



HAL
open science

Contributions en mathématiques de l'imagerie

Pierre Weiss

► **To cite this version:**

Pierre Weiss. Contributions en mathématiques de l'imagerie. Analyse numérique [math.NA]. Université toulouse 3 Paul Sabatier, 2019. tel-02309309

HAL Id: tel-02309309

<https://hal.science/tel-02309309>

Submitted on 9 Oct 2019

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.



MÉMOIRE

En vue de l'obtention de

HDR DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *25/09/2019* par :

PIERRE WEISS

Contributions en mathématiques de l'imagerie.

JURY

ALEXANDRE D'ASPREMONT
JÉRÔME BOLTE
ALBERT COHEN
JULIE DELON
JALAL FADILI
JÉRÔME IDIER
JONAS KAHN

DR CNRS Paris
Pr. Université de Toulouse
Pr. Sorbonne
Pr. Paris Descartes
Pr. Université de Caen
DR CNRS Nantes
CR CNRS Toulouse

Rapporteur
Tuteur
Rapporteur
Examinatrice
Examineur
Rapporteur
Examineur

École doctorale et spécialité :

MITT : Domaine Mathématiques : Mathématiques appliquées

Unité de Recherche :

Institut de Mathématiques de Toulouse (UMR 5219)

Institut des Technologies Avancées en sciences du Vivant (USR 3505)

Tuteur d'habilitation :

Jérôme Bolte

Rapporteurs :

Albert Cohen, Alexandre d'Aspremont et Jérôme Idier

Table des matières

1	Introduction générale	11
1.1	Guide de lecture	11
1.2	Pourquoi une habilitation?	11
1.3	Mon profil scientifique	12
1.4	Le paradis du monde académique	15
1.4.1	Soutiens financiers	16
1.4.2	Locaux et mobilier	16
1.4.3	Congrès, communautés et réseaux d'influence	17
1.4.4	Recherche libre d'accès et industrie	18
1.4.5	👁 Deep-learning	18
1.5	Mon CV détaillé	18
2	Optimisation convexe	27
2.1	Introduction	27
2.2	Dualité et forte convexité	27
2.2.1	Une méthode de descente accélérée	28
2.2.2	Quelques exemples	31
2.3	Programmation dynamique et régressions isotones	33
2.4	Résolution de problèmes en dimension infinie	35
2.4.1	Théorèmes de représentation	35
2.4.2	👁 Nouvelles méthodes de discrétisation	39
3	Échantillonnage compressé	41
3.1	Introduction	41
3.1.1	Un premier théorème	41
3.1.2	Ses limites	42
3.2	Contributions théoriques	43
3.2.1	👁 L'échantillonnage de l'homme saoul	44
3.2.2	L'échantillonnage du voyageur de commerce	45
3.2.3	L'échantillonnage par blocs	46
3.3	Génération de schémas d'échantillonnage	49
3.3.1	Optimisation de probabilité de tirage	50
3.3.2	Optimisation de mesures	50
3.4	👁 Ouverture	53

3.4.1	Reconstruction dans des trames	53
3.4.2	Échantillonnage déterministe	54
3.4.3	Compréhension de la convergence	54
4	Approximation et apprentissage d'opérateurs	57
4.1	Introduction	57
4.2	Approximation d'opérateurs de flou	57
4.2.1	Modélisation	57
4.2.2	Décomposition dans des bases d'ondelettes	60
4.2.3	Décomposition en convolution-produit	70
4.3	Identification d'opérateurs	73
4.3.1	Hypothèses	73
4.3.2	Un estimateur	73
4.3.3	Garanties théoriques	75
4.3.4	Résultats	75
4.4	👁 Ouverture	76
4.4.1	Décomposition rapide d'opérateurs	76
4.4.2	Apprentissage d'espaces d'opérateurs	76
4.4.3	Défloutage aveugle	79
4.4.4	Noyaux portés par des mesures	79
4.4.5	Super-résolution	79
5	Imagerie computationnelle	81
5.1	Imagerie par Résonance Magnétique	81
5.1.1	Introduction	81
5.1.2	Conception de bobines	81
5.1.3	Conception de pulses radio-fréquence	84
5.1.4	Projection sur des espaces de courbes	85
5.1.5	👁 Échantillonnage compressé en IRM	86
5.1.6	Correction des effets de filtrage ADC	90
5.2	Microscopie computationnelle	92
5.2.1	Introduction	92
5.2.2	Débruitage structuré	93
5.2.3	Évaluation et correction multi-vue de l'atténuation	98
6	Autres travaux	101
6.1	Lignes de niveau des images	101
6.1.1	Invariance aux changements d'illumination	101
6.1.2	Mesures de qualité invariantes	103
6.1.3	👁 Ouverture	106
6.2	Segmentation d'objets ellipsoïdaux	106
6.2.1	Segmentation supervisée	107
6.2.2	Processus ponctuels marqués	108
6.2.3	Tenseur de structure	108
6.2.4	Applications en biologie	111
6.2.5	👁 Ouverture	112

7 Synthèse des pistes de recherche	115
8 Références	117
8.1 Mes articles	117
8.1.1 Preprints	117
8.1.2 Chapitres de livre	117
8.1.3 Journal articles (international)	117
8.1.4 Journal articles (national)	121
8.1.5 Conference articles (international)	121
8.1.6 Conference articles (national)	124
8.2 Articles additionnels	125

Remerciements

Mon jury

L'habilitation à diriger les recherches - sous ses airs d'impartialité - permet en réalité au candidat de choisir ses évaluateurs. Je tiens ici à justifier ce choix en espérant - sinon ajouter un éclat supplémentaire à leur réputation - au moins leur donner le sourire, car relire ce travail est chronophage et peu gratifiant. Mon jury est hexagonal pour deux raisons : la réduction des émissions de CO₂ et notre pays où se trouve la fine fleur des chercheurs de mon domaine. Les six membres ont volontiers accepté ce travail significatif, ce qui a réduit le caractère anxiogène de l'exercice.

Je commence par remercier Jérôme Bolte, mon tuteur, qui par sa bonhomie, sa pugnacité et sa recherche de la vérité, réussit depuis plusieurs années à embarquer des chercheurs talentueux à ses côtés et à éclaircir significativement la géométrie d'objets bien plus complexes qu'ils n'y paraissent au premier abord.

Vient le tour des rapporteurs. Alexandre d'Aspremont est un chercheur avec qui je n'ai pas encore eu l'occasion d'échanger. J'ai par contre découvert de nouveaux concepts dans plusieurs des travaux que j'ai pu lire de lui (e.g. conditionnement de Renegar, distance de Banach-Mazur,...), grâce à l'élégance de son écriture et de ses preuves. Il est en réalité assez rare d'être touché par des travaux et je suis donc heureux de le compter dans mon jury. Albert Cohen est une légende vivante en analyse harmonique, en théorie de l'approximation, en traitement du signal et en équations aux dérivées partielles. Il ne fait nulle doute que son action a eu un impact considérable sur la trajectoire de nombreux chercheurs et que je me suis aussi fait emporter malgré moi par ce flot. Jérôme Idier était déjà présent lors de ma soutenance de doctorat et assure donc une certaine constance dans ma vie de chercheur. J'apprécie beaucoup sa capacité à travailler sur des problèmes résistants (e.g. ses travaux récents en illumination structurée pour la microscopie), les résultats qui émergent de cette attitude, ainsi que son indépendance aux grandes modes scientifiques.

Les nombreux échanges que j'ai pu avoir avec mes deux examinateurs Julie Delon et Jalal Fadili rendraient leur présence dans ce jury inacceptable, s'ils ne jouaient pas un rôle important dans l'évaluation au niveau national de la communauté mathématiques-imagerie. Leur implication et intégrité à tous les deux est remarquable, et je devrais probablement m'en servir un peu plus comme exemple (pour l'implication du moins) à l'avenir. Julie sait transmettre

de l'énergie et des idées partout où elle va grâce à sa vision débridée du monde. Elle me semble manier la recherche, l'enseignement et l'évaluation de ses pairs avec une efficacité hors du commun. Jalal Fadili a été un exemple pour moi depuis mon doctorat : j'ai effectivement commencé par travailler sur des thèmes qui lui étaient chers et il a fini par résoudre des problèmes sur lesquels je bloquais, comme la convergence linéaire d'algorithmes d'optimisation au bout d'un temps fini. Sa production scientifique est remarquable et il a su se renouveler systématiquement depuis maintenant 15 ans. Il fait donc partie pour moi de ces figures fraternelles qui nous maintiennent dans le droit chemin. Jonas Kahn quant à lui est un ami et collaborateur de choc dont je vais reparler dans quelques lignes. Sa présence dans ce jury est liée à ma grande estime de lui, mais aussi à des règles administratives de dernière minute.

Mes collègues

La capacité d'un enseignant-chercheur à se réaliser me semble être un phénomène fortement aléatoire et dépendant notamment des rencontres et échanges qu'il peut avoir. Retranscrire mon trajet chaotique de façon attrayante est difficile, mais je tiens tout de même à remercier ici quelques personnes qui m'ont semblé jouer un rôle clé dans mon parcours depuis mon doctorat. Je m'excuse pour les oublis.

Trois collègues et amis se distinguent : Jérôme Fehrenbach qui m'a pris par la main quand j'ai été recruté comme maître de conférence, avec qui nous avons pu créer l'équipe PRIMO, et qui a toujours donné beaucoup d'énergie pour résoudre les problèmes concrets qui nous ont été proposés. Jonas Kahn que j'ai rencontré fortuitement lors d'une semaine MIFOBIO en microscopie, qui s'est immédiatement emparé des problèmes sur lesquels je butais et qui les a souvent hachés tel un légionnaire en me faisant découvrir de nouveaux concepts mathématiques et en enrichissant mon langage. J'espère que nous continuerons à nous fréquenter longtemps à bicyclette ou devant des tableaux noirs. Frédéric de Gournay, qui réussit à se faire complètement envahir par certains problèmes qu'il rencontre et à proposer des solutions d'outre-tombe. S'il réduisait les stupéfiants et rédigeait son HDR, il ne fait pas de doute qu'il deviendrait un des scientifiques les plus en vogue de Toulouse.

Puis, par ordre chronologique depuis ma thèse viennent Laure Blanc-Féraud (et son soutien indéfectible), Mila Nikolova (qui par son exemple m'a donné envie de découvrir Hong-Kong), Xiaoming Yuan (et l'optimisation), Raymond Chan (et les randonnées et échanges spiritueux), Tiejong Zeng (et son accueil), Michael Ng (et la liberté qu'il m'a laissée), Michael Poole (les bobines de gradient et le consulting), Mohamed Masmoudi (et mon recrutement), Gabriel Peyré (par l'exemple qu'il donne à la communauté et toutes les idées qu'ils disperse librement à tout vent), Aude Rondepierre (et nos échanges en tous genres), Alain Bonnafé (et la philosophie d'une vie), Bernard Ducommun (et son travail de remodelage de l'ITAV), Valérie Lobjois (son sérieux et son excitation pour tout nouveau problème ou méthodologie), Corinne Lorenzo (et les bières session), François de Vieilleville (avec qui j'espère avoir une expérience de collabora-

tion industrielle fructueuse), Nathalie Pinkerton (et le lancement de l'ultimate à l'ITAV), Rémy Touron (et ses saltos sans égal pour rattraper les disques), Camille Libet (qui donne envie d'aller voir l'administration plus souvent en trouvant des solutions à tout problème avec le sourire), Philippe Ciuciu (pour m'avoir permis de continuer à travailler sur les IRM), Fabrice Gamboa (pour m'avoir donné quelques clés de travail en équipe), Nicolas Boulant (pour sa générosité), Nicolas Chauffert et Claire Boyer (mes deux premiers doctorants qui m'ont insufflé un goût prononcé de l'encadrement par leur sympathie et leur sérieux), Wenxing Zhang (mon premier post-doctorant et sa culture traditionnelle chinoise), Ludivine Guillaume (ses petits plats, son implication dans la vie de laboratoire et une potentielle collaboration industrielle), Gabriele Steidl (qui me sert aussi d'exemple et d'inspiration), Jérémie Bigot (pour avoir co-encadré Paul Escande et Claire Boyer avec moi en me laissant beaucoup de liberté, tout en apportant quelques idées essentielles), Paul Escande (mon troisième doctorant, toujours heureux de prendre son ex co-directeur à défaut), Emmanuel Soubies (que j'ai trouvé remarquable en stage de M1 et toujours aussi remarquable comme nouveau collègue au CNRS), Etienne Corman (qui avec les deux précédents m'a donné une vision complètement biaisée de la qualité des étudiants de l'INSA), Maximilian März (et son goût des matrices aléatoires et de la géométrie de l'échantillonnage), Carole Lazarus (son sérieux, son implication et sa force scientifique tant en physique qu'en mathématiques), Léo Lebrat et Valentin Debarnot (qui sont de bons compagnons de route sur des thèmes aussi bien scientifiques, échiquéens et sportifs), Thomas Mangeat (dont la capacité à réaliser des expériences de microscopie n'a d'égal que sa grammaire), Renaud Morin (avec qui nous pourrions peut-être co-encadrer un doctorant plutôt que de taper dans des ballons de volley), Grégoire Quinet (qui serait le king s'il ne faisait pas 1m50), Vincent Duval (et sa capacité impressionnante à casser des problèmes d'analyse ardue et à rédiger de façon limpide avec des diagrammes commutatifs), Antonin Chambolle (et ses petites phrases anodines qui montrent qu'il a tout compris avant tout le monde), Didier Henrion (sans qui le séminaire commun ROLLMOPS n'existerait pas), Milan Korda (qui va peut-être me faire découvrir le monde du contrôle optimal en co-encadrant une thèse), Bernard Mariamé (sa philosophie de vie et nos parties échiquéennes endiablées).

Il faut aussi remercier le menu fretin des nombreux étudiants : Matthieu Bouyrie, Thomas Bonnafont et Camille Castera (à qui je souhaite de se réaliser dans leur doctorat et après), Léo Mouly, Guillaume de Brito, Benjamin Font, Guillaume Rauffet, Jean Eymerie, Bastien Kovac (l'ambiance festive des étés, les interfaces Fiji, l'ultimate et bien plus encore), Gabriel Bathie (à qui je souhaite de devenir une référence académique de l'informatique théorique dans quelques années) et Pierre-Jean Benard, Sandrine Bouguen, Ali Rouwanne et Maxime Taisant (pour leur support et leur implication), Landry Duguet (qui si j'en juge à sa détente devrait réussir à monter haut).

Ma famille

Je remercie tout d'abord mes parents qui n'ont de cesse de répondre présents lorsque je les sollicite ainsi que ma belle-mère Nannie qui réussit parfois à alléger notre quotidien.

Survivre à un chercheur au contact parfois rugueux et sans cesse hanté par des questions sibyllines est une vraie prouesse. Faire des enfants avec cette engeance est une vraie gageure. Je dédie ce manuscrit (que j'aurais aimé plus étincelant) à ma compagne Alice et à mes enfants Anouk et Isidore qui rendent tous ces problèmes scientifiques insignifiants.



FIGURE 1 – Anouk et Isidore sur leur papa.

Chapitre 1

Introduction générale

1.1 Guide de lecture

Cette HDR est composée de 5 chapitres scientifiques et d’une introduction sur ma vision de la recherche académique. Les 3 premiers chapitres (optimisation convexe, échantillonnage compressé, approximation d’opérateurs) contiennent des contributions de nature plutôt théorique tandis que les 2 derniers (imagerie computationnelle et autres travaux) décrivent plus la résolution de problèmes concrets, en utilisant les outils des 3 premiers chapitres. Ces chapitres peuvent être lus de manière à peu près indépendante. J’ai choisi d’indiquer mes pistes de recherche par le symbole  tout au long du document et aussi à la fin de chaque chapitre.

1.2 Pourquoi une habilitation ?

La vertu principale que devrait revêtir une habilitation serait de préserver les doctorants d’un encadrement déficient, en validant la capacité à suivre une thèse. En pratique, ce diplôme me semble plus sanctionner la capacité à mener des travaux de recherche, qu’à réellement encadrer des étudiants.¹ De plus, j’ai déjà eu l’occasion de “diriger des recherches” à de nombreuses occasions, et ce diplôme ne m’a pas paru indispensable jusqu’à présent. Ces réflexions, ma paresse naturelle et ma préférence pour la rédaction d’articles m’ont mené à repousser cette tâche jusqu’à maintenant.

Me voici cependant au pied du mur : j’ai en effet obtenu une dérogation pour encadrer la thèse de Valentin Debarnot. Or pour obtenir cette dérogation, je me

1. Le diplôme est un peu paradoxal, car il faut habiliter pour encadrer des étudiants et encadrer des étudiants pour habiliter. A moins d’avoir eu la chance de croiser un “parrain” habilité, proposant un co-encadrement, le précieux titre est donc difficilement accessible. J’ai eu la grande chance de co-encadrer 3 thèses, mais malheureusement, je croise de nombreux collègues brillants n’ayant pas pu faire leurs preuves dans leurs 5 à 10 premières années d’exercice.

suis engagé à soutenir mon HDR avant sa soutenance, ce qui lui permettra à son tour d'obtenir son diplôme. Je décide donc de m'engager aujourd'hui vers ce sommet académique. L'atteindre me permettra - je l'espère - de pouvoir (co)-encadrer plus facilement des thèses à l'avenir et de renforcer mon dossier pour postuler au titre de directeur de recherche.

Avant d'entamer la partie scientifique de ce manuscrit, je vais tenter de faire un travail d'introspection pour donner une rapide description de ma vision de la recherche académique. Je finirai cette introduction en fournissant un CV détaillé, pour une évaluation rapide de mon dossier.

1.3 Mon profil scientifique

Les origines J'ai toujours vu mon père radio-amateur construire des appareils de haute précision le soir en rentrant de son travail d'ingénieur en électronique. Ma mère quant à elle, est une insatiable créatrice, peignant des heures, écrivant des poèmes ou déracinant des arbres à mains nues. De mon côté, j'ai commencé à programmer des jeux vidéos au collège (d'abord sur Commodore 64 puis du Qbasic sur PC) pour m'évader à mon tour.

Ingénieur ou chercheur ? Ces éléments m'ont donné un goût pour la résolution de problèmes très concrets, et le besoin - pour me sentir à l'aise - de vérifier les théories que je développe avec l'outil informatique. En ce sens, je suis donc un ingénieur : je pars d'un problème ou d'une observation palpable, puis j'essaie de le résoudre ou de l'expliquer. Ceci ne manque pas d'une dose d'aléa quand les problèmes sont trop coriaces et mène souvent à des révisions d'objectifs à la baisse. En général, cette démarche me permet tout de même de valoriser les travaux en ciblant des journaux en ingénierie ou en mathématiques appliquées suivant le caractère plus ou moins fondamental du produit final.

Mon parcours et mes thématiques de recherche J'ai effectué mes études à l'INSA de Lyon, puis de Toulouse, dans la filière "Génie Mathématique et Modélisation". À mon époque, la mécanique des fluides et des structures dominaient ma filière et j'ai malheureusement fini par les prendre en grippe. J'ai donc cherché des disciplines alternatives et la présence du CNES à Toulouse m'a permis de faire quelques projets en traitement des images. Un stage de 3 mois avec Stéphane Mallat dans la start-up Let-It-Wave m'a encouragé à poursuivre dans cette voie, et à m'inscrire au Master MVA de l'ENS Cachan lors d'une année supplémentaire. Après une thèse mêlant mathématiques et traitement d'images, je suis devenu un spécialiste du domaine et les problèmes ont commencé à affluer naturellement. Aujourd'hui, l'essentiel de mes travaux sont dans ce thème, même si je n'avais pas spécialement souhaité cette trajectoire au départ. J'aimerais à l'avenir embrasser d'autres disciplines, ce qui devrait être possible en menant préalablement des travaux de natures plus théorique et généraliste.

Mon laboratoire Je suis actuellement affecté à l'ITAV (Institut des Technologies Avancées en sciences du Vivant). C'est un laboratoire d'environ 10 ans construit sur les ruines d'AZF, cette entreprise d'engrais appartenant à Total ayant explosé en 2001. L'idée de départ de ce lieu était de faire venir des pointures internationales sur la vaste thématique du cancer. Ce fut un rêve irréaliste car Toulouse-Métropole n'a donné aucun moyen réel pour attirer des scientifiques internationaux.

En pratique le laboratoire et surtout son directeur actuel Bernard Ducommun ont donc dû trouver des moyens d'attirer des équipes locales pour éviter que le lieu ne périclite faute de moyens et de personnel. L'idée initiale a été de faire du laboratoire un lieu interdisciplinaire. C'est dans cette optique que j'ai eu la chance de monter une des premières équipes en 2012 avec Jérôme Fehrenbach. On trouve aujourd'hui une équipe de biologistes, une équipe plus orientée vers l'optique, un peu de chimie, des mathématiques et de l'informatique. Les interactions entre disciplines existent réellement et c'est donc un petit succès. La motivation principale du site a cependant changé il y a 3-4 ans, en partie à cause de pressions politiques : l'objectif est plus la valorisation de travaux de recherches en créant des start-up. La majorité du personnel présent dans le bâtiment est aujourd'hui employé par de petites entreprises. L'ITAV disparaîtra dans 1 an face à des charges trop lourdes imposées par la ville. On m'a proposé de créer une équipe-projet dans un nouveau laboratoire transdisciplinaire sur le vieillissement appelé RESTORE et j'ai demandé mon rattachement principal à l'Institut de Mathématiques de Toulouse.

En pratique, je passe l'essentiel de mon temps à l'ITAV, mais je suis aussi très lié à mon laboratoire d'origine (l'Institut de Mathématiques de Toulouse) d'où viennent tous les membres de l'équipe. J'en suis membre associé à 20%.

Mes motivations Je m'éloigne fortement de l'image d'Épinal d'un chercheur passionné, capable de passer des mois devant un problème, malgré les difficultés et revers. En vérité, je m'essouffle assez rapidement et finis par procrastiner quand je suis seul. J'ai découvert un besoin crucial de collaborateurs pour échanger et garder l'énergie nécessaire à l'élaboration d'idées complexes. Mes échanges avec des chercheurs permanents, ayant un bagage scientifique riche, des idées alternatives, et souffrant souvent des mêmes maux que moi, sont une vraie aubaine et m'ont apporté énormément. L'interaction avec des disciplines exotiques (biologie ou optique) est aussi motivante, car elle permet de toucher des sujets nouveaux. De plus l'élaboration de modèles, théories ou algorithmes simples a souvent entraîné une vraie reconnaissance de mes collaborateurs qui n'ont pas la formation nécessaire pour faire ce travail. Les séances de travail avec des doctorants - qui prennent souvent la forme de semaines entières sur une question précise - sont des moments particulièrement privilégiés de ma vie de chercheur, que j'apprécie énormément, et qui me motivent à persévérer dans la voie académique. De plus, contrairement à beaucoup de chercheurs plus établis, les doctorants ont souvent gardé une vision plus idéaliste de la recherche sans aigreur ni considération pour les tractations politiques du domaine.

En somme la recherche est avant tout pour moi une aventure humaine, de création ou de compréhension commune.

Enseignement ou encadrement ? Après ma thèse, j'ai commencé par enseigner à l'INSA de Toulouse en tant que maître de conférence pendant 3 ans, puis j'ai eu la chance d'être recruté au CNRS. Après 6 ans dans cet environnement, je voue une grande admiration à mes collègues restés au front et capables - dans ces conditions - de maintenir une activité de recherche de qualité. L'enseignement est pour moi un plaisir lorsque je vois des regards brillants après l'explication de théorèmes ou de phénomènes. Il peut devenir réellement désagréable lorsque je sens une classe hostile, ne souhaitant pas réellement comprendre. Je ne sais pas si c'est lié à ma personnalité ou à l'aléa des promotions, mais les deux situations précédentes se sont produites de façon à peu près équilibrée. Dans tous les cas, j'ai trouvé que les 192h d'enseignement réglementaires réparties arbitrairement sur l'année étaient trop pour moi, créant souvent des frustrations lorsque je devais abandonner des questions scientifiques.

L'encadrement d'étudiants en projet, en stage ou en thèse quant à lui a quasiment systématiquement été très enrichissant. La plupart du temps, j'ai pu rapidement considérer les étudiants comme de vrais collègues de travail et parfois des amis. Leur niveau d'étude ou leur pedigree est - de mon point de vue - loin d'être proportionnel à leur capacité à apporter une contribution. Cette observation m'a amené à recruter des étudiants de tous bords pour des périodes de 3 à 6 mois, avec notamment 10 étudiants en IUT informatique (développement Java pour des interfaces sous les logiciels libres Icy ou Fiji), 3 étudiants en L3, 6 étudiants en M1, et 5 étudiants de M2. Il y a bien eu quelques stages ratés et des inadaptations au monde professionnel parfois cocasses, mais dans l'ensemble, ce fût riche et passionnant.

Même si le niveau de connaissance est souvent à ma faveur, le fait de devoir expliquer posément une théorie ou un algorithme, pour qu'elle soit exploitée par la suite, m'a permis de me forcer à mieux tourner certains articles et surtout d'accélérer la prise de recul en m'évitant certaines tâches parfois abrutissantes de test ou de codage.

Pour conclure cette partie, je dois donc bien me rendre à l'évidence : enseigner à un petit nombre d'élèves triés avec soin est une expérience humaine forte, tandis que j'ai bien plus de mal à trouver les mots pour sortir les masses hagardees de leur torpeur.

Le suivi de doctorants J'ai eu le privilège de co-encadrer officiellement 3 doctorats : ceux de Nicolas Chauffert, Claire Boyer et Paul Escande. Toutes ces personnes ont fourni un travail conséquent, avec un vrai goût pour la recherche et des talents divers. J'ai beaucoup apprécié nos échanges à tous points de vue et je les considère aujourd'hui comme des proches.

Après son doctorat, Nicolas a préféré devenir professeur agrégé avec l'objectif d'enseigner en classes préparatoires bien que son dossier le rende attrayant académiquement. Claire a obtenu un poste de maître de conférence à l'univer-

sité Paris 6 et Paul vient d'être recruté chercheur CNRS à Marseille. Bien que le suivi de 3 doctorats en parallèle a parfois été intense, il m'a énormément apporté, car j'ai souvent pu réexploiter les idées développées avec les uns, pour les besoins d'un autre.

Fort de cette expérience, j'ai cherché à recruter d'autres étudiants, mais je me suis heurté à de sérieux soucis, me permettant de mieux prendre conscience de la chance que j'ai eue. Premièrement, le vivier d'étudiants en mathématiques appliquées est faible. Trouver un étudiant qu'on estime pouvoir suivre pendant 3 ans est donc une tâche ardue. Même lorsqu'on le trouve, la recherche d'une bourse doctorale devient une des tâches les plus chronophages du métier. Ainsi, sur la dizaine de demandes de bourses déposées dans les 3 dernières années, une seule a abouti : celle de Valentin Debarnot. Que les autres reposent en paix.

Face à l'adversité, j'ai cherché des possibilités d'encadrement alternatives et j'ai eu la chance de pouvoir travailler assez intensément avec 3 doctorants supplémentaires. Carole Lazarus a soutenu sa thèse au CEA en septembre 2018. Elle a poursuivi les travaux de Nicolas Chauffert sur la génération de schémas d'échantillonnage sur des IRM. Nous collaborons encore aujourd'hui, puisqu'elle a été recrutée à Hyperfine-research, la société d'imagerie médicale pour laquelle je suis consultant. De plus, lors d'une invitation à l'université de Berlin pour un séminaire, j'ai pu discuter avec les nombreux doctorants du groupe de Gitta Kutyniok. Après ces discussions, j'ai proposé à Maximilian März de venir travailler avec moi, ce qui a été accepté par sa directrice et par lui. Nous nous voyons désormais régulièrement et travaillons sur 3 articles en ce moment. Finalement, j'ai pu inviter Axel Flinth (du même groupe) à une conférence et nous avons rapidement² rédigé un article accepté récemment à *Information&Inference*. Axel nous a rejoint en octobre 2018 pour effectuer un post-doctorat de un ou deux ans, dans le cadre de l'ANR jeune chercheur OMS portée par Jonas Kahn.

1.4 Le paradis du monde académique

Un étudiant modèle obtient les meilleures notes de ses promotions pendant toute sa scolarité. Grâce à ce comportement exemplaire, mélange de talent, d'endurance et d'abnégation, il réussit à décrocher une bourse de thèse, qui va lui permettre de rejoindre le monde merveilleux de ses professeurs. Avec un peu de chance, il ne sortira pas de sa béatitude lors de son doctorat, pouvant alors prétendre à un poste académique, seul refuge face aux affres du monde socio-industriel environnant. La pression forte qui existe lors du recrutement va souvent le rendre fortement attaché à son poste et lui faire oublier partiellement certains aspects désagréables de sa profession.

Bien qu'il existe de nombreux lieux remarquables, les conditions réelles du monde académique en France et dans le monde me semblent cependant loin d'être idéales, et il faut garder une certaine dose d'idéalisme pour y survivre.

2. trop rapidement en réalité, car nous avons tous les deux trop de pression et une seule semaine commune, ce qui a donné des boutons à un relecteur.

1.4.1 *Soutiens financiers*

Je passe une bonne partie de mon temps à déposer des demandes de financement (stages, thèses, projets). Je ne pense pas que ce soit une mauvaise activité pour un chargé de recherche au CNRS qui peut profiter de ce temps pour ficeler des projets à plus long terme. Les taux de succès sont cependant réellement attristants et les retours des relecteurs souvent peu sérieux. Sur les 8 dernières années, 2 projets ANR sur 12 ont été acceptés malgré une implication forte de ma part. Malheureusement, ce sont des projets de biologie pour lesquels j'ai eu essentiellement une activité de service. En réalité mes propres recherches n'ont jamais été soutenues, si ce n'est récemment avec l'ANR Jeune Chercheur OMS (Optimization on Measures Spaces) portée par Jonas Kahn.

La perspective de ne plus avoir de financement est parfois effrayante. Il faut effectivement payer des frais informatiques à l'ITAV, pour permettre de respecter les conditions de locations imposées par la ville. Pour la première fois, je me suis retrouvé incapable de payer les frais associés à une publication acceptée il y a peu (que le laboratoire a accepté de prendre en charge). J'ai déjà payé les frais de voyage de 3 conférences de ma poche. Je sais que je ne pourrai pas recruter d'étudiants en stage (et donc probablement en thèse) cette année³. Les bourses de doctorat ou de post-doctorat semblent inaccessibles à moins que l'étudiant vienne avec un financement. En général j'accueille les doctorants extérieurs à la maison pour réduire les coûts, mais aussi parce que ça génère des moments conviviaux. En somme la tension financière est réelle et génère - de mon point de vue - un stress assez important.

Je n'ai mentionné ici qu'une situation personnelle. La situation de nombreuses universités semble aujourd'hui critique avec une incapacité de remplacer les personnels partant à la retraite. L'indépendance des universités mêlée à l'impossibilité de faire payer les formations rend la situation réellement tendue. En parallèle de cet appauvrissement, on voit de plus en plus une concentration des financements sur très peu de centres ou même d'individus. Des règles telles que "il faut avoir une ERC/tant de doctorants/un historique de collaboration avec entreprise pour pouvoir candidater à tel financement" renforcent cet effet et sont pourtant décidées par des scientifiques.

Peut-on raisonnablement espérer des améliorations de notre métier à l'avenir? Je l'espère, mais je n'ai pas l'impression que les gouvernements y prêtent de l'attention en France.

1.4.2 *Locaux et mobilier*

J'ai eu l'occasion de travailler à divers endroits et de visiter de nombreux laboratoires. Après les conditions rêvées de l'INRIA Sophia-Antipolis, j'ai pu expérimenter divers établissements et mesurer des disparités énormes. Devoir porter des gants en hiver et quitter son bureau à 16h en été pour éviter des chaleurs supérieures à 45° n'est pas acceptable. Je suis aujourd'hui dans un

3. J'ai écrit ces lignes trop tôt, car 2 mois plus tard, j'ai eu la chance de recruter Alban Gossard et peut-être Corbinian Schlosser, deux étudiants qui semblent bourrés de qualités.

laboratoire (l'ITAV) ayant des conditions très privilégiées, avec l'air conditionné, des bureaux spacieux et j'ai pu installer un serveur dans une baie refroidie pour l'équipe. J'espère pouvoir retourner dans des conditions de travail favorables lors de sa fermeture. Dans tous les cas, je reste perplexe face à la situation de certains collègues et il n'est pas surprenant que de plus en plus d'industries réussissent à recruter directement dans le vivier des professeurs ou des maîtres de conférences, notamment dans le domaine de l'intelligence artificielle.

1.4.3 *Congrès, communautés et réseaux d'influence*

De mon point de vue, un aspect important et assez négatif du monde académique est l'existence ou l'émergence de réseaux d'influence et de communautés scientifiques partiellement fermées. La formation d'une telle communauté peut avoir lieu ainsi : un joli résultat ou une discipline émerge (e.g. la théorie des shearlet, la théorie de l'échantillonnage compressé, la programmation semi-infinie, les réseaux de neurones convolutionnels...). Quelques pionniers apportent les briques essentielles à la théorie et des résultats importants émergent. Ceci motive la création de nouvelles conférences ou colloques spécialisés. Les personnes qui participent aux premières conférences finissent par bien se connaître grâce à des discussions scientifiques, mais aussi grâce à des discussions stratégiques sur le pilotage du domaine. Les chercheurs ayant l'esprit politique le plus aiguisé réussissent à devenir les maîtres de la discipline (e.g. éditeurs associés, réseaux européens, organisateurs de congrès) et peuvent finir par avoir le droit de vie et de mort sur de nouveaux travaux. Ils ont souvent peu de temps et il est difficile de ne pas avoir un a priori plutôt favorable sur des travaux de chercheurs que l'on connaît déjà et une attitude moins bienveillante sur de nouveaux arrivants qui pourraient par ailleurs finir par réduire leur influence. Il est fréquent que je me demande comment un papier a pu être accepté dans tel ou tel journal et la réponse devient assez claire après analyse des auteurs et de l'éditeur correspondant.

Se fondre dans un réseau d'influence est souvent possible, rien qu'en allant à des conférences et en discutant avec les chercheurs. Cependant, ceci suppose de pouvoir se déplacer, ce qui peut être impossible pour des chercheurs de pays africains par exemple, et qui devient très difficile pour de nombreux pays européens (Espagne, Italie notamment et aussi certains chercheurs français). Je pense que ce mécanisme peut accélérer la concentration des richesses sur les "happy few", qui me semble assez alarmant avec l'émergence des ERC et des initiatives d'"excellence".

De mon côté, je pense être maintenant relativement bien intégré au réseau "mathématiques et imagerie", porté par des journaux tels que JMIV ou SIAM Imaging Science. En pratique, ceci signifie qu'il devient presque trop facile d'avoir des papiers acceptés dans ces journaux tandis que j'ai souvent des retours arides et agressifs dans d'autres communautés. Tout n'est pas noir cependant : je trouve qu'il y a une ambiance assez familiale et bienveillante une fois qu'on a rejoint une communauté, ce qui permet de réduire le stress et de bénéficier d'entraide.

1.4.4 *Recherche libre d'accès et industrie*

Beaucoup de chercheurs considèrent la recherche libre et reproductible (i.e. la diffusion des codes) comme essentielle. J'ai tendance à abonder dans ce sens. Cependant, cette démarche présente aussi certaines contradictions. La première est liée au renouveau industriel français et européen tant demandé par nos tutelles : un code libre d'accès profite à tous de la même manière alors qu'on nous pousse fortement au patriotisme lors du dépôt de projets. De plus, une fois un code publié, il ne peut plus être breveté ce qui limite d'autant son intérêt mercantile.

A l'avenir, je me rapprocherais volontiers de certaines industries locales ou européennes pour les aider dans des projets de recherche qui m'intéressent. Publier des articles a d'ailleurs un intérêt tout relatif et probablement moins d'impact sociétal que la confection d'un produit manufacturé. Ceci sera d'autant plus vrai, si ces collaborations me permettent de financer certaines missions. Malheureusement, mes essais de rapprochement industriel ont été jusqu'à présent infructueux et mes seules opportunités sont venues d'entreprises américaines qui ont cherché à me débaucher.

1.4.5 *Deep-learning*

Il me semble aujourd'hui difficile de continuer à faire du traitement d'images sans m'intéresser de près au domaine naissant de l'apprentissage profond. Ce courant est effrayant pour plusieurs raisons. Premièrement, la méthodologie que j'ai adoptée jusqu'à présent, qui consiste à essayer de comprendre et de modéliser mes propres sens pour reproduire des tâches avec un ordinateur semble aujourd'hui partiellement révolue : je ne suis probablement pas assez compétent pour me comprendre moi-même. Deuxièmement, de nombreux outils que j'ai pu développer ou que j'ai fini par comprendre sont devenus caduques dans les 5 dernières années. De façon plus importante, on sent qu'on commence à toucher aux limites du cerveau humain : les ordinateurs peuvent aujourd'hui résoudre des tâches longtemps pensées comme le pré gardé des humains, de manière significativement plus efficace que leur créateur. Ainsi, on doit envisager des modifications profondes des professions qualifiées (e.g. médecins, chercheurs, ingénieurs), alors que jusqu'à présent la révolution industrielle avait surtout modifié le travail ouvrier.

Un chercheur dans ma catégorie ne peut pas se permettre d'ignorer ce domaine, et il est clair que vais commencer à l'aborder dans certains de mes projets à venir.

1.5 Mon CV détaillé

Pierre Weiss

Institut des Technologies Avancées du Vivant
 ITAV, USR 3505
 Université de Toulouse, CNRS
 1 place Pierre Potier
 31106 - Toulouse
 France

Phone : +0033 - 5 - 82 - 99 - 10 - 25
 Email : pierre.weiss@itav-recherche.fr
 Webpage

Birthdate : 14/12/1981

Living with a partner and two children.

Employment

- 2017 - present : part-time consultant for 4-Catalyzer (deep learning for medical diagnosis), USA.
- 2012 - present : Researcher CNRS, CR2, ITAV (USR 3505).
- 2009 - 2012 : Associate Professor, department of mathematical modeling, INSA de Toulouse.
- 2008 - 2009 : Post doctoral fellow at Hong Kong Baptist University, ICM (with Michael Ng).

Education

- 2005-2008 : PhD in Mathematics and Image Processing at INRIA, Sophia Antipolis.
 - *Dissertation* : Algorithmes rapides d'optimisation convexe. Application à la reconstruction d'images et à la détection de changements.
 - *Committee* : Gilles Aubert, Laure Blanc-Féraud, Patrick Combettes, Antonin Chambolle, Henri Maitre, Lionel Moisan.
- 2004 - 2005 : Master MVA (Mathématiques, Vision, Apprentissage) at ENS Cachan.
- 2003 - 2004 : Master of applied mathematics, Université Paul Sabatier de Toulouse.
- 1999 - 2004 : Engineering studies at INSA Lyon and Toulouse. Specialization in mathematical modeling

Fields of Research Interest

Applied mathematics : numerical optimization, inverse problems, learning, complexity theory in optimization, approximation theory (functions and operators), compressed sensing, numerical harmonic analysis.

Signal/image processing : computational imaging, magnetic resonance imaging, microscopy (in a wide sense), remote sensing, biomedical image analysis.

Modelling : optics, medical devices, cell biology.

Publications

Voir chapitre 8.

Patents

- Method and apparatus for accelerated magnetic resonance imaging.

Softwares

- Contrast Invariant SNR https://github.com/pierre-weiss/contrast_invariant_snr.
- VSNR CUDA, Matlab and standalone C++ https://github.com/pierre-weiss/VSNR_CUDA_MEX.
- VSNR CUDA, plugin for Fiji https://github.com/pierre-weiss/VSNR_2D-3D_GPU.
- VSNR, main Java interface for Fiji https://www.math.univ-toulouse.fr/~weiss/Codes/VSNR/VSNR2_plugin.jar.
- FitEllipsoid, plugin for Icy https://github.com/pierre-weiss/FitEllipsoid_Icy.
- Multiview Attenuation Computation <https://github.com/pierre-weiss/MAEC>.
- MRI, trajectory projections <http://chauffertn.free.fr/codes.html>.
- MRI, trajectory generation using TSP <http://chauffertn.free.fr/codes.html>.

Visits & Summer schools

I overall tend to avoid traveling too much for family and global warming issues.

- February 2016 : 1 week visit at DTU, Denmark, February, 2016.
- September 2014 : conférencier invité à l'école d'analyse harmonique à Marseille (web).
- April 2011 : conférencier école de printemps Mathématique-Image à Martel (web).
- May 2011 : one month invited stay at Hong Kong Baptist University to work on theoretical aspects of the ADM-method with X. Yuan.

Conference and Seminar Presentations

- ICIAM 2019 (in memory of Mila Nikolova), Valencia, Spain, July 2019 (invited)
- OSA topical meetings Maths & Imaging, Munchen, Germany, June 2019.
- Séminaire MOKAPLAN, Janvier 2019 (invited).
- Journées SFRBT, Avancées de la recherche biomédicale à Toulouse, December 2018 (invited).
- ITWiST 2018, Marseille, November 2018.
- OSA topical meetings Maths & Imaging, Orlando, June 2018.

- Changsha Conference on mathematics and image processing, Changsha, April 2018 (invited).
- IPIP Conference, Beijing, April 2018 (invited).
- ITAV, Toulouse, February 2018.
- Journées SIGMA-SMAI, Paris, IHP, November 2017 (invited).
- Generative models, parameter learning and sparsity, Cambridge, UK, November 2017 (invited).
- Workshop MOA & MIA, Bordeaux, October 2017 (invited).
- Variational Methods, New Optimization Techniques, Cambridge, UK, September 2017 (invited).
- Imaging & Applied Optics Congress, San Francisco, USA, June 2017 (committee).
- SPARS, Lisbon, Portugal, June 2017.
- Imaging & Applied Optics Congress, Heidelberg, Germany, July 2016 (invited).
- Colloquium, Technische Universität Berlin, June 2016 (invited).
- Seminar University of Louvain, March 2016 (invited).
- Visit at DTU, Denmark, February, 2016 (invited).
- Conference MIA, Paris, January 2016 (invited).
- Seminar at Cambridge (Applied and Computational Analysis), November 2015 (invited).
- Workshop GDR 2588, Paris, Institut Pasteur, November 2015.
- Conference GRETSI, Lyon, September 2015.
- Séminaire Image, MAP5, Paris, April 2015 (invited).
- Groupe de Travail Maths-Bio, IMT, Toulouse, February 2015 (invited).
- Winter school on computational harmonic analysis, October 2014, Marseille (invited).
- Conference ITWIST 2014 (plenary speaker), August 2014.
- Workshop on optimization at HKBU, Hong Kong, May 2014 (invited).
- Conference SIAM Imaging science, Hong Kong, May 2014.
- Séminaire SPOT, ENSEEIHT, décembre 2013 (invited).
- Journées MIP, IMT, décembre 2013 (invited).
- Séminaire au laboratoire I3S, Sophia-Antipolis, Juin 2013 (invited).
- Séminaire au XLim, Limoges, Mars 2013 (invited).
- Conference ICPRAM in Barcelona, February 2013.
- Séminaire ITAV, Toulouse, Février 2013 (invited).
- GIS Genotoul, Toulouse, Novembre 2012 (invited).
- Journées bordelaises d'analyse mathématique des images, Bordeaux, Novembre 2012 (invited).
- Séminaire à l'institut de mathématiques d'Avignon, Avignon, Novembre 2012 (invited).
- Animation d'une table ronde, 6th Imaging The Cell, Toulouse, Juin 2012.
- Workshop I-Twist, Marseille, France, May 2012 (invited).
- Séminaire Image, Université de Toulouse, Avril 2012 (invited).
- Conference ICPRAM, Vilamoura, Portugal, February 2012.
- Third Light Sheet Microscopy Workshop, Toulouse, October 2011.

