, vol. 58, n° 2, pp. 1122-1134, fév. 2012.
- [FUC 04] FUCHS J.-J., « On sparse representations in arbitrary redundant bases », *IEEE Trans. Inf. Theory*, vol. 50, n° 6, pp. 1341-1344, juin 2004.
- [FUC 05] FUCHS J.-J., « Recovery of exact sparse representations in the presence of bounded noise », *IEEE Trans. Inf. Theory*, vol. 51, n° 10, pp. 3601–3608, oct. 2005.

- [GAB 07] GABORIAUD F., DUFRÊNE Y. F., « Atomic force microscopy of microbial cells : Application to nanomechanical properties, surface forces and molecular recognition forces », *Colloids Surf. B.*, vol. 54, n° 1, pp. 10–19, jan. 2007.
- [GAS 09] GASSO G., RAKOTOMAMONJY A., CANU S., « Recovering sparse signals with a certain family of nonconvex penalties and DC programming », *IEEE Trans. Signal Process.*, vol. 57, n° 12, pp. 4686–4698, déc. 2009.
- [GHO 11] GHOLAMI A., HOSSEINI S. M., « A general framework for sparsity-based denoising and inversion », *IEEE Trans. Signal Process.*, vol. 59, n° 11, pp. 5202–5211, nov. 2011.
- [GIO 05] GIOVANNELLI J.-F., COULAIS A., « Positive deconvolution for superimposed extended source and point sources », *Astronomy and Astrophysics*, vol. 439, n° 1, pp. 401–412, août 2005.
- [GIR 14] GIRYES R., NAM S., ELAD M., GRIBONVAL R., DAVIES M. E., « Greedy-like algorithms for the cosparsity analysis model », *Linear Alg. Appl.*, vol. 441, pp. 22-60, jan. 2014.
- [GOL 79] GOLUB G. H., HEATH M., WAHBA G., « Generalized cross-validation as a method for choosing a good ridge parameter », *Technometrics*, vol. 21, n° 2, pp. 215–223, mai 1979.
- [GOU 90] GOUSSARD Y., DEMOMENT G., IDIER J., « A new algorithm for iterative deconvolution of sparse spike trains », *Proc. IEEE ICASSP*, vol. 3, Albuquerque, NM, USA, pp. 1547–1550, avr. 1990.
- [GRA 11] GRASMAIR M., SCHERZER O., HALTMEIER M., « Necessary and sufficient conditions for linear convergence of ℓ_1 -regularization », *Comm. Pure Appl. Math.*, vol. 64, n° 2, pp. 161–182, fév. 2011.
- [GRI 03] GRIBONVAL R., NIELSEN M., « Sparse representations in unions of bases », *IEEE Trans. Inf. Theory*, vol. 49, n° 12, pp. 3320–3325, déc. 2003.
- [GRI 06] GRIBONVAL R., VANDERGHEYNST P., « On the exponential convergence of matching pursuits in quasi-incoherent dictionaries », *IEEE Trans. Inf. Theory*, vol. 52, n° 1, pp. 255–261, jan. 2006.
- [GRI 08a] GRIBONVAL R., NIELSEN M., « Beyond sparsity : Recovering structured representations by ℓ^1 minimization and greedy algorithms », *Adv. Comput. Math.*, vol. 28, n° 1, pp. 23–41, 2008.
- [GRI 08b] GRIBONVAL R., RAUHUT H., SCHNASS K., VANDERGHEYNST P., « Atoms of all channels, unite! Average case analysis of multi-channel sparse recovery using greedy algorithms », *J. Fourier Anal. Appl.*, vol. 14, n° 5-6, pp. 655–687, déc. 2008.
- [GUO 11] GUO Y., Représentations parcimonieuses pour la séparation des signaux : quantification en IRM spectroscopique pour l'étude du métabolisme, thèse de doctorat, Université de Reims Champagne-Ardenne, CReSTIC, Reims, avr. 2011.
- [HAL 13] HALTMEIER M., « Stable signal representation via ℓ^1 -minimization in redundant, non-tight frames », *IEEE Trans. Signal Process.*, vol. 61, n° 2, pp. 420–426, jan. 2013.
- [HAN 79] HANNAN E. J., QUINN B. G., « The determination of the order of an autoregression », *J. R. Statist. Soc. B*, vol. 41, n° 2, pp. 190–195, 1979.

