



HAL
open science

Détection et localisation de particules dans des images PIV via des approches parcimonieuses à grille

Soufiane Ait Tilat

► **To cite this version:**

Soufiane Ait Tilat. Détection et localisation de particules dans des images PIV via des approches parcimonieuses à grille. Traitement du signal et de l'image [eess.SP]. Université Paris-Saclay, 2020. Français. NNT : 2020UPASG059 . tel-03157331v2

HAL Id: tel-03157331

<https://hal.science/tel-03157331v2>

Submitted on 5 Mar 2021

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.

Détection et localisation de particules dans des images PIV via des approches parcimonieuses à grille

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 580, sciences et technologies de l'information
et de la communication (STIC)
Spécialité de doctorat: Traitement du signal et des images
Unité de recherche: Université Paris-Saclay, ONERA, Traitement de
l'information et systèmes, 91123, Palaiseau, France
Réfèrent: Faculté des sciences d'Orsay

**Thèse présentée et soutenue en visioconférence totale, le
16/12/2020, par**

Soufiane AIT TILAT

Composition du jury:

Charles SOUSSEN PR, CentraleSupélec/Université Paris-Saclay	Président
Jean-François GIOVANNELLI PR, Université de Bordeaux/IMS	Rapporteur et examinateur
Laurent DAVID PR, Université de Poitiers/Institut PPRIME	Rapporteur et examinateur
Corinne FOURNIER MCF, Université de Saint-Etienne/Institut Hubert Curien	Examinatrice
Frédéric CHAMPAGNAT HDR, ONERA	Directeur de thèse
Cédric HERZET CR, INRIA	Co-encadrant

Table des matières

Table des matières	i
Liste des figures	iii
Liste des tableaux	vii
1 Introduction	1
1.1 Contexte	2
1.2 Problématique	2
1.3 Objectifs de la thèse et lien avec le projet ANR BECOSE	3
1.4 Démarche et contributions	3
1.5 Plan	4
1.6 Liste des publications	5
1.7 Références	6
2 Contextualisation	7
2.1 Mesure de vitesses en aérodynamique	8
2.2 Méthodes basiques de détection et de localisation de particules	12
2.3 Modèle de formation d'images et notion de dictionnaire	14
2.4 Méthodes de détection/localisation parcimonieuses à grille en PIV & PTV	16
2.5 Bilan	18
2.6 Références	19
3 Méthodes de reconstruction parcimonieuses convexes à grille	23
3.1 Limitations de l'approche NNLS	24
3.2 Régularisation l_1	25
3.3 Dictionnaire d'interpolation de Taylor et régularisation l_1	29
3.4 FISTA CBP	31
3.5 Active set CBP	34
3.6 Résultats numériques	35
3.7 Bilan	36
3.8 Références	37
4 Méthodes de reconstruction gloutonnes à grille	41
4.1 Dictionnaires d'interpolation et régularisation l_0	42
4.2 Continuous NNOMP (CNNOMP)	43
4.3 KCNNOMP : Amélioration du CNNOMP	47
4.4 Bilan	50
4.5 Références	51
5 Résultats numériques	53
5.1 Conditions de simulation et métriques	54
5.2 Méthodes convexes	57
5.3 Méthodes gloutonnes	59

5.4 Méthodes gloutonnes contre méthodes convexes	61
5.5 Bilan	63
5.6 Références	63
6 Conclusions et Perspectives	65
6.1 Résumé de contributions	66
6.2 Perspectives	68
6.3 Références	69
A Annexes	I
A.1 Définitions	I
A.2 Problème à résoudre	I
A.3 Résultats complets	I
A.4 Démonstration	III
A.5 Références	XI

Liste des figures

2.1	Tests aérodynamiques sur une voiture et sur un avion.	8
2.2	Setup expérimental de la PIV2D.	9
2.3	Principe de la méthode PIV2D. La figure est extraite de l'article de BROSSARD et collab. [2009]	10
2.4	Image de 5 particules dont 2 se chevauchent partiellement. L'objectif est de les détecter et de les localiser avec précision. Les croix en bleus (×) correspondent à une détection/localisation idéale.	11
2.5	Appariement temporel de détections à 2 instants successifs. Ici, on a apparié uniquement 2 détections pour avoir une figure claire.	12
2.6	La PSF au point x_p est remplacée par la PSF au nœud de la grille le plus proche. Il s'agit d'une discrétisation d'ordre 0.	15
2.7	Cette figure a été extraite de l'article de CORNIC et collab. [2013] et a été légèrement modifiée pour l'adapter à notre contexte 2D. Les points rouges correspondent aux pixels impactés par la diffraction d'une particule dont la projection est située au point vert. Ici on a représenté la PSF dans sa forme unidimensionnelle.	15
2.8	Dictionnaire H construit via une discrétisation d'ordre 0. Chaque colonne d'indice m représente la contribution d'une particule positionnée dans le m -ième nœud de la grille.	16
2.9	Illustration simple d'une approximation parcimonieuse. La couleur rouge indique les coefficients non nuls de la solution recherchée et les atomes(colonnes) associés dans le dictionnaire H. La figure est extraite du manuscrit de thèse de NGUYEN [2019]	16
3.1	Discrétisation très fine pour séparer 2 particules en chevauchement en absence du bruit. L'approximation utilisée par l'active set NNLS consiste à considérer le nœud de la grille le plus proche de la particule car la distance le séparant d'elle est négligeable.	24
3.2	L'active set NNLS avant seuillage (à gauche) et après seuillage (à droite) en présence du bruit. L'image initiale contient plusieurs fausses détections. Le processus de seuillage permet d'éliminer quasiment tous les faux positifs produits par le bruit, mais élimine certaines bonnes détections ayant de faibles intensités (Voir les particules entourées en jaune).	25
3.3	A gauche la solution active set BP avant seuillage et à droite la solution après seuillage, en présence du bruit. La première solution contient peu de fausses détections par rapport à l'active set NNLS non seuillé de la figure 3.2. L'application de seuillage améliore un peu plus les performances de détection de l'active set BP mais peut potentiellement effacer quelques bonnes détections (voir cercles jaunes).	29
3.4	Principe de la construction de dictionnaire par approximation de Taylor dans le cas 1D. La PSF est approchée en un point continu x_p par une somme pondérée de PSF et de sa dérivée au nœud le plus proche indexé m . Δ est la taille de la grille. ϵ est l'erreur de discrétisation.	30
3.5	Relation linéaire entre la position de la particule x_p et la position du nœud m le plus proche.	30

3.6	Représentation graphique d'une tranche à $x = x_0 > 0$ du cône \mathcal{P} et des sous ensembles qui définissent, par morceaux, la solution dans le cas x positif.	33
3.7	Représentation graphique d'une tranche à $x = x_0 < 0$ des sous ensembles qui définissent, par morceaux, la solution dans le cas x négatif.	33
3.8	\mathcal{P} est un cône pyramidal à base carrée, généré par la combinaison linéaire de quatre vecteurs \mathbf{x} , \mathbf{y} , \mathbf{z} et \mathbf{w} à coefficients positifs. Ces vecteurs sont les colonnes de la matrice $\tilde{\mathbf{P}}_\alpha$	35
3.9	Comparaison des méthodes NNLS, BP et CBP en termes de Fscore en fonction du pas de la grille Δ	36
4.1	La matrice \mathbf{G} est composée de 4 blocs. Chaque bloc est la combinaison linéaire de \mathbf{H} , \mathbf{H}_1 et \mathbf{H}_2 . Elle est de taille $N \times 4M$ avec M le nombre de nœuds de la grille et N le nombre de pixels de l'image acquise. α correspond à la moitié du pas de discrétisation.	42
4.2	Après avoir choisi l'atome \mathbf{g} de la matrice \mathbf{G} , indexé l_1 , réalisant le maximum de corrélation avec le résidu, on ajoute les 3 autres colonnes : l_2 , l_3 et l_4	46
4.3	Les 3 premières itérations de CNNOMP pour séparer deux particules de même intensité, qui se chevauchent et qui ont la même largeur de PSF. Les ronds bleus (o) correspondent aux particules réelles et les croix rouges (x) indiquent les détections estimées. Le pas de la grille Δ est égal à 1 pixel.	47
4.4	Les deux sous-figures 2 et 3 correspondent à la première itération avant et après l'étape de minimisation sur les positions et les intensités et les deux sous-figures 4 et 5 correspondent à la deuxième itération avant et après minimisation sur les positions et les intensités. KCNNOMP parvient à séparer les deux particules qui se chevauchent. En effet, l'étape de minimisation non convexe corrige les positions estimées et réduit drastiquement la valeur absolue maximale de l'image résiduelle de 0,15 à 10^{-7}	49
4.5	Comparaison de résultats de reconstruction entre CNNOMP (à gauche) et KCNNOMP (à droite) sur une image de taille 32×32 en l'absence du bruit. Les ronds bleus correspondent aux vraies particules et les croix rouges correspondent aux détections observées. $\Delta = 1$ pixel et $N_{ppp} = 0.1$. Le cercle jaune entoure une zone à forte interactions de particules. L'algorithme KCNNOMP arrive à séparer les particules et à estimer correctement les positions contrairement à CNNOMP qui fonctionne seulement quand les recouvrements sont très partiels.	50
5.1	Exemple d'images de synthèse utilisées dans nos simulations en présence d'un bruit modéré (SNR = 20). La densité des particules dans ces images varient de 0.01 à 0.1. $\sigma_{psf} = 0.6$ ce qui correspond à des motifs de taille d'environ 4×4 pixel.	55
5.2	Illustration de la notion de vrai positif (TP), faux positif (FP) et faux négatif (FN). Les ronds bleus correspondent aux véritables particules et les croix rouges correspondent aux détections.	55
5.3	Courbes de Precision, de Rappel et de Fscore en fonction du paramètre seuil s à la sortie de l'algorithme NNLS. A l'optimum du Fscore, autour de ≈ 0.3 , on a Precision \geq Rappel.	56
5.4	Évolution du Fscore des méthodes convexes en fonction de Δ . Pour un pas Δ variant de 1 à $\frac{1}{20}$ pixel, nous avons optimisé le paramètre de parcimonie λ et le seuil s pour obtenir un Fscore maximal. Lorsque Δ diminue, Fscore augmente quelle que soit la méthode testée. Le Fscore est presque maximal à $\Delta = \frac{1}{10}$ pour NNLS en absence du bruit et à $\Delta = \frac{1}{3}$ pour toutes les méthodes convexes en présence d'un bruit modéré (SNR = 20).	58

5.5	Comparaison des performances de détection entre les méthodes NNLS (rouge), BP (vert) et CBP (bleu) en fonction de la densité des particules. Les courbes continues correspondent à un SNR de 20 et les courbes en pointillées correspondent à un SNR infini. Les barres d'erreur ont été omises pour éviter de surcharger les figures. Leur demi-largeur est d'environ 0.04. Les performances de détection de NNLS se dégradent nettement en présence du bruit. La régularisation l_1 (BP) améliore la précision de NNLS en éliminant les détections parasites générées par le bruit. CBP conduit à la même performance de détection que BP.	59
5.6	Comparaison de l'erreur de localisation moyenne, avant et après raffinement de localisation, en fonction de la densité d'ensemencement. Lorsque le pas de discrétisation est large, les erreurs de localisation sont importantes. Cependant, le processus de raffinement améliore la précision de localisation et conduit à des erreurs comparables à celles obtenues avec $\Delta = \frac{1}{20}$ pixel.	59
5.7	Évolution de la moyenne et de l'écart-type du Fscore des méthodes gloutonnes en fonction de Δ . Chaque méthode est utilisée avec son paramètre de parcimonie optimal. La sensibilité globale du Fscore en fonction de Δ est négligeable pour KCNNOMP et également faible pour les deux autres méthodes NNOMP et CNNOMP. Lorsque Δ diminue, on observe des comportements antagonistes des 3 méthodes : la performance de détection de NNOMP se détériore, celle de CNNOMP s'améliore et celle de KCNNOMP demeure inchangée. Pour garantir un meilleur compromis entre une qualité élevée de détection et un faible temps de calcul, nous avons choisi $\Delta = \frac{1}{3}$ pixel pour NNOMP et $\Delta = 1$ pixel pour CNNOMP et KCNNOMP.	60
5.8	Comparaison des performances de détection des méthodes gloutonnes en fonction de densité de particules. Les courbes continues correspondent à un SNR de 20 et les courbes pointillées correspondent à un SNR infini. La performance de détection de NNOMP se dégrade considérablement à mesure que la densité d'ensemencement augmente. CNNOMP améliore NNOMP et est relativement robuste aux chevauchements. Sa version modifiée KCNNOMP est plus performante, en particulier en absence du bruit.	61
5.9	Comparaison entre BP (courbe continue) et KCNNOMP (courbe en pointillée) en fonction de la densité des particules. Chaque couleur est associée à un niveau de SNR (noir : 100, rose : 20 et bleu : 10). Pour les SNR bas, BP est plus performant en détection que KCNNOMP et vice versa pour les SNR élevés.	61
5.10	À gauche : comparaison du temps de calcul entre BP, BP avec une étape postérieure de raffinement de position et de KCNNOMP en utilisant les paramètres optimaux. NNLS avec $\Delta = \frac{1}{20}$, adopté par CHEMINET et collab. [2018], a également été tracé ici. À droite : le rapport entre le temps de minimisation non convexe et le temps d'exécution global.	62
A.1	Représentation graphique d'une tranche à $x = x_0 > 0$ du cône \mathcal{P} et des sous ensembles qui définissent, par morceaux, la solution dans le cas x positif	II
A.2	Représentation graphique d'une tranche à $x = x_0 < 0$ du cône polaire \mathcal{P}^* des sous ensembles qui définissent, par morceaux, la solution dans le cas x négatif	III

Liste des tableaux

2.1	Quelques méthodes de mesure du flux d'air.	8
3.1	Solutions du problème de projection sur le cône \mathcal{P} . La première colonne correspond à l'ensemble auquel appartient (u, v, w)	32
5.1	Les paramètres des méthodes convexes et gloutonnes.	57
A.1	Expression de $proj_{\mathcal{P}}(a, b, c)$ en fonctions des dix cas.	I

Chapitre 1

Introduction

Sommaire

1.1 Contexte	2
1.2 Problématique	2
1.3 Objectifs de la thèse et lien avec le projet ANR BECOSE	3
1.4 Démarche et contributions	3
1.5 Plan	4
1.6 Liste des publications	5
1.7 Références	6

1.1 Contexte

La Vélocimétrie par Image de Particules (PIV) est une technique de mesure puissante pour appréhender la dynamique des écoulements. Elle a plusieurs applications industrielles notamment en automobile et en aéronautique afin d'équilibrer le profil aérodynamique avec les designs conçus.