- Conference AFG, Toulouse, September 2011.
- Seminar at NUS, National University of Singapore, May 2011.
- Séminaire Image, Université Paris 6, février 2011.
- Séminaire Applications des Mathématiques, ENS Cachan Ker Lann, février 2011.
- Workshop MIBS - Les dessous de l'image, Toulouse, Novembre 2010.
- Séminaire au LATP, à l'université de Marseille, mars 2010.
- Colloque Méthodes Mathématiques Image, Orléans, avril 2010.
- Séminaire à l'université de Marne la vallée, février 2010.
- Séminaire MIP 2009, Toulouse, France.
- Conference EUSIPCO 2009, Glasgow, Scotland.
- Seminar at Chinese University, Hong Kong, 2009.
- Conference SAMPTA 2009, Marseille, France.
- Workshop MCI 2009, Hong Kong Baptist University, Hong Kong.
- Conference ICASSP 2008, Las Vegas, USA.
- Conference Canum 2007, Praz-sur-Arly, France.
- Conference Code 2007, Institut Henri Poincaré, Paris, France.
- Conference SIAM conference on Imaging Science 2007, Minneapolis, USA.
- Conference MIA (Mathematics and Image Analysis) 2007 - Paris, France.
- Conference GretsI 2007, Université de Troyes, France.

Professional Activities

Management Founder and PI of PRIMO project since 2012, a scientific team now involving 4 permanent members. ITAV, USR 3505, (web).

Reviewer for international journals About 15 reviews/year in

- **Mathematics.** SIAM Journals : Scientific Computing, Imaging Sciences. FoCM. Mathematical Programming. Numerical Algorithms. Inverse Problems. JMIV. Applied Mathematics Computation. Computational and Applied Mathematics. Applied Numerical Mathematics. Optimization Software. International Journal of Computer Mathematics. J. of Zhejiang University.
- **Signal processing.** IEEE transactions : Image Processing, Signal Processing, Signal Processing Letters, Biomedical Engineering, Medical Imaging. Multidimensional Systems and Signal Processing. IPOL. J. of Electronic Imaging, CVIU.
- **Other.** Nature methods, Computer& Graphics, Sensors, Magnetic Resonance Imaging.

Organizer for conferences

- Technical Committee SPARS (2019).
- General chair OSA mathematics in imaging (2019).
- Technical committee SPARS (2019).
- Program chair OSA mathematics in imaging (2018).

- Organizing committee CIMI thematic semester on optimization (2018).
- Organizing a session at ISMP (2018).
- Scientific committee OSA mathematics in imaging (2017).
- Organizing committee GDR 2588 workshop on bioimage informatics (2013, 2014, 2015, 2017).
- Organizing committee CIMI thematic school on Image Processing (2013).
- Program committee ICPRAM (2012-2013).
- Organizing committee AFG2011.

Reviewer for conferences

- Reviewer for GRETSI (2015-2017).
- Reviewer SSVM (2015-2016).
- Reviewer ICIP (2013-2014).
- Reviewer Eusipco (2014-2017).
- Reviewer for Sampta (2013-2017,2019).
- Reviewer for ISBI (2012-2014, 2018-2019).
- Program chair of SITIS conference (2008-2009).

PhD committees

- Examiner : Carole Lazarus, Saclay, (September 2018).
- Examiner : Hoai-Nam Nguyen, Rennes, (December 2017).
- Examiner : Kevin Polisano, Grenoble, (December 2017).
- Examiner : Rémi Abergel, Paris 5 (2016).
- Reviewer : Adriana Gonzalez, Louvain-la-Neuve (2016).
- Examiner : Pierre Machart, Marseille (2013).

Teaching

2018-2019 Around 40 hours at INSA de Toulouse and ISAE.

2017-2018 Around 40 hours at INSA de Toulouse and ISAE.

2016-2017 Around 40 hours at INSA de Toulouse and ISAE.

2015-2016 Around 70 hours at INSA de Toulouse, ISAE and M2R Université Paul Sabatier.

2014-2015 Around 70 hours at INSA de Toulouse, ISAE and M2R Université Paul Sabatier.

2013-2014 Around 60 hours at INSA de Toulouse, ISAE and M2R Université Paul Sabatier.

2012-2013 Around 80 hours at INSA de Toulouse and M2R Université Paul Sabatier.

2009-2012 Between 150 and 200 hours every year at INSA de Toulouse and Université Paul Sabatier in the following disciplines :

Analysis (1st, 2nd year), Inverse problems (3rd year), Numerical analysis (3rd year), Optimization (4th year), Imaging and Bayes approaches (5th year).

2005-2008 ESINSA of Nice-Sophia Antipolis : around 200 hours taught in three years.

Signal processing (3rd year), Mathematics (1st year), Boolean electronics (1st year).

Advising

- **Post-doctoral fellows :**

- Axel Flinth (2018-2019 together with J. Kahn, F. de Gournay, J. Bolte).
- Wenxing Zhang (2014-2015 together with J. Fehrenbach).

- **PhD students :**

- Valentin Debarnot (2017-2020 together with T. Mangeat).
- Paul Escande (2013-2016 together with J. Bigot).
 - Prix de l'académie de Toulouse 2017. Meilleure thèse de l'académie de Toulouse dans les domaines mathématiques-informatiques. Chercheur CNRS à Marseille.
- Claire Boyer (2012-2015 together with J. Bigot).
 - Maître de conférence à l'université Paris 6.
 - Prix Adrien Gaussail 2015. Meilleure thèse en maths-physique de l'académie de Toulouse.
 - Prix de thèse EEA. Meilleure thèse au niveau national dans les domaines de signal, image, vision.
- Nicolas Chauffert (2012-2015 together with P. Ciuciu). Professeur agrégé en classes préparatoires.

- **Master/IUT Students :**

- 2018 Camille Castera (M2), Sandrine Bouguen (L3), Gabriel Bathie (L3).
- 2017 Thomas Bonnafont (M2), Jean Eymerie (IUT).
- 2016 Anh-Tuan Nguyen (M2), Valentin Debarnot (M1), Bastien Kovac (IUT), Vincent Remenant (IUT), Alexandre le Guelvouit (IUT).
- 2015 Benjamin Font (IUT), Guillaume De Brito (IUT), Guillaume Rauffet (IUT), Morgan Gauthier (IUT).
- 2014 Quentin Gey (M1).
- 2013 Paul Escande (M2), Léo Mouly (License), Alexandre Receveur (IUT), Omar Dounia (M1).
- 2012 Nicolas Chauffert (M2), Paul Escande (M1), Emmanuel Soubies (M1), Charlotte Emery (M1), Léo Mouly (IUT).
- 2011 Claire Boyer, Mathieu Bouyrie (M1) (together with P. Ciuciu and B. Ducommun resp.).

Grants

- ANR JCJC OMS (2018-2021), 220kE. PI : J. Kahn (IMT, Toulouse). Role : responsible for a pole.
- Défi Imag'In (2017-2018), 20kE. PI : T. Mangeat & Pierre Weiss (CBI, Toulouse). Role : co-PI.
- RITC OPTIMUS (2015-2017), 95kE. PI : F. De Gournay (IMT, Toulouse). Role : participant.
- Plan Cancer MIMMOSA (2014-2017), 275kE. PI : T. Colin (IMB, Bordeaux). Role : responsible for a pole.
- ANR SPHIM3D (2013-2016), 416kE. PI : B. Ducommun (IMT, Toulouse). Role : responsible for a pole.
- Member of MAORI, PGMO Projects, (2013-2015). Role : participant.
- ANR MOTIMO (2012-2015), 222kE. PI : P. Degond (IMT, Toulouse). Role : participant.
- PEPH Caspa3D (2011-2012), 54kE. PI : B. Ducommun (ITAV, Toulouse). Role : participant.
- ABOA-SPIM (2011-2012) 60kE. PI : C. Lorenzo (ITAV, Toulouse). Role : participant.
- Toulouse University. 1kE for a collaboration with X. Yuan in summer 2010. Role : participant.

———— **Current active collaborations (random order)** ————

Jonas Kahn (CR CNRS Lille), Jérôme Fehrenbach (MCF Toulouse), Frédéric de Gournay (MCF Toulouse), Thomas Mangeat (IR Toulouse), Valérie Lobjois (MCF, Toulouse), Vincent Duval (Chercheur, CNRS), Antonin Chambolle (DR, CNRS), Yohan de Castro (MCF, Paris), Philippe Ciuciu (Chercheur, CEA, Paris), Carole Lazarus (PhD, CEA), Léo Lebrat (PhD, Toulouse), Maximilian März (PhD, Berlin), Axel Flinth (PhD, Berlin), Claire Boyer (MCF, Paris 6), Paul Escande (CR, CNRS).

———— **Miscellaneous** ————

Flute player, Bike rider addict.

Chapitre 2

Optimisation convexe

2.1 Introduction

Lors de ma thèse, j'ai travaillé sur des problématiques d'optimisation convexe en dimension finie, avec notamment des notions de complexité qui ne semblaient pas avoir intéressé les gens du domaine signal/image à cette époque. Par la suite, j'ai préféré rester très en retrait sur cette thématique, car il y a eu un engouement qui me semble démesuré avec aujourd'hui des dizaines de milliers de publications sur ce thème.

Si je résume très grossièrement mes contributions, je pense avoir été le premier à introduire les descentes accélérées à la Nesterov en traitement d'images [37] (avant Beck-Teboulle [BT09a] par exemple, puisque mon papier est resté en révision pendant plus de 2 ans). J'ai aussi été un des premiers à proposer d'utiliser les techniques de type ADMM [35]. Ces deux travaux sont aujourd'hui les plus cités. Ceci étant dit, c'est une contribution qui me semble assez légère car l'essentiel de l'analyse théorique avait été faite par des chercheurs tels que Nesterov-Nemirovskii ou Lions-Mercier et Fortin-Glowinski bien avant. Plus récemment, je me suis intéressé aux problèmes convexes semi-infinis et en moindre mesure à la programmation dynamique.

Dans ce chapitre, je commence par résumer mes contributions théoriques principales dans ce domaine, puis quelques applications originales qui reposent sur ces travaux.

2.2 Dualité et forte convexité

Dans cette partie, E et F sont deux espaces vectoriels de dimension finie $n \in \mathbb{N}$ et $m \in \mathbb{N}$ dont les espaces duaux respectifs sont E^* et F^* . On les munit de leurs normes respectives $\|\cdot\|_E$ et $\|\cdot\|_F$. Je présente un algorithme dual pour minimiser des fonctions composites dont l'une est fortement convexe. Je crois avoir été le premier à introduire cet algorithme accéléré et l'analyse de complexité associée.

2.2.1 Une méthode de descente accélérée

Pour commencer, voici une définition généralisée de la forte convexité.

Définition 1. Une fonction $g : E \rightarrow \mathbb{R}$ est σ -fortement convexe par rapport à $\|\cdot\|_E$ sur un ensemble $E' \subset E$ si pour tout $(x_1, x_2) \in E' \times E'$ et pour tout $t \in [0, 1]$

$$g(tx_1 + (1-t)x_2) \leq tg(x_1) + (1-t)g(x_2) - \frac{\sigma}{2}t(1-t)\|x_1 - x_2\|_E^2. \quad (2.1)$$

Le module de convexité σ_g de g est le plus grand réel positif σ qui satisfait l'inégalité (2.1) pour toute paire $(x_1, x_2) \in E' \times E'$.

Notre objectif dans cette partie est indiqué ci-dessous.

Soit $M : E \rightarrow F^*$ un opérateur linéaire. On souhaite résoudre :

$$\min_{x \in E} p(x) \stackrel{\text{def.}}{=} f(Mx) + g(x), \quad (2.2)$$

sous les hypothèses suivantes :

- $f : F^* \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction convexe fermée.
- $g : E \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction fortement convexe par rapport à $\|\cdot\|_E$.
- Pour assurer l'existence d'une solution on suppose que

$$M^{-1} \cdot \text{ri}(\text{dom}(f)) \cap \text{ri}(\text{dom}(g)) \neq \emptyset$$

où $\text{ri}(D)$ est l'intérieur relatif d'un domaine D .

L'algorithme développé a été motivé par le papier d'Antonin Chambolle [Cha04], qui est lui-même un algorithme de type Uzawa introduit dans les années 60. Je l'ai généralisé et accéléré. L'idée principale est de résoudre le problème dual suivant :

$$\sup_{y \in F} d(y) \stackrel{\text{def.}}{=} -g^*(-M^*y) - f^*(y), \quad (2.3)$$

avec une descente de gradient accélérée. Les fonctions f^* et g^* sont les transformées de Fenchel de f et de g . Le lemme suivant (montré dans [28] explicite les propriétés de régularité du dual ainsi que les liens qui existent entre les variables primales et duales. Il me semble peu connu sous cette forme. Il fait suite à ma thèse et a été motivé par des discussions avec Jalal Fadili et Gabriel Peyré [FP11]. Le formalisme est celui proposé dans le papier de Yurii Nesterov [Nes05].

Lemme 1. Soit $f : F^* \rightarrow \mathbb{R} \cup \{+\infty\}$ et $g : E \rightarrow \mathbb{R} \cup \{+\infty\}$ deux fonctions convexes fermées telles que $M \cdot \text{ri}(\text{dom}(f)) \cap \text{ri}(\text{dom}(g)) \neq \emptyset$. Supposons de plus que g est σ_g -fortement convexe par rapport à une norme $\|\cdot\|_E$. Alors

1. La fonction g^* satisfait $\text{dom}(g^*) = E^*$. Elle est différentiable. La fonction $h = g \circ M^*$ est différentiable à gradient Lipschitz sur F :

$$\|\nabla h(y_1) - \nabla h(y_2)\|_{F^*} \leq L\|y_1 - y_2\|_F, \forall (y_1, y_2) \in F \times F, \quad (2.4)$$

avec $L \leq \frac{\|M^*\|_{F \rightarrow E^*}^2}{\sigma_g}$ et $\|M^*\|_{F \rightarrow E^*}^2 \stackrel{\text{def.}}{=} \sup_{\|y\|_F \leq 1} \|M^*y\|_{E^*}$.

2. Notons $x(y) = \nabla g^*(-M^*y)$. Soit x^* le minimiseur de (2.2) et y^* un minimiseur quelconque de (2.3). Alors pour tout $y \in F$:

$$\|x(y) - x^*\|_E^2 \leq \frac{2}{\sigma_g} (d(y) - d(y^*)). \quad (2.5)$$

Le cas euclidien

Nous commençons par un algorithme édulcoré dans le cas euclidien en utilisant l'algorithme de Beck-Teboulle [BT09a]. On pose donc $E = E^* = \mathbb{R}^n$, $\|x\|_E = \|x\|_{E^*} = \|x\|_2$, $F = F^* = \mathbb{R}^m$ et $\|y\|_F = \|y\|_{F^*} = \|y\|_2$.

Algorithm 1 Descente proximale accélérée duale (cas euclidien)

- ① Choisir $y_0 \in \text{dom}(d)$ et poser $\hat{y}_0 = y_0$.
 - ② Calculer $\tau = \frac{\sigma_g}{\|M\|_{2 \rightarrow 2}^2}$.
 - ③ **for** $k = 0 \dots K$ **do**
 - ④ $y_{k+1} = \text{Prox}_{\tau f^*}(\hat{y}_k + \nabla h(\hat{y}_k))$.
 - ⑤ $\hat{y}_{k+1} = y_k + \frac{k-1}{k+2}(y_k - y_{k-1})$
 - ⑥ **end for**
 - ⑦ $x_k = \nabla g^*(-M^*y_k)$.
-

Le lemme 1 utilisé en conjonction avec les résultats de convergence de [BT09a] mène aux garanties suivantes.

Théorème 2. Soit x^* l'unique minimiseur du problème primal (2.2) et y^* un minimiseur quelconque du problème dual (2.3). Alors l'algorithme 1 assure :

$$\|x_k - x^*\|_2^2 \leq \frac{2\|M\|_{2 \rightarrow 2}^2 \|y_0 - y^*\|_2^2}{\sigma^2 k^2}. \quad (2.6)$$

C'est un résultat assez fort de convergence et on verra plus tard qu'il donne des garanties exploitables dans certaines situations favorables et bien moins satisfaisantes dans d'autres problèmes.

Le cas non euclidien

Le cas de normes non euclidiennes est bien plus complexe, mais aussi plus intéressant d'un point de vue académique. Il a d'ailleurs donné lieu à une publication récente [dGJ18], qui montre que la borne de complexité peut être rendue invariante par transformation affine des coordonnées en choisissant bien les métriques. C'est un résultat similaire à ce qui se passe pour une méthode de Newton, au détail important près que ce sont les itérées et non la borne qui sont invariantes, pour la méthode d'ordre 2.

On restreint ici l'étude à la minimisation de fonctions f qui sont les jauges d'un convexe compact $Y \in F^*$: $f(y) = \inf_{t \geq 0, y \in tY} t$. Cette restriction couvre toutes les normes par exemples. Elle est importante car à ma connaissance, les descente de gradient accélérées avec des opérateurs proximaux reposant sur des divergences de Bregman ne sont pas encore étendues à la minimisation de fonctions arbitraires.

Dans ce cas, notre problème dual se simplifie sous la forme :

$$\inf_{y \in Y} h(y) \stackrel{\text{def.}}{=} g^*(-M^*y), \quad (2.7)$$

et on peut appliquer des descentes de gradient projeté accéléré pour le résoudre. On choisit ici des normes $\|\cdot\|_E$ et $\|\cdot\|_F$ arbitraires ainsi qu'une fonction $d : F \rightarrow \mathbb{R} \cup \{+\infty\}$ de module de convexité $\sigma > 0$. Une telle fonction est appelée *fonction prox* par Nesterov dans [Nes05], mais ce n'est pas une appellation courante à ma connaissance. On suppose sans perte de généralité que $d(y) \geq \frac{1}{2}\|y - y_0\|_F^2$ avec $d(y_0) = 0$. La valeur de L est celle donnée dans le lemme 1 tandis que les suites (α_k) et (τ_k) peuvent être déterminées par récurrence (voir [Nes05 ; dGJ18]).

Algorithm 2 Descente proximale accélérée duale (cas euclidien)

- ① Choisir $y_0 \in \text{dom}(d)$ et poser $\hat{y}_0 = y_0$.
 - ② Calculer $\tau = \frac{\sigma_g}{\|M\|_{2 \rightarrow 2}^2}$.
 - ③ **for** $k = 0 \dots K$ **do**
 - ④ $y_k = \arg \min_{y \in Y} \langle \nabla h(\hat{y}_k), y - \hat{y}_k \rangle + \frac{L}{2} \|y - \hat{y}_k\|_Y^2$.
 - ⑤ $\bar{y}_k = \arg \min_{y \in Y} \frac{L}{\sigma} d(y) + \sum_{i=0}^k \alpha_i [h(\hat{y}_i) + \langle \nabla h(\hat{y}_i), y - \hat{y}_i \rangle]$
 - ⑥ $\hat{y}_{k+1} = \tau_t \bar{y}_k + (1 - \tau_t) y_k$.
 - ⑦ **end for**
 - ⑧ $x_k = \nabla g^*(-M^*y_k)$.
-

Théorème 3. Soit x^* l'unique minimiseur du problème primal (2.2) et y^* un minimiseur quelconque du problème dual (2.7). Alors l'algorithme 2 assure :

$$\|x_k - x^*\|_E^2 \leq \frac{4\|M^*\|_{F \rightarrow E^*}^2 d(y^*)}{\sigma^2 \sigma_g k^2}. \quad (2.8)$$

L'intérêt potentiel de cette nouvelle méthode par rapport à celle reposant sur les normes euclidiennes est qu'on peut espérer améliorer la borne (2.8) en la minimisant par rapport aux normes sur E et F ainsi que sur la fonction prox d .

À ma connaissance un des seuls cas où on peut montrer un vrai intérêt théorique de cette méthode est la minimisation sur des espaces de probabilités discrètes. Le diamètre du simplexe (qui apparaît dans $d(y^*)$) diminue en utilisant

l'entropie pour définir d plutôt qu'une norme euclidienne et la norme matricielle $\|M^*\|_{F \rightarrow E^*}$ peut elle aussi s'améliorer. Nous le verrons dans un exemple ci-après.

2.2.2 Quelques exemples

Pour conclure ce paragraphe, je montre quelques forces et limites des approches duales ci-dessus.

Le modèle de Rudin-Osher-Fatemi

Le modèle ROF a fait long feu dans la communauté du traitement d'images. Il consiste à minimiser la variation totale discrète et un terme d'attache aux données quadratique :

$$\min_{x \in \mathbb{R}^n} \|\nabla x\|_1 + \frac{1}{2} \|x - x_0\|_2^2, \quad (2.9)$$

où $x_0 \in \mathbb{R}^n$ est typiquement une image bruitée et ∇ est un opérateur différentiel discret. En appliquant le théorème 2, on obtient le résultat suivant.

Proposition 1. *L'algorithme 1 génère une suite (x_k) qui satisfait :*

$$\|x_k - x^*\|_2^2 \leq \frac{2\|\nabla\|_{2 \rightarrow 2}^2}{k^2} 2^n. \quad (2.10)$$

A première vue, la borne semble terrible car elle contient la dimension n qui reflète le diamètre d'une boule unité ℓ^∞ . Cependant si on réfléchit en terme d'*erreur relative*, et qu'on se rend compte que $\|x^* - x_0\|_2^2 = O(n)$ (dans une tâche de débruitage par exemple, la norme du bruit est en $O(n)$), on se rend compte que la méthode permet d'obtenir une précision relative satisfaisante en quelques itérations. C'est ce qu'on observe en pratique et qui me semble expliquer le succès de cette méthode.

👁 Régressions isotones

Voyons maintenant un exemple moins satisfaisant. Soit A la matrice d'incidence d'un graphe dirigé acyclique. Une telle matrice définit un ordre partiel entre les nœuds d'un graphe (V, E) . Une fonction $x : V \rightarrow \mathbb{R}$ respecte l'ordre partiel si $Ax \geq 0$. Par exemple, pour un graphe linéaire, un vecteur x est admissible s'il est monotone croissant.

Les régressions isotones ont de nombreuses applications en statistiques [Bar72]. Je présenterai une application originale à l'imagerie plus tard dans ce manuscrit. Imaginons par exemple qu'on souhaite débruiter un vecteur dont on sait qu'il doit être croissant. On peut alors être mené à résoudre un problème de la forme suivante :

$$\min_{x \in \mathbb{R}^n, Ax \geq 0} \frac{1}{2} \|x - x_0\|_2^2. \quad (2.11)$$

L'algorithme de descente sur le problème dual a les garanties suivantes.

Proposition 2. *L'algorithme 1 génère une suite (x_k) qui satisfait :*

$$\|x_k - x^*\|_2^2 \leq \frac{2\|A\|_{2 \rightarrow 2}^2 \|y_0 - y^*\|_2^2}{k^2}. \quad (2.12)$$

La grande difficulté ici est de borner le terme $\|y_0 - y^*\|_2^2$, car la variable duale y ne vit plus sur un espace borné. Dans un papier récent [12], nous montrons que la notion de *valeur propre conique minimale* permet de borner la variable duale, ce qui mène finalement à un taux en $O(n^2/k^2)$. Ceci semble mauvais théoriquement et c'est effectivement le cas en pratique.

Je pense que le fait que les variables duales ne soient pas bornées est une difficulté assez profonde. Le travail sur des ensembles compacts de faible diamètre simplifie considérablement la donne. A ma connaissance, les meilleures bornes de complexité pour résoudre le problème (2.11) reposent sur des algorithmes de points intérieurs avec des techniques de résolution de systèmes linéaires très astucieuses et pour le moment instables [ST04; KRS15].

En perspective de cette partie, j'aimerais creuser un peu plus le rôle des valeurs propres minimales coniques en optimisation de premier ordre. Elles sont apparues récemment dans le domaine de l'échantillonnage compressé [Ame+14] et je pense qu'elles peuvent se révéler très utiles pour des analyses de complexité. Un axe de recherche est simplement leur calcul qui est en général ardu. Pour une application linéaire A , il faut évaluer des quantités de type $\min_{x \in K, \|x\|_2=1} \|Ax\|_2$, où K est un cône convexe. L'autre axe est de comprendre de façon assez générale à quoi ces objets peuvent être utiles, avec une forte suspicion qu'elles sont un outil de choix pour obtenir des garanties de recouvrement exact de solution avec des algorithmes de minimisation de premier ordre efficaces.

Optimisation de probabilités

Pour finir, je présente très brièvement un exemple d'application des algorithmes accélérés duaux à l'optimisation de densité de probabilité discrète pour illustrer l'intérêt de choisir une métrique différente. Lors de la thèse de Claire Boyer, nous avons été confronté à la question suivante.

Question 1. *Supposons qu'on puisse tirer aléatoirement des indicatrices d'ensembles de \mathbb{R}^d (e.g. des lignes). En fixant une loi de probabilité sur l'ensemble des objets admissibles, on peut définir un processus aléatoire et sa mesure d'occupation. Comment choisir la loi de probabilité de telle manière à ce que la mesure d'occupation s'approche d'une mesure cible ?*

La figure 2.1 illustre ce problème. Dans la figure 2.1a, on voit une densité de probabilité cible sur $[0, 1]^2$. Sur la figure 2.1b, on voit la mesure d'occupation empirique d'un processus que nous avons optimisé. Sur la figure 2.1c on voit un tirage i.i.d. de lignes suivant la densité optimisée.

En deux mots, on est ramené à la résolution d'un problème convexe de la forme suivante :

$$\inf_{\pi \in \Delta_m} \|M\pi - p\|_1 + \alpha \mathcal{E}(\pi),$$

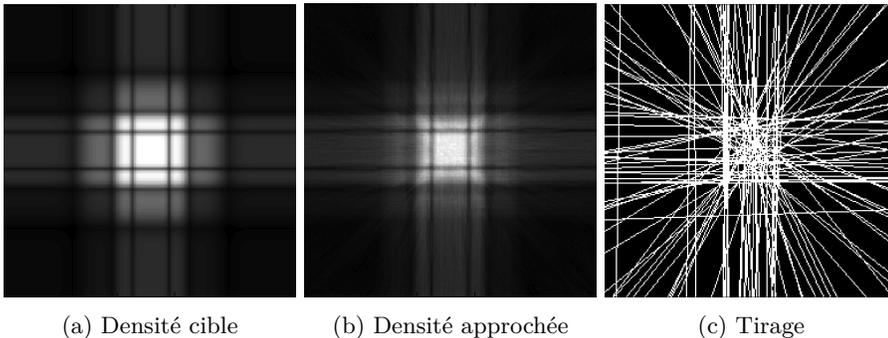


FIGURE 2.1 – Exemple de distribution de tirage apprise. On peut approcher une densité cible (à gauche) en tirant au hasard un ensemble de lignes suivant une densité apprise. La densité approchée est montrée sur la figure du centre, tandis qu’on montre le résultat d’un nombre fini de tirages à droite.

où $p \in \mathbb{R}^n$ est une densité cible, $M : \mathbb{R}^m \rightarrow \mathbb{R}^n$ est une application (pas toujours linéaire) qui permet de définir la mesure d’occupation du processus à partir de la densité de tirage π , et \mathcal{E} est la neg-entropie définie sur le simplexe canonique Δ_m . On a montré dans [28], que le choix optimal pour les normes $\|\cdot\|_E$ et $\|\cdot\|_F$ dépendaient des formes admissibles et vérifié expérimentalement que le choix de la norme euclidienne n’était pas toujours le plus efficace.

2.3 Programmation dynamique et régressions isotones

Récemment, j’ai découvert le monde de la programmation dynamique qui me semble appartenir un peu plus à l’informatique théorique qu’aux mathématiques appliquées. En suivant les travaux de Kolmogorov-Pock-Rolinek [KPR16], nous avons proposé une méthode efficace et unifiée pour résoudre des problèmes de régression isotone sur des graphes linéaires, des arbres ou sur des polyarbres.

Si A est la matrice d’incidence d’un graphe dirigé acyclique, l’idée est de résoudre des problèmes du type :

$$\min_{x \in \mathbb{R}^n, Ax \geq 0} \|x - z\|_p, \quad (2.13)$$

pour $p \in \{1, 2\}$. On cherche le vecteur x le plus proche d’un vecteur z donné qui respecte une relation d’ordre partielle entre ses coefficients.

Pour $p = 2$, nous avons déjà vu dans le paragraphe 2.2.2 que les méthodes de premier ordre peuvent fonctionner très lentement. L’idée de la programmation dynamique est radicalement différente : elle permet d’obtenir une solution exacte en résolvant une suite de problèmes convexe 1D. En notant (V, E) les nœuds (vertices) et arêtes (edges) d’un arbre, le problème (2.13) peut être réécrit sous

la forme :

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^n f_i(x_i) + \sum_{(i,j) \in E} f_{i,j}(x_i - x_j), \quad (2.14)$$

où $f_{i,j}(y) = \begin{cases} 0 & \text{si } y \geq 0 \\ +\infty & \text{sinon.} \end{cases}$, $f_i(x_i) = |x_i - z_i|$ pour $p = 1$ et $f_i(x_i) = (x_i - z_i)^2$ pour $p = 2$.

Notons $M_i(t)$ la valeur optimale de l'énergie restreinte au sous arbre enraciné au nœud i , en fixant la valeur t au nœud i . Cette énergie peut être calculée directement sur les feuilles de l'arbre car elle contient uniquement les termes unaires f_i . Elle peut ensuite être propagée grâce à la formule de récurrence :

$$M_i(t) = f_i(t) + \sum_{j \in \text{child}(i)} (M_j \square f_{i,j})(t), \quad (2.15)$$

où le symbole \square représente une inf-convolution définie par :

$$(M_j \square f_{i,j})(t) = \inf_{s \in \mathbb{R}} M_j(s) + f_{i,j}(t - s). \quad (2.16)$$

L'algorithme de programmation dynamique évalue récursivement les fonctions M_i sur chaque nœud en partant des feuilles jusqu'à la racine. La solution optimale est obtenue par rétro-propagation : la valeur optimale x_r^* au nœud r est donnée par $x_r^* = \arg \min_{t \in \mathbb{R}} M_r(t)$. Le reste des valeurs peut être calculé par la récurrence :

$$x_j^* = \arg \min_{t \in \mathbb{R}} M_j(t) + f_{i,j}(t - x_i), \quad (2.17)$$

pour tous les fils $j \in \text{child}(i)$.

Mettre en pratique ce principe requiert beaucoup d'astuces informatiques que je ne décrirais pas ici (il faut stocker des fonctions continues, évaluer rapidement des inf-convolution, garder en mémoire seulement $O(n)$ coefficients). Du point de vue de l'optimisation, ce travail nous a permis d'améliorer l'état-de-l'art en :

- proposant un algorithme unifié et optimal pour résoudre un ensemble des problèmes attaqués avec des approches très différentes jusqu'à présent. Le lecteur intéressé peut regarder le petit résumé proposé par Quentin Stout par exemple [Sto14].
- On peut aussi résoudre des problèmes jusqu'à présent non traités (e.g. régression sur des polyarbres, i.e. des arbres avec des directions arbitraires entre nœuds adjacents).
- On améliore légèrement les meilleures complexités existantes pour la résolution de problèmes avec attache ℓ^1 sur des graphes linéaires en passant de $O(n \log n)$ à $O(n \log \log n)$.

Les codes sont maintenant disponibles sous GitHub https://github.com/pierre-weiss/contrast_invariant_snr. A ma connaissance un tel outil n'existe pas et est réellement difficile à reproduire (il nécessite des structures de données appelées double ended priority queue qui ne sont pas disponibles dans les librairies standards ou sur GitHub).

2.4 Résolution de problèmes en dimension infinie

Un axe de recherche qui commence à se dégager dernièrement, grâce à une rencontre un peu fortuite avec Axel Flinth [15] (alors doctorant à Berlin) et Michael Unser [UFW17], est celui de la résolution de problèmes inverses en dimension infinie. Suite à ce travail, j’ai pu continuer à creuser cette question en me rapprochant notamment de Vincent Duval et de Frédéric de Gournay qui ont eu un impact important sur la rédaction de l’article [5].

2.4.1 Théorèmes de représentation

Je commence cette présentation par des théorèmes de représentation [5].

Le cadre

Soit E un espace vectoriel de Hausdorff localement convexe. Soit $m \in \mathbb{N}$ et $\Phi : E \rightarrow \mathbb{R}^m$ un opérateur linéaire appelé *opérateur de mesure*. Soit $u \in E$ un signal. On suppose qu’on dispose d’un ensemble de mesures de la forme $y = P(\Phi u)$, où $P : \mathbb{R}^m \rightarrow \mathbb{R}^m$ est une perturbation, telle que de la quantification (1-bit compressed sensing), un module (recouvrement de phase), du bruit additif gaussien,... On souhaite estimer u à partir de y . Une approche qui est devenue très populaire¹ est la suivante :

$$\inf_{u \in E} f(\Phi u) + J_C(u), \quad (2.18)$$

où J_C est la *jauge* d’un ensemble convexe $C \subset E$:

$$J_C(u) = \inf\{\lambda \text{ t.q. } u \in \lambda C, \lambda \geq 0\}, \quad (2.19)$$

et f est une fonction arbitraire convexe ou non-convexe appelée *attache aux données*. Cette fonction doit dépendre du choix des données y et de la perturbation P . Des papiers influents tels que [Cha+12] suggèrent que J_C favorise des solutions u “sparse”, au sens où elles s’expriment comme une combinaison linéaire d’un petit nombre d’atomes élémentaires. Nous avons souhaité expliciter ce principe.

Nos résultats

Sous réserve d’existence, notre résultat principal stipule qu’il existe systématiquement des solutions du problème (2.18) qui sont une combinaison conique d’au plus m points extrémaux ou éléments des rayons extrémaux de C .

1. bien qu’elle perde beaucoup de popularité dernièrement au profit de l’apprentissage profond

On note S^* l'ensemble des solutions de (2.18). On suppose qu'il est non vide. On note $t^* = J_C(u^*)$ avec $u^* \in S^*$ l'infimum dans (2.18) et on définit

$$\delta = \begin{cases} 1 & \text{si } t^* = \inf_{u \in E} J_C(u), \\ 0 & \text{si } t^* > \inf_{u \in E} J_C(u). \end{cases} \quad (2.20)$$

On rappelle que tout convexe C peut être décomposé sous la forme

$$C = C_B + C_K$$

où C_K est l'espace *linéal* de C (i.e. l'ensemble des lignes contenues dans C , qui constitue un sous-espace vectoriel) et $C_B = C \cap B$ où B est un sous-espace complémentaire à C_K . Ainsi, l'espace C_B ne contient aucune ligne.

De la même façon, on peut décomposer S^* sous la forme $S^* = S_K^* + S_B^*$ et on note d la dimension de $\Phi(C_K)$. Notre résultat est valide sous les hypothèses suivantes.

Hypothèse 1.

- C est non vide, convexe et fermé.
- C_B est fermé et localement compact.
- S^* est non vide.

Si f est une fonction quelconque, on obtient le résultat suivant.

Théorème 4. *Sous les hypothèses 1, il existe au moins une solution $u^* \in S^*$ de la forme :*

$$u^* = u_K^* + \sum_{k=1}^r \alpha_k \psi_k, \quad (2.21)$$

avec

- $u_K^* \in C_K$, et $\alpha_k \geq 0$ pour tout k .
- $r \leq m - d + \delta$ et $\psi_k \in \text{ext}(C_B)$ pour tout k ,
- ou $r \leq m - 1 - d + \delta$ et $\psi_k \in \text{rext}(C_B) \cup \text{ext}(C_B)$ pour tout k .

Si f est convexe s.c.i., on obtient un résultat plus fort :

Théorème 5. *Si f est convexe, fermée et sous les hypothèses 1, alors $S_K^* = C_K \cap \ker(\Phi)$. De plus, pour tout $p \in S_B^*$, soit j la dimension de la face de S_B^* contenant p . Alors p peut être écrite comme une combinaison conique de :*

- $m + j - d + \delta$ points extrémaux de C_B ,
- ou $m + j - 1 - d + \delta$ éléments de l'ensemble des des point extrémaux ou des rayons extrémaux de C_B .

Ce résultat décrit toute la structure de l'ensemble des solutions : en prenant $j = 0$ et $j = 1$, on décrit en effet les points extrémaux et les rayons extrémaux

de l'ensemble des solutions. Cette information est suffisante pour reconstruire l'ensemble [Kle57].

Les résultats obtenus dans [5] sont en réalité un peu plus fins, mais je préfère ici cette présentation élémentaire.

Quelques exemples

Norme ℓ^1 et variation totale La norme ℓ^1 définie sur $E = \mathbb{R}^n$ est la jauge J_C de la boule ℓ^1 unité $C = \{u \in \mathbb{R}^n, \|u\|_1 \leq 1\}$. En utilisant cet a priori, notre résultat stipule que certaines solutions de (2.18) sont m -sparse. C'est un des piliers de l'échantillonnage compressé [CRT06].

De façon similaire, soit $E = \mathcal{M}$ est l'espace des mesures de Radon sur un domaine Ω of \mathbb{R}^d . La variation totale (ou masse) d'une mesure $u \in \mathcal{M}$ est notée $J_C(u) = \|u\|_{\mathcal{M}}$, où $C = \{u \in \mathcal{M}, \|u\|_{\mathcal{M}} \leq 1\}$. On peut montrer que $\text{ext}(C) = \{\pm\delta_x, x \in \Omega\}$ et que $\text{rext}(C) = \emptyset$. Dans ce cadre, notre résultat indique que certaines solutions sont une combinaison d'au plus m masses de Dirac, ce qui est un des ingrédients principaux de [CF14; DP15].

Variation totale généralisée Soit $D'(\Omega)$ l'espace des distributions et $E \subset D'(\Omega)$. Soit $L : E \rightarrow \mathcal{M}$ un opérateur linéaire surjectif et continu et $C \stackrel{\text{def.}}{=} \{u \in E, \|Lu\|_{\mathcal{M}} \leq 1\}$. On peut alors définir la fonction de régularisation $J_C(u) = \|Lu\|_{\mathcal{M}}$. Le théorème de Fisher-Jerome [FJ75], qui a été amélioré récemment dans [UFW17; 15] peut être obtenu à partir de nos théorèmes, en remarquant que $C_K = \ker(L)$, que $\text{ext}(C_B) = \{\pm L^+\delta_x, x \in \Omega\}$ et que $\text{rext}(C) = \emptyset$. Ce théorème sur lequel on reviendra un peu plus tard indique que les solutions de (2.18) sont une combinaison linéaire d'au plus m splines.

Par exemple, si on choisit $\Omega = \mathbb{R}^2$ et si L est le bi-laplacien $L = \Delta\Delta$, alors $L^+\delta_{x_i} = \psi(\cdot - x_i) \stackrel{\text{def.}}{=} \|\cdot - x_i\|_2^2 \log(\|\cdot - x_i\|_2)$ est une spline plaque mince et on voit que certaines solutions peuvent s'écrire sous la forme :

$$u^*(x) = \sum_{i=1}^m \alpha_i \psi(x - x_i) + p(x), \quad (2.22)$$

où p est un polynôme de degré inférieur ou égal à 3 et $\alpha_i \in \mathbb{R}$. Ce qui est remarquable ici, est que ces objets sont bien connus dans le cadre des espaces de Hilbert à noyau reproduisant (RKHS). Ils apparaissent par exemple en minimisant des énergies de type

$$\frac{1}{2} \|\Delta u\|_{L^2(\mathbb{R}^2)}^2 + \frac{1}{2} \sum_{i=1}^m \|u(x_i) - y_i\|_2^2, \quad (2.23)$$

ce qui peut servir à faire de l'approximation de fonctions échantillonnées ponctuellement en un ensemble de points (x_i) . Nous obtenons un résultat similaire en remplaçant le terme $\frac{1}{2} \|\Delta u\|_{L^2(\mathbb{R}^2)}^2$ par $\|\Delta\Delta u\|_{\mathcal{M}}$. Les différences principales entre les deux approches sont les suivantes (voir aussi [UFW17]) :

- Dans le cadre des RKHS, les nœuds x_i sont fixés et égaux aux points d'échantillonnage. Dans l'approche par variation totale, les nœuds sont libres. La variation totale permet donc d'obtenir des solutions *adaptatives*, ce qui est une distinction habituelle entre méthodes linéaires [DL93] et méthodes non linéaires [DeV98].
- Dans le cadre des RKHS, le polynôme additionnel correspond au noyau du laplacien et est de degré 1, alors que pour la variation totale, on obtient un polynôme de degré 3 (on peut se ramener à des polynômes de degré 1 en définissant l'espace de minimisation avec précaution).
- La minimisation dans le cadre des RKHS revient simplement à la résolution de systèmes linéaires de dimension finie, tandis que la minimisation de la variation totale généralisée est un problème essentiellement ouvert sur lequel je compte travailler dans les années à venir.

Orthant positif Si $E = \mathbb{R}^n$ et $J_C(u) = \chi_C(u)$, l'indicatrice de l'orthant positif C , alors $\text{ext}(E) = \{0\}$ et $\text{rext}(C) = \{\alpha e_i, \alpha \geq 0, 1 \leq i \leq n\}$. Nos théorèmes indiquent alors qu'au moins une solution de (2.18) est m -sparse. Ce résultat est central pour comprendre l'intérêt de résoudre des problèmes sous contraintes de positivité. C'est aussi un des ingrédients implicites de [DT05].