- [HAU 07] HAUGLAND D., « A bidirectional greedy heuristic for the subspace selection problem », STÜTZLE T., BIRATTARI M., HOOS H. H., Eds., *Engineering stochastic local search algorithms. Designing, implementing and analyzing effective heuristics*, vol. 4638 de *Lect. Notes Comput. Sci.*, Berlin, Allemagne, Springer Verlag, pp. 162-176, sep. 2007.
- [HEN 13] HENROT S., SOUSSEN C., BRIE D., « Fast positive deconvolution of hyperspectral images », *IEEE Trans. Image Process.*, vol. 22, n° 2, pp. 828-833, fév. 2013.
- [HER 07] HERNÁNDEZ-MIER Y., Construction rapide d'images panoramiques applicables à l'exploration cystoscopique et à l'endoscopie de fluorescence en cancérologie, thèse de doctorat, Institut National Polytechnique de Lorraine, CRAN, Nancy, oct. 2007.
- [HER 10] HERZET C., DRÉMEAU A., « Bayesian pursuit algorithms », *Proc. Eur. Sig. Proc. Conf.*, Aalborg, Danemark, pp. 1474-1478, août 2010.
- [HER 13] HERZET C., SOUSSEN C., IDIER J., GRIBONVAL R., « Exact recovery conditions for sparse representations with partial support information », *IEEE Trans. Inf. Theory*, vol. 59, n° 11, pp. 7509-7524, nov. 2013.
- [HOA 05] HOAI AN L. T., TAO P. D., « The DC (Difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems », *Annals of Operations Research*, vol. 133, n° 1-4, pp. 23-46, jan. 2005.
- [HOA 08] HOAI AN L. T., MINH L. H., VAN VINH N., TAO P. D., « A DC Programming approach for Feature Selection in Support Vector Machines learning », *Journal of Advances in Data Analysis and Classification*, vol. 2, n° 3, pp. 259-278, déc. 2008.
- [HUA 11] HUANG S., ZHU J., « Recovery of sparse signals using OMP and its variants : Convergence analysis based on RIP », *Inverse Problems*, vol. 27, n° 3, page 035003, mars 2011.
- [IDI 01] IDIER J., « Convex half-quadratic criteria and interacting auxiliary variables for image restoration », *IEEE Trans. Image Process.*, vol. 10, n° 7, pp. 1001-1009, juil. 2001.
- [JAI 11] JAIN P., TEWARI A., DHILLON I. S., « Orthogonal matching pursuit with replacement », *Proceedings of the Neural Information Processing Systems Conference (NIPS)*, Granada, Espagne, pp. 1-22, déc. 2011.
- [JAN 00] JANSHOFF A., NEITZERT M., OBERDÖRFER Y., FUCHS H., « Force spectroscopy of molecular systems - Single molecule spectroscopy of polymers and biomolecules », *Angew. Chem. Int. Ed.*, vol. 39, n° 18, pp. 3213-3237, sep. 2000.
- [KAK 87] KAK A. C., SLANEY M., *Principles of computerized tomographic imaging*, IEEE Press, New York, 1987.
- [KAR 12] KARAHANOGU N. B., ERDOGAN H., On the theoretical analysis of orthogonal matching pursuit with termination based on the residue, Rapport technique n° 1-7, Department of Electronics Engineering, Sabanci University, Istanbul, oct. 2012.
- [KAS 08] KASAS S., DIETLER G., « Probing nanomechanical properties from biomolecules to living cells », *Pflügers Archiv - European Journal of Physiology, special issue entitled "Atomic force microscopy enters physiology"*, vol. 456, n° 1, pp. 13-27, avr. 2008.