Le principe global de la PIV est comme suit : une caméra acquiert des images de particules injectées dans un écoulement pour visualiser son mouvement. Un algorithme de corrélation est ensuite utilisé pour estimer la carte vectorielle de vitesses. Cependant, la corrélation agit comme un filtre passe-bas qui lisse les gradients de vitesse [SCARANO \[2003\]](#), ce qui limite la résolution spatiale des mesures de vitesses ainsi que toutes les dérivées qui en découlent.

La vélocimétrie par suivi de particules (PTV) contourne ce problème en suivant individuellement chaque particule, ce qui permet d'extraire le vecteur vitesse en tout point de l'espace contenant une particule. [FENG et collab. \[2011\]](#) montrent que la principale source d'erreur de vitesse en PTV, lors de l'utilisation de caméras à haute cadence, est due à l'incertitude sur la position des particules. Ainsi la détection/localisation est une étape déterminante dans la précision des mesures PTV. Cette thèse se focalise sur les aspects bidimensionnels de la détection/localisation, qui intéresse la PTV plane [OHMI et LI \[2000\]](#) et qui constitue aussi une étape intermédiaire dans certaines techniques de PTV tridimensionnelle [FUCHS et collab. \[2016\]](#).

1.2 Problématique

Dans les méthodes de PIV ou de PTV, la résolution spatiale est liée à la densité de marqueurs. L'augmentation de la résolution spatiale est limitée par le chevauchement des images élémentaires de chaque particule. Ce chevauchement rend difficile la séparation de celles-ci via des approches basiques tels que le filtrage adapté [ANGARITA-JAIMES et collab. \[2009\]](#) ou la méthode des maxima locaux [OHMI et LI \[2000\]](#). De plus, il accroît l'erreur de localisation des particules. Pour tenir compte du phénomène de superposition, un modèle d'image additif, basé sur la fonction d'étalement du point PSF, a été utilisé par de nombreux auteurs [CHAMPAGNAT et collab. \[2014\]](#), [CORNIC et collab. \[2015\]](#), [CHEMINET et collab. \[2018\]](#). Ce modèle décrit les enregistrements PIV $\mathbf{i} \in \mathbb{R}^N$ en un pixel $\mathbf{n} \in \mathbb{R}^2$ comme une somme pondérée de P motifs $h: \mathbb{R}^2 \rightarrow \mathbb{R}^+$ identiques décalés dans l'espace et représentant chacun l'image d'une particule par la caméra :

$$\mathbf{i}(\mathbf{n}) = \sum_{p=1}^P h(\mathbf{n} - \mathbf{x}_p) e_p. \quad (1.1)$$

Muni de ce modèle, la détection et la localisation des particules consiste à estimer les intensités des particules e_p , leurs positions \mathbf{x}_p sur l'image, ainsi que le nombre total de particules P dont l'image prédite est la plus proche de l'image observée au sens d'un critère de similarité. Notons que ce critère doit nécessairement s'accompagner d'une pénalisation de l'ordre du modèle (le nombre de particules) car on pourrait ajouter des particules d'intensité arbitrairement faible sans changer notablement l'apparence de l'image associée, ceci quelque soit le niveau de bruit présent sur les images. Cette remarque introduit un ingrédient essentiel dans la résolution du problème : la nécessaire parcimonie du modèle expliquant l'image observée.

Ce problème est difficile en raison de la dépendance non linéaire entre la position \mathbf{x}_p de la particule et son image $\mathbf{i}(\mathbf{n})$ en un pixel \mathbf{n} , de plus l'ordre du modèle est inconnu. Il existe deux catégories de méthodes pour le résoudre selon que l'on considère ou non une approximation du modèle d'imagerie.

La première préserve l'aspect continu de la position des particules, au prix de difficultés d'implémentation. Par exemple la méthode Sliding Frank-Wolfe de [DENOYELLE et collab. \[2019\]](#) requiert la résolution d'un sous problème non convexe afin de déplacer continûment les particules reconstruites à chaque itération.

La seconde s'appuie sur une grille discrète : la position des particules est discrétisée sur une grille fixe de nœuds, on aboutit ainsi à une relation linéaire entre l'image et une distribution d'intensité de particules indexée sur la même grille, via un dictionnaire. Les atomes sont les images de particules localisées sur les nœuds et d'intensité unité. Ces méthodes à grille aboutissent typiquement à un problème quadratique sous contrainte de parcimonie [CORNIC et collab. \[2015\]](#), [BARBU et HERZET \[2016\]](#), celle-ci pouvant même se restreindre à la positivité seulement [CHEMINET et collab. \[2018\]](#). Elles présentent plusieurs limitations notamment la présence de détections multiples liées à l'erreur de discrétisation [DUVAL et PEYRÉ \[2017\]](#) et l'incapacité de séparer les chevauchements quand le pas de la grille est de l'ordre de la largeur de la PSF. Des travaux récents de [CHEMINET et collab. \[2018\]](#) indiquent que ces limitations peuvent être surmontées avec des pas extrêmement fins de l'ordre du vingtième de pixel. Cependant, dans ce cas, le problème requiert la manipulation d'un dictionnaire de très grande taille et par conséquent un coût de calcul et un encombrement mémoire très importants. Plusieurs auteurs [BARBU et HERZET \[2016\]](#) et [CORNIC et collab. \[2015\]](#) ont montré la pertinence d'intégrer la parcimonie comme a priori, dans un contexte 3D, afin de promouvoir la qualité des solutions. Cependant, cela demeure problématique. En effet, il est difficile de savoir s'il faut intégrer la parcimonie et la positivité au même temps ou se limiter simplement à la première. Il n'est également pas intuitif de prévoir l'impact de type de parcimonie sur la qualité de détection ou sur la taille du pas de grille qui contrôle globalement la complexité calculatoire des méthodes de reconstruction.

1.3 Objectifs de la thèse et lien avec le projet ANR BECOSE

Les objectifs de ma thèse sont :

1. Concevoir des algorithmes robustes au bruit et ayant une grande capacité de séparation de chevauchements et de précision de localisation.
2. Remédier à l'encombrement mémoire et au coût calculatoire des algorithmes de reconstruction parcimonieuses en PIV et PTV, dûs à la finesse de la grille adoptée, en favorisant une discrétisation grossière.
3. Évaluer l'impact du type de discrétisation et/ou du type de contrainte de parcimonie sur le degré de finesse de la grille.
4. Évaluer l'impact du type de parcimonie sur la qualité de détection et de localisation

Ce travail est soutenu par le projet ANR Beyond Compressive Sensing, appelé aussi BECOSE (N° ANR – 15 – CE23 – 0021). Ce dernier a trois buts principaux : le premier consiste à développer des méthodes parcimonieuses dédiées aux problèmes inverses mal posés en exploitant des informations supplémentaires telles que la non-négativité, ou en construisant des dictionnaires continus. Le second but consiste à enrichir l'analyse théorique de ces méthodes. Enfin, le dernier but est d'évaluer ces méthodes dans le contexte de la vélocimétrie par images de particules (PIV). Cette thèse répond en partie à la première et à la troisième mission de ce programme.

1.4 Démarche et contributions

Pour ce faire, nous avons revisité la méthode de [CHEMINET et collab. \[2018\]](#) et en particulier la stratégie de réglage du pas de la grille et proposé une nouvelle stratégie basée sur l'optimum d'un compromis de détection. Nous avons également introduit une contrainte de pénalisation de type l_1 afin de forcer la parcimonie des solutions et améliorer sa robustesse vis à vis du bruit, tout en gardant l'aspect convexe du problème. Afin d'obtenir des résultats de localisation précis similaires à ceux obtenus avec une grille fine, nous avons ajouté une étape de post-processing permettant de raffiner la localisation sous pixel initialement estimée par notre algorithme, en tolérant une augmentation légère de temps de calcul. Nous nous sommes aussi inspirés des travaux d'[EKANADHAM et collab. \[2011\]](#) faits dans un contexte 1D sur le Continuous Basis Pursuit (CBP) pour les extrapoler au cas 2D. Le principe est une linéarisation par morceaux de la relation entre l'image acquise

\mathbf{i} et la distribution des positions des centres de motifs de particules, via une approximation de Taylor d'ordre 1. Cette discrétisation fait intervenir deux dictionnaires dérivés supplémentaires associés à deux autres variables contraintes linéairement. Cela introduit des difficultés algorithmiques par rapport au cas 1D. Nous l'avons abordé au départ avec une contrainte de parcimonie l_1 en s'appuyant sur une méthode de gradient proximal. Cette dernière requiert le calcul d'un opérateur proximal [PARIKH et BOYD \[2014\]](#) dont la résolution analytique est indispensable pour traiter des images avec une taille et une densité de particules importantes. Dans notre cas, le calcul de l'opérateur proximal est équivalent à une projection sur un cône pyramidal à base carrée. On l'a résolu géométriquement et on a vérifié les formules obtenues en s'appuyant sur les conditions d'optimalité. Cependant, le coût calculatoire de cette méthode proximale s'est révélée trop coûteuses en pratique pour notre problème. Ainsi on a opté pour une méthode de programmation quadratique à savoir la méthode de l'active set [NOCEDAL et WRIGHT \[1999\]](#). L'implémentation a été grandement simplifiée en reparamétrisant le problème inverse de telle sorte à avoir uniquement des contraintes de positivité au lieu de contraintes linéaires.

Ma contribution suivante a été de combiner les discrétisations et reparamétrisations mentionnés précédemment avec un algorithme parcimonieux glouton de type Orthogonal matching pursuit (OMP) sous contrainte de positivité. On a ensuite implémenté l'algorithme en s'inspirant des travaux de [NGUYEN et collab. \[2017\]](#) sur la méthode OMP non-négatif (NNOMP). On a observé que cette nouvelle technique l_0 échoue quand l'occurrence de chevauchement augmente et ce quel que soit le pas de discrétisation utilisé, contrairement aux méthodes convexes qui s'améliore avec la finesse de la grille. On a analysé expérimentalement ce défaut et on a proposé d'introduire une étape supplémentaire non convexe à l'algorithme courant. Cette étape permet de sortir de la grille et sélectionner des positions continues, ce qui améliore la précision de l'algorithme dans la gestion des chevauchements.

1.5 Plan

Ce manuscrit sera composé, après ce chapitre introductif, de 5 autres chapitres : un chapitre de mise en contexte, 3 chapitres contenant mes contributions et un chapitre final récapitulant les principales conclusions et les perspectives de mes travaux. Nous décrivons plus en détails le contenu de ces chapitres dans les paragraphes suivants.

Le chapitre 2 met en contexte mes travaux de thèse. La première section présente le principe de la PIV et sa principale limitation, due à l'étape de corrélation, à savoir le lissage des gradients, et propose la PTV comme alternative pour améliorer la résolution spatiale tout en mettant l'accent sur l'importance d'une détection/localisation précise. La seconde section montre l'incapacité des méthodes de détection à base de seuillage et de corrélation à gérer les chevauchements dans des images à haute densité de particules. La troisième section présente un modèle de formation d'images favorisant une détection parcimonieuse. La dernière section présente le principe d'une approximation parcimonieuse et décrit quelques méthodes parcimonieuses à grille récemment introduites dans le domaine de la PIV tomographique et de la PTV tout en soulignant leurs limitations en pratique dûes à la discrétisation fine adoptée.

Le chapitre 3 expose nos contributions sur les méthodes convexes parcimonieuses non négatives. La première section montre les limitations de l'approche des moindres carrés non négatifs (NNLS). La seconde section présente la version régularisée de type l_1 de ce problème, suivi d'un récapitulatif succinct des méthodes usuelles destinées à le résoudre, et également d'une description plus détaillée de notre technique active set qu'on a développée. La troisième section décrit notre nouvelle méthode de discrétisation afin de linéariser la relation entre l'image acquise par la caméra et la distribution des positions de particules. Cette discrétisation donne lieu à 2 dictionnaires dérivés, supplémentaires, associés à deux autres variables contraintes linéairement. La quatrième section décrit la méthode proximale utilisée pour la résolution du problème inverse in-

duit. La section 5 traite une re-paramétrisation de ce problème en utilisant que des contraintes de positivité afin d'implémenter une version d'active set plus rapide que notre méthode proximale. Enfin, dans la dernière section, une comparaison entre les 3 méthodes convexes précédemment décrites est effectuée, afin de choisir la méthode réalisant le meilleur compromis entre qualité de détection et temps de calcul.

Le chapitre 4 décrit 2 méthodes parcimonieuses non-négatives gloutonnes qu'on a développées en utilisant la nouvelle discrétisation associée à une pénalisation de type l_0 . La première section présente le problème d'optimisation induit. La seconde section donne un aperçu des principales méthodes présentées dans la littérature pour adresser ce problème. Elle décrit également l'algorithme de résolution qu'on a développé, ainsi que quelques résultats qualitatifs montrant ses limitations en présence d'interactions entre particules dans l'image. La dernière section est une modification de cet algorithme glouton. Elle consiste à ajouter une étape supplémentaire de minimisation non-convexe afin de se détacher de cette grille de discrétisation "figée" et corriger les positions estimées à chaque itération.

Le chapitre 5 est un chapitre de simulations et de commentaires de résultats. La première section décrit tous les détails concernant la génération de nos données synthétiques et les métriques de performances de détection et de localisation usuelles. La deuxième section présente la stratégie de réglage utilisée pour comparer fidèlement les méthodes l_1 ainsi que les résultats obtenus. La section 4 a la même structure que la précédente et vise la comparaison des méthodes l_0 avec le même critère de réglage. Enfin, les meilleures méthodes de chaque catégorie sont comparées en fonction de la densité de particules, du rapport signal sur bruit et du temps de calcul.

Le chapitre 6 est un récapitulatif des approches proposées et des principaux résultats obtenus. Il comprend également les futures perspectives de nos travaux.

1.6 Liste des publications

Les travaux qui sont discutés dans ce manuscrit ont donné lieu à 2 articles de conférence dont les résultats ont été présentés oralement aux congrès GRETSI et ISPIV2019. Une troisième communication, sous forme de résumé, a été proposée au congrès LXLASER2020, mais l'évènement a été annulé en raison de la crise de la covid19.

1. Soufiane Ait Tilat, Cédric Herzet, Frédéric Champagnat. *Détection et localisation en imagerie de particules par approximation de Taylor de dictionnaires continus*. GRETSI 2019, Aug 2019, Lille, France.
2. Soufiane Ait Tilat, Frédéric Champagnat, Cédric Herzet. *A new sparsity based particle image reconstruction approach for particle detection*. ISPIV 2019 - 13th International Symposium on Particle Image Velocimetry, Jul 2019, Munich, Germany.
3. Soufiane Ait Tilat, Frédéric Champagnat, Cédric Herzet. *What is a good grid-size for grid-based sparsity particle image reconstruction approaches in PIV?* LXLASER 2020, Lisbon, Portugal.