De la même façon, si $E = \mathbb{R}^{m \times m}$ et que C est le cône des matrices semi-définies positives, notre théorème indique que certaines solutions sont de rang m .²

👁 **Variation totale du gradient** On s'intéresse ici à la variation totale du gradient [ROF92]. Etant donné l'ensemble $E \stackrel{\text{def.}}{=} L^{d/(d-1)}(\mathbb{R}^d)$, on définit J_C comme la variation totale du gradient, *i.e.*

$$\begin{aligned} J_C(u) &\stackrel{\text{def.}}{=} \|Du\|_{\mathcal{M}^d} \\ &= \sup \left(\int u \text{div}(\phi) dx, \phi \in C_c^1(\mathbb{R}^d)^d, \sup_{x \in \mathbb{R}^d} \|\phi(x)\|_2 \leq 1 \right), \end{aligned}$$

où D est le gradient au sens des distributions et $\|\cdot\|_{\mathcal{M}^d}$ est une variation totale vectorielle. Si $F \subset \mathbb{R}^d$ a une mesure finie, on définit son périmètre par $P(F) \stackrel{\text{def.}}{=} J_C(\mathbb{1}_F)$.

Le résultat suivant est dû à Fleming [Fle57].

Proposition 3 ([Fle57]). *Les points extrémaux de*

$$C = \{u \in L^{d/(d-1)}(\mathbb{R}^d), \|Du\|_{\mathcal{M}^d} \leq 1\}, \quad (2.24)$$

² \triangleleft Ici, je préfère indiquer au lecteur assidu que le résultat annoncé est nettement sous-optimal. Nos théorèmes les plus forts indiquent que les points extrémaux de S_B^* vivent sur une face de dimension au plus $m - 1$ de C_B . Ces $m - 1$ -faces sont elles-mêmes contenues dans l'enveloppe convexe de points extrémaux et rayons extrémaux de C_B , ce qui justifie le théorème 5. Cependant, il se peut que la structure des k -faces soit plus favorables. En particulier, les m -faces du cône SDP contiennent des matrices de rang $O(\sqrt{m})$ [Dat10, Sec. 2.8.2.2].

quotienté par l'espace vectoriel des constantes, sont les fonctions $u = \pm \mathbb{1}_F / P(F)$, où F est un ensemble simple et $P(F) < +\infty$.

De façon informelle, les ensembles simples sont des ensembles simplement connectés (*i.e.* une composante connexe sans trou).

En combinant ce résultat avec nos théorèmes, on peut conclure que les solutions de (2.18) sont la somme de m fonctions indicatrices. Ceci explique des phénomènes tels que le “staircasing” des solutions.

👁 La variation totale du gradient peut être définie par la formule de la co-aire : on intègre le périmètre des ensembles de niveau le long de tous les niveaux. Nous travaillons actuellement sur la généralisation de notre résultat au cas où on remplacerait la notion de périmètre par des fonctions sous-modulaires quelconques. On souhaite aussi décrire précisément les k -faces de la boule unité associée à la variation totale du gradient.

2.4.2 👁 Nouvelles méthodes de discrétisation

Imaginons que l'on souhaite minimiser une énergie du type

$$\inf_{u \in \mathcal{M}(\Omega)} \|u\|_{\mathcal{M}} + f(\Phi u), \quad (2.25)$$

où $\Phi : \mathcal{M} \rightarrow \mathbb{R}^m$ est une application linéaire continue et Ω est un domaine compact de \mathbb{R}^d . On suppose que $(\Phi u)_i = \langle \phi_i, u \rangle$, où ϕ_i est une application continue. Le problème (2.25) est un problème de programmation semi-infinie (son dual consiste à minimiser une fonction de \mathbb{R}^m soumise à un nombre infini de contraintes). Ce problème a reçu beaucoup d'attention dernièrement dans le domaine du traitement du signal [BP13; CF14; DP15].

L'approche de très loin la plus répandue pour résoudre (2.25) consiste à discrétiser le problème de la façon suivante. On fixe un ensemble de points $X = (x_i)_{1 \leq i \leq n} \in \Omega^n$ et on remplace $\mathcal{M}(\Omega)$ par l'espace $\mathcal{M}(X) = \{u = \sum_{i=1}^n \alpha_i \delta_{x_i}\}$. On obtient alors le problème suivant en dimension finie :

$$\inf_{u \in \mathcal{M}(X)} \|u\|_{\mathcal{M}} + f(\Phi u) = \inf_{\alpha \in \mathbb{R}^n} \|\alpha\|_1 + f\left(\sum_{i=1}^m \alpha_i \Phi(x_i)\right), \quad (2.26)$$

où $(\Phi u)_j = \langle \phi_j, u \rangle = \sum_{i=1}^n \alpha_i \phi_j(x_i)$. Si f est convexe fermée, il peut être résolu par les procédures standards de l'optimisation convexe (points intérieurs ou autres méthodes de premier ordre). Dans [15], nous avons montré le résultat suivant.

Théorème 6. *Une solution $u^* = \sum_{i=1}^n \alpha_i^* \delta_{x_i}$ obtenue en résolvant (2.26) est aussi solution du problème en dimension infinie suivant :*

$$\inf_{u \in \mathcal{M}(\Omega)} \|u\|_{\mathcal{M}} + f(\Phi_n u), \quad (2.27)$$

où Φ_n est une approximation de Φ linéaire par morceau, sur la triangulation de Delaunay associée à l'ensemble des points (x_i) .

Ce résultat donne du recul sur la procédure de discrétisation usuelle. Nous avons aussi montré qu’il pouvait s’appliquer plus généralement à n’importe quel problème de minimisation d’énergie de type $\|Lu\|_{\mathcal{M}(\Omega)}$ où L est un opérateur linéaire continu surjectif. Enfin, nous avons justifié l’utilisation de procédures de “parcimonification” des solutions discrètes, en regroupant des masses de Dirac adjacentes dans la triangulation.

De mon point de vue, l’intérêt principal de ce résultat est à venir. Il permet en effet d’envisager (et de justifier) des procédures d’optimisation multi-résolution : on commence par résoudre (2.26) avec un maillage grossier, et on va affiner le maillage uniquement sur certaines parties de la triangulation en fonction de certains critères associés au problème dual. Je m’intéresse aussi actuellement à des problématiques d’optimisation de mesures discrètes sous forme continue : on optimise la position et l’amplitude de masses de Dirac suivant un flot continu. Cette approche semble très prometteuse numériquement et théoriquement [CB18]. J’y reviendrais un peu plus tard.

Le Graal de cette piste de recherche serait - à long terme - d’obtenir des solveurs de problèmes inverses dont la complexité dépend de la parcimonie de la solution (nombre de spikes pour la variation totale, longueur des sauts pour la variation totale du gradient) et pas de la finesse d’un maillage paramétré par n .

Chapitre 3

Échantillonnage compressé

3.1 Introduction

Comme on l’a vu dans le chapitre précédent, minimiser des jauges dont le convexe sous-jacent possède des singularités (les atomes) favorise les solutions parcimonieuses au sens où elles s’expriment comme une combinaison conique d’un petit nombre de points extrémaux. En pratique, cette idée a été explorée dès les années 75 [FJ75] et a commencé à devenir très populaire avec l’introduction de la variation totale du gradient par exemple [ROF92]. Au début de ma thèse en 2007, ces résultats devaient essentiellement leur popularité à leurs bonnes performances empiriques. David Donoho - un des pionniers de ce domaine - a obtenu les premières preuves que sous des conditions restrictives, ces a priori permettent de reconstruire *exactement* les signaux mesurés. Ces résultats souffraient cependant de ce qu’on appelle le “goulot quadratique” : un signal s -parcimonieux peut être reconstruit à partir de s^2 mesures.

3.1.1 Un premier théorème

L’échantillonnage compressé tel qu’on le connaît aujourd’hui est né au départ de ma thèse [CRT06 ; Don06]. En utilisant des outils probabilistes telles que les inégalités de concentration, le goulot quadratique a pu être cassé. Je donne ici un théorème sous sa forme la plus élémentaire [FR13].

Théorème 7. *Soit $s \in \mathbb{N}$ et soit $A \in \mathbb{R}^{m \times n}$ une matrice dont les coefficients sont tirés suivant des lois normales indépendantes. Soit $s < n$ un entier.*

Il existe deux constantes C_1 et C_2 telles que si $m > C_1 s \log(en/s) + C_2 \log(2/\delta)$, alors avec probabilité supérieure à $1 - \delta$, le problème suivant :

$$\inf_{x \in \mathbb{R}^n, Ax = Ax_0} \|x\|_1 \tag{3.1}$$

admet pour solution unique le vecteur x_0 pour tout x_0 s -parcimonieux.

Ce théorème indique qu'à un facteur logarithmique près, le nombre de mesures nécessaires pour reconstruire un signal dépend de sa dimension intrinsèque s (sa parcimonie) et non pas de la dimension ambiante n . Le prix à payer est l'utilisation de méthodes de restauration non linéaires. Il est important de noter que ce théorème permet de traiter le cas de mesures entachées de bruit : la reconstruction n'est plus exacte, mais l'erreur de reconstruction est de l'ordre du niveau de bruit.

3.1.2 Ses limites

Le théorème que je viens de donner est typique d'un grand nombre de résultats du domaine. Sous des hypothèses très restrictives, on est capable de recouvrer uniformément tous les signaux parcimonieux. C'est un résultat puissant, mais les conditions sont si fortes que je ne connais en fait aucune application réelle qui peut en bénéficier. Ma contribution dans ce domaine a été d'essayer de relâcher les hypothèses pour pouvoir attaquer des problèmes complexes d'échantillonnage, notamment dans le domaine de l'imagerie par résonance magnétique (IRM).

Je décrirai l'IRM plus en détail par la suite et je donne juste un avant goût ci-dessous. En première approximation, une image d'IRM peut être vue comme un signal $u : [0, 1]^d \rightarrow \mathbb{C}$. On peut mesurer ce signal uniquement le long de courbes lisses dans le k -space, qui correspond à la transformée de u . Formellement, soit $\lambda : [0, T] \rightarrow \mathbb{R}^d$ une courbe d'échantillonnage. L'IRM permet alors de mesurer

$$y_i = \hat{u}(\lambda(i\Delta t)) + b_i, \quad (3.2)$$

où Δt est un pas d'échantillonnage et b_i est un bruit additif. On peut alors se poser la question suivante.

Question 2 (Le problème d'échantillonnage en IRM). *Étant données les contraintes de la machine et la structure des images de l'IRM, comment faire un design optimal de trajectoires λ ? Plus précisément, comment maximiser le SNR de l'image reconstruite pendant un temps d'échantillonnage T fixé ?*

Ce problème me semble extrêmement complexe : il faut optimiser conjointement une courbe d'échantillonnage λ ainsi qu'un algorithme de reconstruction pour une classe de signaux (uniformément ou en moyenne par exemple). Je pense qu'il y a encore beaucoup de questions ouvertes associées. Malheureusement, l'avènement du deep-learning rend ces questions moins populaires dernièrement au profit de rendement à court terme. Je soulève quelques difficultés techniques ci-dessous.

1. Une image d'IRM n'est pas parcimonieuse dans le domaine des pixels, mais elle peut être approchée par une image parcimonieuse dans un domaine transformé, tel que le domaine des ondelettes. La parcimonie dans ce domaine est cependant assez faible : on peut difficilement espérer reconstruire une image d'IRM décente avec moins de 10% de ses coefficients en ondelette par exemple. Ainsi, la valeur $C_1 \log(en/s)$ dans (7) ne

doit pas dépasser 10... Malheureusement, ces constantes sont fréquemment mal maîtrisées lorsqu'on souhaite reconstruire *uniformément* toute une classe de signaux avec la propriété d'isométrie restreinte¹. On verra que les théorèmes qu'on a établi assurent la reconstruction d'un signal unique, ce qui permet un bien meilleur contrôle des constantes, qui reste malheureusement trop pessimiste.

2. Le théorème 7 est annoncé avec une matrice aléatoire gaussienne. Dans notre cas, il y a deux difficultés importantes : i) on sous-échantillonne ici une transformée déterministe, la transformée de Fourier et ii) les échantillons ne peuvent pas être choisis de manière arbitraire, mais doivent satisfaire la contrainte d'appartenir à une courbe régulière. Le point i) a été sérieusement abordé par des chercheurs tels que Candès, Romberg, Tao [CRT06], Candès, Plan [CP11a], Rauhut, Foucart [FR13], ou encore Adcock, Hansen, Poon [Adc+17]. Notre contribution se situe essentiellement sur la résolution du point ii).

3.2 Contributions théoriques

Dans cette partie, on part d'une matrice orthogonale $A_0 = \begin{pmatrix} a_1^* \\ \vdots \\ a_n^* \end{pmatrix} \in \mathbb{C}^{n \times n}$.

On souhaite retrouver un signal $x \in \mathbb{R}^n$ *s*-parcimonieux (contenant au plus n coefficients non nuls). Par exemple en radio-interférométrie, on peut supposer que A_0 est une transformée de Fourier discrète. En IRM, A_0 serait le produit d'une transformée de Fourier discrète et d'une transformée en ondelettes inverse, puisque les images sont compressibles dans le domaine des ondelettes et qu'elles sont mesurées dans le domaine de Fourier.

Pour mesurer x , on suppose qu'on peut tirer un ensemble de m lignes $(a_{i_1}, \dots, a_{i_m})$ de la matrice A_0 et construire ainsi la matrice $A = \begin{pmatrix} a_{i_1}^* \\ \vdots \\ a_{i_m}^* \end{pmatrix} \in$

$\mathbb{C}^{m \times n}$. Pour obtenir des résultats théoriques, on est particulièrement intéressés par le cas où les indices (i_1, \dots, i_m) sont tirés au hasard. La nouveauté principale de nos travaux est cependant que ces échantillons ne sont pas i.i.d. Nous utiliserons la définition suivante :

Définition 2. Soit $p \in \mathbb{R}^n$ une distribution de probabilité discrète. La cohérence de la matrice A_0 relative à la distribution p est définie par

$$K(A_0, p) = \max_{i \in \{1, \dots, n\}} \frac{\|a_i\|_\infty}{p_i}. \quad (3.3)$$

Nous utiliserons souvent des distributions de tirage optimales.

1. On voit souvent des constantes dépassant allègrement les 10^5 et des facteurs polylogarithmiques.

Lemme 8. *La distribution p^* qui minimise la cohérence $K(A_0, p)$ est définie par*

$$p_i^* = \frac{\|a_i\|_\infty}{\sum_{i=1}^n \|a_i\|_\infty}. \quad (3.4)$$

Nous montrons une distribution optimale sur la figure 3.1a pour une matrice A_0 définie comme le produit d'une transformée de Fourier et d'une transformée en ondelettes Symlet à 6 moments nuls.

3.2.1 L'échantillonnage de l'homme saoul

Notre première contribution est partie d'une idée assez naïve. Comme les échantillons doivent appartenir à une courbe continue, nous avons pensé définir des trajectoires à partir d'une chaîne de Markov : on effectue une marche aléatoire sur l'espace de Fourier discrétisé.

Voici le type de théorème obtenu.

Théorème 9 (Échantillonnage de l'homme saoul [29]). *Soit $X = (X_i)_{i \in \mathbb{N}}$ une chaîne de Markov réversible sur l'ensemble $\{1, \dots, n\}$. Soit $P \in \mathbb{R}^{n \times n}$ sa matrice de transition et $p \in \mathbb{R}^n$ sa distribution stationnaire (i.e. $p = pP$). Soit $\lambda_2(P)$ la deuxième plus grande valeur propre de P et $\epsilon(P) = 1 - \lambda_2(P)$ son trou spectral.*

Si

$$m \geq \frac{12}{\epsilon(P)} K^2(A_0, p) s^2 \log(2n^2/\delta), \quad (3.5)$$

tout vecteur x_0 s -parcimonieux est la solution unique du problème (3.1).

La différence essentielle avec une approche par tirage indépendant des vecteurs de mesure est le terme $\frac{1}{\epsilon(P)}$. Il indique le temps caractéristique nécessaire pour que deux valeurs de la chaîne de Markov puissent être considérées comme quasi indépendantes. En pratique, cette approche est très mauvaise. Si la matrice de transition P n'autorise que des déplacements au plus proche voisin sur une grille à d dimensions, on peut montrer en utilisant des inégalités de Cheeger, que $\epsilon(P) = O(n^{-1/d})$. Ainsi le nombre de mesures doit être multiplié par $n^{1/d}$ par rapport à une approche par tirages indépendants, ce qui est bien trop grand. Le souci vient du fait que la chaîne de Markov prend énormément de temps à bien couvrir l'espace. Ceci est illustré sur la figure 3.1c : la chaîne de Markov est loin d'approcher la distribution stationnaire indiquée sur la gauche, bien qu'on ait autorisé quelques saut sporadiques.

 En ouverture, le lecteur attentif, aura probablement remarqué le goulot quadratique (le terme s^2 à la place de s). Je pense qu'on peut le casser. Lors de la thèse de Nicolas Chauffert, nous avons essayé de développer des inégalités de concentration sur des chaînes de Markov à valeurs matricielles. Jusqu'à présent de tels résultats ne sont disponibles que pour des chaînes de Markov à valeurs réelles [Lez98] ou vectorielles [Kar+07]. Malheureusement, cette tentative a été un échec jusqu'à présent et nous gardons tous les deux à l'esprit ce problème. Un tel outil permettrait de transformer le s^2 en s dans l'inégalité (3.5), ce qui présente un intérêt académique avant tout.

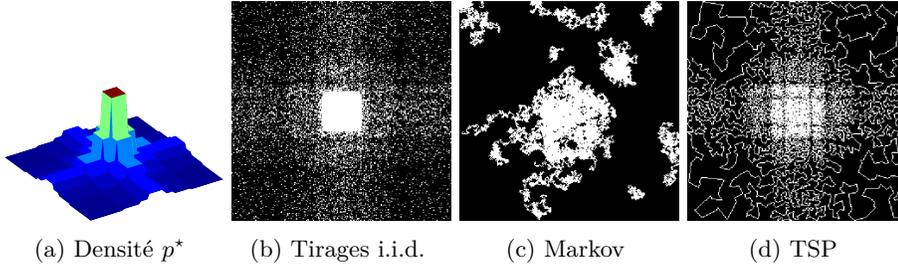


FIGURE 3.1 – Exemples d'échantillonnages aléatoires (densité cible, processus ponctuel, échantillonnage de l'homme saoul et échantillonnage du voyageur de commerce (n'ayant pas bu)).

3.2.2 L'échantillonnage du voyageur de commerce

Le problème de l'échantillonnage de l'homme saoul est le temps de couverture de l'espace, qui est capturé par l'inverse du trou spectral $1/\epsilon(P)$. Pour accélérer la convergence, nous avons essayé la stratégie suivante : tirer quelques points au hasard suivant une certaine densité de probabilité $q : [0, 1]^2 \rightarrow \mathbb{R}_+$, puis rejoindre les échantillons par une polyligne déterminée par la résolution du problème du voyageur de commerce (le chemin le plus court qui passe par tous les points échantillonnés). On appelle ce chemin, le chemin TSP par la suite (pour Travelling Salesman Problem). La question qu'on s'est posée est la suivante.

Question 3 (Le problème du TSP). *Comment choisir la densité de probabilité $q : [0, 1]^d \rightarrow \mathbb{R}_+$ pour que la mesure empirique du chemin TSP soit égal à une densité cible p^* ?*

Soit $q : [0, 1]^d \rightarrow \mathbb{R}$ une densité de probabilité. Soit $(X_i)_{i \in \mathbb{N}}$ un ensemble de vecteurs aléatoires indépendants de $[0, 1]^d$ tirés suivant la loi q (un processus de Poisson). Soit $\gamma_N : [0, 1] \rightarrow \mathbb{R}^d$ la fonction qui paramètre le chemin parcouru par le TSP sur les points $\{X_1, \dots, X_N\}$ à vitesse constante. Soit μ_N sa mesure image définie pour tout borélien $B \subset [0, 1]^d$ par

$$\mu_N(B) = \lambda_{[0,1]}(\gamma_N^{-1}(B)),$$

où $\lambda_{[0,1]}$ est la mesure de Lebesgue sur $[0, 1]$.

Le théorème suivant, illustré sur la figure 3.1d répond à la question 3.

Théorème 10 (Échantillonnage du TSP [29]). *Soit $p : [0, 1]^d \rightarrow \mathbb{R}_+$ une densité de probabilité. Posons $q = \frac{p^{(d-1)/d}}{\int_{[0,1]^d} p(x) dx}$. Alors la mesure image μ_N converge en distribution vers p .*

En pratique, cet échantillonneur fonctionne assez bien. Nous l'avons cependant abandonné au profit de méthodes plus avancées décrites par la suite.

3.2.3 L'échantillonnage par blocs

Lors de la thèse de Claire Boyer, nous avons cherché à comprendre pourquoi il est possible d'échantillonner des images en tirant un ensemble d'ensembles d'échantillons groupés par blocs, tels que des lignes horizontales dans le plan de Fourier. Cette problématique est illustrée sur la figure 3.2 : on peut reconstruire un texte incrusté dans le cerveau à une résolution pixellique avec seulement 13% des coefficients de Fourier mesurés. Les “grands” théorèmes d'échantillonnage compressé ne permettent d'expliquer ce phénomène et je pense que c'est une contribution peu suivie, mais intéressante.

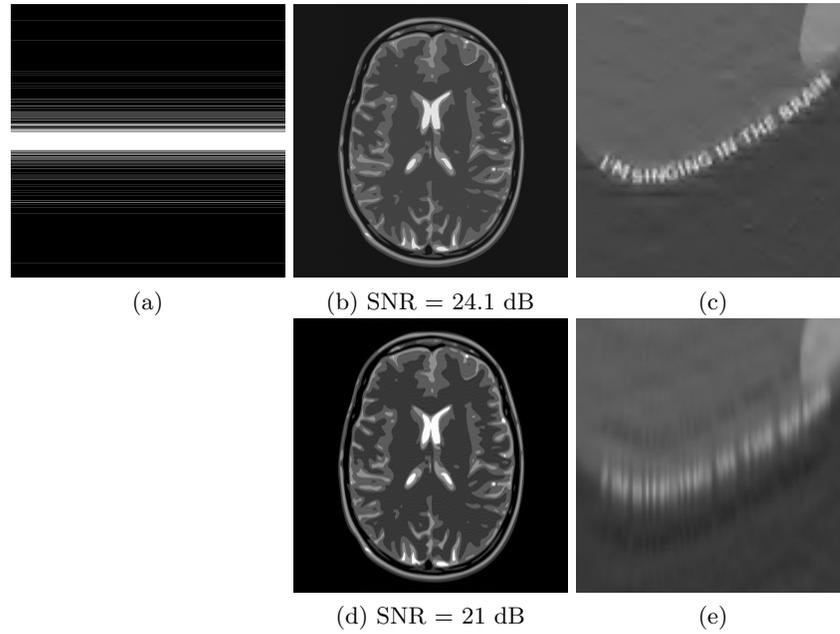


FIGURE 3.2 – Exemple de reconstruction d’une image 2048×2048 à partir de ligne horizontales de mesures. (a) Schéma d’échantillonnage par lignes horizontales (13% des mesures). (b) Reconstruction par minimisation ℓ_1 et son zoom (c). (d) Reconstruction par pseudo-inverse et un zoom (e).

Pour expliquer ce phénomène, nous avons travaillé en parallèle et donc suivi d’assez près les travaux du groupe Adcock, Hansen, Poon [Adc+17]. Nous avons en particulier montré qu’il est *impossible* de reconstruire uniformément toute la classe des signaux s -parcimonieux (ce qui empêche l’utilisation de techniques reposant sur la propriété d’isométrie restreinte (RIP)). Nos résultats sont donc non uniformes, au sens où ils permettent de reconstruire un signal donné, et que la capacité de reconstruction dépend de la *structure* du signal, et en particulier de *son support*.

Un résultat générique

Je présente ici une version simplifiée de notre résultat principal [18]. Nous avons notamment systématiquement montré que les résultats étaient stables par ajout de bruit sur les données, mais je préfère passer ceci sous silence.

Comme dans la partie 3.2.1, on part d'une matrice $A_0 \in \mathbb{C}^{n \times n}$ orthogonale dont les lignes sont notées a_i^* . On fixe une partition² $(I_k)_{1 \leq k \leq M}$ de $\{1, \dots, n\}$. À partir de cette partition, on construit des sous-matrices de mesures (les blocs) $B_k = (a_i^*)_{i \in I_k} \in \mathbb{C}^{|I_k| \times n}$.

On fixe alors un nombre de blocs de mesures $m \in \mathbb{N}$ et on construit une matrice aléatoire de mesures comme suit :

$$A = \frac{1}{\sqrt{m}} \left(\frac{1}{\sqrt{\pi_{K_l}}} B_{K_l} \right), \quad (3.6)$$

où les indices K_l sont des copies i.i.d. d'une variable aléatoire K telle que $\mathbb{P}(K = k) = \pi_k$, avec $\pi \in \mathbb{R}^M$.

Notre objectif est de reconstruire un signal x de support $\text{supp}(x) = S \subset \{1, \dots, n\}$. Nous avons alors dû introduire plusieurs quantités pour caractériser la qualité d'un système de mesure (i.e. d'un triplet $(A_0, (I_k)_{1 \leq k \leq M}, \pi)$) pour reconstruire le signal x .

Définition 3.

$$\Theta(S, \pi) \stackrel{\text{def.}}{=} \max_{1 \leq k \leq M} \frac{1}{\pi_k} \|B_k^* B_{k,S}\|_{\infty \rightarrow \infty} = \max_{1 \leq k \leq M} \max_{1 \leq i \leq n} \frac{\|e_i^* B_k^* B_{k,S}\|_1}{\pi_k}, \quad (3.7)$$

$$\Upsilon(S, \pi) \stackrel{\text{def.}}{=} \max_{1 \leq i \leq n} \sup_{\|v\|_{\infty} \leq 1} \sum_{k=1}^M \frac{1}{\pi_k} |e_i^* B_k^* B_{k,S} v|^2, \quad (3.8)$$

$$\Gamma(S, \pi) \stackrel{\text{def.}}{=} \max(\Upsilon(S, \pi), \Theta(S, \pi)). \quad (3.9)$$

La quantité Θ peut être vue comme une extension de la notion de cohérence à des blocs de mesures. La quantité Υ est difficile à calculer et à interpréter. De façon très grossière, elle permet de prendre en compte des "interférences" entre les blocs. Malheureusement, elle semble indispensable (à moins de faire des hypothèses de signes aléatoires sur les coefficients de x) à l'établissement du théorème suivant.

Théorème 11. *Soit $S \subset \{1, \dots, n\}$ un ensemble d'indices de cardinal $s > 16$. Soit x un signal S -parcimonieux. On construit la matrice de mesures A suivant le modèle (3.6). Soit $\epsilon \in]0, 1[$.*

Si le nombre de mesures m satisfait :

$$m \geq C \cdot \Gamma(S, \pi) \log(s) \log\left(\frac{n}{\epsilon}\right), \quad (3.10)$$

où C est une constante universelle, alors x est solution du problème (3.1) avec probabilité supérieure à $1 - \epsilon$.

2. Le théorème couvre aussi le cas de blocs qui se recouvrent.

Le système Fourier-Haar

Le théorème (11) et en particulier la quantité Γ peuvent paraître sibyllins. C'est pourtant un résultat fort lorsqu'on l'applique à des systèmes particuliers. Je montre ci-dessous qu'il permet d'expliquer le résultat de la figure 3.2.

On suppose qu'on souhaite reconstruire une image carrée de taille $\sqrt{n} \times \sqrt{n}$ où \sqrt{n} est une puissance de 2. La matrice $A_0 = F\Psi$ est le produit d'une transformée de Fourier discrète $F \in \mathbb{C}^{n \times n}$ et d'une transformée en ondelettes de Haar orthogonales $\Psi \in \mathbb{C}^{n \times n}$. Le vecteur x représente donc les coefficients d'ondelettes d'une image. On suppose que les blocs B_k correspondent à des lignes horizontales du plan de Fourier, comme dans la figure 3.2. On effectue une partition dyadique $(\tau_j)_{0 \leq j \leq J}$ de l'ensemble $\{1, \dots, \sqrt{n}\}$, avec $J = \log_2(\sqrt{n}) - 1$. Ainsi $\tau_0 = \{1\}$, $\tau_1 = \{2\}$, $\tau_2 = \{3, 4\}$, $\tau_3 = \{5, 6, 7, 8\}$ et ainsi de suite. On définit des ensemble $\Omega_{\ell, \ell'} = \tau_\ell \times \tau_{\ell'}$. Nous renvoyons le lecteur à la figure 3.3 pour une illustration. Enfin, on définit une fonction $j : \{1, \dots, \sqrt{n}\} \rightarrow \{0, \dots, J\}$ par $j(u) = j$ si $u \in \tau_j$.

Définition 4. Soit $S = \text{supp}(x)$. On définit :

$$s_\ell^c := \max_{0 \leq \ell' \leq J} \max_{k \in \tau_{\ell'}} |S \cap \Omega_{\ell, \ell'} \cap C_k|, \quad (3.11)$$

où C_k représente la k -ième ligne verticale (voir figure 3.3).

Les quantités s_ℓ^c représentent le cardinal du support maximal dans les sous-bandes $(\ell, \ell')_{0 \leq \ell' \leq J}$.

En appliquant le théorème (11) à ce système particulier, on peut démontrer le résultat suivant.

Théorème 12. Soit $x \in \mathbb{C}^n$ un signal S -parcimonieux. Soit $\varepsilon \in (0, 1)$. Soit $(\pi_k)_{1 \leq k \leq \sqrt{n}}$ une densité constante par niveau, i.e. $\pi_k = \tilde{\pi}_{j(k)}$. Si le nombre m de lignes horizontales satisfait

$$m \gtrsim \max_{0 \leq j \leq J} \frac{2^{-j}}{\tilde{\pi}_j} \sum_{r=0}^J 2^{-|j-r|/2} s_r^c \ln(s) \ln\left(\frac{n}{\varepsilon}\right),$$

alors x est l'unique solution du problème (3.1) avec probabilité supérieur à $1 - \varepsilon$.

En particulier, si

$$\pi_k = \frac{2^{-j(k)} \sum_{r=0}^J 2^{-|j-r|/2} s_r^c}{\sum_{\ell=1}^{\sqrt{n}} 2^{-j(\ell)} \sum_{r=0}^J 2^{-|j(\ell)-r|/2} s_r^c},$$

alors

$$m \gtrsim \sum_{j=0}^J \left(s_j^c + \sum_{\substack{r=0 \\ r \neq j}}^J 2^{-|j-r|/2} s_r^c \right) \cdot \ln(s) \ln\left(\frac{n}{\varepsilon}\right)$$

donne une reconstruction exacte avec probabilité supérieure à $1 - \varepsilon$.

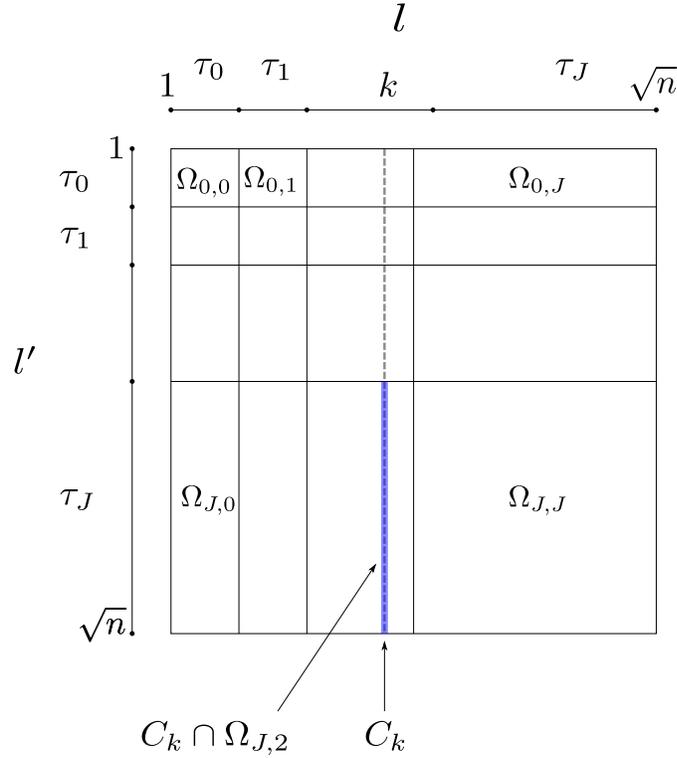


FIGURE 3.3 – Illustration des quantités intervenant pour le système Fourier-Haar avec des lignes horizontales.

Nous aurions probablement été incapables d'identifier les quantités s_l^c sans avoir préalablement montré le théorème (11). Le théorème 12 indique ces quantités a priori anodines jouent un rôle primordial pour la reconstruction d'images à partir de lignes horizontales en Fourier. Nous avons pu valider expérimentalement ce résultat, comme le montre la figure 3.4. A ma connaissance, très peu de résultats en échantillonnage compressé permettent d'expliquer ce qui se passe sur des systèmes de mesures réels. Je trouve donc ce résultat assez remarquable bien qu'il soit difficile à digérer.

3.3 Génération de schémas d'échantillonnage

Dans la section précédente, nous avons cherché à donner des garanties théoriques de reconstruction exacte à partir de tirages aléatoires d'échantillons ou de blocs d'échantillons. En pratique, les quantités émergent de ces analyses sont difficiles à exploiter. Par exemple, pour trouver une densité de probabilité optimale, il faudrait minimiser la quantité $\Gamma(S, \pi)$ définie en (3.9) par rapport

à tous les supports S admissibles.

En parallèle des travaux théoriques précédents, j'ai donc cherché à adopter des approches plus heuristiques de génération de schémas d'échantillonnages efficaces.

3.3.1 Optimisation de probabilité de tirage

Un premier exemple d'application a été présenté dans la section 2.2.2 : nous avons cherché à optimiser une densité de probabilité sur des espaces de blocs admissibles, de telle manière à ce que la mesure empirique des blocs tirés se rapproche d'une densité cible discrète. Je ne décrirais pas plus longuement cette approche que nous avons finalement abandonnée après avoir fait l'observation suivante :

Observation 1. *Si d'un point de vue théorique, les schémas d'échantillonnage les plus efficaces sont aléatoires, il n'en n'est pas de même expérimentalement. On observe que des schémas déterministes sont plus efficaces. La cause probable est que l'aléa a tendance à produire des amas (clusters) de points. Ceci est aussi en accord avec les théories d'échantillonnage avec reconstruction linéaire (e.g. Beurling).*

3.3.2 Optimisation de mesures

Dans ce paragraphe, je vais décrire deux techniques qui ont été motivées par l'observation 1 et par le théorème 10. Pour générer un schéma d'échantillonnage, nous avons souhaité satisfaire les principes suivants [23] :

1. Un schéma doit avoir une densité cible donnée p ,
2. Un schéma doit couvrir l'espace rapidement (i.e. éviter les clusters, ou avoir un temps de mélange rapide),
3. Un schéma doit être admissible, au sens où les échantillons doivent satisfaire un ensemble de contraintes dictées par la machine.

Ces trois principes nous ont mené à considérer le problème de *projection* suivant :

$$\inf_{\mu \in \mathcal{A}} \text{dist}(\mu, \pi), \quad (3.12)$$

où π est une mesure cible à densité $p : [0, 1]^d \rightarrow \mathbb{R}_+$, \mathcal{A} est un sous-ensemble de mesures de probabilité admissibles et $\text{dist}(\cdot, \cdot)$ est une distance entre mesures de probabilités.

L'exemple le plus simple d'espace \mathcal{A} est l'espace des mesures discrètes à m points :

$$\mathcal{M}(\Omega^m) \stackrel{\text{def.}}{=} \left\{ \mu = \frac{1}{m} \sum_{i=1}^m \delta_{x_i}, x_i \in [0, 1]^d \right\}. \quad (3.13)$$

Cet espace montre qu'on ne peut pas choisir des distances arbitraires. Par exemple, si dist est la variation totale de la différence $\mu - p\pi$ par exemple, on obtient $\text{dist}(\mu, \pi) = \|\mu - \pi\|_{\mathcal{M}} = 2$ pour toute mesure $\mu \in \mathcal{M}_m$!

Nous devons donc choisir des distances qui métrisent la convergence faible. Nous avons considéré deux possibilités qui sont les distances de convolutions (aussi appelées “maximum norm discrepancies”, “kernel norms” ou “blurred SSD”) ou la distance de transport W_2 . Nous dédions un paragraphe à chacune d'entre elle ci-dessous.

Distances de convolution

On travaille ici sur le tore $\Omega = \mathbb{T}^d$, ce qui va nous permettre de travailler avec des séries de Fourier. On pourrait aussi travailler sur \mathbb{R}^d avec quelques précautions supplémentaires.

Définition de la distance Soit $h \in L^2(\Omega)$ un noyau de convolution *continu*. On définit une distance d_h entre deux mesures μ_1 et μ_2 comme suit :

$$d_h(\mu_1, \mu_2) = \|h \star (\mu_1 - \mu_2)\|_{L^2(\Omega)}, \quad (3.14)$$

où le symbole \star est le produit de convolution. Soit $\xi \in \mathbb{Z}$. On note $\hat{h}(\xi)$ la valeur de la série de Fourier de h en ξ . Le théorème suivant indique quelques propriétés utiles.

Théorème 13 (Propriétés des distances de convolution [19]).

- Si h est continu, $h \star \mu \in L^2$ pour tout $\mu \in \mathcal{M}(\Omega)$. L'application d_h est donc bien définie sur l'espace des mesures de Radon.
- L'application $\mu \mapsto \sqrt{d_h(\mu, \mu)}$ est une norme si et seulement si la série de Fourier \hat{h} ne s'annule pas sur \mathbb{Z}^d .
- Soit $(\mu_N)_{N \in \mathbb{N}}$ une suite de mesures. Si \hat{h} ne s'annule pas sur \mathbb{Z}^d , alors

$$\lim_{N \rightarrow \infty} d_h(\mu_N, \mu_N) = 0 \Leftrightarrow \mu_N \xrightarrow{N \rightarrow \infty} 0. \quad (3.15)$$

On dit que la distance d_h métrise la convergence faible.

- Si h est une fonction L -Lipschitz, alors pour tout couple $(\mu_1, \mu_2) \in \mathcal{M}(\Omega) \times \mathcal{M}(\Omega)$,

$$d_h(\mu_1, \mu_2) \leq LW_1(\mu_1, \mu_2) \quad (3.16)$$

où W_1 est la distance de transport définie par :

$$W_1(\mu_1, \mu_2) \stackrel{\text{def.}}{=} \sup_{f, \text{Lip}(f) \leq 1} \mu_1(f) - \mu_2(f). \quad (3.17)$$

- Supposons que $\text{dist} = d_h$ et que la série de Fourier \hat{h} ne s'annule pas sur \mathbb{Z}^d . Le problème (3.12) admet une solution si \mathcal{A} est faiblement compact (e.g. s'il est faiblement compact et borné en norme de variation totale).

Discrétisation D'un point de vue numérique, il est nécessaire d'approcher l'ensemble \mathcal{A} par un espace de mesures paramétrées par un nombre fini de variables. La proposition suivante montre que $\cup_{m \in \mathbb{N}} \mathcal{M}_m$ est dense dans l'espace des probabilités.

Proposition 4. *Soit \mathcal{M}_Δ l'ensemble des mesures de probabilité sur Ω et h un noyau Lipschitz. On a alors l'inégalité suivante :*

$$\sup_{\pi \in \mathcal{M}_\Delta} \inf_{\mu \in \mathcal{M}(\Omega^m)} d_h(\mu, \pi) \leq L \left(\sqrt{d}/2 + 1 \right) \frac{1}{m^{1/d} - 1} = O(m^{-1/d}). \quad (3.18)$$

On peut alors montrer facilement le résultat suivant.

Proposition 5. *Pour tout espace \mathcal{A} de mesures non vide, on peut construire un espace approximant \mathcal{M}_n de la forme :*

$$\mathcal{M}_n \stackrel{\text{def.}}{=} \left\{ \mu = \frac{1}{m} \sum_{i=1}^m \delta_{x_i}, x_i \in C_n \right\}, \quad (3.19)$$

où $C_n \subset \Omega^n$ est un espace de contraintes, tel que la distance de Hausdorff

$$\mathcal{H}(\mathcal{A}, \mathcal{M}_n) \stackrel{\text{def.}}{=} \max \left(\sup_{\pi \in \mathcal{A}} \inf_{\mu \in \mathcal{M}_n} d_h(\mu, \pi), \sup_{\mu \in \mathcal{M}_n} \inf_{\pi \in \mathcal{A}} d_h(\mu, \pi) \right) = O \left(m^{-1/d} \right). \quad (3.20)$$

Résolution numérique Pour résoudre le problème (3.12), on commence par le remplacer par sa version discrétisée :

$$\arg \min_{\mu \in \mathcal{M}_n} d_h(\mu, \pi). \quad (3.21)$$

L'idée est ensuite d'utiliser de la programmation non linéaire pour optimiser les positions $(x_i)_{1 \leq i \leq m}$. La proposition suivante indique l'énergie à minimiser.

Proposition 6. *Soit x^* une solution du problème suivant*

$$\min_{(x_i)_{1 \leq i \leq m} \in C_n} J(x) \stackrel{\text{def.}}{=} \underbrace{\frac{1}{n} \sum_{i=1}^n \int_{\Omega} H(x - x_i) d\pi(x)}_{\text{Attraction}} - \underbrace{\frac{1}{2n^2} \sum_{i=1}^n \sum_{j=1}^n H(x_i - x_j)}_{\text{Répulsion}}, \quad (3.22)$$

où H est défini par sa série de Fourier comme $\hat{H} = |\hat{h}|^2$.

La mesure $\mu^* = \frac{1}{n} \sum_{i=1}^n \delta_{x_i^*}$ est solution du problème de projection (3.21).

Pour résoudre ce problème on utilise la descente de gradient projeté suivante :

$$x^{(k+1)} \in \Pi_{C_n}(x^{(k)} - \tau_k \nabla J(x^{(k)})), \quad (3.23)$$

où $\Pi_{C_n}(\cdot)$ est le projecteur sur C_n . Nous avons montré la convergence de cet algorithme vers un point critique en suivant les travaux de Jérôme Bolte notamment [ABS13].

Ce principe d'apparence assez simple cache de vraies difficultés numériques que je décris rapidement ci-dessous :

- En pratique le noyau H doit avoir une singularité en 0 pour que l'algorithme se comporte bien. Pour assurer la convergence, on le lisse légèrement à l'origine, ce qui implique que la constante de Lipschitz du gradient de J est grande. On doit donc prendre des pas de descente τ_k petits. L'approche qui s'est révélée la plus efficace est une technique de choix de pas de type Barzilei-Borwein.
- On souhaite pouvoir optimiser des centaines de milliers de particules x_i . Le calcul du gradient doit donc être rapide, alors qu'une approche naïve coûte $O(n^2)$ opérations. Il faut utiliser des *techniques de sommation rapide* de type méthodes multi-polaires pour réduire ce coût.
- On doit travailler sur des stations multi-cœurs pour atteindre le nombre de particules souhaitées. Les codes sont en $C++$ interfacés avec Matlab.
- Définir C_n et son projecteur peut être difficile en pratique.