- [KON 09] KONDRA S., LAISHRAM J., BAN J., MIGLIORINI E., FOGGIA V. D., LAZZARINO M., TORRE V., RUARO M. E., « Integration of confocal and atomic force microscopy images », *J. Neurosci. Methods*, vol. 177, n° 1, pp. 94-107, fév. 2009.
- [KOR 82] KORMYLO J. J., MENDEL J. M., « Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes », *IEEE Trans. Inf. Theory*, vol. 28, pp. 482-488, mai 1982.
- [KOW 09] KOWALSKI M., « Sparse regression using mixed norms », *Appl. Comp. Harmonic Anal.*, vol. 27, n° 3, pp. 303-324, sep. 2009.
- [KOW 10] KOWALSKI M., GRAMFORT A., « *A priori* par normes mixtes pour les problèmes inverses. Application à la localisation de sources en M/EEG », *Traitement du Signal*, vol. 27, n° 1, pp. 53-78, 2010.
- [KOW 13] KOWALSKI M., SIEDENBURG K., DÖRFLER M., « Social sparsity! Neighborhood systems enrich structured shrinkage operators », *IEEE Trans. Signal Process.*, vol. 61, n° 10, pp. 53-78, mai 2013.
- [KUZ 07] KUZNETSOVA T. G., STARODUBTSEVA M. N., YEGORENKOV N. I., CHIZHIK S. A., ZHDANOV R. I., « Atomic force microscopy probing of cell elasticity », *Micron.*, vol. 38, n° 8, pp. 824-833, déc. 2007.
- [LAH 12] LAHDACHI S., Modélisation et analyse quantitative des images force-volume de cellules, rapport interne, rapport de stage de deuxième année de l'École des Mines de Nancy, CRAN, Université de Lorraine, Nancy, sep. 2012.
- [LAR 05] LARRAINZAR E., O'GARA F., MORRISSEY J. P., « Applications of autofluorescent proteins for *in situ* studies in microbial ecology », *Annual Review of Microbiology*, vol. 59, pp. 257-277, oct. 2005.
- [LEK 09] LEKKA M., LAIDLER P., « Applicability of AFM in cancer detection », *Nature Nanotechnology*, vol. 4, n° 2, pp. 72-73, fév. 2009.
- [LEV 02] LEVEAU J. H. J., LINDOW S. E., « Bioreporters in microbial ecology », *Current Opinion in Microbiology*, vol. 5, n° 3, pp. 259-265, juin 2002.
- [LIN 07a] LIN D. C., DIMITRIADIS E. K., HORKAY F., « Robust strategies for automated AFM force curve analysis – I. Non-adhesive indentation of soft, inhomogeneous materials », *J. Biomech. Eng.*, vol. 129, n° 3, pp. 430-440, juin 2007.
- [LIN 07b] LIN D. C., DIMITRIADIS E. K., HORKAY F., « Robust strategies for automated AFM force curve analysis – II. Adhesion-influenced indentation of soft, elastic materials », *J. Biomech. Eng.*, vol. 129, n° 3, pp. 904-912, déc. 2007.
- [LIU 12] LIU E., TEMLYAKOV V. N., « The orthogonal super greedy algorithm and applications in compressed sensing », *IEEE Trans. Inf. Theory*, vol. 58, n° 4, pp. 2040-2047, avr. 2012.
- [LOR 09] LORENZ D. A., TREDE D., « Greedy deconvolution of point-like objects », *Signal Processing with Adaptive Sparse Structured Representations (SPARS workshop)*, Saint-Malo, pp. 1-6, avr. 2009.
- [LOR 11] LORENZ D. A., SCHIFFLER S., TREDE D., « Beyond convergence rates : exact recovery with the Tikhonov regularization with sparsity constraints », *Inverse Problems*, vol. 27, n° 8, page 085009, nov. 2011.

-
- [MAI 09] MAILHÉ B., GRIBONVAL R., BIMBOT F., VANDERGHEYNST P., « A low complexity Orthogonal Matching Pursuit for sparse signal approximation with shift-invariant dictionaries », *Proc. IEEE ICASSP*, Taipei, Taiwan, pp. 3445-3448, avr. 2009.
- [MAL 93] MALLAT S., ZHANG Z., « Matching pursuits with time-frequency dictionaries », *IEEE Trans. Signal Process.*, vol. 41, n° 12, pp. 3397-3415, déc. 1993.
- [MAL 05] MALIOUTOV D. M., CETIN M., WILLSKY A. S., « Homotopy continuation for sparse signal representation », *Proc. IEEE ICASSP*, vol. V, Philadelphia, PA, USA, pp. 733-736, mars 2005.
- [MAL 10] MALEKI A., DONOHO D. L., « Optimally tuned iterative reconstruction algorithms for compressed sensing », *IEEE J. Sel. Top. Signal Process.*, vol. 4, n° 2, pp. 330-341, avr. 2010.
- [MAR 02] MARZANI F. S., VOISIN Y., LEW YAN VOON L. C., DIOU A., « Calibration of a three-dimensional reconstruction system using a structured light source », *Optical Engineering*, vol. 41, n° 2, pp. 484-492, fév. 2002.
- [MAR 04] MARLER R. T., ARORA J. S., « Survey of multi-objective optimization methods for engineering », *Structural and Multidisciplinary Optimization*, vol. 26, n° 6, pp. 369-395, avr. 2004.
- [MAZ 11] MAZET V., « Joint Bayesian decomposition of a spectroscopic signal sequence », *IEEE Signal Process. Lett.*, vol. 18, n° 3, pp. 191-184, mars 2011.
- [MAZ 13] MAZET V., SOUSSEN C., DJERMOUNE E.-H., « Décomposition de spectres en motifs paramétriques par approximation parcimonieuse », *Actes 24^e coll. GRETSI*, Brest, sep. 2013.
- [MCE 12] MC EWEN G. D., Raman microspectroscopy, atomic force microscopy, and electric cell-substrate impedance sensing for characterization of bio-interfaces : Molecular, bacteria, and mammalian cells, thèse de doctorat, Utah State University, 2012.
- [MEN 83] MENDEL J. M., *Optimal seismic deconvolution*, Academic Press, New York, USA, 1983.
- [MIL 02] MILLER A. J., *Subset selection in regression*, Chapman and Hall, Londres, UK, 2^{ème} édition, avr. 2002.
- [MIR 05] MIRANDA-LUNA R., Contribution au mosaïquage d'images endoscopiques d'organes creux pour l'aide au diagnostic du cancer en phase précoce, thèse de doctorat, Institut National Polytechnique de Lorraine, Nancy, mars 2005.
- [MO 12] MO Q., SHEN Y., « A remark on the restricted isometry property in orthogonal matching pursuit », *IEEE Trans. Inf. Theory*, vol. 58, n° 6, pp. 3654-3656, juin 2012.
- [MOG 08] MOGHADDAM M., GRUBER A., WEISS Y., AVIDAN S., « Sparse regression as a sparse eigenvalue problem », *Information Theory and Applications Workshop*, pp. 219-225, jan. 2008.
- [MOH 09] MOHIMANI G. H., BABAIE-ZADEH M., JUTTEN C., « A fast approach for over-complete sparse decomposition based on smoothed ℓ^0 norm », *IEEE Trans. Signal Process.*, vol. 57, n° 1, pp. 289-301, jan. 2009.