1.7 Références

- ANGARITA-JAIMES, N., M. G. ROCA, C. TOWERS, N. READ et D. TOWERS. 2009, «Algorithms for the automated analysis of cellular dynamics within living fungal colonies», *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, vol. 75, p. 768–80. [2](#)
- BARBU, I. et C. HERZET. 2016, «A new approach for volume reconstruction in tomoPIV with the alternating direction method of multipliers», *Measurement Science and Technology*, vol. 27, p. 104 002 (11pp). [3](#)
- CHAMPAGNAT, F., P. CORNIC, A. CHEMINET, B. LECLAIRE, G. LE BESNERAIS et A. PLYER. 2014, «Tomographic PIV : Particles versus blobs», *Measurement Science and Technology*, vol. 25, n° 8, p. 084 002. [2](#)
- CHEMINET, A., J. KRAWCZYNSKI et P. DRUAULT. 2018, «Particle image reconstruction for particle detection in particle tracking velocimetry», *Measurement Science and Technology*, vol. 29, n° 12, p. 125 202. [2](#), [3](#)
- CORNIC, P., F. CHAMPAGNAT, A. CHEMINET, B. LECLAIRE et G. LE BESNERAIS. 2015, «Fast and efficient particle reconstruction on a 3D grid using sparsity», *Experiments in Fluids*, vol. 56, n° 3, 62, ISSN 0723-4864. [2](#), [3](#)
- DENOYELLE, Q., V. DUVAL, G. PEYRÉ et E. SOUBIES. 2019, «The sliding frank–wolfe algorithm and its application to super-resolution microscopy», *Inverse Problems*, vol. 36, n° 1, p. 014 001. [2](#)
- DUVAL, V. et G. PEYRÉ. 2017, «Sparse spikes super-resolution on thin grids II : the continuous basis pursuit», *Inverse Problems*, vol. 33, n° 9, p. 095 008. [3](#)
- EKANADHAM, C., D. TRANCHINA et P. SIMONCELLI. 2011, «Recovery of sparse translation-invariant signals with continuous basis pursuit», *IEEE Transactions on Signal Processing*, vol. 59, n° 10, p. 4735–4744. [3](#)
- FENG, Y., J. GOREE et B. LIU. 2011, «Errors in particle tracking velocimetry with high-speed cameras», *The Review of scientific instruments*, vol. 82, p. 053 707. [2](#)
- FUCHS, T., R. HAIN et C. J. KÄHLER. 2016, «Double-frame 3D-PTV using a tomographic predictor», *Experiments in Fluids*, p. 1–5. [2](#)
- NGUYEN, T. T., C. SOUSSEN, J. IDIER et E.-H. DJERMOUNE. 2017, «An optimized version of non-negative OMP», dans *XXVIIe Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017*, Juan-les-Pins, France. [4](#)
- NOCEDAL, J. et S. WRIGHT. 1999, *Numerical Optimization*, Springer-Verlag, New-York. [4](#)
- OHMI, K. et H.-Y. LI. 2000, «Particle-tracking velocimetry with new algorithms», *Measurement Science and Technology*, vol. 11, n° 6, p. 603–616. [2](#)
- PARIKH, N. et S. BOYD. 2014, «Proximal algorithms», *Found. Trends Optim.*, vol. 1, n° 3, p. 127–239. [4](#)
- SCARANO, F. 2003, «Theory of non-isotropic spatial resolution in piv», *Experiments in Fluids*, vol. 35, p. 268–277. [2](#)

Chapitre 2

Contextualisation

Sommaire

2.1 Mesure de vitesses en aérodynamique	8
2.1.1 Dispositif expérimental de PIV et contraintes associées	9
2.1.2 PIV	9
2.1.3 PTV	11
2.1.4 Erreurs d'estimation de vitesse en PTV	12
2.2 Méthodes basiques de détection et de localisation de particules	12
2.2.1 Détection via seuillage	13
2.2.2 Détection via corrélation	13
2.3 Modèle de formation d'images et notion de dictionnaire	14
2.3.1 Modèle de formation d'images	14
2.3.2 Approximation d'ordre 0 et construction de dictionnaire	14
2.4 Méthodes de détection/localisation parcimonieuses à grille en PIV & PTV	16
2.4.1 Approximation parcimonieuse	16
2.4.2 Méthodes de détection/localisation parcimonieuses à grille	17
2.5 Bilan	18
2.6 Références	19

2.1 Mesure de vitesses en aérodynamique

Les essais aérodynamiques sont essentiels pour concevoir des engins de transport efficaces : ils sont souvent décisifs pour le succès dans les courses automobiles. En effet un bon design aérodynamique améliore la tenue en route des voitures en agissant sur leurs portances quelque soit les circonstances (virages, sol à basse adhérence...) et facilite leur maîtrise en situation d'urgence. Il est également très utile pour faire des économies de carburant des avions en atténuant leurs traînées. De plus, l'optimisation des flux d'air améliore le confort à bord en amoindrissant les nuisances sonores, ainsi que les mouvements complexes ressentis à vive allure. La figure 2.1 présente un exemple de tests aérodynamiques pratiqués en industrie automobile et aéronautique.



FIGURE 2.1 – Tests aérodynamiques sur une voiture et sur un avion.

Les principales méthodes modernes de mesure du flux d'air peuvent être classées en deux catégories : la première catégorie est basée sur une mesure ponctuelle et a été largement déployée dans le passé notamment l'Anémométrie à fil chaud, l'Anémométrie à ultrasons (UA) et l'Anémométrie laser Doppler (LDA) [SUN et ZHANG \[2007\]](#) (voir le tableau 2.1). Cependant, ces méthodes de mesure recueillent seulement des informations ponctuelles, et les outils de mesures sont souvent intrusifs et peuvent parfois perturber l'écoulement d'air local. En revanche, la seconde catégorie repose sur une mesure globale et dense des caractéristiques de l'écoulement, en fournissant des cartes de champs construites à partir de plusieurs points, notamment la vélocimétrie par images de particules (PIV) [GAO et collab. \[2013\]](#).

Classe	Méthode	Type de mesure
Ponctuel	Anémométrie à fil chaud	Thermo-électrique
	Anémométrie laser Doppler (LDA)	Optique
	Vélocimétrie acoustique Doppler	Acoustique
Global	Vélocimétrie par images de particules (PIV)	Optique + traceurs

TABEAU 2.1 – Quelques méthodes de mesure du flux d'air.

Au cours de ces dernières décennies, grâce aux progrès des caméras numériques, des lasers à très haut taux de répétition et de la puissance de calcul des GPU, ces techniques ont été largement déployées dans l'industrie automobile et aéronautique afin de valider le profil aérodynamique des engins conçus via les outils de simulations de la mécanique de fluide numérique (CFD, computational fluid dynamics). En effet, une simulation CFD requiert la connaissance de paramètres souvent difficiles à estimer, afin de résoudre efficacement les équations modélisant le mouvement d'un écoulement. En outre, le traitement numérique exige une grande capacité de calcul notamment lors du traitement des équations de Navier-Stokes à haut nombre de Reynold (Re). La PIV permet ainsi de contourner ces limitations. Dans cette section, on analysera le protocole expérimental de la PIV et les contraintes associées. On décrira aussi le principe de la PIV bi-dimensionnelle (PIV2D) et les techniques d'extension au cas tri-dimensionnel (3D), et on montrera la principale limitation de cette approche à savoir la diminution de la résolution spatiale.

On présentera ensuite la vélocimétrie par suivi de particules (PTV) comme une alternative de mesure pour surmonter ce défaut et on décrira ses principales étapes. Enfin, on mettra l'accent sur la nécessité d'un ensemencement dense et une détection/localisation précise afin de garder une bonne résolution du champ de vitesse estimé.

2.1.1 Dispositif expérimental de PIV et contraintes associées

La figure 2.2, extraite du livre de RAFFEL et collab. [2007], montre un exemple de configuration dans une soufflerie. L'écoulement étudié est ensemencé de traceurs passifs. Une fine portion de ce dernier est éclairée par un laser de telle sorte que la lumière diffusée par les particules à chaque impulsion est captée par une caméra. Ces clichés consécutifs sont ensuite traités pour estimer les champs de vitesse dans le plan illuminé.

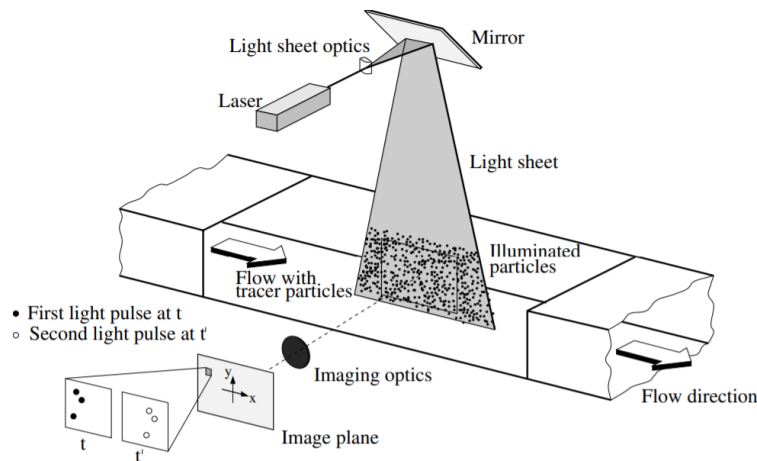


FIGURE 2.2 – Setup expérimental de la PIV2D.

Pour que la configuration ci-dessus donne des résultats précis, l'ensemencement utilisé pour l'écoulement doit respecter certaines conditions : les traceurs doivent être suffisamment légers pour suivre fidèlement l'écoulement et ne pas le perturber. Ils doivent aussi être suffisamment réfléchissants pour être visibles sur les images RAFFEL et collab. [2007]. En outre, la densité de particules traceuses par pixel N_{ppp} doit être élevée afin de résoudre les structures à petite échelle. Parmi les traceurs couramment utilisés pour suivre les flux d'air dans une expérience PIV, on peut mentionner les gouttelettes d'huile, les particules DEHS (Di-Methyl-Hexyl-Sebacate) et les bulles de savon emplies d'hélium (HFSB) RAFFEL et collab. [2007]. Le système d'éclairage est composé d'une source de lumière et d'un système optique capable de positionner avec précision une couche verticale très fine de lumière sur la zone de l'écoulement étudiée afin d'imager les particules se déplaçant dans le même plan. La source lumineuse est un laser qui produit périodiquement deux impulsions lumineuses très courtes. Les images de particules sont souvent acquises par une caméra spécialisée, synchronisée avec le laser, et capable de capturer 2 clichés successifs dans un laps de temps très court.

2.1.2 PIV

La première partie de cette sous-section donne une explication intuitive de la procédure PIV2D, ainsi qu'un aperçu succinct sur quelques techniques d'extension au cas 3D. Enfin, la dernière partie montre les principales limitations cette approche.

Principe de la PIV2D et extension en 3D

La procédure d'estimation du champ de vitesse bidimensionnel est comme suit : chaque image consécutive est divisée en sous-domaines rectangulaires appelées fenêtres d'interrogation. La corrélation entre deux fenêtres prises à 2 instants successifs génère une surface présentant un pic. La

localisation de son maximum correspond au déplacement local de particules le plus probable à l'intérieur de la fenêtre d'interrogation. En répétant cette opération sur la totalité de la 2ème image et en décalant la surface d'interrogation de la 1ère, nous pouvons extraire une estimée du champ global de déplacement des particules et déduire la carte vectorielle globale des vitesses [BROSSARD et collab. \[2009\]](#), [SCHON et collab. \[2007\]](#) (voir la figure 2.3).

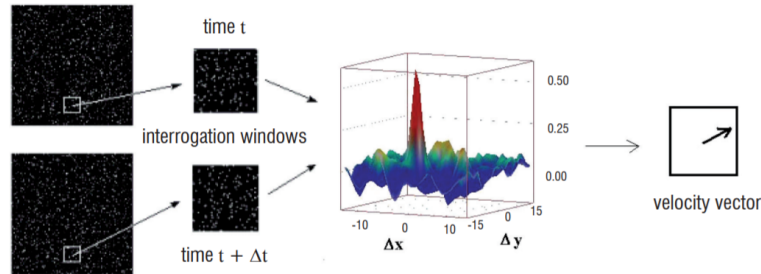


FIGURE 2.3 – Principe de la méthode PIV2D. La figure est extraite de l'article de [BROSSARD et collab. \[2009\]](#).

L'inconvénient de la PIV2D classique est qu'elle n'enregistre que la projection du vecteur vitesse dans le plan illuminé, donc la composante de vitesse perpendiculaire à ce plan n'est pas mesurée. Pour remédier à ce problème et capturer les 3 composantes à la fois, la PIV stéréoscopique [WILLERT \[1997\]](#), [ARROYO et GREATED \[1991\]](#), [SCHON et collab. \[2007\]](#) a été introduite. Cette dernière utilise deux caméras qui imagent simultanément les particules sous deux angles différents. Les projections sur chaque plan image font apparaître des déplacements légèrement différents, d'où la possibilité de reconstruire la troisième composante de vitesse manquante, mais uniquement sur le plan défini par la nappe laser, on parle alors de PIV2D – 3C (PIV2D à trois composantes).

Une autre manière pour capter le mouvement de l'écoulement en 3D est d'éclairer successivement différents plans du volume étudié grâce à un miroir rotatif et puis effectuer les étapes de la PIV bidimensionnelle sur ces plans [PRENEL et BAILLY \[2006\]](#). Ainsi, une analyse quasi volumétrique peut être réalisée sur le volume cible. Cette technique requiert des caméras à haute cadence ainsi que des systèmes mécaniques de balayage très rapides. En outre, le volume d'intérêt doit rester relativement petit si l'on désire capturer l'ensemble des plans dans un délai très court.

La PIV tomographique a été proposée il y a une quinzaine d'années par [ELSINGA et collab. \[2006\]](#) pour surmonter ces limitations. Elle est maintenant largement adoptée, elle utilise plusieurs caméras, typiquement 4, chacune orientée dans une direction différente. En pratique, la reconstruction 3D du champ de vitesse est obtenue en deux étapes. Une première étape de reconstruction tomographique produit à chaque instant de mesure une distribution volumique de particules dans laquelle chaque particule est représentée par un amas de quelques voxels de diamètre. Les distributions volumiques instantanées sont ensuite corrélées, fournissant ainsi un champ de déplacement 3D. Une analyse de ces phases de reconstruction et leurs principales limitations est détaillée dans le papier de [SCARANO \[2012\]](#).