Résultats La dynamique de l'algorithme est illustrée sur la figure 3.5 pour différentes initialisations. Comme on peut le voir, le résultat final est toujours plaisant d'un point de vue visuel.

Finalement, une application de cet algorithme à la génération de schémas d'échantillonnage est présentée dans la figure 5.7.

Transport optimal

Suite à un conseil de Gabriel Peyré lors de la soutenance de thèse de Nicolas Chauffert, nous avons décidé d'essayer de remplacer la distance à noyau d_h par la distance de transport W_2^2 lors de la thèse de Léo Lebrat. Après plusieurs essais sur des distances de transport discrètes ou discrètes régularisées [Ben+15], nous avons finalement opté pour du transport semi-discret [De +12a ; Lé15]. Le manuscrit étant déjà bien long, cette partie étant technique et ayant réalisé une vidéo de démonstration, je renvoie le lecteur intéressé à cette url <https://youtu.be/FjdNLjk1xHM>, pour quelques résultats.

Je tiens simplement à noter ici que c'est un problème délicat d'un point de vue théorique et numérique, qui tient à la géométrie computationnelle. Léo Lebrat et son directeur de thèse Frédéric de Gournay ont produit l'essentiel du code (alors que je contribue d'habitude largement à cette partie).

3.4 Ouverture

Ce chapitre contient plusieurs perspectives à plus ou moins long terme.

3.4.1 *Reconstruction dans des trames*

Nous avons focalisé notre attention sur des régularisations avec des transformées orthogonales jusqu'à présent. En pratique, il est souvent bien plus efficace d'effectuer des régularisation sur des trames redondantes (e.g. des ondelettes invariantes par translation). Lorsqu'on utilise des transformées redondantes,

on peut considérer une régularisation par analyse [Nam+13] ou par synthèse [RSV08]. Ces deux domaines me semblent très mal compris aujourd'hui.

Des chercheurs tels que Candès, Eldar, Needell et Randall [Can+11] ont bien produit des théorèmes de reconstruction reposant sur la D-RIP par exemple. Cependant - et malgré un nombre très important de citations - il est facile de se convaincre qu'ils n'ont qu'un intérêt très limité et qu'ils ne peuvent pas expliquer les résultats observés en pratique.

3.4.2 Échantillonnage déterministe

Les schémas d'échantillonnage obtenus en minimisant une fonctionnelle peuvent être considérés comme déterministes. En pratique, ils fonctionnent significativement mieux que les schémas par tirages i.i.d. suggérés par les théorèmes les plus avancés de la littérature. Mon sentiment est que les techniques de preuve d'échantillonnage compressé actuelles sont sous-optimales et qu'il est possible que des hypothèses et approches radicalement différentes donnent de meilleurs résultats. Je n'ai pas de vraie piste sérieuse ici, mais je compte rester sur le qui-vive, car je pense que de belles découvertes restent à venir.

3.4.3 Compréhension de la convergence

Dans le cadre du projet ANR OMS, nous essayons actuellement de comprendre pourquoi les dynamiques illustrées sur la figure 3.5 mènent toujours à des points critiques agréables. Ce phénomène semble indiquer que l'essentiel des points critiques de la fonction (3.22) sont concentrés autour d'un certain minimum global ou que la dynamique particulière se rapproche de la dynamique continue. C'est un problème qui semble aussi réellement difficile. Cependant, des travaux récents autour des flots de gradient [CB18] donnent quelque espoir de progresser. Une semaine de retraite dans les montagnes avec mes collaborateurs talentueux (J. Bolte, J. Kahn, F. de Gournay, A. Flinth) n'a pas suffi à réellement dégager de piste pour avancer. Il nous reste 3 ans pour faire des progrès...

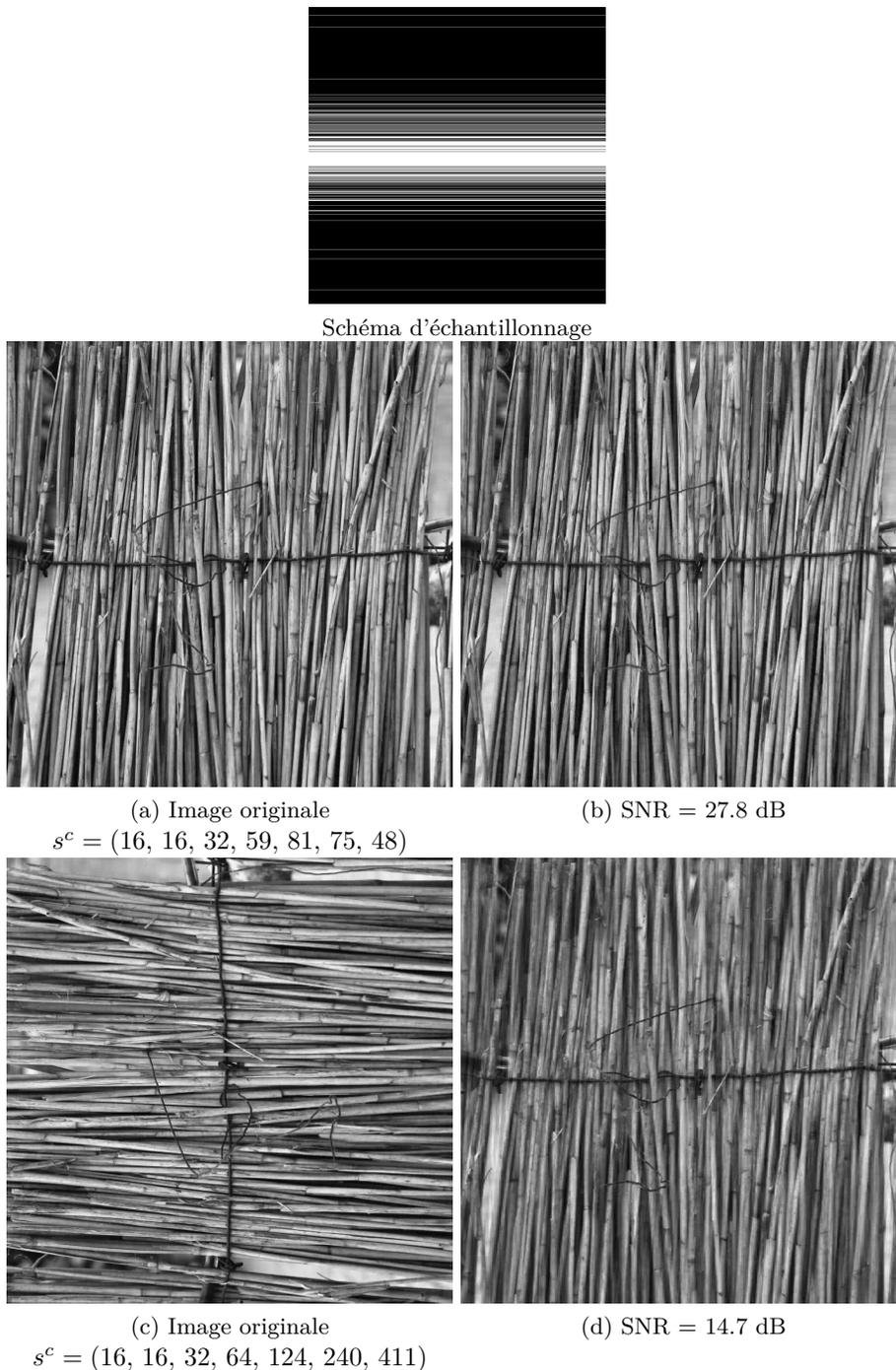


FIGURE 3.4 – Exemple de reconstruction d'une image de taille 2048×2048 réelle mesurée dans le domaine de Fourier. (a) (c), images de références : (c) est identique à (a) mais tournée de 90° . les valeurs du vecteur $s^c = (s_j^c)_{1 \leq j \leq 7}$ sont indiquées pour les deux images. Il est important d'observer que les quantités s_j^c sont plus grandes pour l'image (b). Pour la reconstruction, on utilise un schéma d'échantillonnage composé de 9.8 % des mesures. Les images (b) (d), correspondent à une reconstruction ℓ_1 . On tourne les images dans (d) pour faciliter la comparaison. Observer que (b) est de qualité nettement supérieure à (d). C'est exactement ce que prédit le théorème 12.

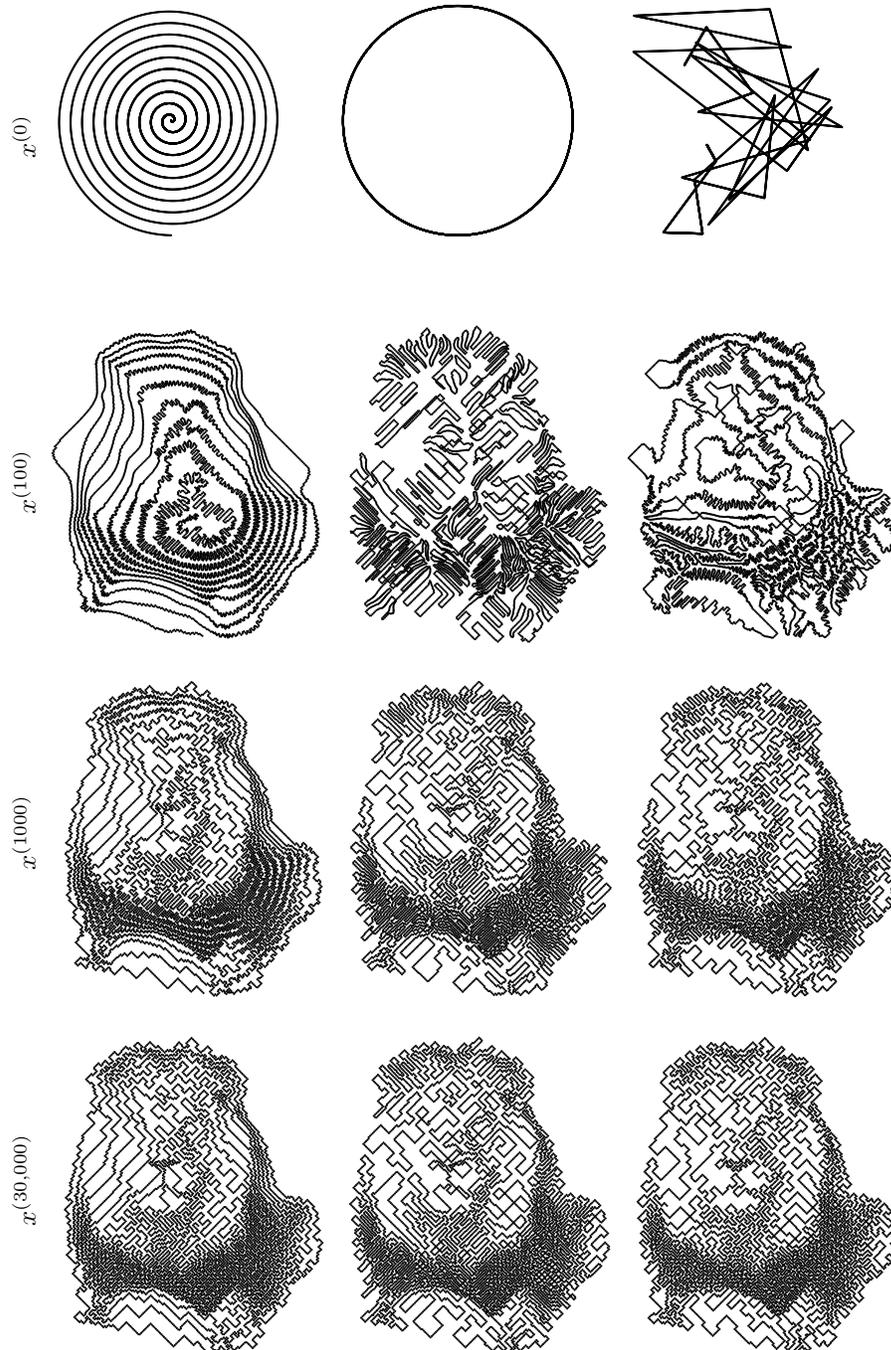


FIGURE 3.5 – Projection d'une image lion sur un espace de courbes à vitesse bornée avec 8,000 particules. La figure montre les itérées $x^{(k)}$ pour plusieurs valeurs de k .

Chapitre 4

Approximation et apprentissage d'opérateurs

4.1 Introduction

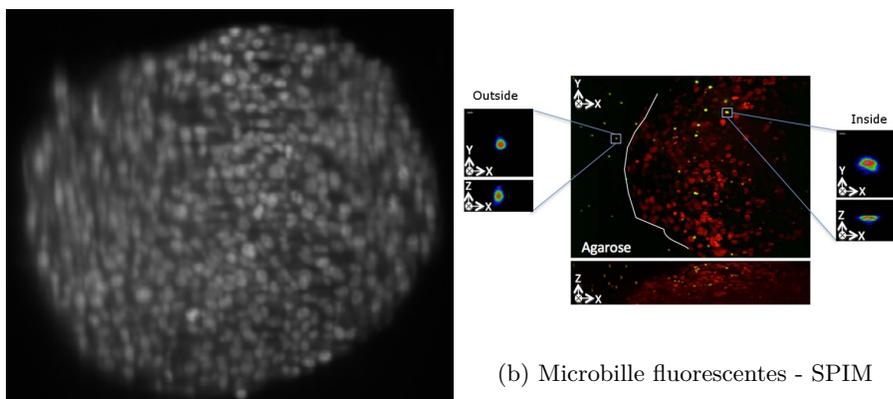
L'un des projets interdisciplinaires qui m'a le plus séduit pour créer l'équipe projet PRIMO a été la restauration d'images SPIM (Selective Plane Illumination Microscope). Je ne souhaite pas rentrer dans les détails de fonctionnement du microscope, mais simplement indiquer que la géométrie d'une feuille de lumière qui vient exciter des molécules fluorescentes crée un flou qui varie spatialement. Ce problème est loin d'être limité à la microscopie SPIM. Il est illustré pour différents imageurs sur la figure 4.1. Malheureusement, je dois bien avouer que nous sommes encore assez loin d'avoir fourni un outil clé en main aux biologistes, malgré 4 années de travail. Je commence à avoir un espoir qu'on arrive à des algorithmes utilisables en routine sur les plateformes d'imagerie d'ici 2 ans.

Cette problématique a pris une grosse part de mon temps de recherche ces 5 dernières années avec la thèse de Paul Escande, et aujourd'hui, celle de Valentin Debarnot. J'ai aussi déposé 5 demandes d'ANR (JCJC et bilatérale avec Hans Feichtinger) sur ce thème qui ont toutes été rejetées malgré des évaluations systématiquement positives. La dernière signalait le fait que je n'obtiendrais pas d'ANR jeune chercheur dans ma carrière.

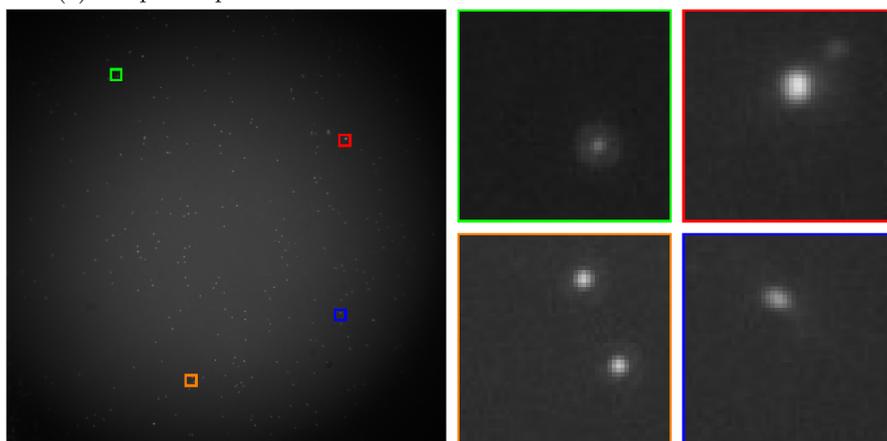
4.2 Approximation d'opérateurs de flou

4.2.1 Modélisation

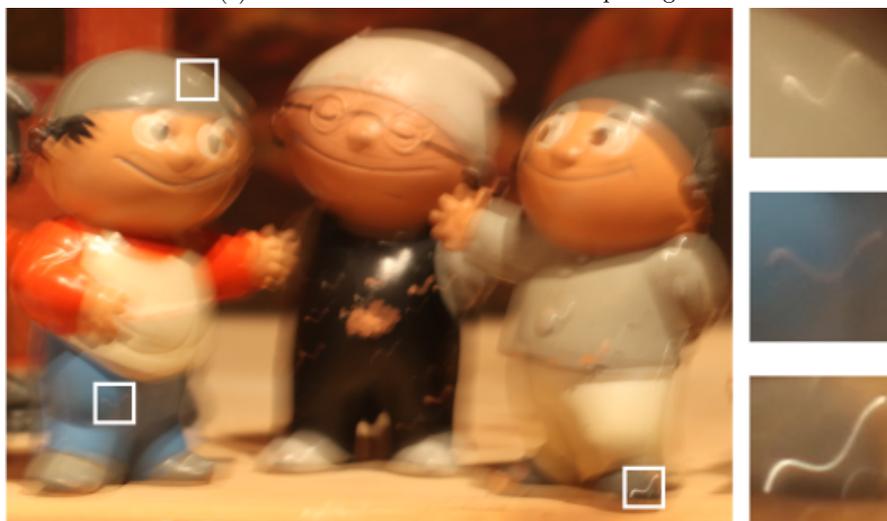
Les résultats énoncés dans cette partie sont valides à la fois sur le domaine $\Omega = \mathbb{R}^d$ et sur le tore $\Omega = \mathbb{T}^d$ avec des conditions de bord périodiques. Il est plus simple de définir les ondelettes avec des conditions de bord périodiques et nous présenterons donc l'essentiel des résultats dans ce cadre.



(a) Coupe de sphéroïde - SPIM



(c) Microbilles fluorescentes - Champs large



(d) Photographie conventionnelle

FIGURE 4.1 – Différents exemples de flous variables spatialement.

Dans tous nos travaux, nous avons modélisé les opérateurs de flou H comme des opérateurs linéaires intégraux bornés $H : L^2(\Omega) \rightarrow L^2(\Omega)$:

$$\forall x \in \Omega, \quad Hu(x) = \int_{\Omega} K(x, y)u(y)dy. \quad (4.1)$$

La fonction $K : \Omega \times \Omega \rightarrow \mathbb{R}$ est appelée noyau de l'opérateur et définit la fonction de répartition du point (PSF) $K(\cdot, y)$ au point $y \in \Omega$. Pour une image $u \in L^2(\Omega)$, l'image Hu correspond à une version floue de u . On propose de donner une définition simple des opérateurs de flou comme suit.

Définition 5 (Opérateurs de flou [26]). *Soit $M \in \mathbb{N}$ et $f : [0, 1] \rightarrow \mathbb{R}_+$ une fonction décroissante bornée. Un opérateur intégral est appelé opérateur de flou dans la classe $\mathcal{A}(M, f)$ s'il satisfait les propriétés suivantes :*

1. Son noyau $K \in W^{M, \infty}(\Omega \times \Omega)$;
2. Ses dérivées partielles K satisfont :

(a)

$$\forall |\alpha| \leq M, \forall (x, y) \in \Omega \times \Omega, \quad |\partial_x^\alpha K(x, y)| \leq f(\|x - y\|_\infty), \quad (4.2)$$

(b)

$$\forall |\alpha| \leq M, \forall (x, y) \in \Omega \times \Omega, \quad |\partial_y^\alpha K(x, y)| \leq f(\|x - y\|_\infty). \quad (4.3)$$

Une telle définition peut choquer les physiciens qui sont habitués à utiliser les lois de propagation des ondes pour déterminer analytiquement les formes de PSFs. Mon avis est que ces modèles sont très utiles pour avoir une bonne idée de ce qu'il faut attendre. Cependant, il me semble presque impossible de rendre compte de la diversité des phénomènes physiques qui peuvent altérer ces modèles idéalisés. De plus, on va voir que les hypothèses de la définition 5 permettent d'obtenir des garanties fortes de compressibilité des opérateurs.

Je tiens à préciser dès à présent les difficultés associées à de tels opérateurs :

- On souhaite travailler avec des images à n pixels, où n varie typiquement entre 1 million et 10 milliards. Si on se contente d'une discrétisation naïve d'un opérateur H sous la forme d'une matrice, cette dernière ne pourra pas être stockée sur un ordinateur (e.g. une matrice de taille $10^{10} \times 10^{10}$ représente probablement plus que toutes les capacités de stockage actuelles).
- En pratique, on ne connaît jamais réellement l'opérateur. On a seulement un accès indirect à travers son action sur des masses de Dirac par exemple (voir figure 4.1). Dans ces conditions, il faut développer des algorithmes de reconstruction d'opérateurs.
- Il est fréquent qu'une calibration d'un microscope, suivie d'un défloutage à opérateur connu soit insuffisante. En effet, beaucoup de paramètres expérimentaux peuvent varier (température, défocalisation, tilts d'éléments optiques). Lorsqu'on souhaite restaurer une image, on est donc en présence d'un problème inverse aveugle.

4.2.2 Décomposition dans des bases d'ondelettes

Dans le papier [26], nous avons supposé qu'on dispose d'un opérateur H connu et nous avons cherché à exploiter les idées de décomposition d'opérateurs dans des bases d'ondelettes [CM97 ; BCR91a] pour pouvoir le représenter efficacement.

Théorie

Soit $(\psi_\lambda)_{\lambda \in \Lambda}$ une base d'ondelettes orthogonales de $L^2(\mathbb{T}^d)$ en produit tensoriel, construite à partir d'une ondelette mère $\psi \in L^2(\mathbb{T})$ à M moments nuls et à support compact : $\text{supp}(\psi) = [-c, c]$. Le coefficient $\lambda = (j, m, e)$ encode simultanément l'échelle de l'ondelette $j \in \mathbb{N}$, un paramètre de translation $m \in \mathbb{Z}^d$, et son type $e \in \{0, 1\}$ (ondelette mère ou d'échelle). On note aussi $\Psi : L^2(\mathbb{T}^d) \rightarrow \ell^2$ la décomposition d'une image en ses coefficients d'ondelettes. Le support d'une ondelette est noté $I_\lambda = \text{supp}(\psi_\lambda)$.

Puisque H est un opérateur borné de $L^2(\mathbb{T}^d)$, il peut être décomposé sous la forme $H = \Psi\Theta\Psi^*$, où $\Theta : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$ est une matrice infinie décrivant H dans le domaine des ondelettes. La matrice Θ est caractérisée par ses coefficients :

$$\theta_{\lambda, \mu} = \Theta[\lambda, \mu] = \langle H\psi_\lambda, \psi_\mu \rangle. \quad (4.4)$$

Le résultat suivant fournit une borne sur leurs amplitudes (elle améliore légèrement les résultats de Meyer [CM97], en exploitant les propriétés supplémentaires de régularité du noyau).

Théorème 14 (Amplitude des coefficients d'ondelettes[26]). *Soit $f_{\lambda, \mu} = f(\text{dist}(I_\lambda, I_\mu))$ et supposons que :*

- *L'opérateur H appartient à la classe $\mathcal{A}(M, f)$ (voir la définition 5).*
 - *L'ondelette mère est à support compact avec M moments nuls.*
- Pour tout $\lambda = (j, m, e) \in \Lambda$ et $\mu = (k, n, e') \in \Lambda$, avec $e, e' \neq 0$:*

$$|\theta_{\lambda, \mu}| \leq C_M 2^{-(M + \frac{d}{2})|j-k|} 2^{-\min(j, k)(M+d)} f_{\lambda, \mu}, \quad (4.5)$$

où C_M est une constante indépendante de λ et de μ .

En somme, les coefficients de Θ décroissent exponentiellement par rapport à la différence d'échelle et comme une fonction de la distance entre les supports des ondelettes.

Discrétisation

Pour obtenir une représentation finie d'un opérateur, on peut tronquer la représentation à une échelle J . On note $\Theta^{(J)}$ la matrice infinie définie par :

$$\Theta^{(J)}[\lambda, \mu] = \begin{cases} \theta_{\lambda, \mu} & \text{si } |\lambda| \leq J \text{ et } |\mu| \leq J, \\ 0 & \text{sinon.} \end{cases} \quad (4.6)$$

Cette matrice contient au plus N^2 coefficients non nuls, où $N = 1 + \sum_{j=0}^{J-1} (2^d - 1)2^{dj}$ représente le nombre d'ondelettes gardées. L'opérateur $H^{(J)} = \Psi\Theta^{(J)}\Psi^*$ est une approximation de approximation de H . L'erreur de discrétisation $\|H^{(J)} - H\|_{X \rightarrow Y}$ peut être contrôlée, où $\|\cdot\|_{X \rightarrow Y}$ est une norme d'opérateur entre deux espaces métriques X et Y . On renvoie le lecteur intéressé à [26] pour plus de détails.

Le théorème suivant est une variation de [BCR91b]. De façon formelle, il stipule que $H^{(J)}$ peut être approchée par une matrice contenant $O(N)$ coefficients.

Théorème 15 (Représentation compacte dans des bases d'ondelettes[26]). *Soit $0 \leq \eta \leq \log_2(N)^{-(M+d)/d}$. Soit $\Theta_\eta^{(J)}$ la matrice obtenue en remplaçant par 0 tous les coefficients de $\Theta^{(J)}$ tels que*

$$2^{-\min(j,k)(M+d)} f_{\lambda,\mu} \leq \eta. \quad (4.7)$$

Soit $H_\eta^{(J)} = \Psi\Theta_\eta^{(J)}\Psi^$. Sous les hypothèses du théorème (14), le nombre de coefficients nécessaire pour que $\|H^{(J)} - H_\eta^{(J)}\|_{2 \rightarrow 2} \leq \epsilon$ est inférieur à*

$$C'_M N \epsilon^{-\frac{d}{M}} \quad (4.8)$$

où $C'_M > 0$ est indépendant de N .

Ce théorème indique que la représentation et le calcul d'un produit avec H peuvent être obtenus avec une précision ϵ en utilisant seulement $O(N\epsilon^{-\frac{d}{M}})$ opérations. La régularité du noyau M est exploitée automatiquement.

Application en dimension finie

Dans un cadre discret, ces résultats peuvent être exploités ainsi. Étant donnée une matrice $H \in \mathbb{R}^{N \times N}$ représentant une version discrète de H , on effectue le changement de base :

$$\Theta = \Psi^* H \Psi, \quad (4.9)$$

où $\Psi \in \mathbb{R}^{N \times N}$ est une transformée isotrope en ondelettes séparables discrète. De la même façon que dans le cas continu, la matrice Θ est concentrée autour de la diagonale des sous-bandes d'ondelette (voir la figure 4.2).

Malheureusement, le calcul de Θ est très coûteux, ce qui peut expliquer que nous ayons été les premiers à étudier ce genre de décomposition pour le traitement d'images. Calculer l'ensemble des coefficients $(\langle H\psi_\lambda, \psi_\mu \rangle)_{\lambda,\mu}$ est un algorithme en $O(N^3)$! On doit d'abord appliquer H à chacune des N ondelettes discrètes. Le coût d'un produit $H\psi_\lambda$ est $O(N^2)$ pour une matrice H arbitraire de taille $N \times N$. Calculer $(\langle H\psi_\lambda, \psi_\mu \rangle)_{\lambda,\mu}$ pour tout μ et un λ fixé coûte $O(N)$ opérations. La complexité globale est donc dominée par le calcul des N vecteurs $H\psi_\lambda$.

En pratique, il est donc difficile de calculer Θ pour de grandes matrices. Ce résultat peut tout de même servir lorsqu'on doit calibrer un appareil une fois pour toute. Dans ce cas, on peut effectuer un calcul très lourd une bonne fois

pour toute, puis utiliser une représentation compressée par la suite. Ce principe est similaire aux méthodes “reduced order basis” par exemple, qui permettent de faire de la simulation de fluides sur smartphone. On verra aussi que les temps de calcul sont fortement réduits pour des opérateurs de *convolution* (voir le paragraphe 4.2.2).

Illustration

Pour illustrer ces résultats, on considère l'opérateur 1D de noyau

$$K(x, y) = \frac{1}{\sigma(y)\sqrt{2\pi}} \exp\left(-\frac{(x-y)^2}{2\sigma^2(y)}\right), \quad (4.10)$$

où $\sigma(y) = 4 + 10y$. Toutes les PSFs sont des gaussiennes de variance variant linéairement. La matrice est représentée en échelle linéaire (resp. échelle log scale) sur la figure 4.2 haut-gauche (resp. haut-droit). La matrice exprimée dans le domaine des ondelettes est significativement plus parcimonieuse que dans l'espace. Elle a une structure particulière par bande capturée par le théorème 14.

Décomposition de convolutions

Lorsque H est un produit de convolution avec un noyau h , le calcul de la représentation devient très rapide. On présente ci-dessous des résultats dans un cadre discret, bien que l'extension au cas continu soit directe. La matrice Θ peut être décomposée en sous-bandes :

$$\Theta = \left(\Theta_{j,k}^{e,e'}\right)_{j,k}, \quad \text{avec } \Theta_{j,k}^{e,e'} = \left(\left\langle H\psi_{j,m}^e, \psi_{k,n}^{e'} \right\rangle\right)_{m \in \mathcal{T}_j, n \in \mathcal{T}_k}. \quad (4.11)$$

Par exemple, pour un signal 1D avec $J = 2$, la matrice Θ peut être décomposée comme sur la figure 4.3, gauche. On décrit ci-dessous la structure des sous-bandes $\Theta_{j,k}^{e,e'}$.

Définition 6 (Opérateur de translation). *Soit $a \in \mathbb{R}^N$ une image à d -dimensions et $m \in \mathbb{Z}^d$ un paramètre de translation. L'image translatée $b = \tau_m(a)$ est définie pour tout i_1, \dots, i_d par :*

$$b[i_1, \dots, i_d] = a[i_1 - m_1, \dots, i_d - m_d] \quad (4.12)$$

avec des conditions de bord circulaires.

Définition 7 (Matrices rectangulaires circulantes). *Soit $A \in \mathbb{R}^{2^j \times 2^k}$ une matrice rectangulaire. Elle est dite circulante si et seulement si :*

- Si $k \geq j$: il existe $a \in \mathbb{R}^{2^k}$, tel que pour tout $0 \leq l \leq 2^j - 1$,

$$A[l, :] = \tau_{2^k - j_l}(a).$$

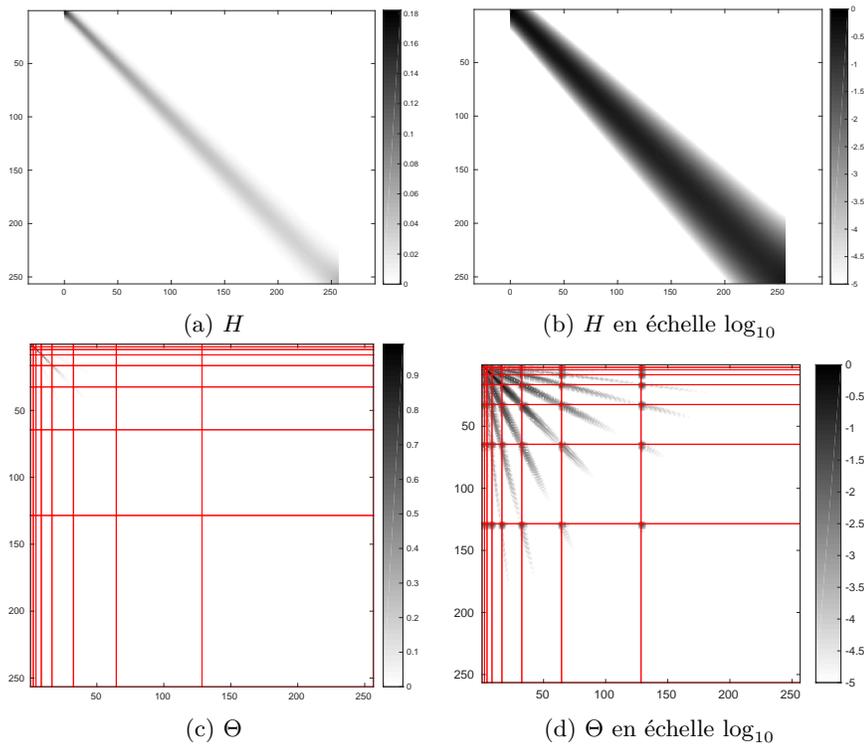


FIGURE 4.2 – Illustration de la compression d'opérateurs de flou variable dans le domaines des ondelettes. On utilise ici des ondelettes de Daubechies avec 10 moments nuls et une décomposition à l'échelle $J = 7$.

— Si $k < j$: il existe $a \in \mathbb{R}^{2^j}$, tel que pour tout $0 \leq l \leq 2^k - 1$,

$$A[:, l] = \tau_{2^j - k_l}(a).$$

Voici un exemple de matrice circulante 4×8 :

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\ a_7 & a_8 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ a_5 & a_6 & a_7 & a_8 & a_1 & a_2 & a_3 & a_4 \\ a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_1 & a_2 \end{pmatrix}.$$

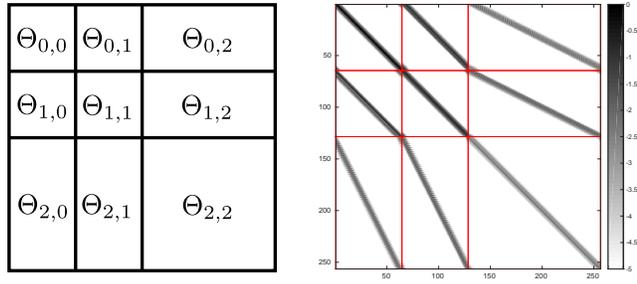


FIGURE 4.3 – Gauche : structure de Θ . Droite : Θ en échelle \log_{10} quand H est une convolution avec un noyau gaussien.

Le théorème 16 indique que toutes les sous-bandes de Θ sont circulantes pour des opérateurs de convolution. C'est illustré sur la figure 4.3, droite.

Théorème 16 (Structure circulante de Θ). *Soit H une matrice de convolution et $\Theta = \Psi^* H \Psi$ sa représentation en ondelettes. Alors, pour tout $j, k \in [0, J]$ et $e, e' \in \{0, 1\}^d$, les sous-matrices $\Theta_{j,k}^{e,e'}$ sont circulantes.*

D'un point de vue pratique la conséquence du théorème 16 est que pour obtenir Θ , il suffit de calculer une ligne et une colonne de chaque matrice $\Theta_{j,k}^{e,e'}$. On peut le faire en calculant $(2^d - 1)J$ transformées en ondelettes (voir l'algorithme 3). La complexité est réduite à $O((2^d - 1)JN)$ opérations à la place de $O(N^3)$ opérations pour des opérateurs génériques.

Stratégies de seuillage

Le théorème 15 indique qu'on peut obtenir de bonnes approximations parcimonieuses de Θ . Cependant, la stratégie de seuillage indiquée par le théorème n'est pas efficace. Dans cette partie, on propose donc des stratégies plus efficaces. On considère la norme d'opérateur suivante :

$$\|H\|_{X \rightarrow Y} = \sup_{\|u\|_X \leq 1} \|Hu\|_Y, \quad (4.13)$$

Algorithm 3 Un algorithme de calcul de Θ pour des opérateurs de convolution

- ① **Entrée** : $h \in \mathbb{R}^N$, le noyau de convolution de H .
 - ② **Sortie** : Θ , la représentation en ondelettes de H
 - ③ **for** $(j, e) \in [0, J] \times \{0, 1\}^d$ **do**
 - ④ Calculer l'ondelettes ψ_λ avec $\lambda = (j, e, 0)$.
 - ⑤ Calculer l'ondelette floue $H\psi_\lambda$ et $H^*\psi_\lambda$.
 - ⑥ Calculer $(\langle H\psi_\lambda, \psi_\mu \rangle)_\mu$ en utilisant une transformée en ondelette.
 - ⑦ Calculer $(\langle H^*\psi_\lambda, \psi_\mu \rangle)_\mu$ en utilisant une transformée en ondelette.
 - ⑧ **for** $(k, e') \in [0, J] \times \{0, 1\}^d$ **do**
 - ⑨ **if** $k \geq j$ **then**
 - ⑩ $\Theta_{j,k}^{e,e'}$ est la matrice circulante de colonne : $(\langle H\psi_\lambda, \psi_{k,n}^{e'} \rangle)_n$
 - ⑪ **else**
 - ⑫ $\Theta_{j,k}^{e,e'}$ est la matrice circulante de ligne : $(\langle H^*\psi_\lambda, \psi_{k,n}^{e'} \rangle)_n =$
 $(\langle \psi_\lambda, H\psi_{k,n}^{e'} \rangle)_n$
 - ⑬ **end if**
 - ⑭ **end for**
 - ⑮ **end for**
-

où $\|\cdot\|_X$ et $\|\cdot\|_Y$ sont deux normes sur \mathbb{R}^N . Pour obtenir une représentation K -parcimonieuse Θ_K de Θ on peut résoudre :

$$\min_{\Theta_K, K\text{-sparse}} \|H_K - H\|_{X \rightarrow Y}, \quad (4.14)$$

où $H_K = \Psi\Theta_K\Psi^*$. L'approche la plus naïve de calcul de Θ_K est la suivante :

$$\Theta_K[\lambda, \mu] = \begin{cases} \Theta[\lambda, \mu] & \text{si } |\Theta[\lambda, \mu]| \text{ appartient au } K \text{ plus grandes valeurs de } |\Theta|, \\ 0 & \text{sinon.} \end{cases} \quad (4.15)$$

Cette stratégie correspond à la minimisation de (4.14), en posant $\|\cdot\|_X = \|\Psi^*\cdot\|_1$ et $\|\cdot\|_Y = \|\Psi^*\cdot\|_\infty$.

La norme ℓ^1 des coefficients d'ondelettes n'est pas adaptée à la description des images. Les images naturelles sont souvent modélisées comme des éléments d'espaces de Besov ou de fonctions à variation bornée [Coh+99; AK06]. Ces espaces peuvent être caractérisés par la décroissance de leurs coefficients d'ondelettes [Coh03] à travers les sous-bandes. Cette observation motive le choix $\|\cdot\|_X = \|\Sigma\Psi^*\cdot\|_1$ où $\Sigma = \text{diag}(\sigma) \in \mathbb{R}^{N \times N}$ est une matrice diagonale et où $\sigma \in \mathbb{R}^N$ est constant par niveau. On est ainsi ramenés au problème suivant :

$$\min_{\Theta_K, K\text{-sparse}} \sup_{\|\Sigma x\|_1 \leq 1} \|(\Theta - \Theta_K)x\|_\infty. \quad (4.16)$$

Sa solution est donnée par :

$$\Theta_K[\lambda, \mu] = \begin{cases} \Theta[\lambda, \mu] & \text{if } |\sigma_\mu \Theta[\lambda, \mu]| \text{ est parmi les } K \text{ plus grandes valeurs de } |\Theta\Sigma|, \\ 0 & \text{sinon.} \end{cases} \quad (4.17)$$

Les poids σ_μ doivent être adaptés à la classe d'image à reconstruire. En pratique le choix $\sigma_\mu = 2^{-k}$ pour $\mu = (k, e', n)$ semble bon. Ils peuvent aussi être appris à partir d'un ensemble d'images représentatives d'une classe. On a aussi proposé des algorithmes gloutons pour $\|\cdot\|_Y = \|\cdot\|_2$ dans [26], mais les résultats ne semblent pas meilleurs que la stratégie simple ci-dessus. L'importance de la stratégie de seuillage pour la résolution de problèmes inverses est illustrée sur la figure 4.4.

Application à la restauration temps-réel

Supposons qu'on observe une image u_0 suivant le modèle suivant

$$u_0 = Hu + b, \quad (4.18)$$

où $u \in \mathbb{R}^N$ est l'image nette à retrouver, $b \sim \mathcal{N}(0, \sigma^2 I_N)$ un bruit Gaussien déviation standard σ et $H \in \mathbb{R}^{N \times N}$ est un opérateur de flou connu.

Une formulation variationnelle standard pour recouvrer u consiste à résoudre

$$\min_{x \in \mathbb{R}^N} E(x) = F(x) + G(x). \quad (4.19)$$

Dans cette équation, $F(x)$ est un terme d'attache aux données quadratique :

$$F(x) = \frac{1}{2} \|Ax - u_0\|_2^2. \quad (4.20)$$

Le terme de régularisation $G(x)$ est défini par :

$$G(x) = \|x\|_{1,w} = \sum_{i=1}^N w[i] |x[i]|. \quad (4.21)$$

Le vecteur de poids $w \in \mathbb{R}_+^N$ est un paramètre de régularisation qui peut varier suivant les sous-bandes d'ondelette. Le problème (4.19) consiste à trouver une image Ψx consistante avec les données observées u_0 avec une représentation x parcimonieuse.

Ce type d'énergie est populaire [FN03; NCB04; VU08; VU09], bien que des résultats plus convaincants soient obtenus avec des trames d'ondelettes Ψ telles que des transformées en ondelettes redondantes [SNM03; Cha+07]. Nous nous focalisons ici sur le cas d'une transformation Ψ orthogonale, ce qui va nous permettre de définir des algorithmes bien plus rapides.

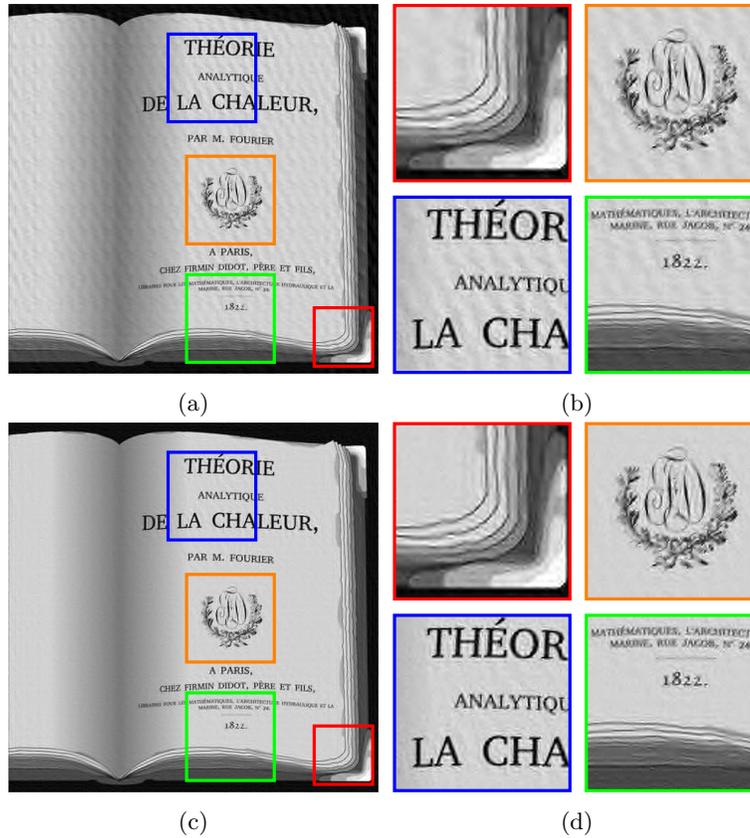


FIGURE 4.4 – Un exemple de déconvolution pour illustrer l'importance de la stratégie de seuillage. Les matrices Θ ont été construites avec le même nombre de coefficients, correspondant à 57 opérations par pixel. En haut : résultat avec une stratégie de seuillage naïve, pSNR = 23.71dB. En bas : une stratégie de seuillage à poids pSNR = 24.07dB. Le flou est celui de la figure 4.5a.