- [MOU 06] MOUSSAOUI S., BRIE D., MOHAMMAD-DJAFARI A., CARTERET C., « Separation of non-negative mixture of non-negative sources using a Bayesian approach and MCMC sampling », *IEEE Trans. Signal Process.*, vol. 54, n° 11, pp. 4133–4145, nov. 2006.
- [MOU 08] MOUSSAOUI S., HAUKSDÓTTIR H., SCHMIDT F., JUTTEN C., CHANUSSOT J., BRIE D., DOUTÉ S., BENEDIKTSSON J. A., « On the decomposition of Mars hyperspectral data by ICA and Bayesian positive source separation », *Neurocomputing*, vol. 71, n° 10–12, pp. 2194–2208, juin 2008.
- [MOU 10] MOURAD N., REILLY J. P., « Minimizing nonconvex functions for sparse vector reconstruction », *IEEE Trans. Signal Process.*, vol. 58, n° 7, pp. 3485–3496, juil. 2010.
- [NAM 13] NAM S., DAVIES M. E., ELAD M., GRIBONVAL R., « The cospase analysis model and algorithms », *Appl. Comp. Harmonic Anal.*, vol. 34, n° 1, pp. 30–56, jan. 2013.
- [NAT 95] NATARAJAN B. K., « Sparse approximate solutions to linear systems », *SIAM J. Comput.*, vol. 24, n° 2, pp. 227–234, avr. 1995.
- [NEE 09] NEEDELL D., TROPP J. A., « CoSaMP : Iterative signal recovery from incomplete and inaccurate samples », *Appl. Comp. Harmonic Anal.*, vol. 26, n° 3, pp. 301–321, mai 2009.
- [NIK 10] NIKOLOVA M., NG M. K., TAM C.-P., « Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction », *IEEE Trans. Image Process.*, vol. 19, n° 12, pp. 3073–3088, déc. 2010.
- [NIK 13] NIKOLOVA M., « Description of the minimizers of least squares regularized with ℓ_0 norm. Uniqueness of the global minimizer », *SIAM J. Imaging Sci.*, vol. 6, n° 2, pp. 904–937, mai 2013.
- [NOC 06] NOCEDAL J., WRIGHT S. J., *Numerical optimization*, Springer texts in Operations Research and Financial Engineering, Springer Verlag, New York, USA, 2ème édition, juil. 2006.
- [PAN 09] PANKAJAKSHAN P., Déconvolution aveugle en imagerie de microscopie confocale à balayage laser, thèse de doctorat, Université de Nice Sophia-Antipolis, Nice, déc. 2009.
- [PAT 93] PATI Y. C., REZAIIFAR R., KRISHNAPRASAD P. S., « Orthogonal matching pursuit : Recursive function approximation with applications to wavelet decomposition », *Proc. 27th Asilomar Conf. on Signals, Systems and Computers*, vol. 1, pp. 40–44, nov. 1993.
- [PEL 13] PELEG T., ELAD M., « Performance guarantees of the thresholding algorithm for the cospase analysis model », *IEEE Trans. Inf. Theory*, vol. 59, n° 3, pp. 1832–1845, mars 2013.
- [PEY 11] PEYRÉ G., « A review of adaptive image representations », *IEEE J. Sel. Top. Signal Process.*, vol. 5, n° 5, pp. 896–911, sep. 2011.
- [PLO 12] PLODINEC M., LOPARIC M., MONNIER C. A., OBERMANN E. C., ZANETTI-DALLENBACH R., OERTLE P., HYOTYLA J. T., AEBI U., LIM R. Y. H., SCHOENENBERGER C. A., « The nanomechanical signature of breast cancer », *Nature nanotechnology*, vol. 7, pp. 757–765, 2012.
- [PLU 07] PLUMBLEY M. D., « On polar polytopes and the recovery of sparse representations », *IEEE Trans. Inf. Theory*, vol. 53, n° 9, pp. 3188–3195, sep. 2007.