Limitation de la PIV

L'algorithme de corrélation vise à suivre un groupe de particules dans une fenêtre d'interrogation entre 2 instants adjacents. Il en résulte une estimation du champ de déplacement moyennée sur l'étendue de la fenêtre de corrélation. Cette opération agit comme un filtre passe bas ce qui lisse les gradients de vitesse et conduit à une résolution spatiale réduite [SCARANO \[2003\]](#), [NOGUEIRA et collab. \[2005\]](#). Ainsi, toutes les valeurs des grandeurs physiques déduites à partir des vitesses estimées, tels que la vorticit , le cisaillement ou l' nergie cin tique turbulente sont peu pr cises. [TANAKA et EATON \[2007\]](#), [SAARENINNE et PIIRTO \[2012\]](#), [KRAWCZYNSKI et collab. \[2010\]](#).

2.1.3 PTV

Une alternative possible pour surmonter le problème de filtrage spatial, lié à la corrélation est l'utilisation de la vélocimétrie par suivi de particules (PTV). Comme son nom l'indique, les particules sont suivies individuellement, ce qui permet d'extraire pour chaque particule le vecteur de déplacement ainsi que sa trajectoire sur de longues périodes. Le protocole expérimental est quasiment le même que celui de la PIV. La seule différence réside dans la quantité de traceurs introduite dans l'écoulement puisque la densité utilisée en PTV est très inférieure à celle utilisée en PIV [FUCHS et collab. \[2017\]](#). L'analyse PTV consiste en 2 étapes essentielles que sont la détection/localisation des particules dans les 2 clichés successifs et leur association temporelle. Une explication de ces deux points est présentée ci-dessous.

Détection/Localisation des centroïdes des images de particules

La partie gauche de la figure 2.4 représente une image de 5 particules, dont deux se chevauchent partiellement. Le but est de les détecter et les localiser précisément.

Les techniques de détection usuelles segmentent les images, pour extraire des motifs traduisant a priori l'existence d'une ou plusieurs particules, via des algorithmes de seuillage [OHMI et LI \[2000\]](#) ou de corrélation [ANGARITA-JAIMES et collab. \[2009\]](#). Les positions sous-pixel des centroïdes sont ensuite déduites en utilisant la méthode de la moyenne pondérée ou l'interpolation gaussienne du pic d'intensité [OUELLETTE et collab. \[2006\]](#). D'autres méthodes plus sophistiquées, robustes aux effets de bruit de mesure et aux chevauchements s'appuient sur un modèle physique et fournissent des détections plus précises [CHEMINET et collab. \[2018\]](#). Une description détaillée des méthodes phares de détection sera présentée dans la section 2.2.

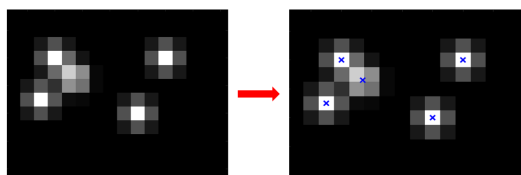


FIGURE 2.4 – Image de 5 particules dont 2 se chevauchent partiellement. L'objectif est de les détecter et de les localiser avec précision. Les croix en bleus (x) correspondent à une détection/localisation idéale.

Appariement temporel des détections

La figure 2.5 décrit la deuxième étape de la PTV à savoir l'appariement temporel. En effet, après la détection/localisation de particules à 2 instants adjacents, un appariement est ensuite effectué afin de repérer les détections correspondantes aux mêmes particules, ce qui permettra de déduire les déplacements locaux. La difficulté intrinsèque des techniques d'appariement dépend essentiellement de la densité d'ensemencement et du déplacement moyen des particules entre 2 images consécutives : une mesure quantitative de cette difficulté est obtenue en calculant le rapport p entre l'espacement moyen entre particules et la distance moyenne parcourue durant ce laps de temps. Si $p \gg 1$, le suivi est relativement facile, si $p \ll 1$ le suivi devient pratiquement impossible [MALIK et collab. \[1993\]](#). Dans le cas des écoulements peu ensemencés, l'application de l'algorithme du "plus proche voisin" est suffisante. En cas de fortes densités d'autres approches sont utilisées notamment l'approche prédicteur-correcteur appelée "PIV super resolution" [KEANE et collab. \[1999\]](#).

Le principe de cette dernière repose sur l'utilisation d'algorithmes de corrélation ne nécessitant pas l'identification des particules en première étape. On détecte et on localise les particules dans la première image, pour chaque détection on interpole le déplacement obtenu par corrélation au point où est localisée cette particule, ceci fournit une prédiction de l'endroit où devrait se trouver la particule dans la seconde image. On associe alors la particule de la seconde image la plus

proche de cette position à la particule de la première image. D'autres approches cherchent une similitude spatiale de particules dans une certaine zone de l'image entre deux instants consécutifs en définissant une fonction de similitude tenant compte de la distance moyenne entre particules et les variations d'intensité lumineuse [OHMI et LI \[2000\]](#). Ces algorithmes peuvent utiliser plusieurs images consécutives. Ainsi, il devient possible de reconstituer des vecteurs vitesses et des vecteurs accélérations à plusieurs instants.

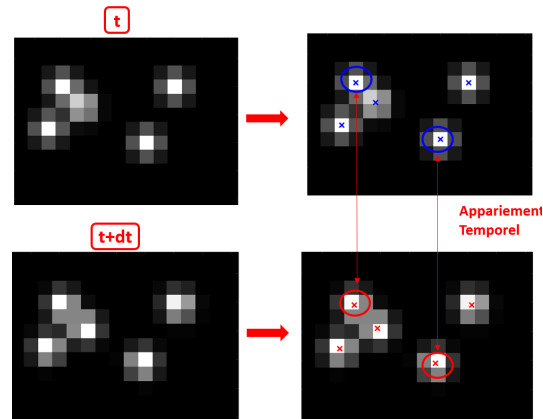


FIGURE 2.5 – Appariement temporel de détections à 2 instants successifs. Ici, on a apparié uniquement 2 détections pour avoir une figure claire.

Comme dans le cas de la PIV, il est possible d'étendre la PTV en 3D en gardant le même dispositif expérimental conçu pour la PIV tomographique. Cela se fait en 3 étapes : la première étape est une détection 2D des centroïdes des particules. La deuxième étape est une identification des détections associées à la même particule afin d'effectuer des triangulations permettant d'estimer la position 3D à chaque instant. Finalement, un appariement temporel est effectué pour estimer le déplacement local des particules [MAAS et collab. \[1993\]](#), [MALIK et collab. \[1993\]](#).

2.1.4 Erreurs d'estimation de vitesse en PTV

[FENG et collab. \[2011\]](#) ont identifié deux types d'erreurs d'estimation de vitesses en PTV : la première erreur provient de l'accélération des particules. En effet, on considère souvent en PTV que la vitesse est constante entre deux images successives. Or, quand cette hypothèse n'est pas vérifiée, on commet inévitablement des erreurs. La seconde résulte des erreurs commises lors de l'estimation de la position de la particule. Ces deux types d'erreurs se propagent et affectent toutes les quantités physiques calculées à partir des vitesses estimées. Par exemple, la détermination de la température cinétique se fait en calculant le carré de la fluctuation moyenne de la vitesse des particules. Ainsi, une estimation peu précise de la vitesse dégradera fortement le résultat. [FENG et collab. \[2011\]](#) ont également montré que l'erreur totale de vitesse est principalement affectée par l'incertitude sur la position des particules pour des intervalles entre clichés dt très petits, alors qu'elle est affectée par l'accélération des particules pour des dt plus larges. Aujourd'hui, les caméras scientifiques à haute cadence sont de moins en moins coûteuses. Ainsi, dans la plupart des expérimentations PTV, on considère uniquement l'incertitude sur la position des particules et on néglige l'effet de l'accélération inconnue sur la vitesse.

2.2 Méthodes basiques de détection et de localisation de particules

Comme évoqué précédemment, une bonne détection/localisation est une condition sine qua non pour aboutir à une estimation de champs de vitesses précise. La section suivante est une description de quelques méthodes phares utilisées en PTV2D. Elles reposent sur des techniques de seuillage et de corrélation.

2.2.1 Détection via seuillage

Dans un contexte aérodynamique à faible grandissement, les particules sont de petite taille et leur image géométrique est petite devant le pixel. Leurs images sur la caméra sont alors caractérisées par des phénomènes de diffraction. Chaque particule entraîne une tâche lumineuse dans l'image dont la variation d'intensité peut être approchée par une fonction gaussienne bidimensionnelle (2D) [RAFFEL et collab. \[2007\]](#). Pour des ensemencements de faible densité, la détection/localisation des particules est généralement effectuée après une binarisation à seuil unique de l'image (*Single-value threshold : SVT*), suivie d'un ajustement gaussien appliqué aux pixels proches des pics d'intensités. La SVT regroupe les pixels adjacents sous forme d'un blob représentant la particule, à condition que la valeur de l'intensité du pixel soit supérieure à un niveau de seuil donné. Cette méthode est rapide et facile à mettre en œuvre. Cependant, elle présente des inconvénients en pratique, notamment son incapacité à traiter des particules qui se chevauchent ou des particules avec des luminosités très différentes. Afin de contourner ces limitations, une segmentation à seuils multiples peut être envisagée. Cette dernière consiste à utiliser des niveaux de seuil variables et appropriés à chaque zone locale de l'image. Toutefois, ces valeurs de seuils locaux sont difficiles à fixer, notamment quand la luminosité de deux particules voisines est trop différente de telle sorte qu'une seule des deux satisfait le seuil local fixé. [OHMI et LI \[2000\]](#) ont proposé une méthode de binarisation dynamique (*Dynamic threshold binarization : DTB*) en ajustant localement la valeur du seuil jusqu'à ce que le niveau de gris moyen de chaque particule soit inférieur à un niveau de contraste pré-réglé. Cet algorithme traite efficacement les particules avec des luminosités irrégulières et parfois les particules qui se chevauchent partiellement en détectant le point de selle entre deux pics d'intensité voisins. Bien qu'il s'agisse d'une amélioration des autres méthodes de binarisation, la DTB tronque souvent l'image des particules, ce qui modifie leurs tailles. Une modification de cette méthode a été proposée par [MIKHEEV et ZUBTSOV \[2008\]](#). Elle utilise le niveau d'intensité maximal des particules au lieu de la moyenne pour la comparaison avec le contraste prédéfini. [CARDWELL et collab. \[2011\]](#) ont constaté que cette modification n'est pas performante pour discriminer les particules en chevauchement et ont proposé une segmentation alternative, un peu plus robuste aux chevauchements, basée sur des opérations morphologiques de dilatation et d'érosion (*Erosion/dilation thresholding : EDT*) afin de protéger les maxima locaux représentant des détections potentielles.

2.2.2 Détection via corrélation

Une autre approche de détection est la corrélation par masques de particules (*Particle Mask Correlation Method : PMCM*) [TAKEHARA et ETOH \[1998\]](#) : une corrélation croisée entre un modèle de particule bidimensionnel et la distribution d'intensité locale en tout point de l'image originale. Le résultat en chaque point est un "plan de corrélation" qui indique la présence de la particule dans l'image originale. Lorsque ce résultat est normalisé entre 1 et -1 , un pic aigu d'intensité 1 indique une correspondance exacte. Dans l'approche PMCM, un modèle gaussien est utilisé. Le processus de corrélation croisée est suivi d'un seuillage pour éliminer les fausses détections ayant des pics inférieurs à 0.7. Cette technique a l'avantage d'identifier correctement les centres de particules, indépendamment de leurs tailles ou du niveau d'intensité maximal des pics mais elle est sensible au bruit et aux chevauchements. De plus les images de particules saturées ne peuvent pas être détectés car ils génèrent de faibles coefficients de corrélation. Une amélioration de cette méthode a été proposée par [ANGARITA-JAIMES et collab. \[2009\]](#) et consiste à utiliser des filtres de corrélation en cascade (*Correlation-Cascade Method : CCM*), ce qui permet de localiser les centroïdes avec des pics plus étroits réduisant ainsi la distance de séparation critique à laquelle deux particules superposées sont résolues. Cependant, cette technique donne des résultats très insuffisants en cas de densités élevées.

2.3 Modèle de formation d'images et notion de dictionnaire

Afin de tenir compte de ce problème de superposition de motifs de particules, les techniques de détection/localisation récentes s'appuient sur des modèles de formation d'image. Cette thèse est illustrée principalement par un modèle de taches quasi gaussiennes de petites dimensions par rapport au pixel, mais les principes présentés sont très généraux et ne se limitent pas à ce modèle. Les sous-sections suivantes présenteront l'équation continue et discrétisée de ce modèle.

2.3.1 Modèle de formation d'images

Il est difficile de concevoir un modèle réaliste de formation d'images PIV car cela nécessite de tenir compte de plusieurs paramètres physiques. Par exemple, dans une configuration optique typique, l'intensité lumineuse varie en raison inverse de la distance à la source. En conséquence, l'intensité des particules n'est pas uniforme. En outre, elle est fortement impactée par les propriétés physiques des traceurs notamment leur différence de diamètre et de leur angle entre la lumière incidente et la direction de la caméra, qui influencent la quantité de lumière diffusée. Cela peut être évalué en utilisant le modèle de Mie [BOHREN et HUFFMAN \[1998\]](#). Enfin, en raison des propriétés intrinsèques des lentilles des caméras, la projection d'une particule sur une image a un impact sur un ensemble de pixels formant ainsi une tache de diffraction au lieu d'une projection ponctuelle. Cette dernière est décrite par la fonction d'étalement du point (PSF). Un modèle physique, couramment utilisé en PIV [RAFFEL et collab. \[2007\]](#), [CHAMPAGNAT et collab. \[2014\]](#), qui décrit la formation d'images est présenté par l'équation (2.1) :

$$i(\mathbf{n}) = \sum_{p=1}^P h(\mathbf{n} - \mathbf{x}_p) e_p. \quad (2.1)$$

En effet, l'intensité $i(\mathbf{n})$ en un pixel \mathbf{n} est la somme de PSF $h: \mathbb{R}^2 \rightarrow \mathbb{R}^+$ identiques, pondérés et décalés dans l'espace, représentant les particules dans l'image. e_p est l'intensité d'une particule p et \mathbf{x}_p sa position associée. P désigne le nombre de particules total dans l'image. Nous utilisons, dans les simulations numériques de cette thèse, le modèle de PSF séparable, standard en PIV [RAFFEL et collab. \[2007\]](#), c'est-à-dire $h(x, y) = h(x) \times h(y)$. Nous utilisons une seule PSF unidimensionnelle pour les deux directions, résultant de la convolution d'une fonction gaussienne g modélisant le flou de diffraction et de défocalisation et une fonction porte π modélisant l'intégration spatiale sur la surface du détecteur :

$$h(x) = (\pi * g)(x) = \frac{1}{2} \left[\operatorname{erf} \left(\frac{x + 0.5}{\sqrt{2}\sigma_{psf}} \right) - \operatorname{erf} \left(\frac{x - 0.5}{\sqrt{2}\sigma_{psf}} \right) \right] \quad (2.2)$$

avec σ_{psf} l'écart type de la fonction gaussienne g . Ce modèle suppose une représentation parcimonieuse des particules dans l'image. Il privilégie également une approche ponctuelle des détections, ce qui permet de faire du tracking par la suite.