Algorithm 4 Descente proximale accélérée

- ① **Entrée** : Un point de départ $x^{(0)} = y^{(1)}$, $\tau = 1/\|A\|_2^2$ et *Nit*.
 - ② **for** $k = 1$ à *Nit* **do**
 - ③ Calculer $\nabla F(y^{(k)}) = A^*(Ay^{(k)} - u_0)$. ▷ 99.35''
 - ④ $x^{(k)} = \text{Prox}_{\tau G}(y^{(k)} - \tau \nabla F(y^{(k)}))$. ▷ 2.7''
 - ⑤ $y^{(k+1)} = x^{(k)} + \frac{k-1}{k+2}(x^{(k)} - x^{(k-1)})$. ▷ 1.1''
 - ⑥ **end for**
-

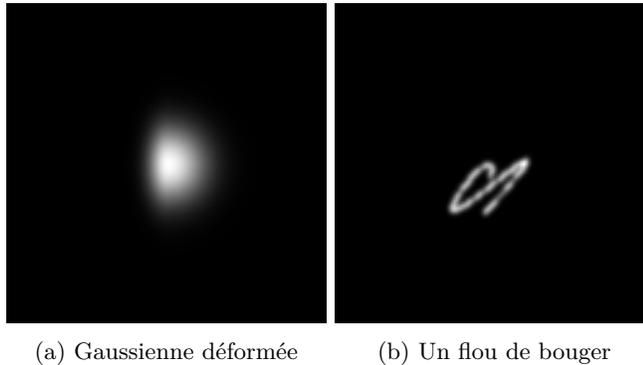


FIGURE 4.5 – Exemples de PSFs.

Algorithmes d'optimisation standards Pour résoudre le problème convexe (4.19), on peut utiliser des algorithmes de descente proximaux, voir par exemple [BT09b; CP11b]. Un exemple populaire est la descente proximale accélérée décrite dans l'algorithme 4.

On suppose que H est un noyau de convolution : $Hu = h \star u$, où h est le flou de bouger de la figure 4.5b. En pratique, l'image se stabilise au bout de 500 itérations. Le temps de calcul sur une station de travail est de 103'' et le résultat est montré sur la figure 4.6. Le profilage du code indique les temps de calcul indiqués sur la droite de l'algorithme 4. On peut voir que 96% du temps est passé à évaluer le gradient, ce qui nécessite le calcul de deux transformées de Fourier et de deux transformées en ondelette rapide. Pour réduire les temps de calcul, on peut donc chercher à

- Accélérer le calcul du gradients.
- Utiliser des algorithmes de minimisation plus sophistiqués.

Le deuxième point a déjà été longuement abordé dans la littérature, mais le premier me semble bien plus original. De plus, on va voir que les décompositions en ondelettes permettent de résoudre les deux.

Les idées principales Dans [14], nous avons proposé 3 idées pour accélérer les calculs. Premièrement, la fonction F dans l'équation (4.20) peut être approchée par une fonction F_K telle que ∇F_K soit peu coûteux. Deuxièmement, on caractérise précisément la structure de la Hessienne de F_K , pour pouvoir construire des préconditionneurs efficaces. Enfin, on implémente les algorithmes sur des cartes graphiques (GPU).

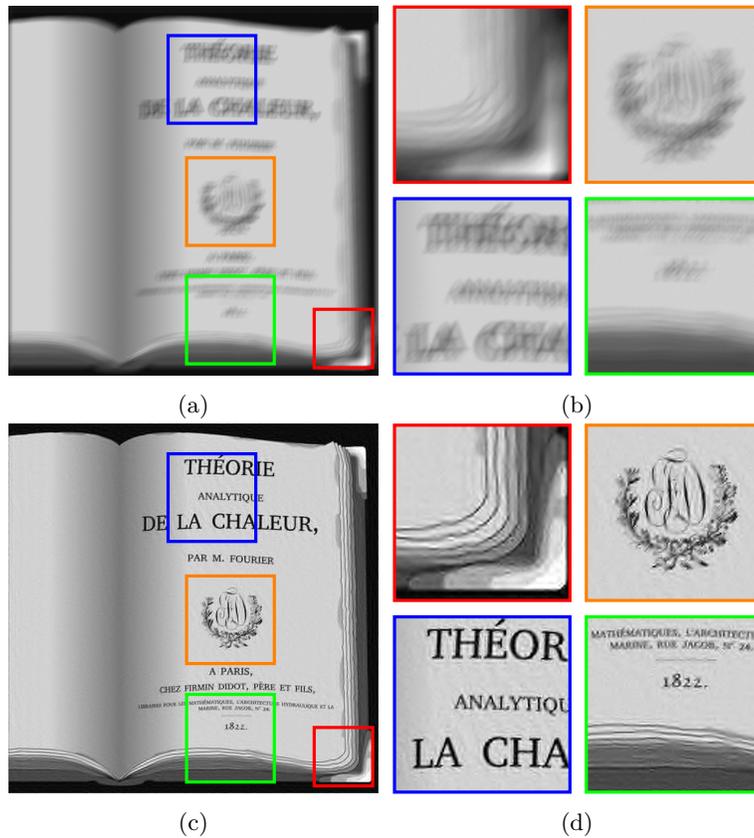


FIGURE 4.6 – Un exemple de déconvolution. Une image de livre est floutée avec un flou de bouger. Le pSNR de l'image dégradée est 17.85dB. Le problème (4.19) est résolu avec un opérateur exact, 500 itérations et des ondelettes Symmlet 6 décomposées 6 fois. Le pSNR de l'image restaurée est 24.14dB.

En exploitant l'orthogonalité de Ψ , on peut écrire :

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|H\Psi x - u_0\|_2^2 + \|x\|_{1,w} \quad (4.22)$$

$$= \min_{x \in \mathbb{R}^N} \frac{1}{2} \|\Psi^*(H\Psi x - u_0)\|_2^2 + \|x\|_{1,w} \quad (4.23)$$

$$= \min_{x \in \mathbb{R}^N} \frac{1}{2} \|\Theta x - x_0\|_2^2 + \|x\|_{1,w}, \quad (4.24)$$

où $x_0 = \Psi^* u_0$ est la décomposition en ondelettes de u_0 . Le problème (4.24) est défini uniquement dans le domaine des ondelettes, contrairement au problème (4.19). On approche ensuite (4.24) par

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|\Theta_K x - x_0\|_2^2 + \|x\|_{1,w}. \quad (4.25)$$

où Θ_K est une version tronquée de Θ . En posant $F_K(x) = \frac{1}{2} \|\Theta_K x - x_0\|_2^2$, le gradient de F_K est

$$\nabla F_K(x) = \Theta_K^* (\Theta_K x - x_0), \quad (4.26)$$

ce qui peut être calculé par du calcul des produits matriciels parcimonieux. Le calcul du gradient approché (4.26) coûte en général bien moins cher que le calcul exact $\nabla F(x)$. Dans les cas favorables de noyaux de convolution lisses, l'évaluation de $\nabla F_K(x)$ demande environ 2 opérations par pixel, ce qui peut réduire les temps de calcul d'un facteur 10 par rapport aux transformées rapides usuelles !

On peut ensuite remarquer que la Hessienne $H_{F_K}(x) = \Theta_K^* \Theta_K$ a une structure quasi diagonale, avec des coefficients diagonaux qui décroissent rapidement. On peut donc définir des *préconditionneurs diagonaux* efficaces, ce qui réduit le nombre d'itérations d'une méthode itérative. En pratique, le préconditionnement accélère le calcul d'un facteur 2 à 5.

Finalement, le calcul massivement parallèle sur une (petite) carte NVIDIA K20c accélère encore le temps de calcul d'un facteur 10. En combinant ces trois idées, on obtient des accélérations qui varient d'un facteur 4 à un facteur 250 pour des problèmes de déconvolution. *De telles facteurs font vraiment la différence entre une méthode utilisable en routine et une méthode à portée essentiellement académique.* Ils augmentent avec la résolution de l'image et la régularité du noyau h . Notons aussi que toutes ces idées s'appliquent directement aux flous variables, avec des gains encore plus significatifs. Une petite vidéo illustre ces résultats ici <https://youtu.be/oHnGNRc9Qeo>.

4.2.3 Décomposition en convolution-produit

La décomposition dans des bases d'ondelettes présente de nombreux avantages, mais possède plusieurs faiblesses :

- Il est difficile de calculer la matrice Θ
- Il est difficile d'interpréter les coefficients.

— La représentation semble peu adaptée à l'identification d'opérateurs à partir de données indirectes.

Face à ces faiblesses, nous avons décidé d'analyser une technique que nous avons nommé convolution-produit avec Paul Escande [20]. Ce type de décomposition a une longue histoire dans différentes disciplines [TF92 ; NO98 ; FR05 ; GV06 ; Hry+10 ; Hir+10 ; Den+15], mais à ma connaissance, ses propriétés théoriques d'approximation n'avaient pas été étudiées. C'est maintenant chose faite !

Soit H un opérateur intégral de la forme 5. Une approximation par convolution-produit H_m consiste à approcher Hu par la formule

$$Hu \simeq H_m u \stackrel{\text{def.}}{=} \sum_{k=1}^m u_k \star v_k \odot u, \quad (4.27)$$

où $u_k \in L^2(\Omega)$ est un noyau de convolution et $v_k \in L^2(\Omega)$ est une fonction de multiplication. Le symbole \odot représente la multiplication terme à terme pour des matrices ou un simple produit pour des fonctions. On note K_m le noyau intégral de l'opérateur H_m . L'observation suivante est centrale pour comprendre l'intérêt des convolution-produit.

Proposition 7. *Soit $S_m(x, y)$ la réponse impulsionnelle variable (RIV) de H_m définie pour tout $(x, y) \in \Omega \times \Omega$ par*

$$S_m(x, y) = K_m(x + y, y). \quad (4.28)$$

Si H_m est de la forme (4.27), alors sa RIV s'écrit

$$S_m(x, y) = \sum_{k=1}^m u_k(x)v_k(y). \quad (4.29)$$

Effectuer une décomposition en convolution-produit revient donc à approcher la RIV par un tenseur de rang m .

De plus la structure de convolution-produit est favorable au calcul numérique puisqu'elle peut être effectuée avec des transformées de Fourier rapides. Dans le papier [20], nous étudions précisément la complexité numérique de ces décompositions pour différents choix des fonctions u_k et v_k . Je ne tiens pas à décrire ici tous les résultats d'approximation. L'image 4.7 peut cependant aider à comprendre en un clin d'œil l'intérêt de ces décompositions pour des chercheurs habitués aux matrices hiérarchiques.

Je donne juste un résultat élémentaire qui sera utilisé dans la partie suivante.

Définition 8. *Soit $(\phi_k)_{k \in \mathbb{N}}$ une base orthogonale de $L^2(\Omega)$. On définit un sous-espace \mathcal{E}^r comme l'ensemble des fonctions $u \in L^2(\Omega)$ telles que*

$$\|u\|_{\mathcal{E}^r}^2 \stackrel{\text{def.}}{=} \sum_{k \in \mathbb{N}} w[k] |\langle u, \phi_k \rangle|^2 < +\infty, \quad (4.30)$$

avec $w[k] = (1 + k^2)^{r/d}$.

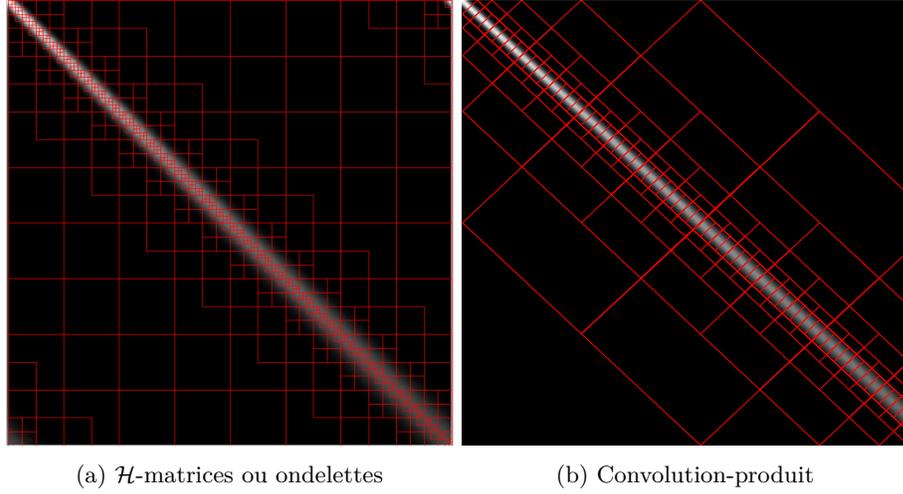


FIGURE 4.7 – La décomposition en convolution-produit permet de décomposer le domaine de façon plus adaptée à la structure des opérateurs.

Ce genre d'ensemble est standard en théorie de l'approximation. Il permet par exemple de redéfinir les espaces de Sobolev $W^{r,2}$, si la base ϕ_k est une base d'ondelettes. Il permet notamment d'établir le résultat suivant.

Proposition 8. Soit $f \in \mathcal{E}^r$ et $f_N = \sum_{k=1}^N \langle f, \phi_k \rangle \phi_k$. Alors

$$\|f - f_N\|_{L^2}^2 \leq c \|f\|_{\mathcal{E}^r}^2 N^{-2r/d} \quad (4.31)$$

où c est une constante universelle.

Une conséquence immédiate est la suivante.

Proposition 9. Soit S la RIV d'un opérateur H telle que

$$S(\cdot, y) \in \mathcal{E}^r, \forall y \in \Omega. \quad (4.32)$$

En choisissant $u_k = \phi_k$ et $v_k(y) = \langle S(\cdot, y), \phi_k \rangle$ dans l'équation (4.27), l'opérateur H_m de noyau K_m satisfait

$$\|K - K_m\|_{L^2(\Omega \times \Omega)}^2 = O\left(\|K\|_{L^2(\Omega \times \Omega)}^2 m^{-2r/d}\right). \quad (4.33)$$

Cette proposition indique que si les réponses impulsionnelles du système peuvent être exprimées simultanément dans une même base (ϕ_k) , alors l'approximation par convolution-produit devrait bien fonctionner. C'est une hypothèse raisonnable en optique qu'on va exploiter pour identifier un opérateur à partir de quelques PSF.

4.3 Identification d'opérateurs

Dans cette partie, je résume brièvement les résultats principaux de [17]. Comme on peut le voir dans la figure 4.1, il est assez facile d'observer les PSFs d'un système optique en quelques points du champs de vue. On s'est donc posé la question suivante.

Question 4. *Comment estimer un opérateur H à partir de son action sur quelques masses de Dirac δ_{y_i} ?*

4.3.1 Hypothèses

Pour pouvoir construire un estimateur, il est nécessaire de supposer une certaine régularité sur les PSFs, comme dans la proposition 9, mais aussi sur les variations des PSFs d'un point à un autre de l'espace. Cette remarque motive d'introduire l'espace suivant.

Définition 9. *Soient A_1 et A_2 deux constantes positives. Soit $r > \frac{d}{2}$ et $s > \frac{d}{2}$. La boule $\mathcal{E}^{r,s}(A_1, A_2)$ est définie comme l'ensemble des opérateurs linéaires intégraux H dont la RIV S appartient à $L^2(\mathbb{R}^d \times \mathbb{R}^d)$ avec :*

$$\text{Variation régulières} \quad \int_{x \in \mathbb{R}^d} \|S(x, \cdot)\|_{H^s(\mathbb{R}^d)}^2 dx \leq A_1 \quad (4.34)$$

$$\text{Régularité des PSFs} \quad \int_{y \in \mathbb{R}^d} \|S(\cdot, y)\|_{\mathcal{E}^r(\mathbb{R}^d)}^2 dy \leq A_2 \quad (4.35)$$

On suppose qu'on observe n PSFs à travers leur projection sur un l'espace de dimension N défini par $\text{vect}(\phi_1, \dots, \phi_N)$. On a donc accès à une famille de vecteurs $(F_i^\epsilon)_{1 \leq i \leq n}$ définie pour tout $k \in \{1, \dots, N\}$ par

$$F_i^\epsilon[k] = \langle S(\cdot, y_i), \phi_k \rangle + \epsilon_i[k], \quad 1 \leq i \leq n, \quad (4.36)$$

où ϵ_i est un vecteur aléatoire gaussien dont les composantes sont indépendantes, de moyenne nulle et de variance σ^2 .

4.3.2 Un estimateur

On se propose de construire un estimateur de la fonction $F : \Omega \rightarrow \mathbb{R}^N$ définie pour tout $k \in \mathbb{N}$ par

$$F(y)[k] = \langle S(\cdot, y), \phi_k \rangle. \quad (4.37)$$

Cette fonction est appelée IRC pour Impulse Response Coefficients. Elle permet de reconstruire intégralement la RIV S (car (ϕ_k) est une base orthogonale) et par suite l'opérateur H .

Définition 10 (IRC admissibles). Soit $\Omega \subseteq \mathbb{R}^d$ un ensemble, $N \leq \infty$ un entier et $\alpha \in]0, 1]$ un réel. L'espace $\mathcal{H}(\Omega)$ des IRC admissibles est défini comme l'ensemble des fonctions à valeurs vectorielles $G : \Omega \rightarrow \mathbb{R}^N$ telles que

$$\|G\|_{\mathcal{H}(\mathbb{R}^d)} \stackrel{\text{def.}}{=} \alpha \int_{y \in \Omega} \sum_{k=1}^N w[k] |G(y)[k]|^2 + (1 - \alpha) \sum_{k=1}^N |G(\cdot)|_{H^s}^2 < +\infty, \quad (4.38)$$

où $|\cdot|_{H^s}$ est la semi-norme H^s .

On construit un estimateur \hat{F} de F en résolvant le problème variationnel suivant :

$$\hat{F}_\mu = \arg \min_{F \in \mathcal{H}(\mathbb{R}^d)} \frac{1}{n} \sum_{i=1}^n \|F_i^\epsilon - F(y_i)\|_{\mathbb{R}^N}^2 + \mu \|F\|_{\mathcal{H}(\mathbb{R}^d)}. \quad (4.39)$$

Cette formulation peut être interprétée avec le formalisme de la régression dans des espaces de Hilbert à noyau reproduisant vectoriels [MP04; MP05].

La solution du problème (4.39) peut être calculée très efficacement grâce à des techniques d'approximation reposant sur les fonctions radiales [Wen04].

Proposition 10. La solution du problème (4.39) peut être calculée en $O(Nn^3)$ opérations.

Je ne rentre pas dans les détails ici par souci de concision. La figure 4.8 illustre la méthodologie proposée pour une base (ϕ_k) d'ondelettes orthogonales.

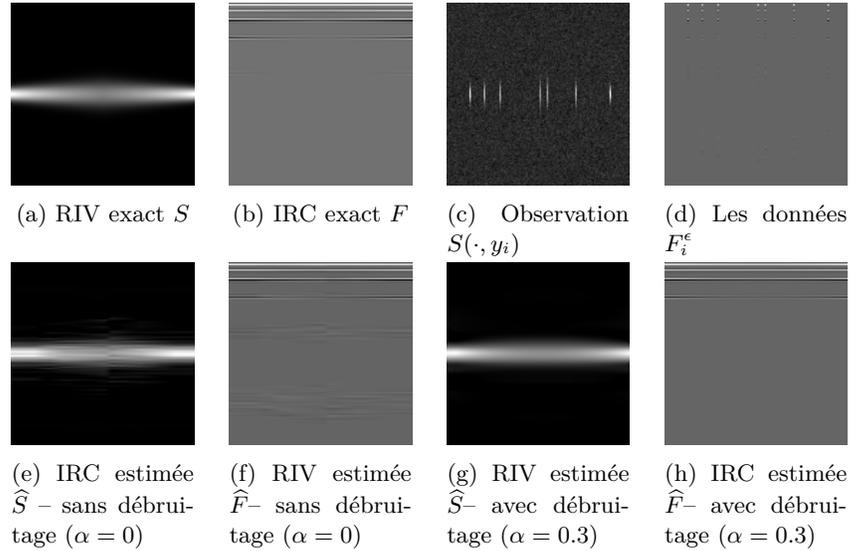


FIGURE 4.8 – Illustration de la méthodologie et d'un résultat sur un problème d'estimation 1D.

4.3.3 Garanties théoriques

Nous donnons maintenant quelques garanties théoriques de notre estimateur. On suppose que les points de mesure (y_i) sont des éléments d'un compact Ω .

Définition 11 (Distance de remplissage). *La distance de remplissage de $Y = \{y_1, \dots, y_n\} \subset \Omega$ est définie par*

$$h_{Y,\Omega} = \sup_{y \in \Omega} \min_{1 \leq j \leq n} \|y - y_j\|_2. \quad (4.40)$$

Définition 12 (Distance de séparation). *La distance de séparation de $Y = \{y_1, \dots, y_n\} \subset \Omega$ est définie par*

$$q_{Y,\Omega} = \frac{1}{2} \min_{i \neq j} \|y_i - y_j\|_2. \quad (4.41)$$

La condition suivante [NW91] est centrale.

Définition 13 (Condition de quasi-uniformité). *Un ensemble $Y = \{y_1, \dots, y_n\} \subset \Omega$ est dit quasi-uniforme par rapport à une constante $B > 0$ si*

$$q_{Y,\Omega} \leq h_{Y,\Omega} \leq Bq_{Y,\Omega}. \quad (4.42)$$

Nous sommes prêts à établir un théorème de convergence.

Théorème 17. *Supposons que $H \in \mathcal{E}^{r,s}(A_1, A_2)$ et que sa RIV S soit échantillonnée suivant la condition de quasi-uniformité de la définition 13. L'estimateur \hat{H} de RIV \hat{S} définie dans l'équation (4.28) satisfait*

$$\mathbb{E} \left(\|H - \hat{H}\|_{HS}^2 \right) \lesssim N^{-\frac{2r}{d}} + (N\sigma^2 n^{-1})^{\frac{2s}{2s+d}} (1 - \alpha)^{-\frac{2d}{2s+d}}, \quad (4.43)$$

pour $\mu \propto (N\sigma^2 n^{-1})^{\frac{2s}{2s+d}} (1 - \alpha)^{\frac{2s-d}{2s+d}}$.

En particulier, si $\mu \propto (\sigma^2 n^{-1})^{\frac{2q}{2q+d}} (1 - \alpha)^{\frac{2s-d}{2s+d}}$ et $N \propto (\sigma^2 n^{-1})^{-\frac{dq}{r(2q+d)}} (1 - \alpha)^{\frac{d^2(r+s)}{2q+d}}$, l'inégalité suivante est valide

$$\mathbb{E} \left(\|H - \hat{H}\|_{HS}^2 \right) \lesssim (\sigma^2 n^{-1} (1 - \alpha)^{-d/s})^{\frac{2q}{2q+d}}, \quad (4.44)$$

avec $1/q = 1/r + 1/s$.

De plus, sous des conditions techniques supplémentaires, nous avons montré que l'estimateur \hat{H} était minimax, ce qui signifie qu'on ne peut pas construire de meilleur estimateur, uniformément sur la classe $\mathcal{E}^{r,s}(A_1, A_2)$.

4.3.4 Résultats

Nous utilisons aujourd'hui cette méthodologie très fréquemment pour des applications en microscopie. La méthode est très efficace et fonctionne quasiment en temps réel. La figure 4.9c présente une image de micro-billes fluorescentes

avec un microscope champs large ainsi que des zooms sur quelques PSFs. La figure 4.9f montre un estimateur de l'opérateur complet. Pour le représenter, on l'applique à un peigne de Dirac. On peut ensuite utiliser cet estimateur pour effectuer du défloutage en supposant que les conditions d'imagerie n'ont pas changé entre la calibration et l'image. La figure 4.10f présente un résultat de défloutage à partir de l'estimateur \hat{H} . Noter l'amélioration de la qualité d'image uniformément sur l'image.

4.4 Ouverture

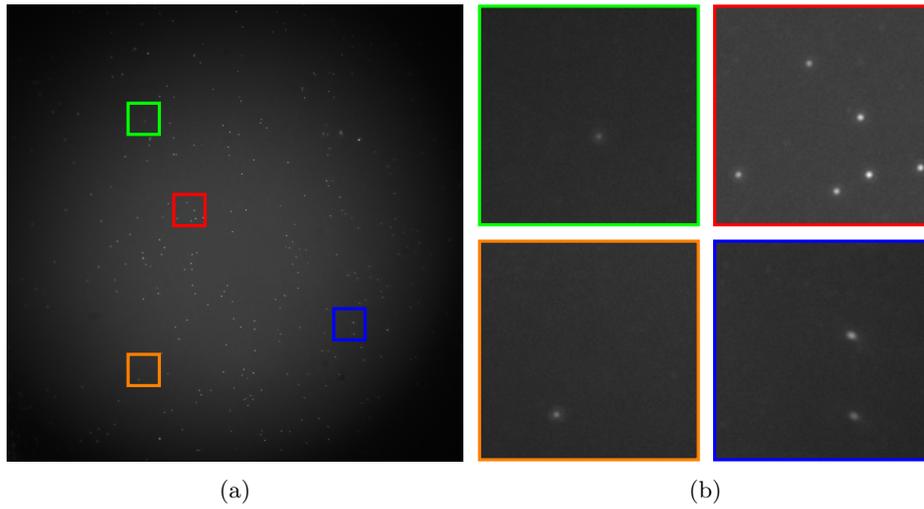
J'ai envie de continuer à explorer cette thématique de recherche dans les années à venir : elle me semble en effet presque vierge et il reste beaucoup de pistes qui ne demandent qu'à être tracées. De plus, j'ai trouvé un collaborateur opticien talentueux en la personne de Thomas Mangeat. Cette entente devrait permettre de valoriser les travaux assez théoriques dans des communautés plus appliquées. J'entends quelque fois des retours du type : est-ce que la portée de ce problème n'est pas limitée ? Je ne pense réellement pas. D'un point de vue pratique, l'augmentation de résolution est un problème critique en optique et beaucoup de méthodes populaires sont heuristiques et me semblent sous-optimales. La compréhension un peu plus fine de la structure des opérateurs me semble aussi essentielle pour faire de réels progrès dans l'accélération des méthodes de résolution de certains problèmes inverses. Finalement, avoir accès à une gamme plus étendue d'opérateurs instationnaires peut permettre d'enrichir des structures telles que les réseaux de neurones convolutionnels.

4.4.1 *Décomposition rapide d'opérateurs*

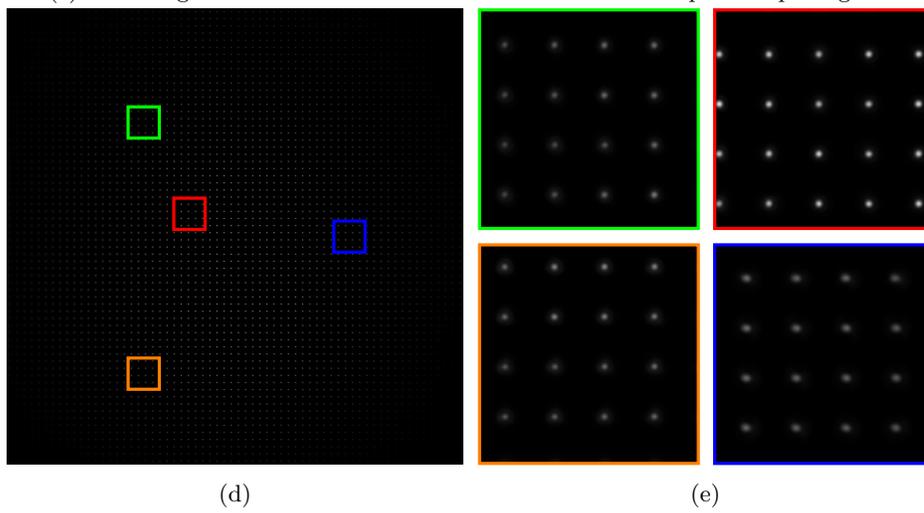
Comme on l'a vu dans la section 4.2.2, la décomposition d'un opérateur dans une base d'ondelettes est difficile numériquement, à part pour les opérateurs de convolution. Nous explorons actuellement la possibilité de *décomposer des opérateurs de convolution-produit rapidement*. Nous pourrions ainsi exploiter simultanément la force des convolution-produit pour l'estimation d'opérateurs et la rapidité des décompositions en ondelettes pour la restauration.

4.4.2 *Apprentissage d'espaces d'opérateurs*

Nous explorons actuellement avec Valentin Debarnot, Thomas Mangeat (ingénieur de recherche CBI) et Paul Escande la possibilité d'estimer un sous-espace d'opérateurs associé à un microscope. L'idée est d'imager des champs de microbilles fluorescentes dans de nombreuses conditions expérimentales différentes (e.g. variations de température, aberrations créées par un SLM, mauvaise focalisation). On peut ensuite utiliser cette information pour identifier un sous-espace d'opérateurs de faible dimension qui capture tous les états possibles d'un microscope. Un tel outil devrait améliorer nettement l'identifiabilité des problèmes de défloutage aveugle.



(c) Une image de micro-billes 2048×2048 avec un microscope champs large.



(f) Estimation d'opérateur à partir de l'image 4.9c.

FIGURE 4.9 – Exemple d'estimation d'opérateur. Collaboration avec Thomas Mangeat.

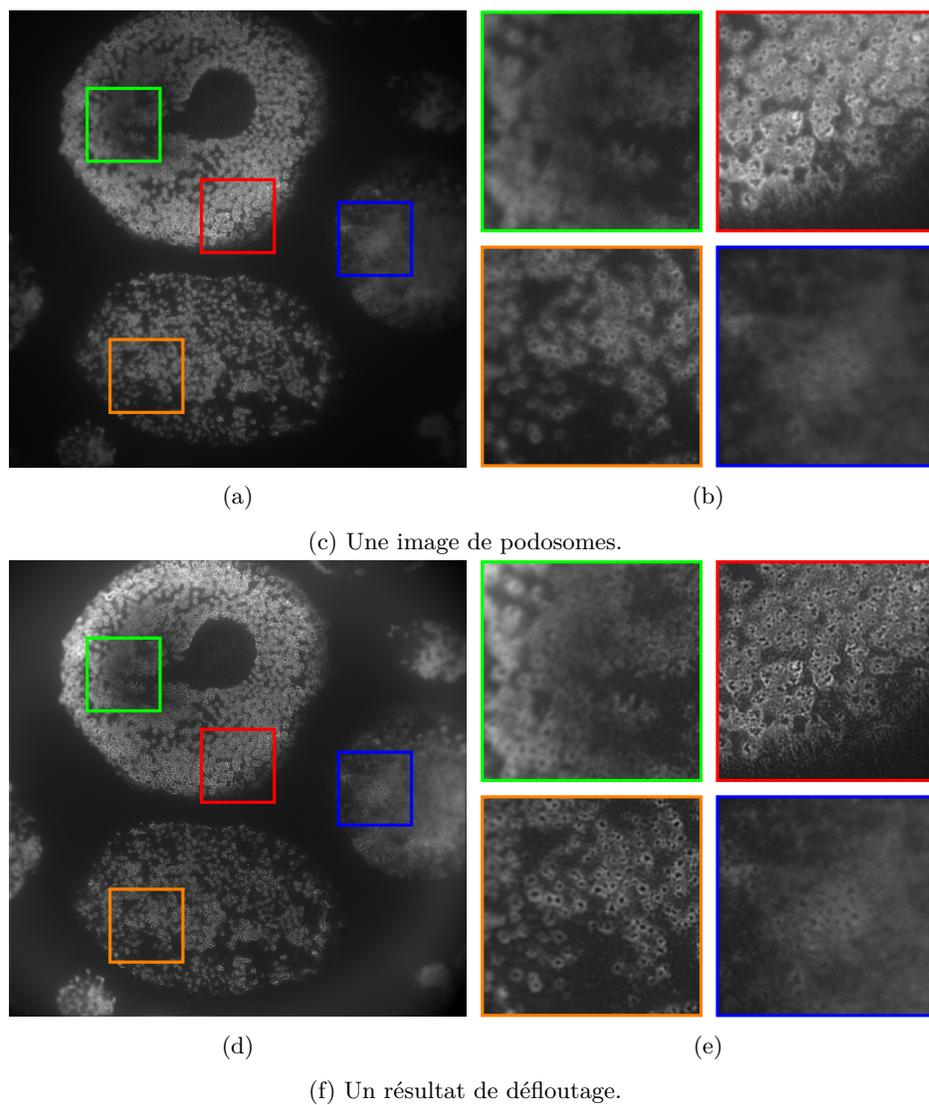


FIGURE 4.10 – Un exemple de défloutage instationnaire en microscopie champs large. Collaboration Thomas Mangeat.

4.4.3 Défloutage aveugle

Je suis de près les évolution récentes en défloutage aveugle [ARR12; Li+18] qui reposent sur des techniques de lifting. J'espère que l'amélioration technique de ces idées, combinées avec l'apprentissage de sous-espaces d'opérateurs, permette de repousser significativement les limites actuelles de résolution de problèmes inverses aveugles. Nous explorons aussi en parallèle des approches plus heuristiques reposant sur l'apprentissage profond. Celles-ci semblent déjà mieux fonctionner que les approches les plus abouties théoriquement, même si nous ne comprenons pas pourquoi.

4.4.4 Noyaux portés par des mesures

L'hypothèse $K \in W^{M,\infty}(\Omega \times \Omega)$ dans la définition 5 est très forte. Elle semble réaliste en optique car la tâche de diffraction par une ouverture circulaire (les tâches de Airy) est C^∞ . Cependant, si on oublie cette diffraction, un flou stationnaire de bouger par exemple est défini par un noyau de convolution qui est une mesure le long d'une courbe. Je pense étudier ces opérateurs dont les noyaux sont des mesures dans mes travaux à venir. C'est un problème d'analyse fonctionnelle qui est presque vierge et qui me semble avoir un intérêt pratique assez fort, par exemple pour mieux comprendre les flous de bouger ou pour mieux comprendre l'échantillonnage en IRM (voir section 5.1.6).

4.4.5 Super-résolution

Finalement, en collaboration avec Thomas Mangeat, nous explorons de nombreuses pistes pour améliorer la résolution des microscopes (champs large, TIRF, SIM, confocal) en exploitant de la métrologie précise des PSFs ainsi que des éléments optiques avancés tels que des SLM. Ce projet est aussi soutenu par de nombreux biologistes qui apportent des problématiques concrètes sur la plateforme d'imagerie du CBI.

Chapitre 5

Imagerie computationnelle

Je décris dans ce chapitre des contributions pratiques en imagerie qui reposent en partie sur les outils développés dans les chapitres précédents.

5.1 Imagerie par Résonance Magnétique

5.1.1 *Introduction*

Ma rencontre avec l'IRM a été fortuite. Lorsque j'étais post-doctorant à Hong-Kong, un autre post-doctorant australien (Michael Poole) m'a contacté pour me demander si je pouvais l'aider à résoudre un problème de conception de bobines. Après des centaines de 'tchat' sur Gmail, je lui ai apporté des connaissances en optimisation convexe et lui sur la physique de l'IRM qui aujourd'hui encore me semble très complexe. Je lui suis très reconnaissant de ces échanges. Par la suite, j'ai été contacté par Philippe Ciuciu qui souhaitait monter une ANR sur la reconstruction d'images en IRM. Ceci m'a permis de continuer à travailler sur ce domaine avec des problèmes de conception de pulses (coll. Nicolas Boulant) et surtout en échantillonnage compressé. Aucun financement n'a malheureusement été accepté sur ce thème, malgré 5 demandes (ANR, Europe).

Je pense aujourd'hui avoir un profil assez rare pour ce domaine, mêlant une connaissance assez bonne de la physique, des mathématiques et du signal. Je pense donc continuer à travailler dans ce domaine riche par la suite, dans un cadre privé ou académique.

Une description précise de la physique de l'IRM me prendrait trop de temps. Je m'excuse d'avance pour un lecteur non averti qui aura probablement du mal à comprendre s'il ne connaît pas la physique suffisamment.

5.1.2 *Conception de bobines*

Mon premier travail dans ce domaine n'est pas du traitement d'images, mais de la conception de bobines qui tient plus à l'électromagnétisme [36 ; 67 ; 66]. La figure 5.1, droite montre comment une bobine de gradient est construite. Nous

avons pu développer des design de bobines qui permettent de réduire significativement les températures maximales dans ces bobines. Une IRM est constituée

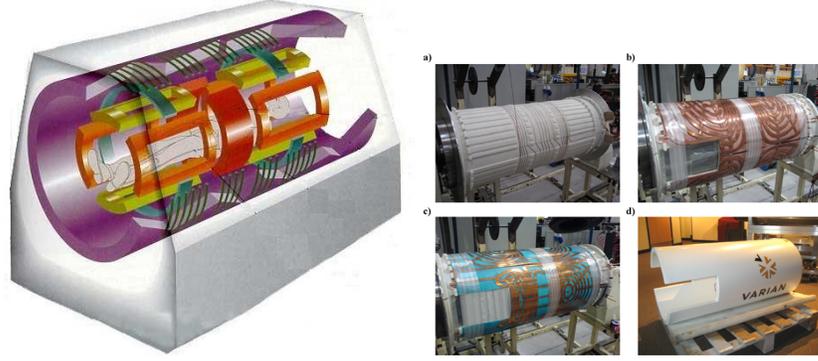


FIGURE 5.1 – A gauche : une IRM est composée de nombreuses bobines imbriquées. Les bobines oranges, rouges et jaunes sont les 3 bobines de gradient. A droite : les étapes de construction d’une bobine de gradient.

de nombreuses bobines (voir figure 5.1, gauche). Nous nous intéressons ici aux 3 bobines de gradient, chacune étant dédiée à une direction de l’espace : X,Y,Z. Une bobine de gradient X a pour rôle de générer un champs magnétique aligné avec le champs principal B_0 . Ce champs a une intensité qui varie linéairement dans la direction X. En faisant varier des courants électriques temporellement dans les bobines X, Y et Z, on peut piloter une trajectoire d’échantillonnage dans l’espace de Fourier.

Le problème qu’on se pose ici est un problème d’optimisation de forme.

Question 5. *Comment trouver une structure de bobine qui - lorsqu’on y fait circuler un courant électrique - génère un champs magnétique donné (ce champs qui varie en X par exemple) ? On souhaite aussi que la structure soit réalisable (e.g. lisse) et qu’elle minimise les échauffements dans la bobine.*

L’optimisation d’une structure est un problème complexe et non convexe en général. Pour ce travail, nous avons choisi une approche de relaxation : à la place d’optimiser une structure (un problème de topologie), on peut optimiser un champ vectoriel continu sur un cylindre, car la bobine finale aura une forme cylindrique. Connaissant le champs vectoriel de courant continu, on peut ensuite extraire ses lignes de champs qui vont permettre de définir la trajectoire des fils de cuivre. Ce principe est illustré sur la figure 5.2.

Le lien entre un champs vectoriel électrique de surface J et du champ magnétique B généré est donné par la loi de Biot-Savart :

$$B(x) = \frac{\mu_0}{4\pi} \int J(x') \times \frac{x' - x}{\|x' - x\|_2^3} dx'.$$

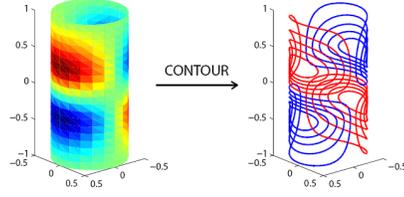


FIGURE 5.2 – Extraction des lignes de champs d'un courant défini continûment.

D'un point de vue numérique, ces équations peuvent être discrétisées par des techniques d'éléments finis (voir figure 5.3). L'opérateur intégral de Biot-

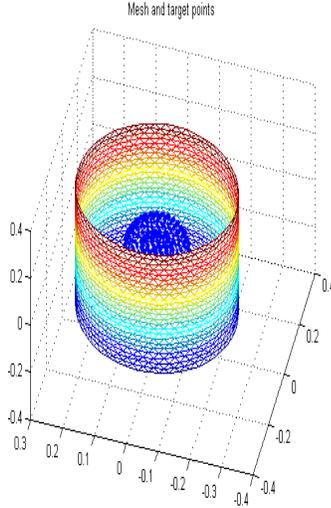


FIGURE 5.3 – Discrétisation par éléments finis. A l'intérieur : points d'échantillonnages du champ magnétique.

Savart (5.1.2) se transforme ainsi en matrice $A \in \mathbb{R}^{3m \times 3n}$ de dimension finie. Les facteurs 3 sont liés au fait qu'on considère des champs vectoriels en dimension 3. Une fois la discrétisation effectuée, la façon traditionnelle de faire du design consiste à résoudre le problème de moindres carrés suivant :

$$\inf_{J \in \mathbb{R}^{3n}, \text{div}(J)=0} \frac{1}{2} \|AJ - B\|_2^2 + \frac{\lambda}{2} \|WJ\|_2^2, \quad (5.1)$$

où $B \in \mathbb{R}^{3m}$ est un champ magnétique cible discrétisé, λ est un paramètre de régularisation et $W \in \mathbb{R}^{3n \times 3n}$ est une application linéaire telle que $\|WJ\|_2^2$ représente l'énergie stockée dans la bobine. La contrainte $\text{div}(J) = 0$ indique qu'aucun courant n'est créé en dehors des sources.

L'idée principale que nous avons explorée avec Michael est de *minimiser la densité maximale de courant* $\|J\|_\infty$, car c'est elle qui détermine les échauffements de la bobine (la densité d'énergie en un point x de la bobine est définie par $|J(x)|R(x)$, où $R(x)$ est la résistivité. Cette dernière est en général constante car la bobine est constituée d'un matériau homogène d'épaisseur constante. Notre problème devient donc le suivant :

$$\inf_{J \in \mathbb{R}^{3n}, \text{div}(J)=0} \frac{1}{2} \|AJ - B\|_2^2 + \frac{\lambda}{2} \|J\|_\infty. \quad (5.2)$$

En pratique, j'avais défini des codes d'optimisation qui reposaient sur des méthodes de premier ordre préconditionnées. Le nombre d'éléments finis n étant cependant souvent inférieur à 10000, nos derniers codes reposent sur des techniques de points intérieurs qui sont plus fiables et rapides.