-
- [POL 11] POLYAKOV P., SOUSSEN C., DUAN J., DUVAL J. F., BRIE D., FRANCIUS G., « Automated force volume image processing for biological samples », *PLoS ONE*, vol. 6, n° 4, page e18887, avr. 2011.
- [PUS 10] PUSTELNIK N., Méthodes proximales pour la résolution de problèmes inverses. Application à la Tomographie par Emission de Positrons, thèse de doctorat, Université de Paris-Est, Marne-la-Vallée, déc. 2010.
- [RAM 12] RAMÍREZ I., SAPIRO G., « Universal regularizers for robust sparse coding and modeling », *IEEE Trans. Image Process.*, vol. 21, n° 9, pp. 3850–3864, sep. 2012.
- [RAO 03] RAO B. D., ENGAN K., COTTER S. F., PALMER J., KREUTZ-DELGADO K., « Subset selection in noise based on diversity measure minimization », *IEEE Trans. Signal Process.*, vol. 51, n° 3, pp. 760–770, mars 2003.
- [RAU 08] RAUHUT H., SCHNASS K., VANDERGHEYNST P., « Compressed sensing and redundant dictionaries », *IEEE Trans. Inf. Theory*, vol. 54, n° 5, pp. 2210–2219, mai 2008.
- [REB 02] REBOLLO-NEIRA L., LOWE D., « Optimized orthogonal matching pursuit approach », *IEEE Signal Process. Lett.*, vol. 9, n° 4, pp. 137–140, avr. 2002.
- [RIS 83] RISSANEN J., « A universal prior for integers and estimation by minimum description length », *Ann. Statist.*, vol. 11, n° 2, pp. 416–431, 1983.
- [ROS 01] VAN ROSSUM M. C., « A novel spike distance », *Neural Computation*, vol. 13, n° 4, pp. 751–763, avr. 2001.
- [SAH 12a] SAHNOUN S., Développement de méthodes d'estimation modale de signaux multidimensionnels. Application à la spectroscopie RMN, thèse de doctorat, Université de Lorraine, CRAN, Nancy, nov. 2012.
- [SAH 12b] SAHNOUN S., DJERMOUNE E.-H., SOUSSEN C., BRIE D., « Sparse multidimensional modal analysis using a multigrad dictionary refinement », *Eurasip Journal on Advances in Signal Processing*, vol. 2012, n° 60, pp. 1–10, mars 2012.
- [SAR 06] SARDER P., NEHORAI A., « Deconvolution methods for 3-D fluorescence microscopy images », *IEEE Trans. Signal Process. Mag.*, vol. 23, n° 3, pp. 32–45, mai 2006.
- [SCH 07] SCHNASS K., VANDERGHEYNST P., « Average performance analysis for thresholding », *IEEE Signal Process. Lett.*, vol. 14, n° 11, pp. 828–831, nov. 2007.
- [SEL 13] SELESNICK I., BAYRAM I., Sparse signal estimation by maximally sparse convex optimization, Rapport technique, Polytechnic Institute of New York University, fév. 2013.
- [SEN 05] SEN S., SUBRAMANIAN S., DISCHER D. E., « Indentation and adhesive probing of a cell membrane with AFM : Theoretical model and experiments », *Biophysical Journal*, vol. 89, n° 5, pp. 3203–3213, nov. 2005.
- [SEN 10] SEN S., KUMAR S., « Combining mechanical and optical approaches to dissect cellular mechanobiology », *Journal of Biomechanics*, vol. 43, n° 1, pp. 45–54, jan. 2010.
- [SHU 05] SHUKLA R., DRAGOTTI P. L., DO M., VETTERLI M., « Rate distortion optimized tree structured compression algorithms for piecewise smooth images », *IEEE Trans. Image Process.*, vol. 14, n° 3, pp. 343–359, mars 2005.