2.3.2 Approximation d'ordre 0 et construction de dictionnaire

Afin de discrétiser le modèle continu de formation d'images, on pourrait s'appuyer sur une approximation de Taylor d'ordre 0 de la fonction PSF. La figure 2.6 schématise cette opération pour une PSF unidimensionnelle. En effet, on pourrait approcher la PSF centrée en un point continu \mathbf{x}_p par une PSF identique centrée autour du nœud le plus proche.

Considérons une grille à M nœuds avec Δ comme pas de discrétisation. Pour une particule p donnée, localisée en \mathbf{x}_p , on associe le nœud le plus proche de la grille, indexé $m \in \mathbb{N}$. La position de ce dernier est dénotée $\mathbf{m}\Delta$ avec \mathbf{m} un vecteur bidimensionnel de \mathbb{Z}^2 . Ainsi, $\boldsymbol{\epsilon}_m = (\epsilon_{1m}, \epsilon_{2m})^t = \mathbf{x}_p - \mathbf{m}\Delta$ est l'erreur de discrétisation. Notons $\mathbf{x}_{nm} = \mathbf{n} - \mathbf{m}\Delta$, alors l'approximation de Taylor d'ordre 0 dans le cas 2D s'écrit :

$$h(\mathbf{n} - \mathbf{x}_p) = h(\mathbf{n} - (\mathbf{m}\Delta + \boldsymbol{\epsilon}_m)) \simeq h(\mathbf{x}_{nm}) \quad (2.3)$$

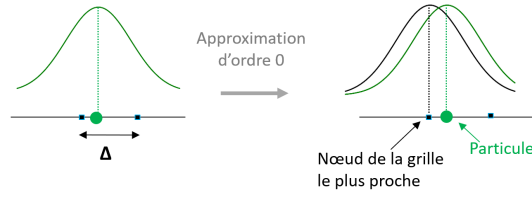


FIGURE 2.6 – La PSF au point x_p est remplacée par la PSF au nœud de la grille le plus proche. Il s'agit d'une discrétisation d'ordre 0.

L'ordre d'approximation dans ce cas est $O(\Delta)$. En évaluant (2.3) au pixel \mathbf{n} , on a :

$$i(\mathbf{n}) \simeq \sum_{m=1}^M e_m h(\mathbf{x}_{nm}). \quad (2.4)$$

L'équation équivalente à (2.4) en forme matricielle est décrite ci-dessous :

$$\mathbf{i} \simeq \mathbf{H}\mathbf{e} \quad \mathbf{H} = (h(\mathbf{x}_{nm}))_{nm} = [h_{nm}] \quad (2.5)$$

où \mathbf{H} est une matrice de taille $N \times M$, appelée dictionnaire, et dont les coefficients h_{nm} représentent la contribution du nœud \mathbf{m} au pixel \mathbf{n} . Ces coefficients sont directement calculés via la PSF comme décrit par l'équation 2.3 : $h_{nm} = h(\mathbf{n} - \mathbf{m}\Delta)$. Il est important de noter qu'avec ce modèle, l'image d'une particule située en un point de la grille (nœud) est associée à une colonne unique de \mathbf{H} et que le nombre de pixels impactés par une particule est limité et dépend essentiellement de la taille de σ_{psf} . Dans nos scripts Matlab développés à l'Onera, la construction des dictionnaires se fait via des matrices creuses et la largeur des fenêtres entourant les positions des nœuds est choisie comme suit :

- une fenêtre de largeur 4×4 pixels, si $\sigma_{psf} \leq 0.65$
- une fenêtre de largeur 6×6 pixels, si $0.65 < \sigma_{psf} \leq 1$
- une fenêtre de largeur 8×8 pixels, si $1 < \sigma_{psf} \leq 1.5$.

La figure 2.7 ci-dessous correspond au cas que nous adopterons dans nos simulations : $\sigma_{psf} = 0.6$. Les points rouges correspondent aux pixels impactés par la diffraction d'une particule dont la projection est située au point vert.

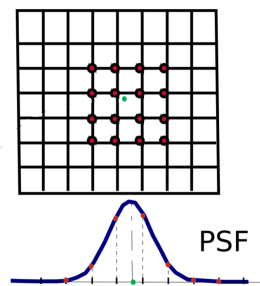


FIGURE 2.7 – Cette figure a été extraite de l'article de CORNIC et collab. [2013] et a été légèrement modifiée pour l'adapter à notre contexte 2D. Les points rouges correspondent aux pixels impactés par la diffraction d'une particule dont la projection est située au point vert. Ici on a représenté la PSF dans sa forme unidimensionnelle.

L'illustration 2.8 montre que les colonnes de la matrice \mathbf{H} représentent simplement la même PSF dans des positions de l'espace décalées.

Comme mentionné précédemment, \mathbf{H} est construit en considérant une liste régulière de positions potentielles de particules appelés nœuds. Cela n'est souvent pas vérifié en pratique avec un nombre fini de nœuds : en effet, une particule physique ne se projettera pas exactement sur le nœud, ce qui conduit à un bruit de modèle non-négligeable et notamment pour des pas de grille

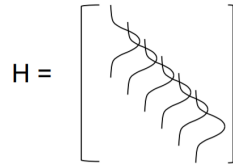


FIGURE 2.8 – Dictionnaire H construit via une discrétisation d'ordre 0. Chaque colonne d'indice m représente la contribution d'une particule positionnée dans le m -ième nœud de la grille.

grossiers. Ainsi, afin de réduire ce bruit de modèle, il faut choisir un pas de discrétisation Δ très fin. Cependant cela a des inconvénients : en effet, le nombre d'atomes du dictionnaire H augmente de $O(\Delta^{-2})$. Donc, l'utilisation d'une grille très fine pour construire H pourrait conduire à un coût de calcul et de mémoire inacceptable, notamment lors du traitement des images mégapixel.

2.4 Méthodes de détection/localisation parcimonieuses à grille en PIV & PTV

La construction de la matrice H requiert l'utilisation d'une discrétisation fine conduisant ainsi à un problème sous-déterminé : H est une matrice contenant beaucoup plus de colonnes que de lignes $M \gg N$. Par conséquent, le système $i = H\mathbf{e}$ admet une infinité de solutions. De nombreuses méthodes ont été proposées dans la littérature PIV afin de résoudre ce problème, en particulier dans le contexte de la reconstruction tomographique, à savoir les algorithmes de reconstruction algébriques : ART [PETRA et collab. \[2008\]](#), MART [ELSINGA et collab. \[2006\]](#) et SMART [ATKINSON et SORIA \[2009\]](#). Cependant, ces méthodes ont quelques inconvénients : en effet, leur vitesse de convergence est lente, en outre ils n'exploitent pas pleinement toutes les informations sur le signal recherché. Par exemple, MART et SMART ne prennent en compte que la non-négativité de la solution alors que ART n'exploite ni la positivité ni la parcimonie. Dans la section suivante, on décrira ce qu'une approximation parcimonieuse est et on présentera quelques méthodes parcimonieuses utilisées dans la littérature PIV et leurs limitations.

2.4.1 Approximation parcimonieuse

Étant donné un signal $\mathbf{e} \in \mathbb{R}^M$, et un dictionnaire redondant $H \in \mathbb{R}^{N \times M}$, on s'intéresse à trouver le plus petit ensemble de colonnes de dictionnaire (également appelés atomes) qui décrit assez bien \mathbf{e} . En d'autres termes, on cherche un vecteur $\mathbf{e} \in \mathbb{R}^M$, qui donne la meilleure approximation $\mathbf{i} \simeq H\mathbf{e}$ et ayant le moins de coefficients non nuls. On dit que \mathbf{e} est une solution parcimonieuse. Une illustration simple de cette définition est donnée par la figure 2.9 :

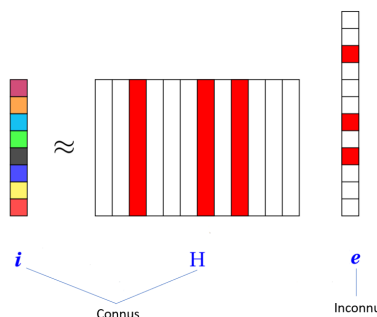


FIGURE 2.9 – Illustration simple d'une approximation parcimonieuse. La couleur rouge indique les coefficients non nuls de la solution recherchée et les atomes (colonnes) associés dans le dictionnaire H . La figure est extraite du manuscrit de thèse de [NGUYEN \[2019\]](#).

Une approximation parcimonieuse conduit souvent à la résolution du problème de minimisation suivant :

$$\min_{\mathbf{e}} \|\mathbf{i} - \mathbf{H}\mathbf{e}\|_2^2 \quad \text{t.q.} \|\mathbf{e}\|_0 \leq K \quad (2.6)$$

où $\|\cdot\|_0$ représente le nombre de coefficients non nul du vecteur \mathbf{e} . On dit que $\mathbf{H}\mathbf{e}$ est une représentation K -parcimonieuse si $\|\mathbf{e}\|_0 = K$. Trouver la solution exacte de 2.6 est un problème NP-complet, c'est-à-dire qu'il nécessite généralement une recherche combinatoire sur l'ensemble de l'espace de solution. Pour des problèmes de dimensionnalité moyenne à élevée, les approches combinatoires sont insolubles d'où le recours à des procédures heuristiques par exemple, des méthodes basées sur la relaxation et les algorithmes gloutons. Une description de ces approches sera présentée dans la section suivante.

En PIV, on intègre souvent la non-négativité du vecteur intensité \mathbf{e} de la distribution des particules comme second a priori. Ainsi, le problème inverse dans ce cas est défini comme ci-dessous :

$$\min_{\mathbf{e} \geq 0} \|\mathbf{i} - \mathbf{H}\mathbf{e}\|_2^2 \quad \text{t.q.} \|\mathbf{e}\|_0 \leq K \quad (2.7)$$

2.4.2 Méthodes de détection/localisation parcimonieuses à grille

Depuis l'avènement de la PIV, plusieurs procédures de reconstruction parcimonieuses ont été étudiées dans la littérature. Les procédures gloutonnes visent à résoudre des problèmes de type l_0 (2.6) en opérant une succession d'optimisations locales, c'est-à-dire en sélectionnant ou en désélectionnant, à chaque itération, les atomes du dictionnaire conduisant à la plus forte diminution de l'écart $\|\mathbf{i} - \mathbf{H}\mathbf{e}\|_2^2$. On peut distinguer plusieurs approches par la façon dont on choisit les vecteurs de la décomposition parcimonieuse : forward, backward et forward-backward. La première catégorie comprend les schémas classiques qui sélectionnent à chaque itération un atome, notamment la méthode des moindres carrés orthogonaux (Orthogonal least squares, OLS) CHEN et collab. [1989], la méthode du Matching pursuit (MP) MALLAT et ZHIFENG ZHANG [1993] et la méthode des MP orthogonaux (Orthogonal matching pursuit, OMP) PATI et collab. [1993]. La seconde catégorie suit une procédure itérative qui part d'un support contenant tous les atomes et désélectionne successivement les moins pertinents. Ses faibles performances dans le cadre d'un dictionnaire sous-déterminé rendent l'utilisation de ce type de méthodes inadaptées à notre situation. Enfin, des procédures forward-backward permettent d'optimiser localement en ajoutant et/ou en retirant des atomes du support courant par exemple : Compressive sampling matching pursuit (CoSaMP) NEEDELL et TROPP [2009] et Subspace pursuit (SP) DAI et MILENKOVIC [2009]. La mise en œuvre de ces étapes selon un rythme alterné permet de corriger les erreurs éventuelles : une fois que nous avons sélectionné un faux atome, il sera toujours possible de le désélectionner. Une évaluation numérique approfondie de quelques méthodes de cette famille (OMP, SP ...) a été présentée par BARBU et collab. [2011] afin de reconstruire une distribution volumique de particules. Des travaux similaires sur le même problème 3D ont été réalisés par CORNIC et collab. [2015] afin de montrer les avantages de CoSaMP avec un coefficient de parcimonie K fixé d'avance. Ces algorithmes gloutons peuvent également prendre compte de non-négativité. Par exemple, une extension non négative de l'OMP (Non-negative orthogonal matching pursuit, NNOMP) a été introduite pour la première fois par BRUCKSTEIN et collab. [2008]. Elle sélectionne les atomes ayant le maximum de corrélation avec le résidu et met à jour les coefficients associés aux atomes sélectionnés à l'itération courante, en résolvant un problème de moindres carrés non négatifs (Non-negative least squares, NNLS) au lieu d'un simple problème de moindres carrés non contraints (Unconstrained least squares, ULS). En outre, contrairement au problème ULS, le NNLS n'a pas de solution analytique. Ainsi, NGUYEN et collab. [2017] ont proposé une implémentation itérative de ce problème afin d'accélérer le calcul de la solution NNLS à chaque itération et par conséquent aboutir à un algorithme de reconstruction rapide.

Des solutions parcimonieuses ont été également recherchées avec des procédures de relaxation convexes. Dans la pratique, la norme l_1 est souvent préférée comme substitut convexe à la norme l_0 . Nous mentionnons, par exemple, les travaux de BARBU et HERZET [2016] sur la reconstruction 3D avec l'ADMM : une technique très préconisée par la communauté du traitement de

signal, pour aborder les problèmes d’optimisation à grande échelle. Leur méthode permet d’intégrer simultanément la non-négativité et la parcimonie et a plusieurs avantages à savoir un traitement plus efficace des observations bruitées et une complexité par itération du même ordre que les méthodes algébriques mais avec un taux de convergence plus important. Plus récemment, une reconstruction avec uniquement une contrainte de positivité a été étudiée par CHEMINET et collab. [2018] dans un cadre 2D. Cela aboutit à un problème de moindres carrés sous contrainte de positivité NNLS. L’étude avait les mêmes objectifs que cette thèse à savoir la détection et la localisation sous-pixel de toutes les projections des particules dans l’espace image, mais son approche se limitait à intégrer uniquement la contrainte de positivité et résoudre le problème NNLS induit, via l’algorithme d’active set de LAWSON et HANSON [1974].

2.5 Bilan

La vélocimétrie par image de particules ou (PIV) selon son acronyme anglais, est une technique de mesure de champ de vitesse qui s’est développée depuis une trentaine d’années. Le principe consiste à découper 2 images successives en fenêtres d’interrogation et de les corrélérer afin d’estimer le déplacement d’une zone de fluide. Cependant, cette étape agit comme un filtre passe bas qui lisse les gradients de vitesse, ce qui détériore la résolution spatiale des mesures de vitesses ainsi que toutes les dérivées spatiales qui en découlent. La vélocimétrie par suivi de particules (PTV) dépasse cette limitation en suivant individuellement l’évolution temporelle de la position de chaque particule, ce qui permet l’estimation du vecteur vitesse à chaque instant et à chaque point de l’écoulement contenant une particule. Il est à noter qu’en PTV, l’erreur sur la vitesse d’une particule lors de l’utilisation de caméras à haute cadence, est essentiellement liée à l’incertitude sur la position des particules. Ainsi, une bonne détection de la présence d’une particule et de l’estimation de sa position est une condition déterminante afin de garantir des mesures précises en PTV. Parmi les techniques usuelles pour détecter et localiser les particules, on peut mentionner celles basées sur une procédure de seuillage à savoir la DTB et l’EDT. D’autres techniques à base de corrélation sont également utilisées notamment la méthode (PMCM) et sa version améliorée (CCM) comprenant 2 filtres mis en série.