La figure 5.4 montre quelle est la différence entre des designs standards et les designs que nous avons proposés. Il est important de noter l'écartement quasi constant entre les fils de cuivre de la bobine pour notre design (5.2), tandis que cette distance diminue pour le design (5.1). Nos bobines ont été construites en pratique et une entreprise japonaise semblait souhaiter les commercialiser. Je ne sais pas ce qui est advenu au final. Toujours est-il que les expériences physiques menées sur les bobines ont été très positives. La figure 5.5 montre la température de nos bobines pour une efficacité équivalente. Comme on peut le voir, celle-ci est réduite de 40 degrés pour nos designs, ce qui est très significatif!

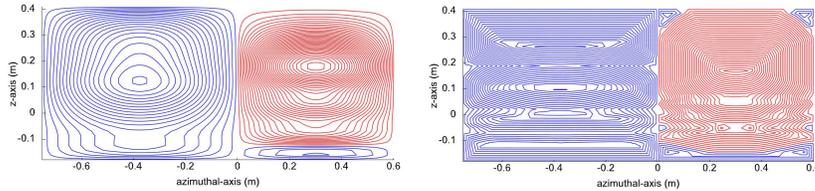


FIGURE 5.4 – Comparaison entre les bobines usuelles (à gauche) et la bobine que nous avons créée (à droite). Ici, les bobines ont été déroulées pour permettre une meilleure visualisation.

5.1.3 Conception de pulses radio-fréquence

Nicolas Boulant, chercheur au CEA a eu la grande sympathie de me proposer de travailler sur un problème de génération de pulses radio-fréquence (RF) [31]. Ceux-ci sont générés par des antennes RF localisées autour de l'objet à imager. Elles servent à exciter les spins et à les faire basculer dans le plan orthogonal au champ magnétique B_0 . Je ne rentrerais pas dans les détails de ce travail ici, car i) je n'étais vraiment pas leader sur ce projet et tous mes remerciements vont à Nicolas Boulant et Andrés Hoyos-Idrobo et ii) je souhaite limiter la taille de mon manuscrit et mon temps de rédaction.

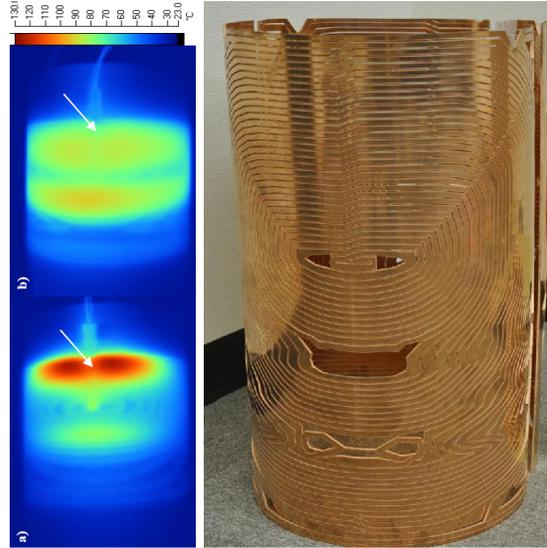


FIGURE 5.5 – A gauche : comparaison du fonctionnement des bobines de gradient usuelles (en bas) et de nos design (en haut). A droite : une de nos bobines réellement construite.

En deux mots : le problème considéré est un problème d’optimisation non convexe qui - d’un point de vue numérique - ressemble assez à la récupération de phase. Nous avons effectué une étude empirique de plusieurs méthodes d’optimisation et de relaxations : sequential quadratic programming, magnitude squared least square, points intérieurs. Cette étude empirique a eu un impact plutôt élevé, qui me réjouit, mais me surprend aussi un peu.

5.1.4 Projection sur des espaces de courbes

En première approximation, une IRM permet de mesurer la transformée de Fourier \hat{u} d’une image, le long d’une trajectoire $\lambda : [0, T] \rightarrow \mathbb{R}^d$, avec $d \in \{2, 3\}$. La vitesse de la trajectoire $\lambda(t)$ est proportionnelle aux courants G_x, G_y, G_z émis dans les bobines x, y et z dont j’ai parlé au début de ce chapitre :

$$\dot{\lambda} \propto [G_x, G_y, G_z].$$

Le nom bobine de gradient vient d’ailleurs de cette relation.

En pratique les courants ont une amplitude bornée et ne peuvent pas varier trop rapidement, ceci se traduit par les contraintes cinématiques suivantes sur la trajectoire : $\|\dot{\lambda}\|_\infty \leq \alpha$ et $\|\ddot{\lambda}\|_\infty \leq \beta$, où les constantes α et β sont respectivement l’amplitude maximale et le “slew-rate”, définis par le constructeur. De plus, il peut y avoir d’autres contraintes linéaires, du type, partir du centre du k -space $\lambda(0) = 0$ ou de passage au centre à un moment t_0 donné $\lambda(t_0) = 0$, pour des

questions d'écho de gradient. En résumé, un ensemble admissible de courbes peut s'écrire sous la forme :

$$C = \{ \lambda \in (W^{2,\infty}([0, T]))^d, \|\dot{\lambda}\|_\infty \leq \alpha, \|\ddot{\lambda}\|_\infty \leq \beta, A\lambda = b \}, \quad (5.3)$$

où $A : (W^{2,\infty}([0, T]))^d \rightarrow \mathbb{R}^m$ est une transformée linéaire qui définit un ensemble de contraintes.

Etant donnée une trajectoire non admissible λ_0 , on peut souhaiter trouver une autre trajectoire admissible au sens où elle satisfait les contraintes ci-dessus. Lors de la thèse de Nicolas Chauffert, la seule approche existante reposait sur le contrôle optimal : trouver la trajectoire qui passe par un ensemble de points donnés avec une vitesse ou une accélération maximale systématique [LKP08]. Dans [24], nous avons proposé une méthode alternative qui consiste à projeter la trajectoire λ_0 sur l'espace (convexe) des contraintes. Ce travail est assez simple conceptuellement, mais très utile en pratique. C'est notamment une des briques de base pour la génération de trajectoires d'échantillonnage optimales. Sur la figure 5.6 nous présentons un exemple de résultat de projection.

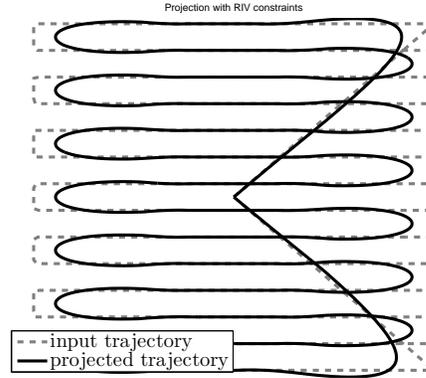


FIGURE 5.6 – Exemple de projection d'une trajectoire EPI en IRM.

5.1.5 Échantillonnage compressé en IRM

L'échantillonnage compressé en IRM fonctionnelle est définitivement un gros morceau de mon activité de recherche les 5 dernières années. Les thèses de Claire Boyer, Nicolas Chauffert et Carole Lazarus ont été essentiellement focalisées sur ce problème, avec des questions à la fois théoriques et pratiques.

Dans cette partie, je décris l'application des idées du chapitre 2 à des IRM 7 Teslas avec 32 canaux de réception du CEA Saclay. Ces travaux ont été obtenus en grande partie grâce au travail de Carole Lazarus [49 ; 46 ; 16].

Pour commencer, la figure 5.7 rappelle le principe de génération des trajectoires SPARKLING (Spreading Projection Algorithm for Rapid K-space samPLING). Étant donnée une trajectoire initiale et une densité cible, les algorithmes de projection du chapitre 2 génèrent des schémas d'échantillonnages

admissibles, au sens où ils satisfont les contraintes de la machine et où ils approchent au mieux la densité. Pour les physiciens, un moyen d'évaluer la qualité d'une trajectoire consiste à regarder la PSF, i.e. la transformée de Fourier de la mesure empirique associée aux points échantillonnés. Empiriquement, on observe que plus la PSF est piquée, meilleurs sont les résultats de reconstruction. Je ne connais pas d'explication théorique formelle à ce phénomène. On voit en tous cas sur la figure 5.7 que l'étape de projection permet d'améliorer nettement la PSF et de mieux couvrir le k -space.

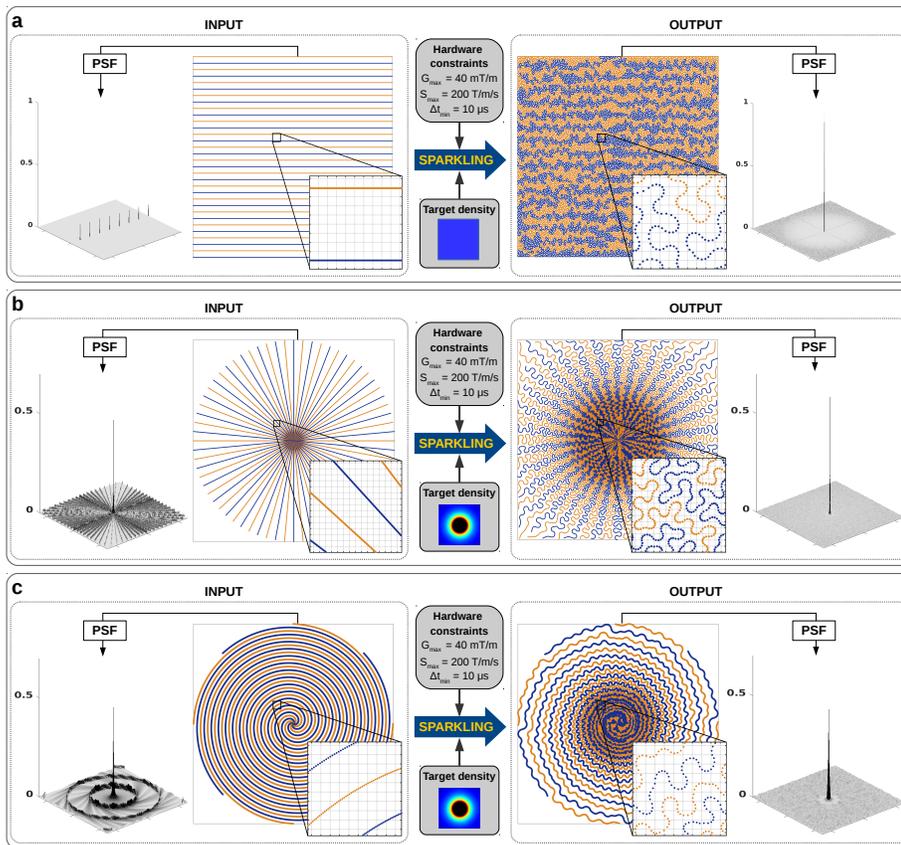


FIGURE 5.7 – Principe de génération des trajectoires SPARKLING.

Un théoricien pur pourrait croire qu'il suffit de donner la trajectoire à la machine pour obtenir des résultats de reconstruction. En réalité, l'obtention de ces résultats est longue et complexe, car il y a beaucoup d'effets à prendre en compte et à contrer pour obtenir un échantillonnage décent. De plus, il faut travailler avec de larges volumes de données et des reconstituteurs facétieux

(estimation des cartes de sensibilité, réglage des paramètres, etc...). Je ne décrirais pas ici ce long processus et me contenterais de montrer les résultats finaux. La figure 5.8 compare des reconstructions issues de schémas d'échantillonnage avec trois facteurs d'accélération différents (10, 15 et 20). Un facteur 10 signifie que l'échantillonnage a pris 10 fois moins de temps que la séquence constructeur. Comme on peut le voir, les schémas radiaux donnent des résultats de mauvaise qualité avec une perte rapide de résolution. Les schémas en spirale ont une résolution assez satisfaisante, cependant ils présentent des effets significatifs de ringing. La trajectoire SPARKLING quant à elle donne des résultats satisfaisants pour toute la gamme de facteurs d'accélération 10-20. Ceci est très encourageant.

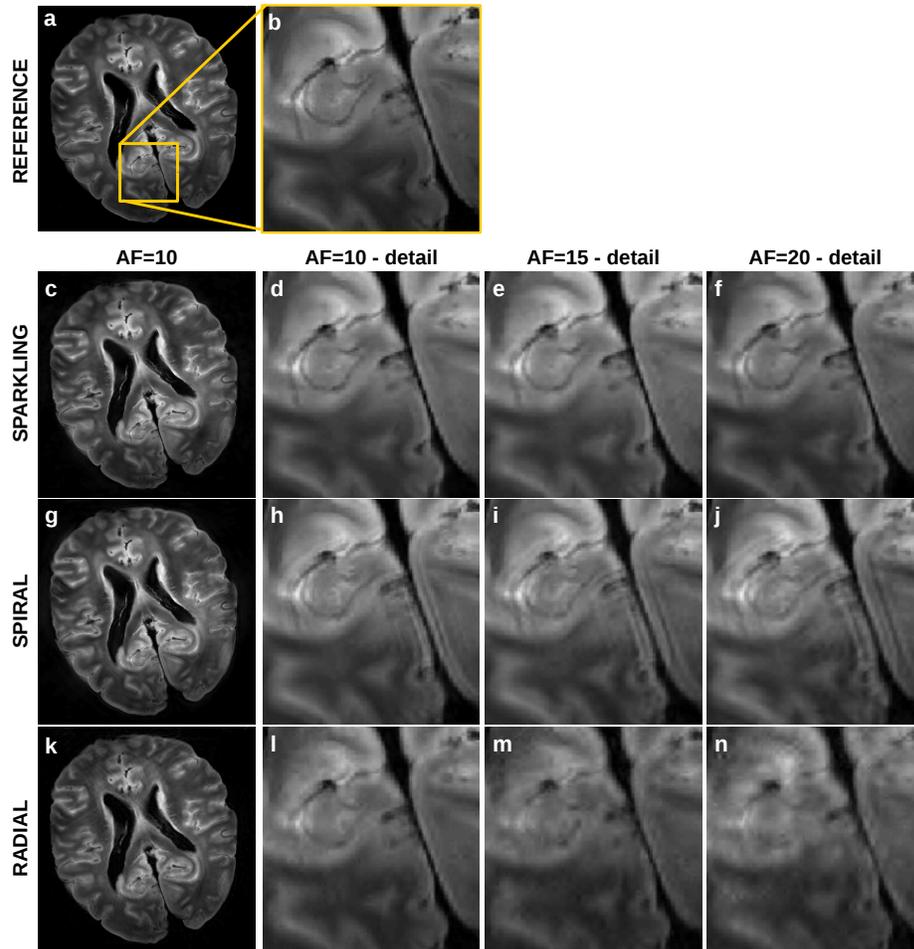


FIGURE 5.8 – Comparaisons de résultats de reconstruction sur des images ex-vivo de cerveau humain.

Expliquer ces résultats semble assez complexe et nous ne possédons pas encore d'explication sûre. Cependant, une piste sérieuse est la suivante : lorsqu'on donne une trajectoire à la machine, elle ne la joue pas exactement pour plusieurs raisons (inhomogénéités du champ principal B_0 , gradients imparfaits, échauffements de bobines, etc...). Si on admet que la déformation est une convolution [Van+17], on peut explorer ses effets en simulation. La figure 5.9 montre les erreurs de trajectoire pour une trajectoire SPARKLING et une trajectoire en spirale. Les erreurs sont de plus faible amplitude et sont bien moins cohérentes pour les trajectoires proposées.

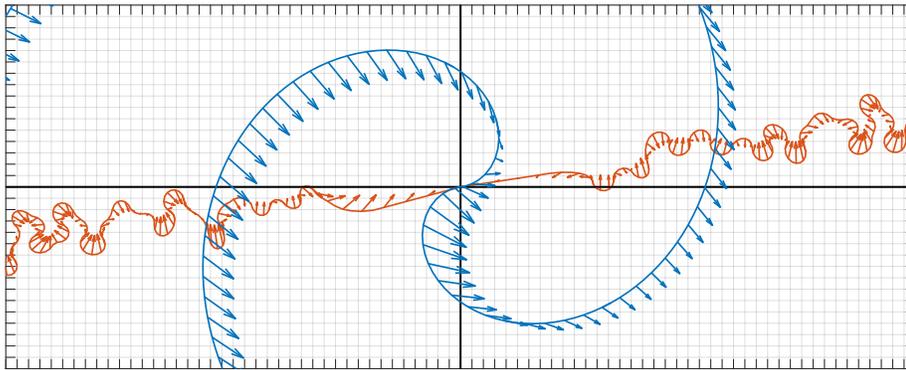


FIGURE 5.9 – Erreurs de trajectoires pour des séquences spirales et SPARKLING.

Ce projet me semble globalement très prometteur d'un point de vue sociétal, puisque nous avons aussi pu valider le concept sur des coupes in-vivo de cerveau humain. Un exemple de résultat est présenté sur la figure 5.10, avec des conclusions similaires à ce qui se passe en ex-vivo. Il est important de noter que ces résultats ne sont pas stables : les inhomogénéités de B_0 en in-vivo ont des effets délétères sur les trajectoires SPARKLING qui semblent moins prononcés sur des séquences constructeurs.

👁 **Perspectives** Il y a encore beaucoup d'idées à explorer autour de ces questions d'échantillonnage en IRM et je trouverais un peu dommage ne pas capitaliser sur ce thème. Il est malheureusement difficile d'implémenter les séquences au CEA depuis le départ de Carole Lazarus, aucune moyen n'a jamais été alloué à Toulouse et j'ai un peu peur de devoir clore ce thème ou de me rapprocher de l'industrie.

Une des questions à explorer est la génération de trajectoires 3D. C'est un sujet déjà assez avancé et je l'illustre simplement avec deux figures. La figure 5.11 montre quelques trajectoires 3D existantes et la figure 5.12 montre les idées qui ont commencé à être explorées pendant la thèse de Carole. En deux mots : ce sont des idées réellement prometteuses sur le papier et sur des images ex-vivo. Malheureusement, nous n'avons pas encore de résultat satisfaisant sur des

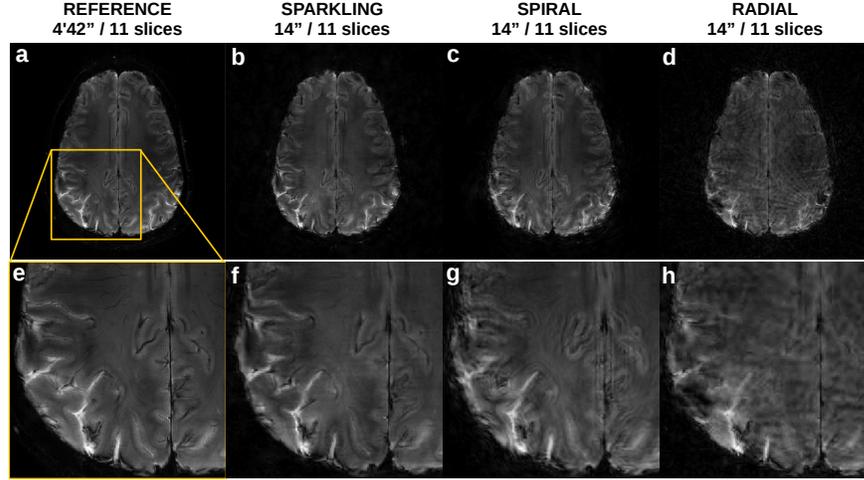


FIGURE 5.10 – Comparaisons de résultats de reconstruction sur des images in-vivo de cerveau humain.

images in-vivo, du fait d'incertitudes plus importantes lorsqu'on passe en 3D.

Je commence aussi à explorer la possibilité de corriger ces erreurs de trajectoires uniquement à partir de mesures d'échantillonnage. C'est un joli problème inverse aveugle non-linéaire qui me semble original et ouvert.

5.1.6 Correction des effets de filtrage ADC

Le projet suivant [2] est issu d'une collaboration avec Carole Lazarus, alors doctorante au CEA Saclay sous la direction de Philippe Ciuciu et de Maximilian März, doctorant à TU Berlin.

Les premières implémentations de schémas d'échantillonnage Sparkling par Carole ont donné des résultats très peu satisfaisants, voir figure 5.13. Identifier la source du problème de reconstruction est malheureusement réellement complexe car énormément de facteurs peuvent potentiellement expliquer un tel effet (erreurs de gradient, inhomogénéité B_0 , effets de filtrages,...).

Nous avons fini par comprendre exactement d'où venait le problème grâce à des techniques de reverse engineering et des tests en simulation numérique. Cette méthodologie semble assez commune chez les spécialistes de l'IRM, car les constructeurs ne donnent pas tous les détails sur leur machine et il faut donc la comprendre avec des astuces de mesure.

Dans notre cas, nous avons compris que le problème venait d'effets de filtrage du signal temporel généré par la machine. A ma connaissance, le modèle utilisé par tous les gens qui souhaitent reconstruire une image est le suivant :

$$y_j = \hat{u}(\lambda(j\Delta t)). \quad (5.4)$$

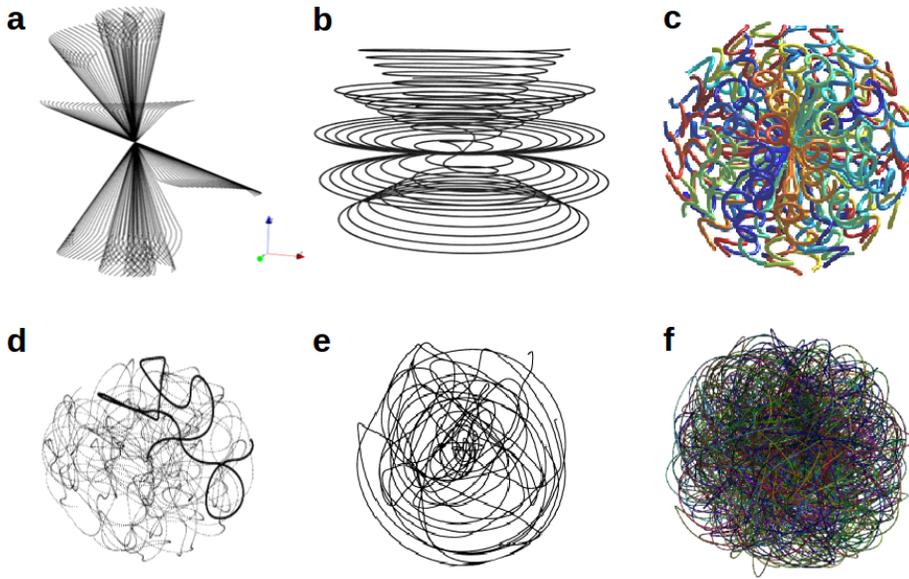


FIGURE 5.11 – Exemples de trajectoires 3D existantes pour l'IRM.

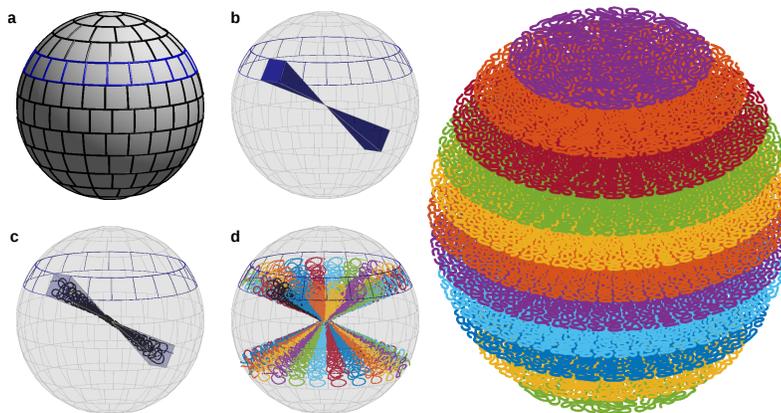


FIGURE 5.12 – Exemples de trajectoires 3D SPARKLING. A gauche : SPARKLING igloo. A droite : stack of SPARKLING.

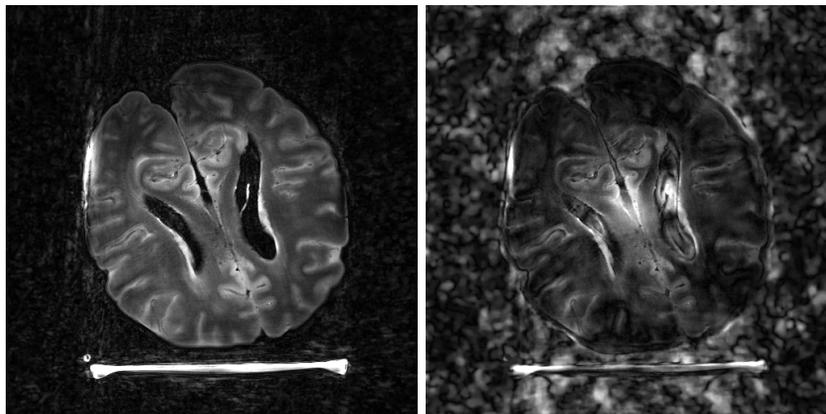


FIGURE 5.13 – A gauche : image obtenue avec une séquence constructeur. A droite : image obtenue à partir d’une séquence sparkling avec une réduction de temps d’un facteur 15, reconstruite avec le modèle usuel (5.4).

On observe la transformée de Fourier de notre signal le long de la trajectoire λ avec un pas d’échantillonnage Δt . Ceci est bien sûr impossible d’un point de la physique et un modèle plus correct est le suivant :

$$y_j = (h \star f)(j\Delta t), \quad (5.5)$$

où h est un filtre de convolution implémenté dans un ADC qui dépend du constructeur de la machine. Malgré de nombreuses demandes aux ingénieurs SIEMENS, il nous a été impossible de connaître le filtre h constructeur et nous avons donc effectué plusieurs travaux.

- Analyse théorique et empirique des effets de filtrage sur la reconstruction d’image.
- Développement de techniques numériques d’estimation du filtre h qui reposent sur le design de séquences de gradient de chirps linéaires.
- Développement de procédures numériques de reconstruction qui prennent en compte l’effet de filtrage (5.5). Ceci nous a permis d’obtenir de bien meilleurs résultats de reconstruction (voir la figure 5.14).

5.2 Microscopie computationnelle

5.2.1 Introduction

J’ai créé l’équipe PRIMO à l’ITAV en 2012 avec Jérôme Fehrenbach. L’objectif principal affiché alors était l’amélioration des microscopes de fluorescence

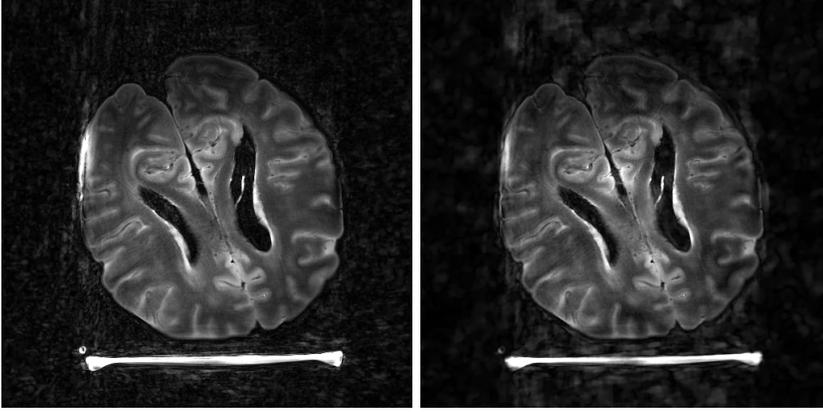


FIGURE 5.14 – A gauche : image obtenue avec une séquence constructeur. A droite : image obtenue à partir d’une séquence sparkling et la prise en compte des effets de filtrage ADC. Ce résultat est à comparer à la figure 5.13, droite.

et notamment le SPIM (Selective Plane Illumination Microscope). Une partie de mes résultats dans ce domaine ont été montrés dans le Chapitre 4. Je montre ci-dessous quelques résultats supplémentaires.

5.2.2 Débruitage structuré

Un travail qui a bien plu à mes collègues biologistes est le débruitage de bruit “structuré”. La meilleure façon d’expliquer le problème est probablement de montrer quelques images. La figure 5.17 est constituée de quelques images prises avec différents types de microscopes. On y voit clairement des raies qui rendent plus difficile l’analyse des images. Ces raies peuvent être liées à différents phénomènes, tels que le parcours rectiligne d’une pointe en microscopie à force atomique, la diffraction d’une feuille de lumière en SPIM, ou encore de l’absorption en microscopie électronique.

Un premier modèle [32] D’un point de vue mathématique, le premier modèle de bruit que nous avons proposé était assez complexe : nous avons fait l’hypothèse que l’image bruitée $u_0 \in \mathbb{R}^n$ était la somme d’une image idéale $u^* \in \mathbb{R}^n$ et d’un bruit $b \in \mathbb{R}^n$ de la forme suivante :

$$b = \sum_{i=1}^m \lambda_i \star h_i, \quad (5.6)$$

où $h_i \in \mathbb{R}^n$ sont des filtres qui vont permettre de reproduire la structure du bruit et $\lambda_i \in \mathbb{R}^n$ sont des réalisations de vecteurs aléatoires dont les composantes sont i.i.d. Ce modèle permet d’obtenir des bruits avec des structures complexes. La figure 5.15 montre un exemple d’une image de disque perturbée par un bruit de

la forme (5.6) avec $m = 3$. Le bruit est cependant *stationnaire*, ce qui est très naturel pour beaucoup d'applications.

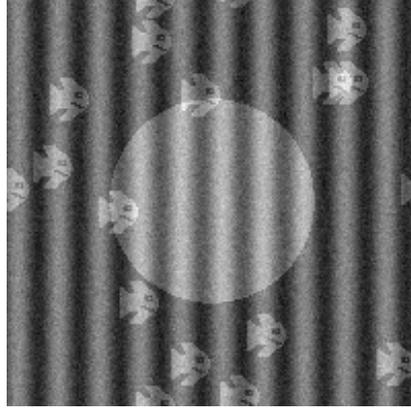


FIGURE 5.15 – Exemple de bruit structuré pouvant être capturé par le modèle (5.6).

Une fois le modèle posé, des raisonnements bayésiens standards mènent à la résolution du problème suivant :

$$\inf_{\lambda=(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^{m \times n}} \left\| u_0 - \sum_{i=1}^m \lambda_i \star h_i \right\|_{TV} + \sum_{i=1}^m f_i(\lambda_i), \quad (5.7)$$

où les fonctions f_i doivent être choisies en fonction de la statistique des processus sous-jacents aux vecteurs λ_i dans le modèle (5.6).

Le problème (5.7) peut être résolu avec des techniques d'optimisation convexes telles que celles présentées dans le premier chapitre avec un taux de convergence en $O(1/k)$, où k est le nombre d'itérations. En pratique, ce modèle simple s'est révélé efficace et versatile, voir figure 5.17.

Une simplification [30] Le choix de la statistique de bruit dans le modèle 5.7 peut être assez complexe, notamment pour des chercheurs en biologie. Nous avons donc cherché à simplifier le modèle. Dans le papier [30], nous avons montré que sous des hypothèses d'étalement du filtre h_i , un bruit du type $b_i = \lambda_i \star h_i$ pouvait être *assimilé à un bruit gaussien* grâce au théorème de Berry-Esseen.

En pratique, ceci signifie que sous de bonnes hypothèses, on peut fixer les fonctions f_i dans le modèle (5.7) à des normes ℓ^2 :

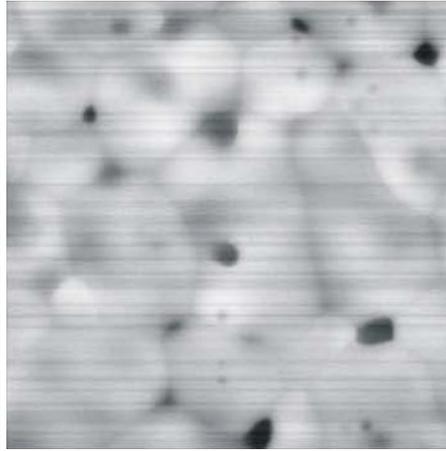
$$\inf_{\lambda=(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^{m \times n}} \left\| u_0 - \sum_{i=1}^m \lambda_i \star h_i \right\|_{TV} + \sum_{i=1}^m \frac{\alpha_i}{2} \|\lambda_i\|_2^2. \quad (5.8)$$

On a ensuite montré que le modèle (5.8) était équivalent au modèle suivant :

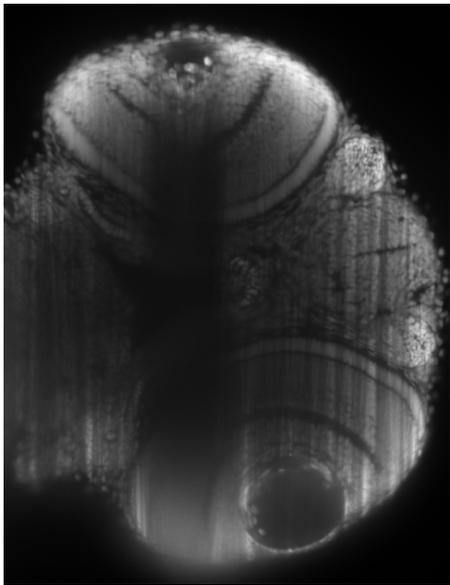
$$\inf_{\lambda \in \mathbb{R}^n} \|u_0 - \lambda \star h\|_{TV} + \frac{\alpha}{2} \|\lambda\|_2^2, \quad (5.9)$$



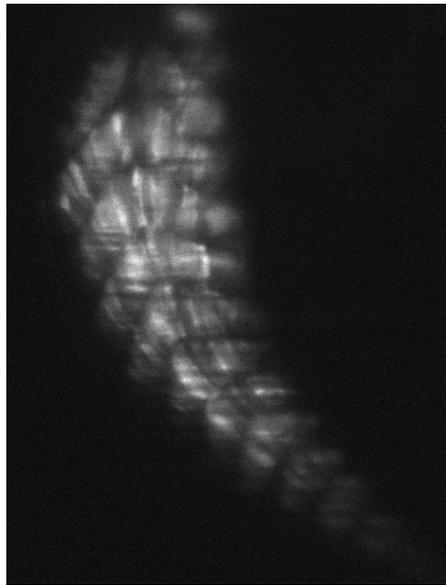
(a) Microscopie électronique



(b) AFM

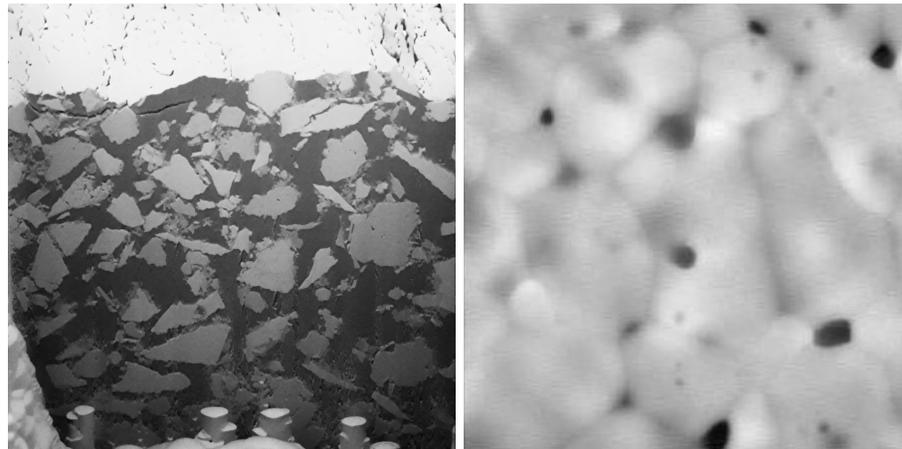


(c) SPIM - Zebrafish



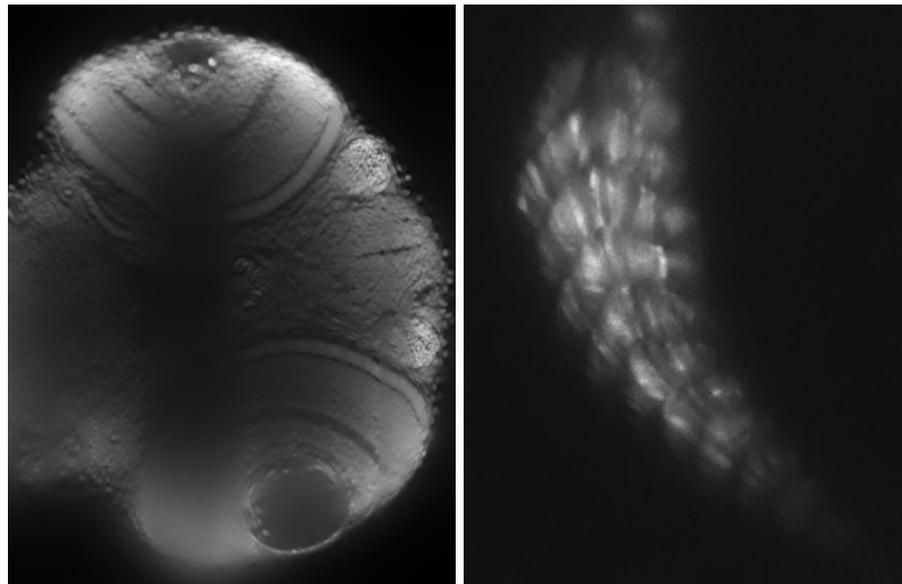
(d) SPIM - Zebrafish

FIGURE 5.16 – Exemple de bruit structuré apparaissant en microscopie. En haut : images de microscopie électronique et de microscope à force atomique. En bas : images de microscopie par feuille de lumière.



(a) Microscopie électronique

(b) AFM



(c) SPIM - Zebrafish

(d) SPIM - Zebrafish

FIGURE 5.17 – Exemple de résultat de débruitage sur les images 5.16.

pour des valeurs explicites de α et h en fonction de (α_i) et (h_i) . De plus, la résolution du problème (5.9) permet de remonter aux vecteurs λ_i optimaux dans le modèle (5.8) et donc aux composantes $b_i = \lambda_i \star h_i$.

Le dernier modèle (5.9) présente plusieurs avantages :

- Il est fortement convexe, ce qui va permettre d'utiliser les accélérations duales présentées dans le premier chapitre.
- On est passé d'une optimisation sur un espace de taille $m \times n$ à une optimisation sur un espace de taille n , avec une réduction à peu près proportionnelle de la complexité.
- Finalement, nous avons donné des bornes supérieures et inférieures sur la norme de la composante b_i en fonction des paramètres α_i , ce qui nous a permis de fortement simplifier la recherche des paramètres de régularisation α_i et donc la tâche de l'utilisateur.

En pratique, le travail [30] a donc permis de simplifier significativement la tâche de l'utilisateur et de réduire les temps de calcul.

Le cas multiplicatif Notre dernier travail dans cette lignée a été l'étude du bruit multiplicatif [21]. Après une analyse théorique, nous sommes arrivés au modèle suivant :

$$\inf_{\lambda \in \mathbb{R}_+^n} \|(u_0 \odot (\lambda \star h))\|_{TV} + \alpha \langle \lambda - \log(\lambda), 1 \rangle, \quad (5.10)$$

où le symbole \odot indique une multiplication point à point. Cette énergie ressemble à ce qui apparaît lorsqu'on travaille avec du bruit de Poisson, mais à ma connaissance, elle n'a jamais été proposée auparavant. Son avantage principal est de mener à un *modèle convexe* pour du bruit multiplicatif. Les comparaisons avec des modèles plus anciens a été à notre avantage, même pour du bruit i.i.d. [AA08; ST10; SO08]. Ceci dit, cette étude devrait être effectuée de manière plus systématique et de toutes façons, les modèles variationnels disparaissent au profit du deep-learning. Un exemple de résultat et une comparaison avec les modèles additifs est donnée dans la figure 5.18.

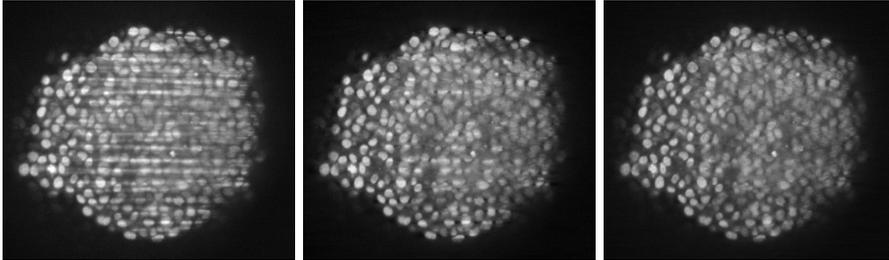


FIGURE 5.18 – Exemple de débruitage multiplicatif. De gauche à droite : image originale de sphéroïde prise avec un microscope SPIM, image traitée avec le modèle additif et image traitée avec le modèle multiplicatif.

👁 **Identification automatique des filtres** Ces travaux sont réellement utilisés en routine sur quelques plateformes d'imagerie. Le choix des filtre h_i reste cependant difficile pour les utilisateurs. En ouverture de cette section, je pense qu'il serait donc intéressant d'un point pratique (et probablement théorique) d'explorer la recherche automatique de filtres. C'est un problème bilinéaire inverse qui s'insère assez bien dans mon projet de recherche actuel.

5.2.3 *Évaluation et correction multi-vue de l'atténuation*

Pour finir ce chapitre, je présente en 2 mots un travail que je trouve assez original publié dans JMIV [7]. J'ai aussi un espoir qu'il soit exploitable en pratique, bien que cette dernière assertion soit incertaine au moment où j'écris.

L'idée a été motivée par des travaux de biologie que j'ai eu l'occasion de voir à l'ITAV. Je me suis rendu compte lors de la rédaction de l'article qu'elle avait déjà été proposée 30 ans auparavant dans le domaine du lidar [Kun87] et un relecteur nous a fait remarquer qu'elle avait été étudiée plus récemment dans le domaine de la microscopie confocale [Sch+13]. Ceci réduit l'originalité du travail, mais les contributions sur la résolution du problème et sa convexification sont tout de même significatives.

La figure 5.19 illustre le problème. En microscopie confocale par exemple, la lumière d'excitation et d'émission est atténuée lors du passage dans l'échantillon par des éléments absorbants dont la densité est inconnue¹. Ceci crée une atténuation des contrastes lorsqu'on pénètre en profondeur.

La question qu'on s'est posée est la suivante.

Question 6. *Est-il possible de retrouver la carte d'atténuation et la carte de densité à partir de deux images prises depuis deux (ou plusieurs) orientations différentes ?*

Nous montrons que la réponse est positive, et que ça peut être réalisé de façon *stable* en résolvant un *problème convexe*.