- [SIR 08] SIRGHI L., PONTI J., BROGGI F., ROSSI F., « Probing elasticity and adhesion of live cells by atomic force microscopy indentation », *European Biophysics Journal*, vol. 37, n° 6, pp. 935-945, mars 2008.
- [SMI 96] SMITH M. S., KOHN R., « Nonparametric regression using Bayesian variable selection », *J. Econometrics*, vol. 75, n° 2, pp. 317-343, déc. 1996.
- [SOU 07] SOUSSEN C., BRIE D., GABORIAUD F., KESSLER C., « Atomic force microscopy imaging and related issues in signal processing : a preliminary work », *Institut franco-allemand pour les applications de la recherche*, Grenoble, IAR Workshop, pp. 1–12, nov. 2007.
- [SOU 08a] SOUSSEN C., BRIE D., GABORIAUD F., KESSLER C., « Modeling of force-volume images in atomic force microscopy », *Proc. IEEE ISBI*, Paris, pp. 1605-1608, mai 2008.
- [SOU 08b] SOUSSEN C., IDIER J., « Reconstruction of three-dimensional localized objects from limited angle X-ray projections : an approach based on sparsity and multigrid image representation », *J. Electr. Imag.*, vol. 17, n° 3, page 033011, août 2008.
- [SOU 11] SOUSSEN C., IDIER J., BRIE D., DUAN J., « From Bernoulli-Gaussian deconvolution to sparse signal restoration », *IEEE Trans. Signal Process.*, vol. 59, n° 10, pp. 4572-4584, oct. 2011.
- [SOU 12] SOUSSEN C., IDIER J., CARCREFF E., SIMON L., POTEL C., « Ultrasonic non destructive testing based on sparse deconvolution », *Journal of Physics : Conference Series*, vol. 353, n° 012018, pp. 1-10, mars 2012.
- [SOU 13a] SOUSSEN C., BRIE D., FRANCIUS G., IDIER J., « Approximation parcimonieuse pour l'analyse d'images force-volume en microscopie de force atomique », IDIER J., GIOVANNELLI J.-F., Eds., *Méthodes d'inversion appliquées au traitement du signal et de l'image*, Paris, Hermès, pp. 53–77, nov. 2013.
- [SOU 13b] SOUSSEN C., GRIBONVAL R., IDIER J., HERZET C., « Joint k -step analysis of orthogonal matching pursuit and orthogonal least squares », *IEEE Trans. Inf. Theory*, vol. 59, n° 5, pp. 3158-3174, mai 2013.
- [SOU 13c] SOUSSEN C., HERZET C., IDIER J., GRIBONVAL R., « Unrecoverable subsets by OMP and Basis Pursuit », *Signal Processing with Adaptive Sparse Structured Representations (SPARS workshop)*, Lausanne, Suisse, pp. 1-1, juil. 2013.
- [SOU 13d] SOUSSEN C., IDIER J., DUAN J., BRIE D., ℓ_2 - ℓ_0 regularization path tracking algorithms, Rapport technique, CRAN – IRCCyN, sep. 2013.
- [STO 12] STOICA P., BABU P., « Sparse estimation of spectral lines : Grid selection problems and their solutions », *IEEE Trans. Signal Process.*, vol. 60, n° 2, pp. 962–967, fév. 2012.
- [TAO 97] TAO P. D., HOAI AN L. T., « Convex analysis approach to D. C. programming : Theory, Algorithm and Applications », *Acta Mathematica Vietnamica*, vol. 22, n° 1, pp. 289-355, 1997.
- [TEM 00] TEMLYAKOV V. N., « Weak greedy algorithms », *Adv. Comput. Math.*, vol. 12, n° 2-3, pp. 213-227, fév. 2000.

-
- [THI 10] THIÉBAUT E., GIOVANNELLI J.-F., « Image reconstruction in optical interferometry », *IEEE Trans. Signal Process. Mag.*, vol. 27, n° 1, pp. 97-109, jan. 2010.
- [TRO 04] TROPP J. A., « Greed is good : Algorithmic results for sparse approximation », *IEEE Trans. Inf. Theory*, vol. 50, n° 10, pp. 2231-2242, oct. 2004.
- [TRO 06] TROPP J. A., « Just relax : Convex programming methods for identifying sparse signals in noise », *IEEE Trans. Inf. Theory*, vol. 52, n° 3, pp. 1030-1051, mars 2006.
- [TRO 07] TROPP J. A., GILBERT A. C., « Signal recovery from random measurements via orthogonal matching pursuit », *IEEE Trans. Inf. Theory*, vol. 53, n° 12, pp. 4655-4666, déc. 2007.
- [TRO 10] TROPP J. A., WRIGHT S. J., « Computational methods for sparse solution of linear inverse problems », *Proc. IEEE, invited paper (Special Issue "Applications of sparse representation and compressive sensing")*, vol. 98, n° 5, pp. 948-958, juin 2010.
- [TRZ 09] TRZASKO J., MANDUCA A., « Highly undersampled magnetic resonance image reconstruction via homotopic ℓ_0 -minimization », *IEEE Trans. Medical Imaging*, vol. 8, n° 1, pp. 106-121, jan. 2009.
- [TUR 11] TUREK J. S., YAVNEH I., ELAD M., « On MMSE and MAP denoising under sparse representation modeling over a unitary dictionary », *IEEE Trans. Signal Process.*, vol. 59, n° 8, pp. 3526-3535, août 2011.
- [VAI 13] VAITER S., PEYRÉ G., DOSSAL C., FADILI J., « Robust sparse analysis regularization », *IEEE Trans. Inf. Theory*, vol. 59, n° 4, pp. 2001-2016, avr. 2013.
- [VAS 10] VASWANI N., LU W., « Modified-CS : Modifying compressive sensing for problems with partially known support », *IEEE Trans. Signal Process.*, vol. 58, n° 9, pp. 4595-4607, sep. 2010.
- [VEL 02] VELEGOL S. B., LOGAN B. E., « Contributions of bacterial surface polymers, electrostatics, and cell elasticity to the shape of AFM force curves », *Langmuir*, vol. 18, n° 13, pp. 5256-5262, mai 2002.
- [VET 02] VETTERLI M., MARZILIANO P., BLU T., « Sampling signals with finite rate of innovation », *IEEE Trans. Signal Process.*, vol. 50, n° 6, pp. 1417-1428, juin 2002.
- [WAH 77] WAHBA G., « Practical approximate solutions to linear operator equations when the data are noisy », *SIAM J. Num. Anal.*, vol. 14, n° 4, pp. 651-667, 1977.
- [WAI 09] WAINWRIGHT M. J., « Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1 -constrained quadratic programming (Lasso) », *IEEE Trans. Inf. Theory*, vol. 55, n° 5, pp. 2183-2202, mai 2009.
- [WAN 04] WANG Y., « Model selection », GENTLE J. E., HÄRDLE W., MORI Y., Eds., *Handbook of Computational Statistics*, vol. 1, Berlin, Springer-Verlag, pp. 437-466, août 2004.
- [WAN 09] WANG J., WAN Z., LIU W., LI L., REN L., WANG X., SUN P., REN L., ZHAO H., TU Q., ZHANG Z., SONG N., ZHANG L., « Atomic force microscope study of tumor cell membranes following treatment with anti-cancer drugs », *Biosensors & bioelectronics*, vol. 25, n° 4, pp. 721-727, sep. 2009.