Bien que ces techniques soient rapides et facilement implémentables, leurs performances se détériorent fortement quand l’occurrence des chevauchement augmentent dans l’image. Afin de tenir compte du phénomène de superposition, plusieurs auteurs de la communauté PIV ont adopté des modèles de formation d’image, entre autres, le modèle basé sur la PSF, présenté dans ce chapitre. Ce modèle suppose que chaque pixel illuminé de l’image acquise, est le résultat de contribution de particules ponctuelles, dont les projections par le système d’imagerie, sont régies par des PSF pondérées. La reconstruction des particules est souvent faite en 2 étapes : premièrement, on discrétise le modèle de formation d’images sur une grille figée de nœuds, via une approximation de Taylor d’ordre 0 pour chaque pixel de l’image. Cela aboutit à une relation matricielle linéaire entre les observations et une distribution d’intensité de particules. Deuxièmement, on inverse le système linéaire en intégrant la contrainte de positivité ou de parcimonie pour extraire une solution unique, par exemple l’algorithme NNLS, NNOMP et COSAMP. Ces approches peuvent reconstruire des particules proches uniquement si le pas de discrétisation de la grille est suffisamment fin. CHEMINET et collab. [2018] ont montré dans leur article, que NNLS requiert au minimum un pas de vingtième de pixel pour gérer les chevauchements et bien localiser les détections. Néanmoins, cela engendre un encombrement mémoire et un temps de calcul important. Le chapitre suivant présentera une première réponse de cette problématique, en proposant une amélioration de la méthode NNLS et une nouvelle technique alternative de discrétisation avec des pas plus grossiers. Cette dernière sera combinée dans un premier temps avec une contrainte de parcimonie de type l_1 . Le problème d’optimisation induit sera ensuite résolu dans un premier temps via une technique proximale, et dans un second temps via une méthode plus rapide de type active set afin de chercher le meilleur réglage conduisant à une meilleure détection et un minimum de temps de calcul.

2.6 Références

- ANGARITA-JAIMES, N., M. G. ROCA, C. TOWERS, N. READ et D. TOWERS. 2009, «Algorithms for the automated analysis of cellular dynamics within living fungal colonies», *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, vol. 75, p. 768–80. [11](#), [13](#)
- ARROYO, M. P. et C. A. GREATED. 1991, «Stereoscopic particle image velocimetry», *Measurement Science and Technology*, vol. 2, n° 12, p. 1181–1186. [10](#)
- ATKINSON, C. et J. SORIA. 2009, «An efficient simultaneous reconstruction technique for tomographic particle image velocimetry», *Experiments in Fluids*, vol. 47, p. 553–568. [16](#)
- BARBU, I. et C. HERZET. 2016, «A new approach for volume reconstruction in tomoPIV with the alternating direction method of multipliers», *Measurement Science and Technology*, vol. 27, p. 104 002 (11pp). [17](#)
- BARBU, I., C. HERZET et E. MÉMIN. 2011, «Sparse models and pursuit algorithms for piv tomography», . [17](#)
- BOHREN, C. et D. R. HUFFMAN. 1998, *Absorption and Scattering of Light by Small Particles*, Wiley Science Paperback Series. [14](#)
- BROSSARD, C., J. MONNIER, P. BARRICAU, F. VANDERNOOT, Y. LE SANT, F. CHAMPAGNAT et G. LE BESNERAIS. 2009, «Principles and applications of particle image velocimetry», vol. 1. [iii](#), [10](#)
- BRUCKSTEIN, A., M. ELAD et M. ZIBULEVSKY. 2008, «On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations», *Information Theory, IEEE Transactions on*, vol. 54, p. 4813 – 4820. [17](#)
- CARDWELL, N. D., P. P. VLACHOS et K. A. THOLE. 2011, «A multi-parametric particle-pairing algorithm for particle tracking in single and multiphase flows», *Measurement Science and Technology*, vol. 22, n° 10, p. 105 406. [13](#)
- CHAMPAGNAT, F., P. CORNIC, A. CHEMINET, B. LECLAIRE, G. LE BESNERAIS et A. PLYER. 2014, «Tomographic PIV : Particles versus blobs», *Measurement Science and Technology*, vol. 25, n° 8, p. 084 002. [14](#)
- CHEMINET, A., J. KRAWCZYNSKI et P. DRUAULT. 2018, «Particle image reconstruction for particle detection in particle tracking velocimetry», *Measurement Science and Technology*, vol. 29, n° 12, p. 125 202. [11](#), [18](#)
- CHEN, S., S. A. BILLINGS et W. LUO. 1989, «Orthogonal least squares methods and their application to non-linear system identification», *International Journal of Control*, vol. 50, n° 5, p. 1873–1896. [17](#)
- CORNIC, P., F. CHAMPAGNAT, A. CHEMINET, B. LECLAIRE et G. LE BESNERAIS. 2015, «Fast and efficient particle reconstruction on a 3D grid using sparsity», *Experiments in Fluids*, vol. 56, n° 3, 62, ISSN 0723-4864. [17](#)
- CORNIC, P., F. CHAMPAGNAT, A. CHEMINET, B. LECLAIRE et G. LEBESNERAIS. 2013, «Computationally efficient sparse algorithms for tomographic piv reconstruction», . [iii](#), [15](#)
- DAI, W. et O. MILENKOVIC. 2009, «Subspace pursuit for compressive sensing signal reconstruction», *IEEE Transactions on Information Theory*, vol. 55, n° 5, p. 2230–2249. [17](#)
- ELSINGA, G., F. SCARANO, B. WIENEKE et B. OUDHEUSDEN. 2006, «Tomographic particle image velocimetry», *Experiments in Fluids*, vol. 41, p. 933–947. [10](#), [16](#)

- FENG, Y., J. GOREE et B. LIU. 2011, «Errors in particle tracking velocimetry with high-speed cameras», *The Review of scientific instruments*, vol. 82, p. 053 707. [12](#)
- FUCHS, T., R. HAIN et C. KÄHLER. 2017, «Non-iterative double-frame 2d/3d particle tracking velocimetry», *Experiments in Fluids*, vol. 58, p. 119. [11](#)
- GAO, Q., H. WANG et G. SHEN. 2013, «Review on development of volumetric particle image velocimetry», *Chinese Science Bulletin*, vol. 58, p. 4541–4556. [8](#)
- KEANE, R., R. ADRIAN et Y. ZHANG. 1999, «Super-resolution particle imaging velocimetry», *Measurement Science and Technology*, vol. 6, p. 754. [11](#)
- KRAWCZYNSKI, J., B. RENOU et L. DANAILA. 2010, «The structure of the velocity field in a confined flow driven by an array of opposed jets», *Physics of Fluids - PHYS FLUIDS*, vol. 22. [10](#)
- LAWSON, C. L. et R. J. HANSON. 1974, *Solving Least Squares problems*, Prentice-Hall, Saddle River, NJ, USA, p. 149–199. [18](#)
- MAAS, H.-G., A. GRUEN et D. PAPANTONIOU. 1993, «Particle tracking velocimetry in three-dimensional flows part 1. photogrammetric determination of particle coordinates», *Experiments in Fluids*, vol. 15, p. 133–146. [12](#)
- MALIK, N., T. DRACOS et D. PAPANTONIOU. 1993, «Particle tracking velocimetry in three-dimensional flows - part ii : Particle tracking», *Experiments in Fluids*, vol. 15, p. 279–294. [11](#), [12](#)
- MALLAT, S. G. et ZHIFENG ZHANG. 1993, «Matching pursuits with time-frequency dictionaries», *IEEE Transactions on Signal Processing*, vol. 41, n° 12, p. 3397–3415. [17](#)
- MIKHEEV, A. V. et V. M. ZUBTSOV. 2008, «Enhanced particle-tracking velocimetry (EPTV) with a combined two-component pair-matching algorithm», *Measurement Science and Technology*, vol. 19, n° 8, p. 085 401. [13](#)
- NEEDEL, D. et J. TROPP. 2009, «Cosamp : Iterative signal recovery from incomplete and inaccurate samples», *Applied and Computational Harmonic Analysis*, vol. 26, n° 3, p. 301 – 321, ISSN 1063-5203. [17](#)
- NGUYEN, T. T. 2019, *Algorithmes gloutons orthogonaux sous contrainte de positivité*. [iii](#), [16](#)
- NGUYEN, T. T., C. SOUSSEN, J. IDIER et E.-H. DJERMOUNE. 2017, «An optimized version of non-negative OMP», dans *XXVIIe Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017*, Juan-les-Pins, France. [17](#)
- NOGUEIRA, J., A. LECUONA et P. RODRÌGUEZ. 2005, «Limits on the resolution of correlation piv iterative methods. fundamentals», *Experiments in Fluids*, vol. 39, p. 305–313. [10](#)
- OHMI, K. et H.-Y. LI. 2000, «Particle-tracking velocimetry with new algorithms», *Measurement Science and Technology*, vol. 11, n° 6, p. 603–616. [11](#), [12](#), [13](#)
- OUELLETTE, N., H. XU et E. BODENSCHATZ. 2006, «A quantitative study of three-dimensional Lagrangian particle tracking algorithms», *Experiments in Fluids*, vol. 40, n° 2, p. 301–313. [11](#)
- PATI, Y. C., R. REZAIIFAR et P. S. KRISHNAPRASAD. 1993, «Orthogonal matching pursuit : recursive function approximation with applications to wavelet decomposition», dans *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, p. 40–44 vol.1. [17](#)
- PETRA, S., C. SCHNÖRR, A. SCHRÖDER et B. WIENEKE. 2008, «Tomographic image reconstruction in experimental fluid dynamics : Synopsis and problems», . [16](#)

- PRENEL, J. et Y. BAILLY. 2006, «Recent evolutions of imagery in fluid mechanics : From standard tomographic visualization to 3d volumic velocimetry», *Optics and Lasers in Engineering*, vol. 44, p. 321–334. [10](#)
- RAFFEL, M., C. WILLERT, S. WERELEY et J. KOMPENHANS. 2007, *Particle Image Velocimetry - A Practical Guide (2nd Edition)*, ISBN 978-3-540-72307-3. [9](#), [13](#), [14](#)
- SAARENINNE, P. et M. PIIRTO. 2012, «Turbulent kinetic energy dissipation rate estimation from piv velocity vector fields», *Experiments in Fluids*, vol. 29, p. S300–S307. [10](#)
- SCARANO, F. 2003, «Theory of non-isotropic spatial resolution in piv», *Experiments in Fluids*, vol. 35, p. 268–277. [10](#)
- SCARANO, F. 2012, «Tomographic PIV : principles and practice», *Measurement Science and Technology*, vol. 24, n° 1, p. 012 001. [10](#)
- SCHON, J.-P., T. FOURNEL et C. FOURNIER. 2007, «Mesures de caractéristiques d'écoulement par imagerie numérique», *Techniques de l'ingénieur Caractérisation des fluides*, vol. base documentaire : TIB540DUO., n° ref. article : r2162. Fre. [10](#)
- SUN, Y. et Y. ZHANG. 2007, «An overview of room air motion measurement : Technology and application», *Hvac& R Research*, vol. 13, p. 929–950. [8](#)
- TAKEHARA, K. et T. ETOH. 1998, «A study on particle identification in ptv particle mask correlation method», *J. Vis.*, vol. 1, n° 3, p. 313–323, ISSN 1343-8875. [13](#)
- TANAKA, T. et J. K. EATON. 2007, «A correction method for measuring turbulence kinetic energy dissipation rate by piv», *Experiments in Fluids*, vol. 42, p. 893–902. [10](#)
- WILLERT, C. 1997, «Stereoscopic digital particle image velocimetry for application in wind tunnel flows», *Measurement Science and Technology*, vol. 8, n° 12, p. 1465–1479. [10](#)

Chapitre 3

Méthodes de reconstruction parcimonieuses convexes à grille

Sommaire

3.1 Limitations de l'approche NNLS	24
3.2 Régularisation l_1	25
3.2.1 Formulation du problème BP	25
3.2.2 Aperçu de quelques méthodes d'optimisation convexe pour la résolution de BP	25
3.2.3 Implémentation de l'active set BP	27
3.2.4 Intérêt de la régularisation l_1	28
3.3 Dictionnaire d'interpolation de Taylor et régularisation l_1	29
3.3.1 Dictionnaire de Taylor	29
3.3.2 Formulation du problème CBP	31
3.4 FISTA CBP	31
3.4.1 Forme explicite de l'opérateur proximal	31
3.4.2 L'algorithme de FISTA CBP	33
3.5 Active set CBP	34
3.5.1 Reparamétrisation	34
3.5.2 Algorithme de l'active set CBP	35
3.6 Résultats numériques	35
3.7 Bilan	36
3.8 Références	37

CHEMINET et collab. [2018] ont proposé dans leur article de reconstruire les centroïdes des images de particules en résolvant un problème de moindres carrés non négatifs (NNLS), via l'active set de LAWSON et HANSON [1974], implémentée dans MatLab sous le nom de lsqnonneg. Ce dernier a une structure gloutonne et présente un niveau de parcimonie plus élevé par rapport à la méthode algébrique SMART BYRNE [2008] quand il est combiné à une extraction de maxima locaux. Les auteurs ont également montré que le degré de parcimonie augmente avec des pas plus fins. En outre les performances de détection et de localisation en cas de hautes densités sont quasiment optimales quand le niveau de bruit est très faible et quand le pas de discrétisation est de l'ordre de $\frac{1}{20}$ pixel. Cependant, cette approche conduit à un encombrement mémoire et un temps de calcul important quand la taille de l'image grandit. En outre, elle est fortement sensible au bruit. Ce chapitre présentera une première réponse à cette problématique. Il est organisé comme suit : la première section est un récapitulatif des limitations de l'approche NNLS. La seconde section présente une amélioration de celle-ci à savoir l'intégration d'une pénalité de parcimonie de type l_1 sur la variable intensité. La troisième section présente une alternative de discrétisation basée sur l'approximation de Taylor d'ordre 1 qui demeure valide même pour des pas de grille larges, et décrit le problème d'optimisation Continuous basis pursuit (CBP) induit. Les sections 4 et 5 exposent deux manières de résoudre numériquement ce problème d'optimisation : une basée sur une méthode proximale et une autre basée sur l'active set. Enfin, la dernière section présente les résultats de performances de détection des approches testées dans ce chapitre.