👁 Nous sommes actuellement en contact avec Juan Cuesta, chercheur à l'université Paris Est-Créteil, pour évaluer l'importance pratique de ce travail en lidar [CF10].

1. Il y a bien d'autres phénomènes que l'absorption dont la diffraction qui est souvent critique.

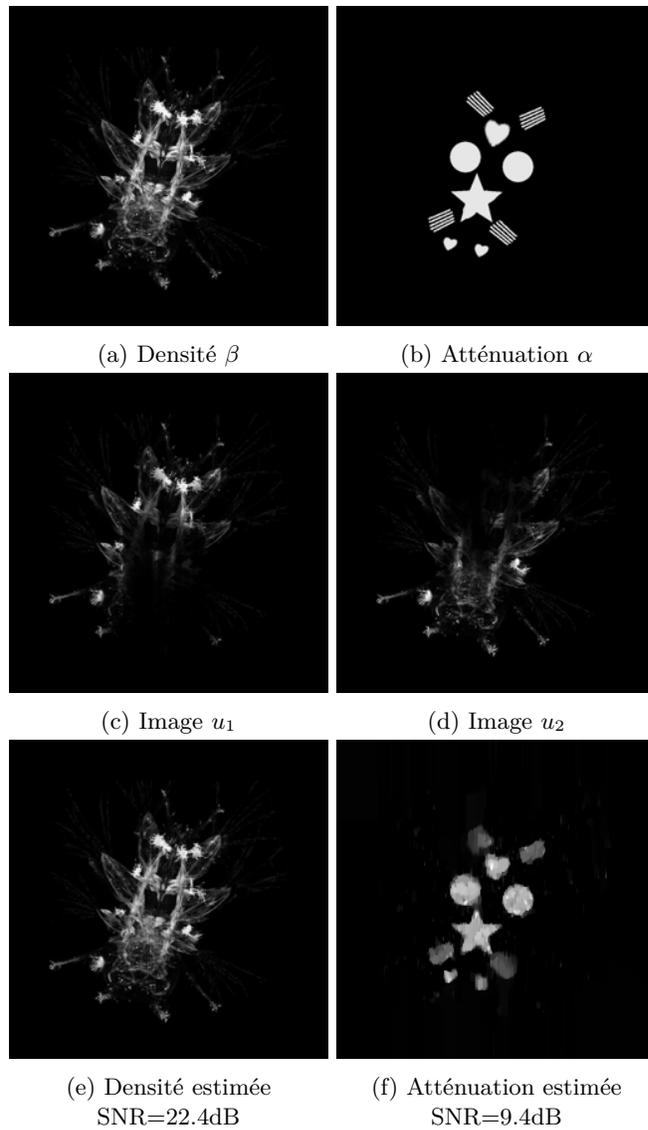


FIGURE 5.19 – Illustration de l'estimation de l'atténuation. Un insecte possède une densité de fluorophores α (Fig. 5.19a) et une carte d'atténuation β (Fig.5.19b). les images mesurées u_1 et u_2 sont affichées sur les Fig. 5.19c et 5.19d). Elles sont atténuées différemment suivant le trajet optique (de haut en bas ou de bas en haut). A partir de ces deux images, notre algorithme renvoie une estimation raisonnable de chacune des cartes (Fig. 5.19e et 5.19f) malgré un bruit de Poisson.

Chapitre 6

Autres travaux

6.1 Lignes de niveau des images

Enfant, mes parents m'ont initié à la randonnée et m'ont laissé assez libre pour en faire seul très jeune. Ceci m'a encouragé à passer des heures devant des cartes IGN pour tracer mon chemin et souvent aussi pour le retrouver. Les cartes topographiques m'ont donc sauvé plusieurs fois. En arrivant au master MVA, j'ai eu d'excellents cours de Lionel Moisan où ces objets sont réapparus. Ceci m'a motivé à les étudier plus finement.

6.1.1 *Invariance aux changements d'illumination*

Mon premier essai sur les cartes topographiques a été motivé par les travaux de Caselles, Coll, Morel [CCM99a]. Dans cet article, les auteurs suggèrent d'utiliser la *carte topographique* d'une image pour construire des algorithmes invariants aux changements de contraste. Cette idée est au centre de toute la morphologie mathématique [Ser82], une discipline riche, complexe et très française.

Le système visuel humain est à peu près invariant aux *changements d'illumination* : on reconnaît une personne ou une scène qu'elle soit sous le soleil ou les nuages. Les changements de contraste sont un modèle mathématique assez simple de changement d'illumination. Ils peuvent avoir plusieurs définitions, mais une constante est qu'ils préservent les *lignes de niveau* d'une image. Nous donnons ci-dessous deux définitions possibles de ligne de niveau. Je ne prends pas de précaution ici sur les espaces fonctionnels.

Définition 14 (Ligne de niveau de type 1). *Soit $u : \Omega \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ une image définie en tout point de l'espace. Les lignes de niveau $\alpha \in \mathbb{R}$ de u sont définies comme les composantes connexes des ensembles $\{x \in \Omega, u(x) = \alpha\}$.*

Cette définition est assez simple. Elle présente le défaut qu'une ligne peut être n'importe quel objet du plan. En particulier, elle n'est pas forcément de co-

dimension 1. La définition suivante est celle généralement préférée en traitement des images.

Définition 15 (Ligne de niveau de type 2). *Soit $u : \Omega \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ une image définie en tout point de l'espace. Les ensembles de niveau inférieurs de u sont définis par $L_\alpha^-(u) = \{x \in \Omega, u(x) \leq \alpha\}$. Les lignes de niveau $\alpha \in \mathbb{R}$ de u sont définies comme les frontières des composantes connexes de $L_\alpha^-(u)$.*

Dans cette deuxième définition, les lignes de niveau sont réellement des objets de co-dimension 1. Les deux définitions coïncident pour des images C^1 dont le gradient ne s'annule pas.

La première question à laquelle j'ai tenté de répondre avec Alexandre Fournier, alors doctorant à l'INRIA Sophia-Antipolis est la suivante.

Question 7. *À quelle condition est-ce que les lignes niveau sont invariantes aux changements d'illumination ?*

Nous y avons répondu dans l'article [34] en utilisant des outils de géométrie différentielle. Pour répondre à cette question, il faut d'abord construire un modèle mathématique qui décrit les effets de variations d'illumination. Nous avons utilisé le modèle lambertien, qui stipule que la luminosité perçue d'une surface dépend de son albédo (i.e. sa capacité à réfléchir la lumière ou à l'absorber) et aussi de l'angle entre la lumière incidente et la normale à la surface. Le théorème formel suivant résume notre contribution principale.

Théorème 18 (Invariance des lignes de niveau à l'illumination). *Sous un modèle d'illumination lambertien, les seules scènes qui produisent des images dont les lignes de niveau de type 1 sont invariantes par illumination sont les scènes développables dont l'albédo varie uniquement de façon orthogonale aux génératrices. Ce résultat est valide seulement pour presque toute illumination.*

Comme une image vaut probablement mieux qu'un long discours, la figure 6.1 présente une surface non développable (une demi-sphère) et une surface développable (un cône). Comme on peut le voir, les lignes de niveau de la sphère changent fortement en fonction de la direction de l'illumination. Celles du cône sont invariantes à presque toute illumination, mais il y en a une critique : si la lumière arrive exactement dans la direction centrale du cône, on ne perçoit qu'une unique ligne de niveau de type 1.

En un sens, ce résultat est plutôt négatif : il montre que les lignes de niveau ne sont pas un très bon outil si on souhaite assurer une invariance aux changements d'illumination. Malheureusement, définir une classe d'images qui pourraient représenter la même scène sous-jacente sous différentes illuminations avec un modèle mathématique simple me semble actuellement hors de portée ¹. En l'absence d'un tel modèle, les lignes de niveaux restent un excellent outil. En

1. Il est possible que les réseaux de neurones aient déjà répondu à cette question...

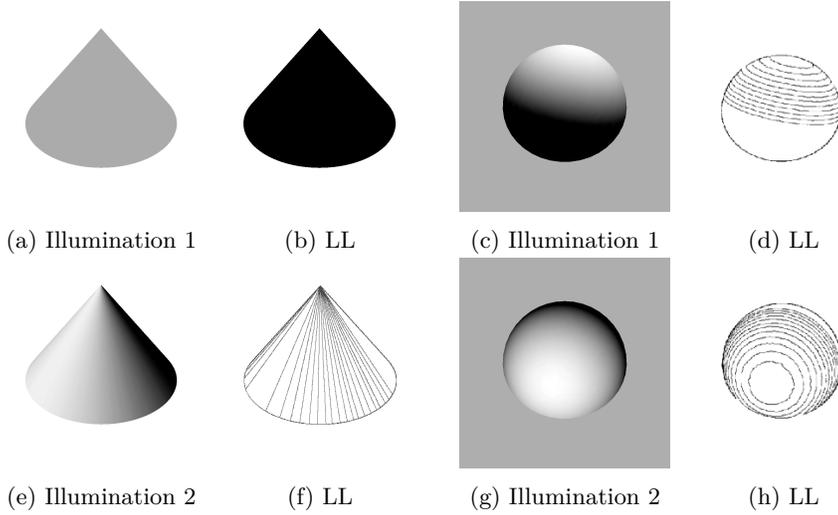


FIGURE 6.1 – Un cône et une demi-sphère sous deux illuminations différentes ainsi que leurs lignes de niveau (LL pour Level-Line).

effet, elles capturent les éléments les plus essentiels des images tels que les occlusions qui créent des discontinuités nettes quasi invariantes à l'illumination et les jonction triples qui sont aussi un invariant de choix en vision par ordinateur.

6.1.2 Mesures de qualité invariantes

Dans l'article [CCM99b], les changements de contraste locaux sont définis comme des opérations préservant les lignes de niveau d'une image, mais aussi la *relation d'ordre* entre ces lignes. C'est un modèle simpliste de changement d'illumination qui est valable pour des illuminations et des scènes très simples (e.g. lumière ambiante (non directionnelle), scène contenant des objets à différents profondeurs avec différents niveaux de gris (albédos)).

Dans un travail récent [12], nous avons exploré comment la notion de contraste local permet de définir une mesure de qualité d'image. Nous présentons ce travail très rapidement ci-dessous dans un cadre discret. On suppose que Ω est un domaine discret de cardinal n . On suppose que $u_0 : \Omega \rightarrow \mathbb{R}$ et $u : \Omega \rightarrow \mathbb{R}$ sont deux images en niveaux de gris.

Définition 16 (Changement de contraste global). *On dit que u diffère de u_0 par un changement de contraste global s'il existe une application croissante $T : \mathbb{R} \rightarrow \mathbb{R}$ telle que $T(u) = u_0$.*

On note $\mathcal{U}_{\text{glo}}(u)$ l'ensemble des images qui diffèrent de u par un changement de contraste global.

Les définitions de changement de contraste local sont plus complexes.

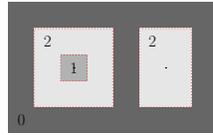
Définition 17 (Changement de contraste local de type 1). Soit (Δ_i) l'ensemble des lignes de niveau de type 1 de u_0 . On dit que u diffère de u_0 par un changement de contraste local de type 1 si $u|_{\Delta_i}$ est constant pour chaque Δ_i et si pour deux lignes de niveau Δ_i et Δ_j liées en 4-connexité, $(u(x) - u(y)) \cdot (u_0(x) - u_0(y)) \geq 0$.

On note $\mathcal{U}_{loc1}(u)$ l'ensemble des images qui diffèrent de u par un changement de contraste local de type 1.

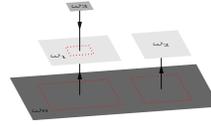
Pour définir correctement la définition de changement de contraste de type 2, il faut introduire l'arbre des formes et la transformée en lignes de niveau [MG00], qui sont des objets assez complexes. La Figure 6.2 illustre cet arbre pour une image simple. L'arbre des formes permet de coder la relation d'inclusion entre les formes $(\omega_i)_{i \in I}$ (définies comme les composantes connexes saturées des ensembles de niveau supérieurs ou inférieurs de u). Elle permet de décomposer une image u sous la forme algébrique suivante² :

$$u_0 = R(u_0)\alpha = \sum_{i \in I} \alpha_i s_i \mathbb{1}_{\omega_i},$$

où les coefficients $s_i \in \{-1, 1\}$ indiquent si la forme provient d'un niveau inférieur ou supérieur et où les coefficients $\alpha_i > 0$ encodent le saut entre formes adjacentes. L'application linéaire $R(u_0) : \mathbb{R}_+^{|I|} \rightarrow \mathbb{R}^n$ est un opérateur de reconstruction.



(a) Une image



(b) Son arbre de formes

FIGURE 6.2 – L'arbre des formes.

Nous sommes prêts à introduire une troisième notion de changement de contraste local.

Définition 18 (Changement de contraste local de type 2). On dit que u diffère de u_0 par un changement de contraste local de type 2 si elle peut s'écrire sous la forme $u = R(u_0)\beta$ pour un certain $\beta \in \mathbb{R}_+^n$.

On note $\mathcal{U}_{loc2}(u_0)$ l'ensemble des images qui diffèrent de u_0 par un changement de contraste local de type 2.

2. Je n'ai pas vu ce genre d'expression dans la littérature sur la FLST (Fast Level Set Transform), voir e.g. le livre [CM09]. De mon point de vue, elle est assez profonde et pratique. Elle reflète le théorème de Gustave Choquet [CM63] qui indique que dans un espace de Banach tel que l'espace des fonctions à variation bornée, on peut représenter une fonction comme une combinaison conique des points extrémaux de la boule unité de l'espace. La proposition 3 indique que les fonctions $\mathbb{1}_{\omega_i}/P(\Omega_i)$ sont justement ces points extrémaux !

Nous avons alors prouvé les résultats suivants :

Proposition 11 (Propriétés des modèles). *Pour toute image u , on a*

- Les ensembles $\mathcal{U}_{glo}(u)$, $\mathcal{U}_{loc2}(u)$ et $\mathcal{U}_{loc1}(u)$ sont convexes et fermés.
- Ils satisfont la propriété d'inclusion suivante :

$$\mathcal{U}_{glo}(u) \subseteq \mathcal{U}_{loc2}(u) \subseteq \mathcal{U}_{loc1}(u). \quad (6.1)$$

Nous sommes maintenant prêts à définir ce qu'on a nommé Contrast Invariant Signal-To-Noise Ratio :

Définition 19 (SNR invariant). *Étant données deux images u_0 et u_1 , on définit*

$$u_{glo} = \arg \min_{u \in \mathcal{U}_{glo}(u_1)} \frac{1}{2} \|u - u_0\|_2^2 \quad (6.2)$$

et

$$SNR_{glo} = -10 \log_{10} (\|u_{glo} - u_0\|_2^2 / \|u_0\|_2^2).$$

Il en va de même pour u_{loc1} , u_{loc2} , SNR_{loc1} et SNR_{loc2} .

C'est une mesure de qualité d'image qui présente des propriétés d'invariance claires et plus riches que le maintenant standard SSIM (Structural Similarity Index Measure). Par exemple SNR_{glo} est une mesure invariante à n'importe quel réhaussement de contraste que ferait un biologiste sous Fiji en ouvrant son image. Nous établissons plusieurs propriétés théoriques de ces modèles dans [12].

D'un point de vue numérique, la résolution du problème (6.2) est assez complexe. En reprenant les algorithmes décrits dans les sections 2.2.2 et 2.3, on peut montrer que le coût de calcul de u_{glo} est en $O(n)$. Le coût de calcul de u_{loc2} est en $O(n \log(n))$, grâce à l'utilisation conjointe de la FLST [MG00] et de programmation dynamique. C'est cependant un enfer à programmer, et vous trouverez les codes ici https://github.com/pierre-weiss/contrast_invariant_snr. Le calcul de u_{loc1} est inexact. Nos algorithmes ont un coût en $O(n^2/\epsilon)$ pour avoir une précision ϵ au pire des cas.

Pour conclure cette partie, je présente rapidement 2 résultats numériques pour montrer ce que représentent les projections. La Figure 6.3 montre comment les différents algorithmes sont capables de corriger les images qui diffèrent d'un changement d'illumination. La Figure 6.4 au contraire montre comment les algorithmes réagissent aux changements entre les différentes scènes. Il est particulièrement intéressant de voir les images des différences : elles contiennent les objets dans une image qui n'appartiennent pas à l'autre.

Les images u_{glo} , u_{loc1} et u_{loc2} sont des projections sur des ensembles convexes. Cette brique de base devrait permettre de définir des algorithmes robustes aux changements d'illumination qui sont nécessaires dans des tâches telles que le recalage d'image multi-modal, la stéréo-vision, le calcul du flot optique ou certains problèmes inverses.

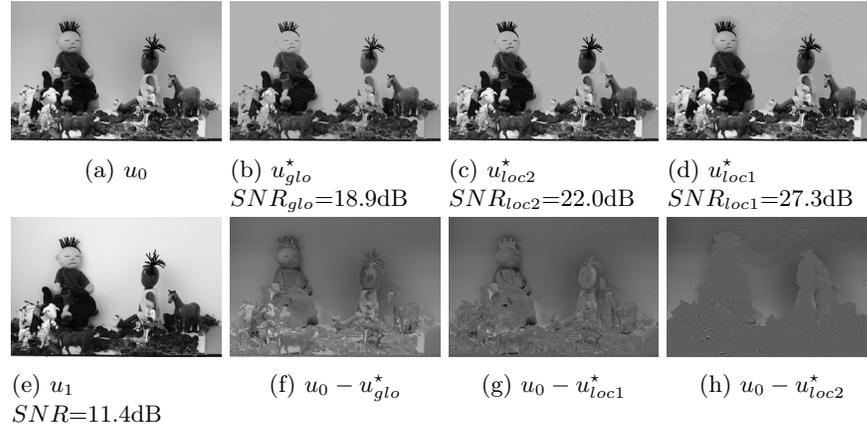


FIGURE 6.3 – Comparaisons des différents algorithmes pour des scènes identiques sous des conditions d’illumination différentes.

6.1.3 Ouverture

Mon sujet de thèse était officiellement le traitement variationnel des images couleurs. La question que je me posais était la suivante.

Question 8. *Quelles sont les relations qui existent entre les différents canaux d’une image ?*

Une réponse naturelle qui a été partiellement validée dans [CCM02a] est que les lignes de niveau des différents canaux sont identiques.

Ceci m’a motivé à développer des algorithmes de restauration qui favorisent l’alignement des lignes de niveau. Je n’ai cependant jamais écrit d’article à ce sujet, bien que j’y aie travaillé sérieusement. Il y a deux raisons qui expliquent ceci : premièrement, les modèles proposés étaient non convexes, ce qui me semblait être rédhibitoire à la grande époque des algorithmes avec garanties fortes. Deuxièmement, j’avais proposé un modèle essentiellement équivalent (et moins bien rédigé) à celui du papier [Bal+06] pour le problème de génération d’images hyper-spectrales haute résolution à partir d’images basse résolution et d’une image panchromatique haute résolution. Ce modèle était convexe...

J’ai aujourd’hui moins peur du monde non convexe, et je pourrais être tenté de réessayer de développer de tels modèles. J’ai d’ailleurs trouvé par hasard le papier suivant [EA14] qui reproduit essentiellement certaines idées que j’avais testées en thèse. Si je me lançais dans ce chantier, j’essaierais de définir de nouvelles notions telles que des cartes topographiques moyennes.

6.2 Segmentation d’objets ellipsoïdaux

Ma présence dans un centre transdisciplinaire de recherche sur le cancer m’a confronté de nombreuses fois à des questions de ce genre.

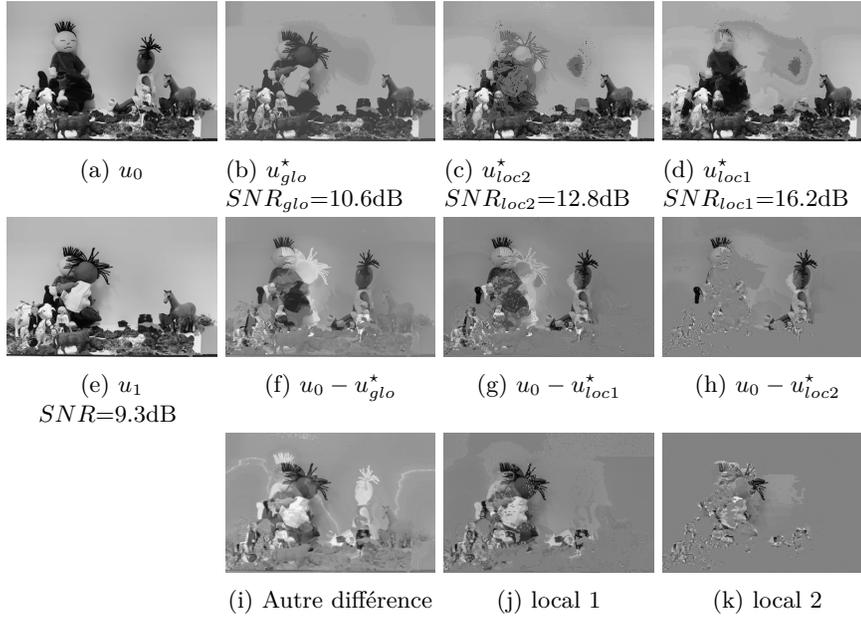


FIGURE 6.4 – Comparaisons des algorithmes pour des scènes différentes sous une illumination identique.

Question 9. *Pierre, peux-tu nous aider à identifier les noyaux dans nos images de micro-tumeur ? Comment peut-on décrire leur géométrie au sein de leur environnement ?*

La réponse à la première question est fréquemment non, car les images sont de bien trop mauvaise qualité (un humain serait incapable de faire la segmentation 3D à cause d'un flou bien trop important en profondeur). Ce genre de réponse à mené mes collègues à utiliser des techniques de transparenisation au BABB pour améliorer la situation, ce qui m'a amené à développer un certain nombre de techniques d'analyse d'objets ellipsoïdaux.

6.2.1 Segmentation supervisée

Nous venons de publier un article dans BMC Bioinformatics [8] qui décrit un plugin Icy [De +12b] pour aider à la segmentation 3D d'objets ellipsoïdaux. Comme une vidéo vaut mieux qu'un long texte, je renvoie le lecteur intéressé à ce lien pour plus d'informations : <https://www.youtube.com/watch?v=MjotgTZi6RQ&feature=youtu.be>. La figure 6.5 donne aussi une idée approximative du plugin.

D'un point de vue technique, ce projet nous a permis de développer des codes efficaces de fitting d'ellipsoïdes, qui peut être exprimé comme un *problème convexe* sur le cône des matrices SDP (voir par exemple [Cal02], entre

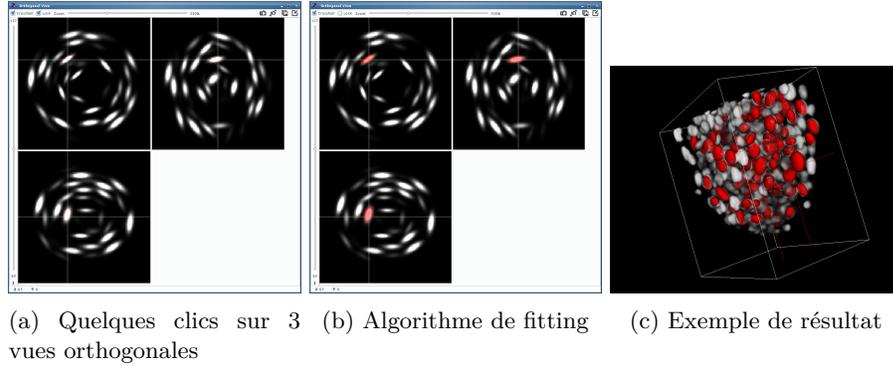


FIGURE 6.5 – Illustration du plugin FitEllipsoid sur la segmentation d’une micro-tumeur.

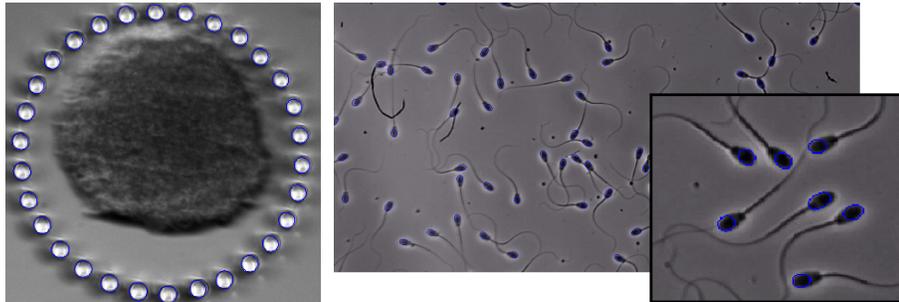
des centaines de références). Ce fut un projet assez amusant et le plugin est aujourd’hui assez unanimement apprécié de ses rares utilisateurs.

6.2.2 *Processus ponctuels marqués*

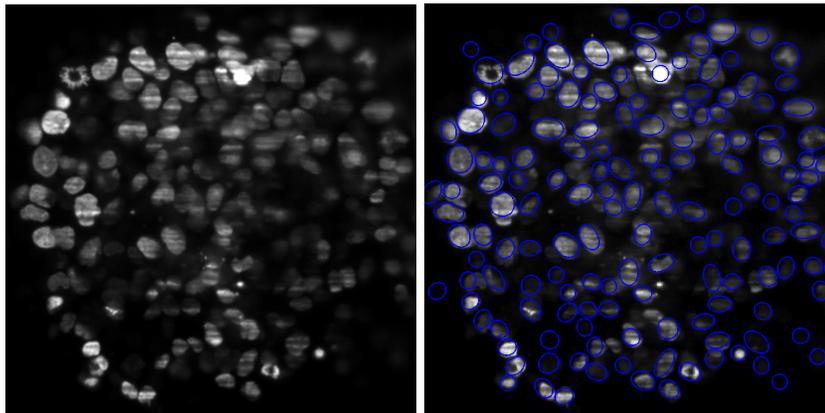
J’ai fait ma thèse à l’INRIA Sophia-Antipolis où sévissait Xavier Descombes et ses processus ponctuels marqués. J’ai donc naturellement proposé à Emmanuel Soubies - lors de son stage de M1 - de travailler sur ces méthodes pour effectuer de la segmentation de noyaux ellipsoïdaux. Nos meilleurs résultats ont été obtenus en suivant [Eld+12] qui mêle des techniques d’esprit MCMC (on jette un tas d’ellipsoïdes aléatoirement sur l’image) suivies de techniques de graph-cuts pour ne garder que le meilleur sous-ensemble admissible (ne se superposant pas) [61]. Je ne souhaite pas rentrer dans les détails de cette approche assez barbare, mais juste mentionner qu’elle nous a permis d’obtenir quelques résultats sympathiques, dont la Figure 6.6 est représentative.

6.2.3 *Tenseur de structure*

Finalement, lorsqu’on a affaire à des tissus denses, constitués localement de noyaux homogènes, on peut voir le tissu comme un continuum et chercher à trouver l’orientation et l’anisotropie locale des noyaux, sans chercher à segmenter chacun d’entre eux individuellement. Dans le travail [25], nous avons démontré que le *tenseur de structure* (voir e.g. [Wei99]) permet de résoudre ce problème précisément. Ces résultats théoriques permettent de mieux comprendre le fonctionnement du tenseur de structure qui est un outil fondamental en traitement des images et aussi de l’utiliser sereinement pour analyser des tissus. Nous avons ainsi développé un plugin Fiji pour faciliter son utilisation pour la communauté des biologistes. Nous présentons deux résultats sur la Figure 6.7.

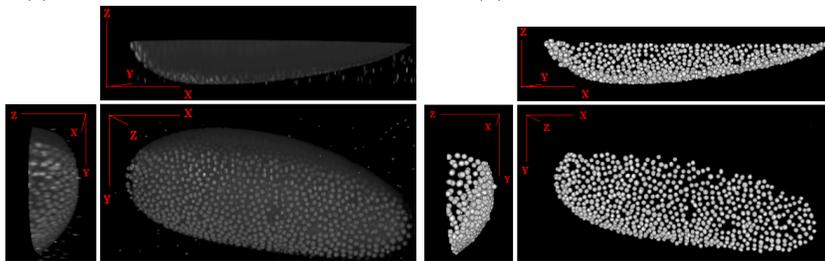


(a) Segmentation de micro-piliers de PDMS (b) Segmentation de têtes de spermatozoïdes



(c) Sphéroïde tumoral 3D au SPIM

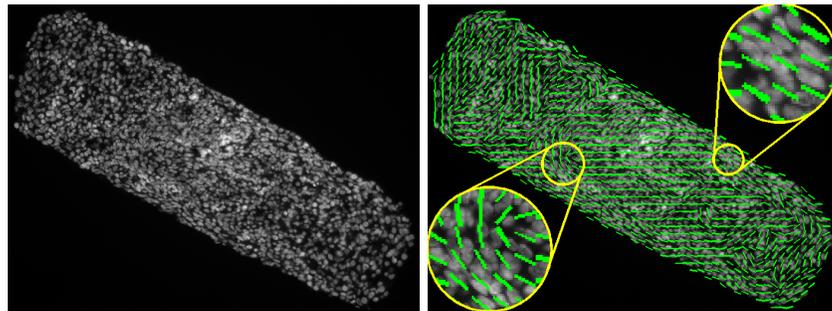
(d) Résultat de segmentation



(e) Embryon 3D de drosophile sur un SPIM

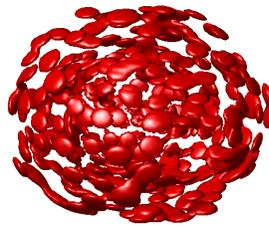
(f) Résultat de segmentation

FIGURE 6.6 – Illustration des algorithmes de segmentation par processus ponctuels marqués et graph-cuts.

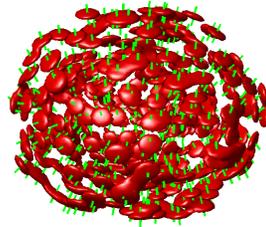


(a) Un sphéroïde tumoral en croissance dans une gouttière

(b) Analyse de l'orientation des noyaux par tenseur de structure



(c) Un sphéroïde synthétique



(d) Analyse de son orientation par tenseur de structure

FIGURE 6.7 – Illustration du tenseur de structure pour l'analyse de l'orientation de noyaux dans des tissus denses. Nous ne représentons par ici l'anisotropie (rapport des axes de ellipsoïdes) qui est elle aussi bien estimée par le tenseur.

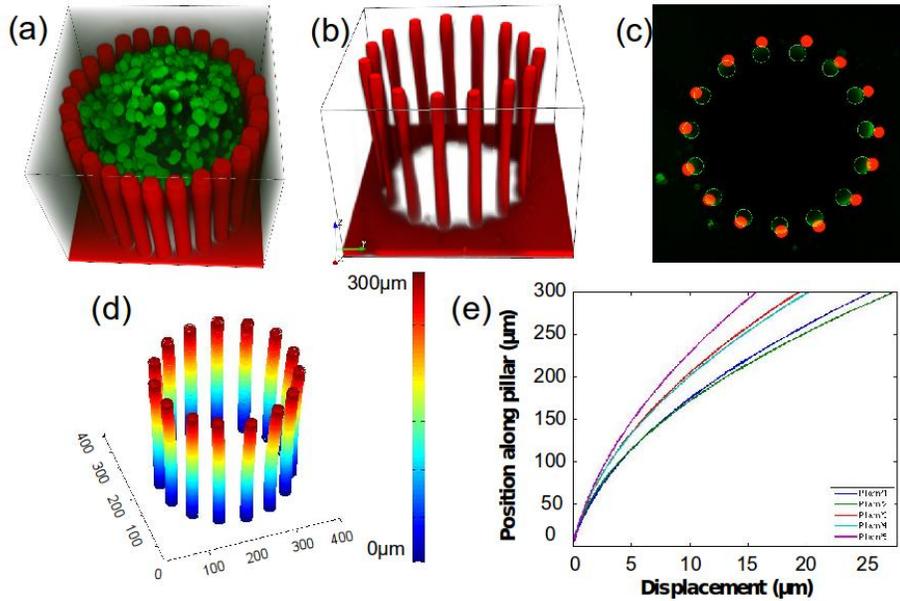


FIGURE 6.8 – Illustration du projet de mesures de forces.

6.2.4 Applications en biologie

L'analyse (semi)-automatisée d'images de microscopes est probablement un des problèmes les plus fréquemment rencontrés dans la communauté de la biologie. Les travaux que je viens de présenter ont été motivés par des questionnements concrets formulés par mes collègues Valérie Lobjois, Bernard Ducommun et Corinne Lorenzo. Je décris brièvement 2 travaux conjoints ci-dessous.

Détection de micro-piliers et mesures de forces Dans le papier [27], quelques chercheurs du LAAS ont construit un dispositif constitué de micro-piliers en PDMS placés le long d'un cercle pour mesurer les forces que peuvent exercer une micro-tumeur sur son environnement extérieur. On place un petit sphéroïde entre ces piliers dans un milieu nutritif et on le laisse croître jusqu'à ce qu'il vienne pousser les piliers. La déflexion des piliers permet de remonter aux forces exercées soit en utilisant des abaques, soit en résolvant un problème inverse basé sur la théorie des poutres (je ne décris pas ce problème ici par souci de concision). Mon rôle ici a été de poser et de résoudre le problème inverse, de retrouver la forme 3D des micro-piliers, grâce à des algorithmes de segmentation, d'effectuer du suivi des piliers dans le temps (du transport optimal a été utilisé, car il peut y avoir de très grands déplacements). Nous avons ainsi pu remonter à des valeurs numériques raisonnables des forces exercées. Le projet est résumé sur la Figure 6.8.

Géométrie dans un sphéroïde tumoral Finalement, dans le papier [13], nous avons cherché à analyser précisément l'orientation des noyaux et l'orientation de la division cellulaire dans des sphéroïdes. Ce genre d'étude peut servir à mieux comprendre la dynamique de croissance d'une tumeur dans un milieu libre ou confiné. Un outil automatisé (pour le moment il est juste supervisé) permet aussi de faire du criblage : on teste un grand nombre de médicaments sur une population de sphéroïdes tumoraux et on analyse les effets de chacun sur la division avec des analyses statistiques. Ce genre de projet nécessite - au delà de compétences techniques - de longues discussions pour bien poser le problème.

La Figure 6.9 présente le genre d'analyse proposée : on s'est focalisé sur les noyaux en périphérie du sphéroïde et on a analysé leur orientation par rapport au bord du sphéroïde. Les conclusions de biologie cellulaire sont les suivantes :

1. En milieu libre, les noyaux sont plutôt allongés dans une direction perpendiculaire à la frontière.
2. En milieu confiné (un bain d'agarose) cet allongement et l'alignement à la frontière décroît.
3. Ce phénomène devient très marqué lorsqu'on ajoute de la Latrunculin au milieu qui tend à désorganiser les filaments d'actine.
4. Les temps de division cellulaires sont allongés dans un milieu confiné.

6.2.5 *Ouverture*

J'ai aujourd'hui trouvé un partenaire en optique compétent en la personne de Thomas Mangeat. Ceci me laisse rêver de faire des progrès assez conséquents sur les techniques de microscopie, notamment en super-résolution. Malheureusement, ce pan de recherche mêlant mathématiques et optique me semble inaccessible à la majorité de biologistes (ce qui explique probablement partiellement le rejet aux ANR).

En pratique, ces derniers peuvent trouver des financements pour acheter des équipements très onéreux de constructeurs tels que Zeiss et le souci qu'ils expriment est plutôt l'analyse automatisée de leurs données. Les quelques outils que nous avons développés peuvent les aider, mais il semble clair que l'apprentissage profond va les rendre caduques assez rapidement. Dans un souci d'aide à cette communauté je pense donc mener une activité partielle de "service", ce qui me familiarisera avec l'apprentissage profond et me permettra peut-être d'avoir de nouvelles idées pour ce domaine naissant. Cette activité pourra notamment avoir lieu dans le cadre du nouveau laboratoire RESTORE sur le vieillissement, dans lequel on m'a proposé de monter une équipe. Je pense que l'intérêt de ce projet dépendra fortement de ma capacité à trouver des financements et à co-encadrer des étudiants avec des biologistes. Traiter des données réelles peut effectivement prendre énormément de temps, ce qui peut peser beaucoup sur des travaux plus fondamentaux.

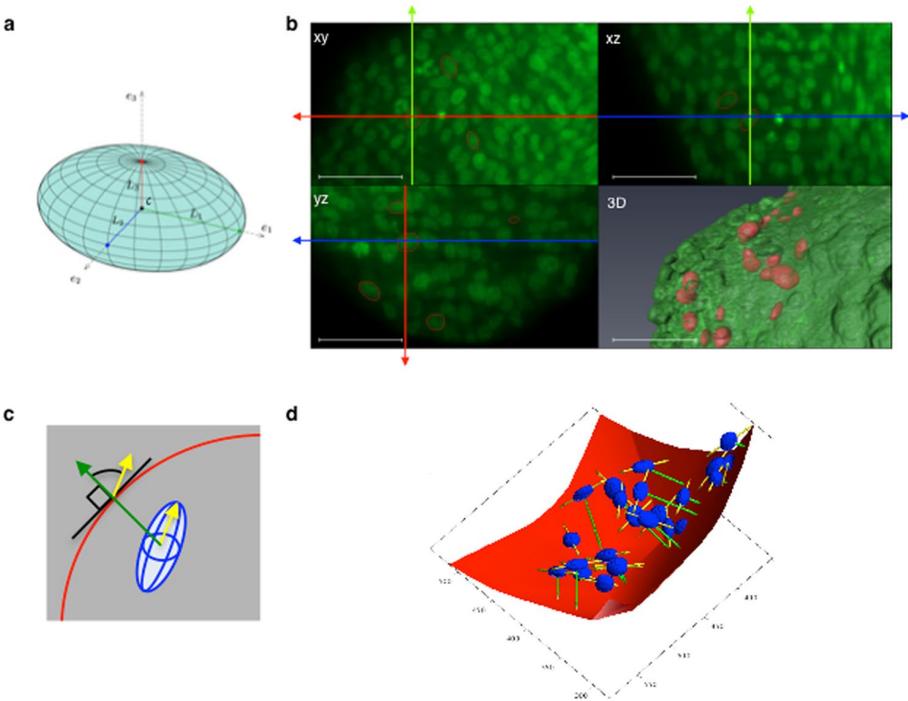


FIGURE 6.9 – Illustration du projet d'analyse de la géométrie des noyaux dans un sphéroïde tumoral.

Chapitre 7

Synthèse des pistes de recherche

Pour conclure cette HDR, je propose une synthèse de mes pistes de recherche actuelles et à venir dans le tableau 7.1. La signification des symboles est la suivante :

- 😊, 😊, 😊, 😊 : projets en cours dont le terme/la probabilité d'émergence dépend du sourire du smiley.
- 😊 Projets en cours mais dont l'issue est encore incertaine.
- 🟢 Projets essentiellement bouclés (article rédigé).
- 🟡 Projets à long terme (mais ayant commencé) ayant une issue complètement incertaine et pouvant déboucher sur des reconversions thématiques.

TABLE 7.1 – Résumé de mes pistes de recherche.

Optimisation sur des espaces de mesures	
😊	Développement de systèmes dynamiques de particules pouvant atteindre la dizaine de million de particules (méthodes multipolaires et calcul parallèle). P. Ciuciu
😊	Analyse des k-faces de la boule unité de la variation totale du gradient. C. Boyer, Y. de Castro, V. Duval, F. de Gournay
😊	Analyse de convergence des algorithmes d'échange en programmation semi-infinie. A. Flinth & F. de Gournay
😊	Analyse de convergence des algorithmes de discrétisation en programmation semi-infinie. A. Flinth & F. de Gournay
😊	Développement de nouvelles méthodes sans grille pour la minimisation de la variation totale du gradient. A. Flinth & F. de Gournay
😬	Analyse de convergence de dynamiques particulières vers un flot de gradient. J. Bolte, F. de Gournay, J. Kahn
Théorie de l'échantillonnage	
😊	Analyse de la reconstruction ℓ^1 dans des trames redondantes (formulation de synthèse). C. Boyer, J. Kahn, M. März
😊	Analyse du rôle des valeurs propres minimales coniques. J. Kahn
😊	Concentration sur des chaînes de Markov à valeurs matricielles. N. Chauffert
😊	Échantillonnage de fonctions u où les mesures sont de la forme $y_i = \int u d\mu_i$ et où μ_i est une mesure singulière supportée sur une courbe.
😬	Apprentissage de densités optimales d'échantillonnage pour l'IRM. F. de Gournay & A. Gossard
😬	Échantillonnage compressé non aléatoire (notamment lien entre les théories d'échantillonnage à la Beurling et les théories d'échantillonnage compressé).
Approximation, estimation et apprentissage d'opérateurs	
😊	Décomposition rapide d'opérateurs de type produit-convolution dans des bases d'ondelettes. P. Escande
😊	Apprentissage de sous-espaces d'opérateurs structurés pour la calibration de systèmes de mesure. V. Debarnot
😊	Analyse d'opérateurs intégraux dont le noyau est une mesure singulière supportée sur une courbe.
😊	Développement de nouveaux réseaux de neurones instationnaires de type Convolutional Neural Networks. V. Debarnot
Imagerie	
😊	Correction automatique des erreurs de gradient en IRM. M. März & C. Lazarus
😊	Échantillonnage compressé Sparkling en IRM 3D. P. Ciuciu, A. Vignaud
😊	Défloutage aveugle instationnaire (lifting et relaxation convexe et apprentissage profond). V. Debarnot & T. Mangeat
😊	Super-résolution avec opérateur instationnaire inconnu en microscopie de type molécule unique (lifting et relaxation convexe et apprentissage profond). V. Debarnot & T. Mangeat
😊	Évaluation multi-vue de l'atténuation. J. Cuesta & V. Debarnot
😊	Analyse automatisée et tracking de cellules (service aux biologistes) V. Debarnot & L. Lebrat

Chapitre 8

Références

8.1 Mes articles

8.1.1 *Preprints*

- [1] FLINTH, AXEL AND DE GOURNAY, FRÉDÉRIC AND WEISS, PIERRE. “On the linear convergence rates of exchange and continuous methods for total variation minimization.” preprint. 2019.
- [2] LAZARUS, CAROLE AND MÄRZ, MAXIMILIAN AND WEISS, PIERRE. “Correcting the side effects of ADC filtering in MR image reconstruction”. preprint. 2018.

8.1.2 *Chapitres de livre*

- [3] Emmanuel SOUBIES, Pierre WEISS et Xavier DESCOMBES. “Graph cut based segmentation of predefined shapes : Applications to biological imaging”. In : *Pattern Recognition Applications and Methods*. Springer, 2015, p. 153–170.