- [WAN 12a] WANG J., KWON S., SHIM B., « Generalized orthogonal matching pursuit », *IEEE Trans. Signal Process.*, vol. 60, n° 12, pp. 6202–6216, déc. 2012.
- [WAN 12b] WANG J., SHIM B., « On the recovery limit of sparse signals using orthogonal matching pursuit », *IEEE Trans. Signal Process.*, vol. 60, n° 9, pp. 4973–4976, sep. 2012.
- [WAS 73] WASSERSTROM E., « Numerical solutions by the continuation method », *SIAM Rev.*, vol. 15, n° 1, pp. 89–119, jan. 1973.
- [WAT 12] WATANABE T., KURAMOCHI H., TAKAHASHI A., IMAI K., KATSUTA N., NAKAYAMA T., FUJIKI H., SUGANUMA M., « Higher cell stiffness indicating lower metastatic potential in B16 melanoma cell variants and in (–)-epigallocatechin gallate-treated cells », *Journal of Cancer Research and Clinical Oncology*, vol. 138, n° 5, pp. 859–866, fév. 2012.
- [WIE 94] WIESENDANGER R., *Scanning probe microscopy and spectroscopy : Methods and applications*, Cambridge Univ. Press, Cambridge, MA, USA, sep. 1994.
- [WIP 10] WIPF D. P., NAGARAJAN S., « Iterative reweighted ℓ_1 and ℓ_2 methods for finding sparse solutions », *IEEE J. Sel. Top. Signal Process. (Special issue on Compressive Sensing)*, vol. 4, n° 2, pp. 317–329, avr. 2010.
- [WOH 03] WOHLBERG B., « Noise sensitivity of sparse signal representations : Reconstruction error bounds for the inverse problem », *IEEE Trans. Signal Process.*, vol. 51, n° 12, pp. 3053–3060, déc. 2003.
- [WU 10] WU Y., MCEWEN G. D., HARIHAR S., BAKER S. M., DEWALD D. B., ZHOU A., « BRMS1 expression alters the ultrastructural, biomechanical and biochemical properties of MDA-MB-435 human breast carcinoma cells : An AFM and Raman microspectroscopy study », *Cancer Letters*, vol. 293, n° 1, pp. 82–91, juil. 2010.
- [XU 11] XU W., HASSIBI B., « Precise stability phase transitions for ℓ_1 minimization : A unified geometric framework », *IEEE Trans. Inf. Theory*, vol. 57, n° 10, pp. 6894–6919, oct. 2011.
- [XU 12] XU W., MEZENCEV R., KIM B., WANG L., MC DONALD J., SULCHEK T., « Cell stiffness is a biomarker of the metastatic potential of ovarian cancer cells », *PLoS ONE*, vol. 7, n° 10, page e46609, oct. 2012.
- [YAO 02] YAO X., WALTER J., BURKE S., STEWART S., JERICHO M. H., PINK D., HUNTER R., BEVERIDGE T. J., « Atomic force microscopy and theoretical considerations of surface properties and turgor pressures of bacteria », *Colloids and Surfaces B : Biointerfaces*, vol. 23, n° 2–3, pp. 213–230, fév. 2002.
- [ZHA 11a] ZHANG T., « Adaptive forward-backward greedy algorithm for learning sparse representations », *IEEE Trans. Inf. Theory*, vol. 57, n° 7, pp. 4689–4708, juil. 2011.
- [ZHA 11b] ZHANG T., « Sparse recovery with orthogonal matching pursuit under RIP », *IEEE Trans. Inf. Theory*, vol. 57, n° 9, pp. 6215–6221, sep. 2011.
- [ZIB 10] ZIBULEVSKY M., ELAD M., « $\ell_1 - \ell_2$ optimization in signal and image processing », *IEEE Trans. Signal Process. Mag.*, vol. 27, n° 3, pp. 76–88, mai 2010.
- [ZOU 06] ZOU H., « The adaptive Lasso and its oracle properties », *J. Acoust. Soc. Amer.*, vol. 101, n° 476, pp. 1418–1429, déc. 2006.