3.1 Limitations de l'approche NNLS

Un paramètre clé du modèle de formation d'images discrétisé est son pas de grille Δ . En fait, les erreurs de ce modèle sont contrôlées par Δ : quand la discrétisation est grossière ($\Delta \approx 1\text{pix}$), l'active set NNLS ne parvient pas à séparer les particules en chevauchement. En revanche, quand la grille est suffisamment fine ($\Delta = \frac{1}{20}$ pixel) cela améliore drastiquement les performances de détection et de localisation. La figure 3.1 montre qu'en absence du bruit, l'adoption d'un pas de discrétisation très fin permet de séparer 2 particules proches quand la distance entre elles n'est pas extrêmement petite pour être considérée comme critique. Néanmoins dans ce cas, les besoins en puissance de calcul et de mémoire explosent car le nombre d'atomes du dictionnaire H augmente de $1/\Delta^2$.

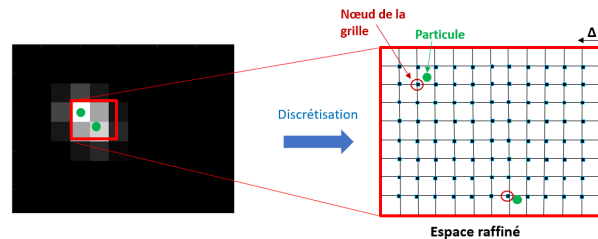


FIGURE 3.1 – Discrétisation très fine pour séparer 2 particules en chevauchement en absence du bruit. L'approximation utilisée par l'active set NNLS consiste à considérer le nœud de la grille le plus proche de la particule car la distance le séparant d'elle est négligeable.

Cette méthode est également sensible au bruit et nécessite une étape de seuillage pour écarter les fausses détections introduites par le bruit. Une illustration qualitative de la solution active set NNLS, en présence du bruit, est représentée dans la figure 3.2. La partie gauche correspond à une solution NNLS sans seuillage sur les intensités et la partie de droite correspond à la même solution avec seuillage. La partie de l'image analysée correspond à une densité d'ensemencement $N_{ppp} = 0,10$ particule par pixel. σ_{psf} est fixé à 0,6 pixel. Δ est égale à $\frac{1}{20}$ pixel. Toutes les particules ont une même valeur d'intensité égale à 1 et sont placées arbitrairement dans l'espace image. Nous indiquons par (o) les particules réelles et (x) les détections observées.

La solution reconstruite est affectée par le bruit. En effet, en présence du bruit, l'algorithme indique des détections un peu partout dans l'image, ce qui baisse drastiquement les performances

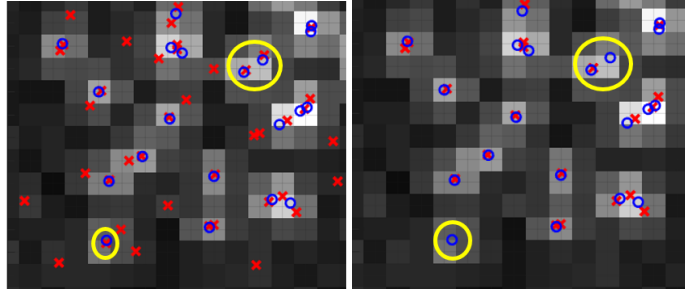


FIGURE 3.2 – L’active set NNLS avant seuillage (à gauche) et après seuillage (à droite) en présence du bruit. L’image initiale contient plusieurs fausses détections. Le processus de seuillage permet d’éliminer quasiment tous les faux positifs produits par le bruit, mais élimine certaines bonnes détections ayant de faibles intensités (Voir les particules entourées en jaune).

de détection. Ainsi, on a introduit une étape supplémentaire de seuillage d’intensité afin éliminer les fausses détections générées par le bruit. Cette procédure a l’inconvénient de systématiquement sousestimer le nombre total de détections notamment en cas de fortes densité d’ensemencement (voir les particules entourées en jaune). De plus, le seuil optimal en détection n’est pas facile à régler car il varie fortement en fonction du niveau de bruit et de la densité des particules.

3.2 Régularisation l_1

3.2.1 Formulation du problème BP

Comme mentionné précédemment, inverser le système $\mathbf{i} \simeq \mathbf{H}\mathbf{e}$ avec uniquement une contrainte de positivité est insuffisant pour traiter des images bruitées. Malgré, sa structure gloutonne, l’active set NNLS ne dispose d’aucun paramètre permettant d’imposer un niveau de parcimonie plus élevé quand le bruit varie. Ainsi, on a introduit un terme de régularisation parcimonieuse de type l_1 [DONOHO \[2006\]](#) comme présenté dans l’équation (3.1) afin d’éliminer encore plus les détections produites par le bruit sans affecter les bonnes détections à faibles intensités :

$$\operatorname{argmin}_{\mathbf{e} \geq 0} \frac{1}{2} \|\mathbf{i} - \mathbf{H}\mathbf{e}\|_2^2 + \lambda \|\mathbf{e}\|_1, \quad (3.1)$$

Cette approche est appelée dans la littérature *Basis Pursuit Denoising* (BPD) [CHEN et collab. \[1998\]](#). Dans ce qui suit, on notera ce problème BP au lieu de BPD. Le paramètre λ est défini par l’utilisateur et contrôle le niveau de parcimonie de la solution recherchée.

Des valeurs importantes de λ produisent des résultats plus parcimonieux. Il est difficile de sélectionner une valeur appropriée de ce paramètre en avance. En conséquence, nous devons souvent résoudre (3.1) de manière répétée pour différents choix de λ afin de sélectionner la valeur garantissant une détection optimale.

3.2.2 Aperçu de quelques méthodes d’optimisation convexe pour la résolution de BP.

La norme l_1 est la mesure de parcimonie la plus utilisée. Elle présente sous certaines conditions de bonnes approximations du problème de parcimonie avec la norme l_0 [CANDÈS et TAO \[2006\]](#), [WRIGHT et collab. \[2009\]](#), [DONOHO \[2006\]](#). De plus, le problème d’optimisation basé sur la norme l_1 peut être résolu en un temps polynomial [FIGUEIREDO et collab. \[2008\]](#). Cette section présente les principales méthodes utilisées en pratique pour résoudre ce problème.

Méthodes proximales

L’une des méthodes les plus populaires pour résoudre un problème (3.1) fait partie de la classe des algorithmes Iterative Shrinkage-Thresholding Algorithms (ISTA) [PARIKH et BOYD \[2014\]](#), où chaque itération implique une multiplication matrice-vecteur impliquant \mathbf{H} et \mathbf{H}^t , suivie d’une

étape de rétrécissement/seuillage DAUBECHIES et collab. [2004], COMBETTES et WAJS [2005], FIGUEIREDO et NOWAK [2003], BLUMENSATH et DAVIES [2008]. Plus précisément, l'étape principale d'ISTA est :

$$\mathbf{e}_{j+1} = \tau_{\lambda\tilde{t}}(\mathbf{e}_j - 2\tilde{t}\mathbf{H}^t(\mathbf{H}\mathbf{e}_j - \mathbf{i})) \quad (3.2)$$

avec \tilde{t} une valeur appropriée du pas de descente et $\tau_{\lambda\tilde{t}}: \mathbb{R}^M \rightarrow \mathbb{R}^M$ un opérateur de rétrécissement défini pour chaque composante d'indice i par :

$$\tau_{\alpha}(e)_i = (|e_i| - \alpha)_+ \operatorname{sgn}(e_i) \quad (3.3)$$

Cet algorithme peut être rattaché aux méthodes proximales itératives forward-backward introduites dans BRUCK [1977] et PASSTY [1979] dans le cadre des *méthodes de splitting*.

L'analyse de vitesse de convergence d'ISTA a été bien étudiée dans la littérature dans divers contextes FIGUEIREDO et collab. [2007], DAUBECHIES et collab. [2004], COMBETTES et WAJS [2005]. L'avantage d'ISTA réside dans sa simplicité, et son adéquation à des problèmes de grande taille. Toutefois sa vitesse de convergence est faible. En effet, l'erreur commise à l'itération j est typiquement $O(\frac{1}{j})$ PARIKH et BOYD [2014]. Le papier de BREDIES et LORENZ [2007] montre que dans certaines conditions, selon la nature de l'opérateur \mathbf{H} , la séquence $(e_j)_j$ produite par ISTA pourrait être très arbitraire conduisant à une vitesse de convergence extrêmement lente. Récemment, des alternative plus rapides ont été proposée à savoir la méthode TWIST BIOUCAS-DIAS et FIGUEIREDO [2007] et la méthode FISTA BECK et TEBoulLE [2009]. Ces méthodes reposent sur le calcul de l'itération suivante, non seulement à partir de la précédente, mais aussi à partir des deux itérations calculées précédemment. Leurs vitesses de convergence sont plus importantes par exemple FISTA a une vitesse de $O(\frac{1}{j^2})$ BECK et TEBoulLE [2009], PARIKH et BOYD [2014].

Méthodes de points intérieurs

Les méthodes de points intérieurs ont été parmi les premières approches développées pour résoudre les problèmes d'approximation parcimonieuse par relaxation convexe. Les premiers algorithmes CHEN et collab. [1998], WRIGHT [1997] appliquent un framework primal-dual où les sous-problèmes internes sont formulés comme des problèmes linéaires de moindres carrés facilement résolubles par des méthodes itératives basées sur des multiplications matrice-vecteur rapides. Une implémentation est disponible dans la boîte à outils SparseLab sous le nom pdco et SolveBP TROPP et WRIGHT [2010]. D'autres méthodes de point intérieur utilisent une approche barrière logarithmique KIM et collab. [2007] après la reformulation de (3.1) en un programme quadratique avec des contraintes linéaires :

$$\operatorname{argmin}_{\mathbf{e}, \mathbf{z} \geq 0} \frac{1}{2} \|\mathbf{i} - \mathbf{H}\mathbf{e}\|_2^2 + \lambda \mathbf{1}^t \mathbf{z} \quad \text{tel que} \quad -\mathbf{z} \leq \mathbf{e} \leq \mathbf{z} \quad (3.4)$$

La méthode est implémentée dans le même toolbox sous le nom *llls* TROPP et WRIGHT [2010]. Elle intègre une variante spéciale de la méthode de Newton appelée méthode de Newton tronquée permettant de compléter les itérations internes de Newton de manière efficiente et rapide KIM et collab. [2007], ZHANG et collab. [2015]. En général, la vitesse de convergence des méthodes de points intérieurs est insensible à la valeur du paramètre de régularisation.

Méthodes d'homotopie

Une autre variante du problème BP est définie comme suit :

$$\operatorname{argmin}_{\mathbf{e} \geq 0} \|\mathbf{i} - \mathbf{H}\mathbf{e}\|_2^2 \quad \text{tel que} \quad \|\mathbf{e}\|_1 \leq \beta. \quad (3.5)$$

Le problème (3.5) est équivalent à (3.1) dans le sens où le chemin des solutions de (3.5) paramétré par un β positif correspond au chemin des solutions de (3.1) lorsque λ varie TROPP et WRIGHT [2010].

Les algorithmes d'homotopie tracent explicitement le chemin des solutions quand le paramètre scalaire β varie dans un intervalle. Ces méthodes exploitent la linéarité par morceaux de la solution en fonction de β , une conséquence directe des conditions d'optimalité de Karush-Kuhn-Tucker (KKT). En effet, en se référant au système KKT, on peut identifier rapidement le prochain point de rupture sur le chemin continu de solutions, c'est à dire la valeur la plus proche de β à laquelle la dérivée de la fonction linéaire par morceaux change.

La méthode d'homotopie de **OSBORNE et collab. [2000]** suit cette approche. Elle commence par une valeur de $\beta = 0$ où la solution de (3.5) est $\mathbf{e} = 0$ puis localise progressivement la valeur suivante la plus élevée de β où une composante de \mathbf{e} passe de zéro à une valeur non nulle, ou vice versa **TROPP et WRIGHT [2010]**. À chaque étape, la méthode met à jour une factorisation QR de la sous-matrice de H correspondante aux composantes non nulles de $\mathbf{e} \in \mathbb{R}_+^M$. Une méthode similaire **EFRON et collab. [2004]** est implémentée sous le nom de `SolveLasso` dans la boîte à outils `SparseLab`. La méthode d'homotopie est efficace en terme de temps de calcul. Elle nécessite environ $2s$ multiplications matrice-vecteur impliquant H et H^t pour identifier s composantes non nulles de \mathbf{e} , ainsi que $O(Ms^2)$ opérations pour mettre à jour les factorisations nécessaires et effectuer d'autres opérations d'algèbre linéaire. Son coût calculatoire est comparable à celui de l'Orthogonal matching pursuit (OMP) **TROPP et WRIGHT [2010]**.

Méthodes d'active set

Les méthodes d'active set sont efficaces pour résoudre les problèmes quadratiques de taille petites et moyennes **NOCEDAL et WRIGHT [1999]**.

Les méthodes primales d'active set passent d'une itération à l'autre en résolvant un sous-problème quadratique dans lequel certaines des contraintes d'inégalité sont imposées comme des égalités **NOCEDAL et WRIGHT [1999]**. Ce sous-ensemble est appelé un "active set" et est désigné au k -ième itération par W_k . En général, un algorithme primal d'active set a la structure suivante **NOCEDAL et WRIGHT [1999]**, **NGUYEN et collab. [2017]** : on commence par un point de départ admissible \mathbf{x}_k , puis on cherche la direction de descente \mathbf{p}_k de la fonction objectif du problème d'égalité défini par l'active set W_k . Si aucune direction ne permet la diminution du critère ($\mathbf{p}_k = 0$), on calcule les multiplicateurs de Lagrange associés à ce problème d'égalité. Si tous les multiplicateurs de Lagrange sont positifs alors la solution \mathbf{x}^* est égale à \mathbf{x}_k et le programme s'arrête. Sinon, on enlève la contrainte correspondante au multiplicateur ayant la plus petite valeur négative et on recommence le procédé de recherche d'une direction de descente. Dans le cas où il y a une direction de descente ($\mathbf{p}_k \neq 0$), on calcule le pas de descente optimal α et puis le point admissible \mathbf{x}_{k+1} suivant :

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \mathbf{p}_k$$

Une des variantes d'active set est la méthode de **LAWSON et HANSON [1974]**. Elle est attrayante en raison de sa simplicité. Elle a été introduite pour la première fois pour traiter des problèmes de moindres carrés sous contraintes de positivité (NNLS). La section suivante présente une adaptation de cette méthode au problème BP défini par l'équation (3.1).