8.1.3 *Journal articles (international)*

- [4] AOUN, LAURENE AND LARNIER, STANISLAS AND WEISS, PIERRE AND HERBULOT, ARIANE AND DUCOMMUN, BERNARD AND VIEU, CHRISTOPHE AND LOBJOIS, VALÉRIE. “ Measure and characterization of the forces exerted by growing multicellular spheroids using microdevice arrays.” In : *PLOS ONE* (2019).
- [5] BOYER, CLAIRE AND CHAMBOLLE, ANTONIN AND DE CASTRO, YOHANN AND DUVAL, VINCENT AND DE GOURNAY, FRÉDÉRIC AND WEISS, PIERRE. “On Representer Theorems and Convex Regularization.” In : *SIAM journal on Optimization* (2019).

- [6] DEBARNOT, VALENTIN AND ESCANDE, PAUL AND WEISS, PIERRE. “A scalable estimator of sets of integral operators.” In : *Inverse Problems* (2019).
- [7] Valentin DEBARNOT, Jonas KAHN et Pierre WEISS. “Multiview Attenuation Estimation and Correction”. In : *Journal of Mathematical Imaging Vision* (2019).
- [8] Jérôme FEHRENBACH, Bastien KOVAC, Ludivine GUILLAUME et Pierre WEISS. “FitEllipsoid : a fast supervised ellipsoid segmentation plugin”. In : *BMC Bioinformatics* (2019).
- [9] LAZARUS, CAROLE AND WEISS, PIERRE AND EL GUEDDARI, LOUBNA AND MAUCONDUIT, FRANCK AND VIGNAUD, ALEXANDRE AND CIUCIU, PHILIPPE. “3D Sparkling trajectories for high-resolution T2*-weighted Magnetic Resonance Imaging.” In : *NMR in biomedicine* (2019).
- [10] LAZARUS, CAROLE AND WEISS, PIERRE AND VIGNAUD, ALEXANDRE AND CIUCIU, PHILIPPE. “SPARKLING : variable-density k-space filling curves for accelerated T2*-weighted MRI”. In : *Magnetic Resonance in Medicine* (2019).
- [11] LEBRAT, LÉO AND DE GOURNAY, FRÉDÉRIC AND KAHN, JONAS AND WEISS, PIERRE. “Optimal transport approximation of densities”. In : *SIAM journal on Imaging Science* (2019).
- [12] Pierre WEISS, Paul ESCANDE, Gabriel BATHIE et Yiqiu DONG. “Contrast Invariant SNR”. In : *International Journal of Computer Vision* (2019). URL : <https://hal.archives-ouvertes.fr/hal-01370251>.
- [13] Annaïck DESMAISON, Ludivine GUILLAUME, Sarah TRICLIN, Pierre WEISS, Bernard DUCOMMUN et Valérie LOBJOIS. “Impact of physical confinement on nuclei geometry and cell division dynamics in 3D spheroids.” In : *Nature Scientific Reports* (2018).
- [14] Paul ESCANDE et Pierre WEISS. “Accelerating l1-l2 deblurring using wavelet expansions of operators”. In : *Journal of Computational and Applied Mathematics* 343 (2018), p. 373–396. ISSN : 0377-0427. DOI : <https://doi.org/10.1016/j.cam.2018.04.063>. URL : <http://www.sciencedirect.com/science/article/pii/S0377042718302668>.
- [15] Axel FLINTH et Pierre WEISS. “Exact solutions of infinite dimensional total-variation regularized problems”. In : *Information & Inference* (2018).
- [16] Carole LAZARUS, Pierre WEISS, Alexandre VIGNAUD et Philippe CIUCIU. “An Empirical Study of the Maximum Degree of Undersampling in Compressed Sensing for T2*-weighted MRI.” In : *Magnetic Resonance Imaging* (2018).

- [17] Jérémie BIGOT, Paul ESCANDE et Pierre WEISS. “Estimation of linear operators from scattered impulse responses”. In : *Applied and Computational Harmonic Analysis* (2017). ISSN : 1063-5203. DOI : <https://doi.org/10.1016/j.acha.2017.12.002>. URL : <http://www.sciencedirect.com/science/article/pii/S1063520317301409>.
- [18] BOYER, CLAIRE AND BIGOT, JÉRÉMIE AND WEISS, PIERRE. “Compressed sensing with structured sparsity and structured acquisition”. In : *Applied and Computational Harmonic Analysis* (2017). ISSN : 1063-5203. DOI : <http://dx.doi.org/10.1016/j.acha.2017.05.005>. URL : <http://www.sciencedirect.com/science/article/pii/S1063520317300489>.
- [19] Nicolas CHAUFFERT, Philippe CIUCIU, Jonas KAHN et Pierre WEISS. “A projection method on measures sets”. In : *Constructive Approximation* 45.1 (2017), p. 83–111.
- [20] Paul ESCANDE et Pierre WEISS. “Approximation of Integral Operators Using Product-Convolution Expansions”. In : *Journal of Mathematical Imaging and Vision* 58.3 (2017), p. 333–348. ISSN : 1573-7683. DOI : 10.1007/s10851-017-0714-8. URL : <https://doi.org/10.1007/s10851-017-0714-8>.
- [21] Paul ESCANDE, Pierre WEISS et Wenxing ZHANG. “A variational model for multiplicative structured noise removal”. In : *Journal of Mathematical Imaging and Vision* 57.1 (2017), p. 43–55.
- [22] Jérémie BIGOT, Claire BOYER et Pierre WEISS. “An analysis of block sampling strategies in compressed sensing”. In : *IEEE Transactions on Information Theory* 62.4 (2016), p. 2125–2139.
- [23] Claire BOYER, Nicolas CHAUFFERT, Philippe CIUCIU, Jonas KAHN et Pierre WEISS. “On the generation of sampling schemes for magnetic resonance imaging”. In : *SIAM Journal on Imaging Sciences* 9.4 (2016), p. 2039–2072.
- [24] Nicolas CHAUFFERT, Pierre WEISS, Jonas KAHN et Philippe CIUCIU. “A projection algorithm for gradient waveforms design in Magnetic Resonance Imaging”. In : *IEEE transactions on medical imaging* 35.9 (2016), p. 2026–2039.
- [25] W. ZHANG, J. FEHRENBACH, A. DESMAISON, V. LOBJOIS, B. DUCOMMUN et P. WEISS. “Structure Tensor Based Analysis of Cells and Nuclei Organization in Tissues”. In : *IEEE Transactions on Medical Imaging* 35.1 (2016), p. 294–306. ISSN : 0278-0062. DOI : 10.1109/TMI.2015.2470093.
- [26] Paul ESCANDE et Pierre WEISS. “Sparse wavelet representations of spatially varying blurring operators”. In : *SIAM Journal on Imaging Sciences* 8.4 (2015), p. 2976–3014.

- [27] Laurène AOUN, Pierre WEISS, Adrian LABORDE, Bernard DUCOMMUN, Valérie LOBJOIS et Christophe VIEU. “Microdevice arrays of high aspect ratio poly (dimethylsiloxane) pillars for the investigation of multicellular tumour spheroid mechanical properties”. In : *Lab on a Chip* 14.13 (2014), p. 2344–2353.
- [28] Claire BOYER, Pierre WEISS et Jérémie BIGOT. “An algorithm for variable density sampling with block-constrained acquisition”. In : *SIAM Journal on Imaging Sciences* 7.2 (2014), p. 1080–1107.
- [29] Nicolas CHAUFFERT, Philippe CIUCIU, Jonas KAHN et Pierre WEISS. “Variable Density Sampling with Continuous Trajectories”. In : *SIAM Journal on Imaging Sciences* 7.4 (2014), p. 1962–1992. DOI : 10.1137/130946642. eprint : <https://doi.org/10.1137/130946642>. URL : <https://doi.org/10.1137/130946642>.
- [30] Jérôme FEHRENBACH et Pierre WEISS. “Processing stationary noise : Model and parameter selection in variational methods”. In : *SIAM Journal on Imaging Sciences* 7.2 (2014), p. 613–640.
- [31] Andres HOYOS-IDROBO, Peter WEISS, Aurélien MASSIRE, Alexis AMADON et Nicolas BOULANT. “On variant strategies to solve the magnitude least squares optimization problem in parallel transmission pulse design and under strict SAR and power constraints”. In : *Medical Imaging, IEEE Transactions on* 33.3 (2014), p. 739–748.
- [32] Jérôme FEHRENBACH, Pierre WEISS et Corinne LORENZO. “Variational algorithms to remove stationary noise : applications to microscopy imaging”. In : *Image Processing, IEEE Transactions on* 21.10 (2012), p. 4420–4430.
- [33] Corinne LORENZO et al. “Live cell division dynamics monitoring in 3D large spheroid tumor models using light sheet microscopy”. In : *Cell division* 6.1 (2011), p. 22.
- [34] Pierre WEISS, Alexandre FOURNIER, Laure BLANC-FÉRAUD et Gilles AUBERT. “On the illumination invariance of the level lines under directed light : Application to change detection”. In : *SIAM Journal on Imaging Sciences* 4.1 (2011), p. 448–471.
- [35] Michael K NG, Pierre WEISS et Xiaoming YUAN. “Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods”. In : *SIAM journal on Scientific Computing* 32.5 (2010), p. 2710–2736.
- [36] Michael POOLE, Pierre WEISS, Hector Sanchez LOPEZ, Michael NG et Stuart CROZIER. “Minimax current density coil design”. In : *Journal of Physics D : Applied Physics* 43.9 (2010), p. 095001.
- [37] Pierre WEISS, Laure BLANC-FÉRAUD et Gilles AUBERT. “Efficient schemes for total variation minimization under constraints in image processing”. In : *SIAM journal on Scientific Computing* 31.3 (2009), p. 2047–2080.

8.1.4 *Journal articles (national)*

- [38] Mikael CARLAVAN, Pierre WEISS et Laure BLANC-FÉRAUD. “Régularité et parcimonie pour les problèmes inverses en imagerie : algorithmes et comparaisons”. In : *Traitement du signal* (2010).

8.1.5 *Conference articles (international)*

- [39] BOYER, CLAIRE AND KAHN, JONAS AND MÄRZ, MAXIMILIAN AND WEISS, PIERRE. “Compressed sensing and the synthesis formulation”. In : *Proceedings SPARS*. 2019.
- [40] DEBARNOT, VALENTIN AND WEISS, PIERRE. “A scalable estimator of space varying blurs - application in super-resolution.” In : *Proceedings SPARS*. 2019.
- [41] Valentin DEBARNOT, Paul ESCANDE, Thomas MANGEAT et Pierre WEISS. “Blind-deblurring : a learning based approach”. In : *Proceedings of OSA, Mathematics in imaging*. 2019.
- [42] FLINTH, AXEL AND WEISS, PIERRE. “Iterative Discretization of Optimization Problems Related to Superresolution.” In : *Proceedings SAMPTA*. 2019.
- [43] Axel FLINTH, Frédéric de GOURNAY et Pierre WEISS. “Eventual linear convergence rate of an exchange algorithm for super-resolution.” In : *Proceedings SPARS*. 2019.
- [44] Carole LAZARUS, Pierre WEISS, Loubna EL GUEDDARI, Alexandre VIGNAUD et Philippe CIUCIU. “3D SPARKLING for accelerated ex vivo T2*-weighted MRI with compressed sensing”. In : *Proceedings 27th Scientific Meeting, International Society for Magnetic Resonance in Medicine*. 2019.
- [45] Valentin DEBARNOT, Paul ESCANDE et Pierre WEISS. “A scalable estimator of sets of integral operators”. In : *Proceedings of iTWIST 2018*. 2018.
- [46] Carole LAZARUS, Pierre WEISS, Loubna EL GUEDDARI, Franck MAUCONDUIT, I. ZEMMOURA, Alexandre VIGNAUD et Philippe CIUCIU. “Distribution-controlled and optimally spread non-Cartesian sampling curves for accelerated in vivo brain imaging at 7 Tesla”. In : *Proceedings 26th Scientific Meeting, International Society for Magnetic Resonance in Medicine*. 2018.
- [47] Valentin DEBARNOT, Jonas KAHN et Pierre WEISS. “Multiview Attenuation Computation and Correction”. In : *SPARS Proceedings*. SPARS. 2017.
- [48] Valentin DEBARNOT, Jonas KAHN et Pierre WEISS. “Multiview Attenuation Computation and Correction”. In : *Mathematics in Imaging*. Optical Society of America. 2017, MTu1C-2.

- [49] Carole LAZARUS, Pierre WEISS, Nicolas CHAUFFERT, Franck MAUCONDUIT, A VIGNAUD et P CIUCIU. “SPARKLING : Novel Non-Cartesian Sampling Schemes for Accelerated 2D Anatomical Imaging at 7T Using Compressed Sensing”. In : *Proceedings 25th Scientific Meeting, International Society for Magnetic Resonance in Medicine, Honolulu, HW, USA*. 2017.
- [50] Nicolas CHAUFFERT, Pierre WEISS, Marianne BOUCHER, Sébastien MÉRIAUX et Philippe CIUCIU. “Variable density sampling based on physically plausible gradient waveform. Application to 3D MRI angiography”. In : *ISBI proceedings*. 2015.
- [51] Jérôme FEHRENBACH, Mila NIKOLOVA, Gabriele STEIDL et Pierre WEISS. “Bilevel Image Denoising Using Gaussianity Tests”. In : *Scale Space and Variational Methods in Computer Vision*. Springer, 2015, p. 117–128.
- [52] Laurene AOUN, Stanislas LARNIER, Pierre WEISS, Ariane HERBULOT, Bernard DUCOMMUN, Valérie LOBJOIS et Christophe VIEU. “PDMS microdevice arrays for measuring the forces exerted by growing multicellular tumor spheroids”. In : *The 18th International Conference on Miniaturized Systems for Chemistry and Life Sciences (MicroTAS 2014)*. 2014.
- [53] Alexis HUCK, François de VIEILLEVILLE, Pierre WEISS et Manuel GRIZONNET. “Hyperspectral pan-sharpening : a variational convex constrained formulation to impose parallel level lines, solved with ADMM”. In : *Proceedings of Whispers*. 2014.
- [54] G LE CORRE, P WEISS, B DUCOMMUN et C LORENZO. “Optical signature of multicellular tumor spheroid using index-mismatch-induced spherical aberrations”. In : *SPIE BiOS*. International Society for Optics et Photonics. 2014, p. 894111–894111.
- [55] Claire BOYER, Jérémie BIGOT et Pierre WEISS. “A theoretical analysis of a sampling strategy among blocks of measurements for compressed sensing.” In : *SampTA-10th Conference International Conference on Sampling Theory and Applications*. 2013.
- [56] Nicolas CHAUFFERT, Philippe CIUCIU, Fabrice GAMBOA et Pierre WEISS. “From variable density sampling to continuous sampling using Markov chains.” In : *SampTA-10th Conference International Conference on Sampling Theory and Applications*. 2013.
- [57] Nicolas CHAUFFERT, Philippe CIUCIU, Jonas KAHN et Pierre WEISS. “Travelling salesman-based variable density sampling”. In : *SampTA-10th Conference International Conference on Sampling Theory and Applications*. 2013, p. 509–512.
- [58] Nicolas CHAUFFERT, Philippe CIUCIU et Peter WEISS. “Variable density compressed sensing in MRI. Theoretical VS heuristic sampling strategies”. In : *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*. IEEE. 2013, p. 298–301.

- [59] Paul ESCANDE, Pierre WEISS et François MALGOUYRES. “Image restoration using sparse approximations of spatially varying blur operators in the wavelet domain”. In : *Journal of Physics : Conference Series*. T. 464. 1. IOP Publishing. 2013, p. 012004.
- [60] Paul ESCANDE, Pierre WEISS et François MALGOUYRES. “Spatially varying blur recovery. diagonal approximations in the wavelet domain”. In : *Proceedings of ICPRAM*. 2013.
- [61] Emmanuel SOUBIES, Pierre WEISS et Xavier DESCOMBES. “A 3D Segmentation Algorithm for Ellipsoidal Shapes-Application to Nuclei Extraction.” In : *ICPRAM*. 2013, p. 97–105.
- [62] Claire BOYER, Philippe CIUCIU, Pierre WEISS et Sébastien MÉRIAUX. “HYR 2 PICS : Hybrid regularized reconstruction for combined parallel imaging and compressive sensing in MRI”. In : *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*. IEEE. 2012, p. 66–69.
- [63] Jérôme FEHRENBACH, Pierre WEISS et Corinne LORENZO. “Variational algorithms to remove stripes : a generalization of the negative norm models”. In : *Proc. ICPRAM*. 2012, p. 1–6.
- [64] Matthieu KOWALSKI, Pierre WEISS, Alexandre GRAMFORT et Sandrine ANTHOINE. “Accelerating ISTA with an active set strategy”. In : *OPT 2011 : 4th International Workshop on Optimization for Machine Learning*. 2011, p. 7.
- [65] F de VIELLEVILLE, P WEISS, V LOBJOIS et D KOUAMÉ. “Alternating direction method of multipliers applied to 3D light sheet fluorescence microscopy image deblurring using GPU hardware”. In : *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE. 2011, p. 4872–4875.
- [66] MS POOLE, P WEISS, H Sanchez LOPEZ, M NG et S CROZIER. “Minimax Current Density Coil Design.” In : *Proceedings of the International Society for Magnetic Resonance in Medicine*. T. 18. 2010.
- [67] MS POOLE, P WEISS, H Sanchez LOPEZ, M NG et S CROZIER. “Temperature Characteristics of Gradient Coils with Minimax Current Density”. In : *Proceedings of the International Society for Magnetic Resonance in Medicine*. T. 18. 2010, p. 564.
- [68] Mikael CARLAVAN, Pierre WEISS, Laure BLANC-FÉRAUD et Josiane ZERUBIA. “Complex wavelet regularization for solving inverse problems in remote sensing”. In : *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*. T. 3. IEEE. 2009, p. III–212.
- [69] Pierre WEISS et Laure BLANC-FÉRAUD. “A proximal method for inverse problems in image processing”. In : *Signal Processing Conference, 2009 17th European*. IEEE. 2009, p. 1374–1378.

- [70] Pierre WEISS, Mikaël CARLAVAN, Laure BLANC-FÉRAUD et Josiane ZERUBIA. “Smoothing techniques for convex problems. Applications in image processing.” In : *SAMPTA*. 2009.
- [71] Alexandre FOURNIER, Pierre WEISS, Laure BLANC-FÉRAUD et Gilles AUBERT. “A contrast equalization procedure for change detection algorithms : applications to remotely sensed images of urban areas”. In : *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE. 2008, p. 1–4.
- [72] Pierre WEISS, Laure BLANC-FÉRAUD, Thomas ANDRÉ et Marc ANTONINI. “Compression artifacts reduction using variational methods : algorithms and experimental study”. In : *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE. 2008, p. 1173–1176.

8.1.6 Conference articles (national)

- [73] DE GOURNAY, FRÉDÉRIC AND KAHN, JONAS AND LEBRAT, LÉO AND WEISS, PIERRE. “Approches variationnelles pour le stippling : distances L2 ou transport optimal?” In : *GRETSI Proceedings*. 2017.
- [74] Carole LAZARUS, Pierre WEISS, Nicolas CHAUFFERT, Alexandre VIGNAUD et Philippe CIUCIU. “SPARKLING : nouveaux schémas d’échantillonnage compressif prospectif pour l’IRM haute résolution”. In : *GRETSI Proceedings*. 2017.
- [75] Claire BOYER, Jérémie BIGOT et Pierre WEISS. “Echantillonnage compressé avec acquisition structurée et parcimonie structurée”. In : *Proceedings of GRETSI*. 2015.
- [76] Nicolas CHAUFFERT, Philippe CIUCIU, Jonas KAHN et Pierre WEISS. “Comment représenter une image avec un spaghetti?” In : *Proceedings of GRETSI*. 2015.
- [77] Paul ESCANDE et Pierre WEISS. “Approximation et Estimation des Opérateurs de Flou Variable”. In : *Proceedings of GRETSI*. 2015.
- [78] Pierre WEISS, Nicolas CHAUFFERT, Philippe CIUCIU et Jonas KAHN. “Sur la génération de schémas d’échantillonnage compressé en IRM”. In : *Proceedings of GRETSI*. 2015.
- [79] François de VIEILLEVILLE, Pierre WEISS, Valérie LOBJOIS et Denis KOUAMÉ. “Restauration 3D en imagerie fluorescente par feuille de lumière avec implantation sur GPU”. In : *Proceedings of GRETSI*. 2011.
- [80] Mikael CARLAVAN, Pierre WEISS, Laure BLANC-FÉRAUD et Josiane ZERUBIA. “Algorithme rapide pour la restauration d’image régularisée sur les coefficients d’ondelettes”. In : *GRETSI 2009*. GRETSI, Groupe d’Etudes du Traitement du Signal et des Images. 2009.

- [81] Mikael CARLAVAN, Pierre WEISS, Laure BLANC-FÉRAUD et Josiane ZERUBIA. “Reconstruction d’images satellitaires à partir d’un échantillonnage irrégulier”. In : *GRETSI*. 2008.
- [82] Pierre WEISS, Laure BLANC-FÉRAUD et Gilles AUBERT. “Sur la complexité et la rapidité d’algorithmes pour la minimisation de la variation totale sous contraintes”. In : *GRETSI, 2007*. GRETSI, Groupe d’Etudes du Traitement du Signal et des Images. 2007.

8.2 Articles additionnels

- [BT09a] Amir BECK et Marc TEBoulLE. “A fast iterative shrinkage-thresholding algorithm for linear inverse problems”. In : *SIAM journal on imaging sciences* 2.1 (2009), p. 183–202.
- [Cha04] Antonin CHAMBOLLE. “An algorithm for total variation minimization and applications”. In : *Journal of Mathematical imaging and vision* 20.1-2 (2004), p. 89–97.
- [FP11] Jalal M FADILI et Gabriel PEYRÉ. “Total variation projection with first order schemes”. In : *IEEE Transactions on Image Processing* 20.3 (2011), p. 657–669.
- [Nes05] Yu NESTEROV. “Smooth minimization of non-smooth functions”. In : *Mathematical programming* 103.1 (2005), p. 127–152.
- [dGJ18] Alexandre D’ASPREMONT, Cristóbal GUZMÁN et Martin JAGGI. “Optimal Affine-Invariant Smooth Minimization Algorithms”. In : *SIAM Journal on Optimization* 28.3 (2018), p. 2384–2405.
- [Bar72] Richard E BARLOW. *Statistical inference under order restrictions; the theory and application of isotonic regression*. Rapp. tech. 1972.
- [KRS15] Rasmus KYNG, Anup RAO et Sushant SACHDEVA. “Fast, Provable Algorithms for Isotonic Regression in all L_p -norms”. In : *Advances in Neural Information Processing Systems*. 2015, p. 2701–2709.
- [ST04] Daniel A SPIELMAN et Shang-Hua TENG. “Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems”. In : *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*. ACM. 2004, p. 81–90.
- [Ame+14] Dennis AMELUNXEN, Martin LOTZ, Michael B MCCOY et Joel A TROPP. “Living on the edge : Phase transitions in convex programs with random data”. In : *Information and Inference : A Journal of the IMA* 3.3 (2014), p. 224–294.
- [KPR16] Vladimir KOLMOGOROV, Thomas POCK et Michal ROLINEK. “Total variation on a tree”. In : *SIAM Journal on Imaging Sciences* 9.2 (2016), p. 605–636.
- [Sto14] Quentin F. STOUT. *Fastest Isotonic Regression Algorithms*. 2014. URL : web.eecs.umich.edu/~qstout/IsoRegAlg_140812.pdf.

- [UFW17] Michael UNSER, Julien FAGEOT et John Paul WARD. “Splines Are Universal Solutions of Linear Inverse Problems with Generalized TV Regularization”. In : *SIAM Review* 59.4 (2017), p. 769–793.
- [Cha+12] Venkat CHANDRASEKARAN, Benjamin RECHT, Pablo A PARRILO et Alan S WILLSKY. “The convex geometry of linear inverse problems”. In : *Foundations of Computational mathematics* 12.6 (2012), p. 805–849.
- [Kle57] V. L. KLEE. “Extremal structure of convex sets”. en. In : *Archiv der Mathematik* 8.3 (août 1957), p. 234–240. ISSN : 0003-889X, 1420-8938. DOI : 10.1007/BF01899998. URL : <https://link.springer.com/article/10.1007/BF01899998> (visité le 01/03/2018).
- [CRT06] Emmanuel J CANDÈS, Justin ROMBERG et Terence TAO. “Robust uncertainty principles : Exact signal reconstruction from highly incomplete frequency information”. In : *IEEE Transactions on information theory* 52.2 (2006), p. 489–509.
- [CF14] Emmanuel J CANDÈS et Carlos FERNANDEZ-GRANDA. “Towards a Mathematical Theory of Super-resolution”. In : *Communications on Pure and Applied Mathematics* 67.6 (2014), p. 906–956.
- [DP15] Vincent DUVAL et Gabriel PEYRÉ. “Exact support recovery for sparse spikes deconvolution”. In : *Foundations of Computational Mathematics* 15.5 (2015), p. 1315–1355.
- [FJ75] S. D. FISHER et J. W. JEROME. “Spline solutions to L1 extremal problems in one and several variables”. In : *Journal of Approximation Theory* 13.1 (1975), p. 73–83.
- [DL93] Ronald A DEVORE et George G LORENTZ. *Constructive approximation*. T. 303. Springer Science & Business Media, 1993.
- [DeV98] Ronald A DEVORE. “Nonlinear approximation”. In : *Acta numerica* 7 (1998), p. 51–150.
- [DT05] David L DONOHO et Jared TANNER. “Sparse nonnegative solution of underdetermined linear equations by linear programming”. In : *Proceedings of the National Academy of Sciences of the United States of America* 102.27 (2005), p. 9446–9451.
- [Dat10] Jon DATTORRO. *Convex optimization & Euclidean distance geometry*. Lulu.com, 2010.
- [ROF92] Leonid I RUDIN, Stanley OSHER et Emad FATEMI. “Nonlinear total variation based noise removal algorithms”. In : *Physica D : nonlinear phenomena* 60.1-4 (1992), p. 259–268.
- [Fle57] WH FLEMING. “Functions with generalized gradient and generalized surfaces”. In : *Annali di Matematica Pura ed Applicata* 44.1 (1957), p. 93–103.

- [BP13] Kristian BREDIES et Hanna Katriina PIKKARAINEN. “Inverse problems in spaces of measures”. In : *ESAIM : Contr Optim Ca.* 19.1 (2013), p. 190–218.
- [CB18] Lenaïc CHIZAT et Francis BACH. “On the global convergence of gradient descent for over-parameterized models using optimal transport”. In : *Advances in neural information processing systems*. 2018, p. 3040–3050.
- [Don06] David L DONOHO. “Compressed sensing”. In : *IEEE Transactions on information theory* 52.4 (2006), p. 1289–1306.
- [FR13] Simon FOUCART et Holger RAUHUT. *A mathematical introduction to compressive sensing*. T. 1. 3. Birkhäuser Basel, 2013.
- [CP11a] Emmanuel J CANDÈS et Yaniv PLAN. “A probabilistic and RIPless theory of compressed sensing”. In : *IEEE transactions on information theory* 57.11 (2011), p. 7235–7254.
- [Adc+17] Ben ADCOCK, Anders C HANSEN, Clarice POON et Bogdan ROMAN. “Breaking the coherence barrier : A new theory for compressed sensing”. In : *Forum of Mathematics, Sigma*. T. 5. Cambridge University Press. 2017.
- [Lez98] Pascal LEZAUD. “Chernoff-type bound for finite Markov chains”. In : *Annals of Applied Probability* (1998), p. 849–867.
- [Kar+07] Vladislav KARGIN et al. “A large deviation inequality for vector functions on finite reversible Markov chains”. In : *The Annals of Applied Probability* 17.4 (2007), p. 1202–1221.
- [ABS13] Hedy ATTOUCH, Jérôme BOLTE et Benar Fux SVAITER. “Convergence of descent methods for semi-algebraic and tame problems : proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods”. In : *Mathematical Programming* 137.1-2 (2013), p. 91–129.
- [Ben+15] Jean-David BENAMOU, Guillaume CARLIER, Marco CUTURI, Luca NENNA et Gabriel PEYRÉ. “Iterative Bregman projections for regularized transportation problems”. In : *SIAM Journal on Scientific Computing* 37.2 (2015), A1111–A1138.
- [De +12a] Fernando DE GOES, Katherine BREEDEN, Victor OSTROMOUKHOV et Mathieu DESBRUN. “Blue noise through optimal transport”. In : *ACM Transactions on Graphics (TOG)* 31.6 (2012), p. 171.
- [Lév15] Bruno LÉVY. “A numerical algorithm for L2 semi-discrete optimal transport in 3D”. In : *ESAIM : Mathematical Modelling and Numerical Analysis* 49.6 (2015), p. 1693–1715.
- [Nam+13] Sangnam NAM, Mike E DAVIES, Michael ELAD et Rémi GRIBONVAL. “The cosparsity analysis model and algorithms”. In : *Applied and Computational Harmonic Analysis* 34.1 (2013), p. 30–56.

- [RSV08] Holger RAUHUT, Karin SCHNASS et Pierre VANDERGHEYNST. “Compressed sensing and redundant dictionaries”. In : *IEEE Transactions on Information Theory* 54.5 (2008), p. 2210–2219.
- [Can+11] Emmanuel J CANDES, Yonina C ELDAR, Deanna NEEDELL et Paige RANDALL. “Compressed sensing with coherent and redundant dictionaries”. In : *Applied and Computational Harmonic Analysis* 31.1 (2011), p. 59–73.
- [BCR91a] Gregory BEYLKIN, Ronald COIFMAN et Vladimir ROKHLIN. “Fast Wavelet Transform and Numerical Algorithm”. In : *Commun. Pure and Applied Math.* 44 (1991), p. 141–183.
- [CM97] Ronald COIFMAN et Yves MEYER. “Wavelets, Calderón-Zygmund and multilinear operators”. In : *Cambridge Studies in Advanced Math* 48 (1997).
- [BCR91b] Gregory BEYLKIN, Ronald COIFMAN et Vladimir ROKHLIN. “Fast wavelet transforms and numerical algorithms I”. In : *Communications on pure and applied mathematics* 44.2 (1991), p. 141–183.
- [Coh+99] Albert COHEN, Ronald DEVORE, Pencho PETRUSHEV et Hong XU. “Nonlinear approximation and the space BV [inline-graphic xmlns : xlink=" http://www.w3.org/1999/xlink" xlink : href=" 01i"/]”. In : *American Journal of Mathematics* 121.3 (1999), p. 587–628.
- [AK06] Gilles AUBERT et Pierre KORNPROBST. *Mathematical problems in image processing : partial differential equations and the calculus of variations*. T. 147. Springer Science & Business Media, 2006.
- [Coh03] Albert COHEN. *Numerical analysis of wavelet methods*. T. 32. Elsevier, 2003.
- [FN03] Mário AT FIGUEIREDO et Robert D NOWAK. “An EM algorithm for wavelet-based image restoration”. In : *Image Processing, IEEE Transactions on* 12.8 (2003), p. 906–916.
- [NCB04] Ramesh NEELAMANI, Hyeokho CHOI et Richard BARANIUK. “ForWaRD : Fourier-wavelet regularized deconvolution for ill-conditioned systems”. In : *Signal Processing, IEEE Transactions on* 52.2 (2004), p. 418–433.
- [VU09] Cédric VONESCH et Michael UNSER. “A fast multilevel algorithm for wavelet-regularized image restoration”. In : *Image Processing, IEEE Transactions on* 18.3 (2009), p. 509–523.
- [VU08] Cédric VONESCH et Michael UNSER. “A fast thresholded Landweber algorithm for wavelet-regularized multidimensional deconvolution”. In : *Image Processing, IEEE Transactions on* 17.4 (2008), p. 539–549.
- [Cha+07] Caroline CHAUX, Patrick L COMBETTES, Jean-Christophe PESQUET et Valérie R WAJS. “A variational formulation for frame-based inverse problems”. In : *Inverse Problems* 23.4 (2007), p. 1495.

- [SNM03] Jean-Luc STARCK, Mai K NGUYEN et Fionn MURTAGH. “Wavelets and curvelets for image deconvolution : a combined approach”. In : *Signal Processing* 83.10 (2003), p. 2279–2283.
- [BT09b] Amir BECK et Marc TEOULLE. “Gradient-based algorithms with applications to signal recovery”. In : *Convex Optimization in Signal Processing and Communications* (2009).
- [CP11b] Patrick L COMBETTES et Jean-Christophe PESQUET. “Proximal splitting methods in signal processing”. In : *Fixed-point algorithms for inverse problems in science and engineering*. Springer, 2011, p. 185–212.
- [Den+15] Loïc DENIS, Eric THIÉBAUT, Ferréol SOULEZ, Jean-Marie BECKER et Rahul MOURYA. “Fast approximations of shift-variant blur”. In : *International Journal of Computer Vision* 115.3 (2015), p. 253–278.
- [FR05] Ralf C. FLICKER et François J. RIGAUT. “Anisoplanatic deconvolution of adaptive optics images”. In : *Journal of the Optical Society of America A* 22.3 (2005), p. 504–513.
- [GV06] Erez GILAD et Jost VON HARDENBERG. “A fast algorithm for convolution integrals with space and time variant kernels”. In : *Journal of Computational Physics* 216.1 (2006), p. 326–336.
- [Hir+10] Michael HIRSCH, Suvrit SRA, Bernhard SCHOLKOPF et Stefan HARMELING. “Efficient filter flow for space-variant multiframe blind deconvolution”. In : *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2010.
- [Hry+10] Tomasz HRYCAK, Saptarshi DAS, Gerald MATZ et Hans G. FEICHTINGER. “Low complexity equalization for doubly selective channels modeled by a basis expansion”. In : *Signal Processing, IEEE Transactions on* 58.11 (2010), p. 5706–5719.
- [NO98] James G. NAGY et Dianne P. O’LEARY. “Restoring images degraded by spatially variant blur”. In : *SIAM Journal on Scientific Computing* 19.4 (1998), p. 1063–1082.
- [TF92] Joel TRUSSELL et Sergei FOGEL. “Identification and restoration of spatially variant motion blurs in sequential images”. In : *Image Processing, IEEE Transactions on* 1.1 (1992), p. 123–126.
- [MP05] C. A. MICCHELLI et M. PONTIL. “On learning vector-valued functions”. In : *Neural Computation* 17.1 (2005), p. 177–204.
- [MP04] C. A. MICCHELLI et M. PONTIL. “Kernels for multi-task learning”. In : *Proceedings of NIPS 2004*. 2004.
- [Wen04] Holger WENDLAND. *Scattered Data Approximation*. T. 17. Cambridge Univ. Press, 2004.

- [NW91] Francis J NARCOWICH et Joseph D WARD. “Norms of inverses and condition numbers for matrices associated with scattered data”. In : *Journal of Approximation Theory* 64.1 (1991), p. 69–94.
- [ARR12] Ali AHMED, Benjamin RECHT et Justin ROMBERG. “Blind Deconvolution Using Convex Programming”. In : *IEEE Inf. Th.* 60.3 (2012), p. 1711–1732.
- [Li+18] Xiaodong LI, Shuyang LING, Thomas STROHMER et Ke WEI. “Rapid, robust, and reliable blind deconvolution via nonconvex optimization”. In : *ACHA* (2018).
- [LKP08] Michael LUSTIG, Seung-Jean KIM et John M PAULY. “A fast method for designing time-optimal gradient waveforms for arbitrary k-space trajectories”. In : *IEEE transactions on medical imaging* 27.6 (2008), p. 866–873.
- [Van+17] S Johanna VANNESJO, Yolanda DUERST, Laetitia VIONNET, Benjamin E DIETRICH, Matteo PAVAN, Simon GROSS, Christoph BARMET et Klaas P PRUESSMANN. “Gradient and shim pre-emphasis by inversion of a linear time-invariant system model”. In : *Magnetic resonance in medicine* 78.4 (2017), p. 1607–1622.
- [AA08] Gilles AUBERT et Jean-Francois AUJOL. “A variational approach to removing multiplicative noise”. In : *SIAM Journal on Applied Mathematics* 68.4 (2008), p. 925–946.
- [SO08] Jianing SHI et Stanley OSHER. “A nonlinear inverse scale space method for a convex multiplicative noise model”. In : *SIAM Journal on Imaging Sciences* 1.3 (2008), p. 294–321.
- [ST10] Gabriele STEIDL et Tanja TEUBER. “Removing multiplicative noise by Douglas-Rachford splitting methods”. In : *Journal of Mathematical Imaging and Vision* 36.2 (2010), p. 168–184.
- [Kun87] Gerard J KUNZ. “Bipath method as a way to measure the spatial backscatter and extinction coefficients with lidar”. In : *Applied optics* 26.5 (1987), p. 794–795.
- [Sch+13] Thorsten SCHMIDT, Jasmin DÜRR, Margret KEUPER, Thomas BLEIN, Klaus PALME et Olaf RONNEBERGER. “Variational attenuation correction in two-view confocal microscopy”. In : *BMC bioinformatics* 14.1 (2013), p. 366.
- [CF10] Juan CUESTA et Pierre H FLAMANT. “Lidar beams in opposite directions for quality assessment of Cloud-Aerosol Lidar with Orthogonal Polarization spaceborne measurements”. In : *Applied optics* 49.12 (2010), p. 2232–2243.
- [CCM99a] V. CASELLES, Bartomeu COLL et Jean-Michel MOREL. “Topographic Maps and Local Contrast Changes in Natural Images”. In : *International Journal of Computer Vision* 33 (1 1999).

- [Ser82] Jean SERRA. *Image analysis and mathematical morphology*. Academic press, 1982.
- [CCM99b] Vicent CASELLES, Bartomeu COLL et Jean-Michel MOREL. “Topographic maps and local contrast changes in natural images”. In : *International Journal of Computer Vision* 33.1 (1999), p. 5–27.
- [MG00] Pascal MONASSE et Frederic GUICHARD. “Fast computation of a contrast-invariant image representation”. In : *IEEE Transactions on Image Processing* 9.5 (2000), p. 860–872.
- [CM09] Vicent CASELLES et Pascal MONASSE. *Geometric description of images as topographic maps*. Springer, 2009.
- [CM63] Gustave CHOQUET et Paul-André MEYER. “Existence et unicité des représentations intégrales dans les convexes compacts quelconques”. In : *Ann. Inst. Fourier (Grenoble)* 13 (1963), p. 139–154.
- [CCM02a] Vicent CASELLES, Bartomeu COLL et Jean-Michel MOREL. “Geometry and color in natural images”. In : *Journal of Mathematical Imaging and Vision* 16.2 (2002), p. 89–105.
- [Bal+06] Coloma BALLESTER, Vicent CASELLES, Laura IGUAL, Joan VERDERA et Bernard ROUGÉ. “A variational model for P+ XS image fusion”. In : *International Journal of Computer Vision* 69.1 (2006), p. 43–58.
- [EA14] Matthias Joachim EHRHARDT et Simon R ARRIDGE. “Vector-valued image processing by parallel level sets”. In : *IEEE Transactions on Image Processing* 23.1 (2014), p. 9–18.
- [De +12b] Fabrice DE CHAUMONT et al. “Icy : an open bioimage informatics platform for extended reproducible research”. In : *Nature methods* 9.7 (2012), p. 690.
- [Cal02] Giuseppe CALAFIORE. “Approximation of n-dimensional data using spherical and ellipsoidal primitives”. In : *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans* 32.2 (2002), p. 269–278.
- [Eld+12] Ahmed Gamal ELDIN, Xavier DESCOMBES, Guillaume CHARPIAT et Josiane ZERUBIA. “Multiple birth and cut algorithm for multiple object detection”. In : *Journal of Multimedia Processing and Technologies* (2012).
- [Wei99] Joachim WEICKERT. “Coherence-enhancing diffusion filtering”. In : *International journal of computer vision* 31.2-3 (1999), p. 111–127.
- [VBB18] Luca VENTURI, Afonso BANDEIRA et Joan BRUNA. “Neural networks with finite intrinsic dimension have no spurious valleys”. In : *arXiv preprint arXiv :1802.06384* 15 (2018).
- [AHA98] Franz AURENHAMMER, Friedrich HOFFMANN et Boris ARONOV. “Minkowski-type theorems and least-squares clustering”. In : *Algorithmica* 20.1 (1998), p. 61–76.

- [Bal+00] C. BALLESTER, E. CUBERO-CASTAN, M. GONZALEZ et J. MOREL. “Contrast Invariant Image Intersection”. In : *Advanced Mathematical Methods in Measurement and Instrumentation* (2000), p. 41–55. URL : citeseer.ist.psu.edu/287102.html.
- [BS81] Robert BUSBY et Harvey SMITH. “Product-convolution operators and mixed-norm spaces”. In : *Transactions of the American Mathematical Society* 263.2 (1981), p. 309–341.
- [CCM02b] V. CASELLES, B. COLL et J.M. MOREL. “Geometry and Color in Natural Images”. In : *J. of Math. Imaging and Vision* 16 (2002), p. 89–107.
- [CS07] Anwei CHAI et Zuowei SHEN. “Deconvolution : A wavelet frame approach”. In : *Numerische Mathematik* 106.4 (2007), p. 529–587.
- [CD12] Jiawei CHIU et Laurent DEMANET. “Matrix probing and its conditioning”. In : *SIAM Journal on Numerical Analysis* 50.1 (2012), p. 171–193.
- [SdB16] Damien SCIEUR, Alexandre D’ASPREMONT et Francis BACH. “Regularized nonlinear acceleration”. In : *Advances In Neural Information Processing Systems*. 2016, p. 712–720.
- [Kai59] Thomas KAILATH. “Sampling models for linear time-variant filters”. In : (1959).
- [KMT16] Jun KITAGAWA, Quentin MÉRIGOT et Boris THIBERT. “A Newton algorithm for semi-discrete optimal transport”. In : *arXiv preprint arXiv :1603.05579* (2016).
- [KP05] Werner KOZEK et Götz E PFANDER. “Identification of operators with bandlimited symbols”. In : *SIAM journal on mathematical analysis* 37.3 (2005), p. 867–888.
- [Mér11] Quentin MÉRIGOT. “A multiscale approach to optimal transport”. In : *Computer Graphics Forum*. T. 30. 5. Wiley Online Library. 2011, p. 1583–1592.
- [PRT08] Götz E PFANDER, Holger RAUHUT et Jared TANNER. “Identification of matrices having a sparse representation”. In : *IEEE Transactions on Signal Processing* 56.11 (2008), p. 5376–5388.
- [Sma98] Steve SMALE. “Mathematical problems for the next century”. In : *The Mathematical Intelligencer* 20.2 (1998), p. 7–15.