Résumé

Ce manuscrit synthétise mon activité de recherche au CRAN entre 2005 et 2013. Les projets menés s’inscrivent dans les domaines des problèmes inverses en traitement du signal et des images, de l’approximation parcimonieuse, de l’analyse d’images hyperspectrales et de la reconstruction d’images 3D. Je détaille plus particulièrement les travaux concernant la conception, l’analyse et l’utilisation d’algorithmes d’approximation parcimonieuse pour des problèmes inverses caractérisés par un dictionnaire mal conditionné. Dans un premier chapitre, je présente les algorithmes heuristiques conçus pour minimiser des critères mixtes ℓ_2 - ℓ_0 . Ce sont des algorithmes gloutons « bidirectionnels » définis en tant qu’extension de l’algorithme *Orthogonal Least Squares* (OLS). Leur développement est motivé par le bon comportement empirique d’OLS et de ses versions dérivées lorsque le dictionnaire est une matrice mal conditionnée. Le deuxième chapitre est une partie applicative en microscopie de force atomique, où les algorithmes du premier chapitre sont utilisés avec un dictionnaire particulier dans le but de segmenter automatiquement des signaux. Cette segmentation permet finalement de fournir une cartographie 2D de différents paramètres électrostatiques et bio-mécaniques. Le troisième chapitre est une partie théorique visant à analyser les algorithmes gloutons OMP (*Orthogonal Matching Pursuit*) et OLS. Une première analyse de reconstruction exacte par OLS en k itérations est proposée. De plus, une comparaison poussée des conditions de reconstruction exacte lorsqu’un certain nombre d’itérations ont déjà été effectuées fournit un éclairage sur le meilleur comportement d’OLS (par rapport à OMP) pour les problèmes mal conditionnés. Dans un quatrième chapitre, je dresse quelques perspectives méthodologiques et appliquées dans le domaine de l’analyse parcimonieuse en lien avec les chapitres précédents.

Mots-clés : approximation parcimonieuse, algorithmes gloutons, *Orthogonal Least Squares*, problèmes inverses, microscopie de force atomique, reconstruction exacte en k itérations.

Abstract

This manuscript is a synthesis of my research activity at CRAN between 2005 and 2013. My research projects deal with inverse problems in signal and image processing, sparse approximation, hyperspectral image analysis, and 3D image reconstruction. I pay specific attention to the development, analysis and utilization of sparse approximation algorithms for inverse problems characterized by ill-conditioned dictionaries. In the first chapter, heuristic algorithms are proposed to minimize mixed ℓ_2 - ℓ_0 cost functions. These are “bidirectional” greedy algorithms defined as extensions of *Orthogonal Least Squares* (OLS). Indeed, empirical comparisons show that OLS and its derived versions behave nicely when the dictionary is an ill-conditioned matrix. The second chapter is an applicative part in atomic force microscopy, where the OLS based algorithms are utilized with a specific dictionary in order to perform automatic segmentation of signals. This segmentation leads to the reconstruction of a set of 2D images representing electrostatic and bio-mechanical properties at the nanoscale. The third chapter is a theoretical study aiming to analyze the greedy algorithms OMP (*Orthogonal Matching Pursuit*) and OLS. A first k -step recovery analysis or OLS is provided. Then, the worst case exact recovery conditions are being thoroughly evaluated for both OMP and OLS when a number of iterations have already been performed. The comparisons validate the better behavior of OLS for problems involving ill-conditioned dictionaries. The fourth chapter sketches a few perspectives, both methodological and applicative, regarding sparse analysis for inverse problems.

Keywords: sparse approximation, greedy algorithms, *Orthogonal Least Squares*, inverse problems, atomic force microscopy, k -step exact recovery conditions.