3.2.3 Implémentation de l'active set BP

Grâce à la contrainte de positivité, (3.1) peut s'écrire comme suit :

$$\underset{\mathbf{e} \geq 0}{\operatorname{argmin}} \|\mathbf{i} - H\mathbf{e}\|_2^2 + \lambda \sum_{m=1}^M e_m. \quad (3.6)$$

Ce problème peut donc être résolu en utilisant des méthodes de programmation quadratique à savoir les méthodes d'active set **NOCEDAL et WRIGHT [1999]**. Notre choix a été porté sur l'algorithme de **LAWSON et HANSON [1974]** car nous disposons déjà d'une implémentation sur Matlab destinée aux problèmes NNLS. Nous proposons ici une description détaillée des principales modifications effectuées afin de l'adapter au problème BP (voir l'algorithme 1).

Les variables d'entrée sont l'image acquise \mathbf{i} , la matrice H et le paramètre de parcimonie λ . La variable de sortie est l'intensité \mathbf{e} . Nous désignons par \mathbf{r} le résidu, $S = \{m \in [1 M], \mathbf{e}_m \neq 0\}$ le support de \mathbf{e} , i le nombre de composantes non nulles de \mathbf{e} . H_S est la sous matrice de H dont les indices appartiennent au support S et \mathbf{e}_S la restriction de \mathbf{e} sur cet ensemble. \mathbf{h}_j correspond à la j ème colonne de H et $\mathbf{1}$ est un vecteur colonne dont toutes les composantes sont égales à 1. À chaque itération, l'algorithme sélectionne l'atome \mathbf{h}_l ayant la plus forte corrélation avec le résidu courant \mathbf{r} . Il ajoute ensuite l'indice associé l au support S puis résout le critère des moindres carrés pénalisés sur ce support. Lorsque certains des coefficients de la solution deviennent négatifs, une boucle interne est répétée jusqu'à ce que tous les indices correspondants à ces coefficients négatifs soient éliminés de S .

Les principales modifications dans le cas BP correspondent aux lignes 2 et 6 et ont pour but d'intégrer le terme de parcimonie l_1 dans la procédure d'optimisation. Noter que dans le cas $\lambda = 0$, on aboutit exactement à l'implémentation `lsqnonneg` adoptée par [CHEMINET et collab. \[2018\]](#), appelée dans ce manuscrit `active set NNLS`.

Algorithm 1 L'active set BP

Input: \mathbf{i}, H, λ
Output: $\mathbf{e} = \text{bp}(\mathbf{i}, H, \lambda)$

- 1: $S \leftarrow \emptyset, \mathbf{e} \leftarrow \mathbf{0}, \mathbf{r} \leftarrow \mathbf{i}$
- 2: $\mu_S \leftarrow H_S^t \mathbf{r} - \lambda \mathbf{1}$
- 3: **while** $\tilde{S} \neq \emptyset$ & $\max \mu_{\tilde{S}} > 0$ **do**
- 4: $l \leftarrow \underset{j \in \tilde{S}}{\text{argmax}} \langle \mathbf{h}_j, \mathbf{r} \rangle$ {retourne l'indice correspondant au produit scalaire maximum entre \mathbf{r} et $H_{\tilde{S}}$.}
- 5: $S \leftarrow S \cup \{l\}$
- 6: $\mathbf{z}_S \leftarrow \underset{\mathbf{e}_S}{\text{argmin}} \|\mathbf{i} - H_S \mathbf{e}_S\|_2^2 + \lambda \sum_{m \in S} e_m$
- 7: $\mathbf{z}_{\tilde{S}} \leftarrow \mathbf{0}$
- 8: **while** $\min \mathbf{z}_S \leq 0$ **do** {boucle interne pour éliminer les composantes négatives de \mathbf{z}_S .}
- 9: $j \leftarrow \underset{m \in S}{\text{argmin}} \left(-\frac{e_m^k}{e_m^k - z_m^k} \right)$ {renvoie les indices où \mathbf{z}_S est négatif. j est l'itération courante de la boucle interne.}
- 10: $\alpha \leftarrow \min \left\{ 1, \min_{m \in S} \left(-\frac{e_m^k}{e_m^k - z_m^k} \right) \right\}$
- 11: $\mathbf{e} \leftarrow \mathbf{e} + \alpha (\mathbf{z} - \mathbf{u})$ {choisit les vecteurs \mathbf{e} sans composantes négatives.}
- 12: $S \leftarrow S - \{j\}$ {l'indice j est supprimé du support courant S .}
- 13: $\mathbf{z}_S \leftarrow \underset{\mathbf{e}_S}{\text{argmin}} \|\mathbf{i} - H_S \mathbf{e}_S\|_2^2 + \lambda \sum_{m \in S} e_m$
- 14: $\mathbf{z}_{\tilde{S}} \leftarrow \mathbf{0}$
- 15: **end while**
- 16: $\mathbf{e} \leftarrow \mathbf{z}$
- 17: **end while**

3.2.4 Intérêt de la régularisation l_1

A l'instar de l'active set NNLS, l'active set BP introduit des détections multiples autour de chaque particule quelque soit le pas de discrétisation utilisé, d'où l'ajout d'une étape d'extraction de maxima locaux afin de garder une unique détection autour de chaque particule. Ici, le réglage du couple (λ, Δ) a été effectué de telle sorte à avoir un compromis entre le nombre de faux positifs et le nombre de faux négatifs. Nous fournirons dans la sous-section 5.2.2 du chapitre 5, tous les détails nécessaires à la compréhension de la procédure de réglage adoptée. Lors de nos premières simulations, on a pu constater qu'il est inutile de trop affiner la grille quand les images sont bruitées que ça soit pour l'active set NNLS ou l'active set BP et qu'il suffit de se contenter d'un pas de $\frac{1}{3}$ pixel au lieu de $\frac{1}{20}$ pixel. Pour ce qui suit, on garde les mêmes simulations que celles utilisées dans la section précédente. La figure 3.3 est une comparaison qualitative entre 2 solutions active set BP : la figure de gauche représente le résultat obtenu sans seuillage et la figure de droite correspond au résultat obtenu avec un seuillage optimal en termes de détection. La zone de l'image analysée comprend une densité $N_{\text{ppp}} = 0.10$. Il est clair que, par rapport aux résultats de l'active set NNLS, le nombre de fausses détections générées par le bruit est sensiblement réduit par l'active set BP. Le résultat s'améliore encore plus quand on seuille les intensités à condition de choisir

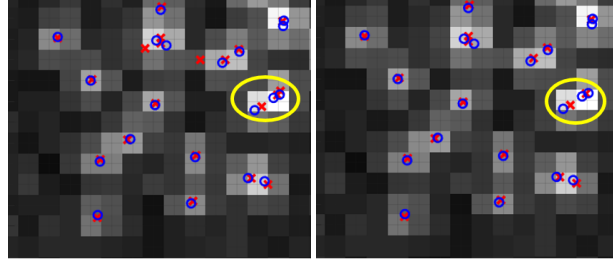


FIGURE 3.3 – A gauche la solution active set BP avant seuillage et à droite la solution après seuillage, en présence du bruit. La première solution contient peu de fausses détections par rapport à l’active set NNLS non seuillé de la figure 3.2. L’application de seuillage améliore un peu plus les performances de détection de l’active set BP mais peut potentiellement effacer quelques bonnes détections (voir cercles jaunes).

minutieusement le bon seuil afin de garder les bonnes détections de faibles densités.

Néanmoins, cette stratégie a le défaut de requérir un post-processing, ce qui se traduit par un ajout de nouveaux paramètres à régler. En outre, le pas de discrétisation demeure relativement petit. Face à ce défi, on s’est proposé de changer le type de discrétisation et opter pour une approximation plus précise même pour des pas plus grossiers. Le chapitre suivant expose les grandes lignes de cette approche.

3.3 Dictionnaire d’interpolation de Taylor et régularisation l_1

3.3.1 Dictionnaire de Taylor

La technique de discrétisation décrite dans la section 2.1 du chapitre 2 est grossière : elle consiste à approcher l’image d’une particule p ayant une position continue \mathbf{x}_p par l’image de la même particule située au nœud le plus proche. Cela correspond à une approximation de Taylor d’ordre 0 de la fonction PSF au point continu \mathbf{x}_p . Nous nous attendons à une meilleure approximation du modèle de formation d’images continu si nous augmentons l’ordre de l’approximation de la PSF. En outre, il serait possible d’avoir un dictionnaire avec peu d’atomes car les espacements entre nœuds pourraient être plus larges vu la qualité de l’approximation. Nous décrivons ci-dessous une technique de discrétisation alternative introduite par [EKANADHAM et collab. \[2011a\]](#) dans le cas de signaux 1D. Elle est obtenue via un développement de Taylor d’ordre 1 de la fonction PSF. Le principe est d’approximer l’image de la particule en un point continu \mathbf{x}_p de l’image par une combinaison linéaire de la fonction PSF et de sa dérivée au nœud le plus proche. La figure 3.4 illustre cette approximation dans un cas 1D. L’image d’une particule peut être approximée fidèlement par une combinaison linéaire de la même image de particule et de sa dérivée au nœud de la grille le plus proche.

Nous généralisons cette approximation au cas 2D. Nous gardons les mêmes notations que celles employées dans la section 2.3.1 du chapitre 2 sur l’approximation de Taylor d’ordre 0. M est le nombre de nœuds de la grille. $\mathbf{x}_p = (x_{1p}, x_{2p}) \in \mathbb{R}^2$, est la position continue de la particule. Le nœud le plus proche associé est indexé m et sa position est dénotée $\mathbf{m}\Delta$, $\mathbf{m} = (m_1, m_2) \in \mathbb{N}^2$. \mathbf{n} est le vecteur des coordonnées d’un pixel de l’image, dont l’indice est noté n , cf Section 2.3.1. Le décalage horizontal entre le centre de la particule et le nœud le plus proche est noté ϵ_{1m} et le décalage vertical est noté ϵ_{2m} . Ces variables sont liées linéairement comme décrit ci-dessous :

$$x_{1p} = m_1\Delta + \epsilon_{1m} \quad \text{et} \quad x_{2p} = m_2\Delta + \epsilon_{2m}. \quad (3.7)$$

Ainsi, l’approximation de Taylor du premier ordre de la fonction PSF peut être exprimée comme suit :

$$\begin{aligned} h(\mathbf{n} - \mathbf{x}_p) &\simeq h((\mathbf{n} - \mathbf{m}\Delta) - \boldsymbol{\epsilon}) \\ &\simeq h(\mathbf{x}_{nm}) - \epsilon_{1m} \frac{\partial h}{\partial x}(\mathbf{x}_{nm}) - \epsilon_{2m} \frac{\partial h}{\partial y}(\mathbf{x}_{nm}). \end{aligned} \quad (3.8)$$

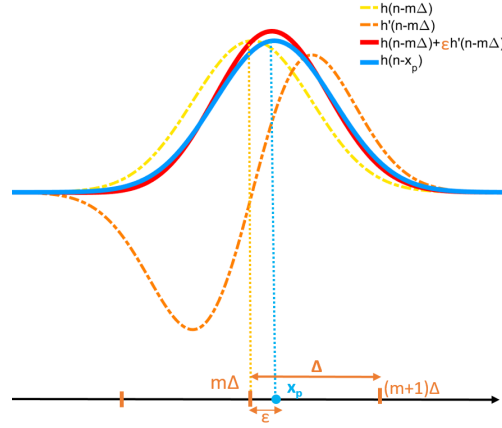


FIGURE 3.4 – Principe de la construction de dictionnaire par approximation de Taylor dans le cas 1D. La PSF est approchée en un point continu \mathbf{x}_p par une somme pondérée de PSF et de sa dérivée au nœud le plus proche indexé m . Δ est la taille de la grille. ϵ est l'erreur de discrétisation.

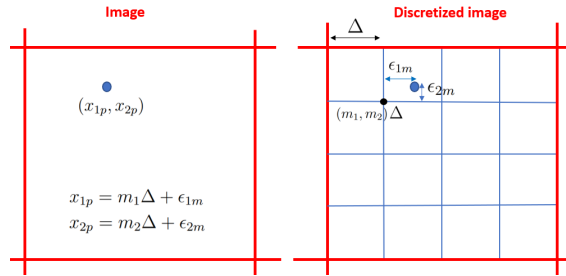


FIGURE 3.5 – Relation linéaire entre la position de la particule \mathbf{x}_p et la position du nœud \mathbf{m} le plus proche.

avec $\mathbf{x}_{nm} = \mathbf{n} - \mathbf{m}\Delta$, $|\epsilon_{1m}| \leq \frac{\Delta}{2}$ et $|\epsilon_{2m}| \leq \frac{\Delta}{2}$.

En intégrant (3.8) pour chaque position de particule \mathbf{x}_p dans l'équation du modèle d'imagerie (2.1), on aboutit à :

$$\mathbf{i}(\mathbf{n}) \simeq \sum_{m=1}^M \mathbf{e}_m h(\mathbf{x}_{nm}) - d_{1m} \frac{\partial h}{\partial x}(\mathbf{x}_{nm}) - d_{2m} \frac{\partial h}{\partial y}(\mathbf{x}_{nm}) \quad (3.9)$$

avec

$$d_{1m} = \epsilon_{1m} e_m \quad \text{et} \quad d_{2m} = \epsilon_{2m} e_m \quad (3.10)$$

Cette procédure introduit deux variables spatiales supplémentaires d_{1m} et d_{2m} , contraintes dans un intervalle 2D, en fonction du pas de la grille Δ et de l'intensité de la distribution des particules \mathbf{e} . Cela se traduit géométriquement par l'appartenance du triplet (e_m, d_{1m}, d_{2m}) à un cône pyramidal à base carré \mathcal{P} , défini comme suit :

$$\mathcal{P} = \left\{ (x, y, z) \in \mathbb{R}^3 \mid |y| \leq \frac{\Delta}{2} x, |z| \leq \frac{\Delta}{2} x \right\}. \quad (3.11)$$

La contrainte sur d_{1m} et d_{2m} est essentielle et a 2 objectifs : garantir que l'approximation de Taylor ne soit utilisée que lorsque celle-ci est précise et éviter d'expliquer l'image d'une même particule par des jeux de coefficients associés à des nœuds de grille voisins. Enfin, l'équation (3.9) conduit à l'expression matricielle suivante :

$$\mathbf{i} \simeq \mathbf{H}\mathbf{e} + \mathbf{H}_1 \mathbf{d}_1 + \mathbf{H}_2 \mathbf{d}_2, \quad \begin{cases} \mathbf{H} = [h(\mathbf{x}_{nm})]_{nm} \\ \mathbf{H}_1 = [-\frac{\partial h}{\partial x}(\mathbf{x}_{nm})]_{nm} \\ \mathbf{H}_2 = [-\frac{\partial h}{\partial y}(\mathbf{x}_{nm})]_{nm} \end{cases} \quad (3.12)$$

Où $(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2)$ sont des vecteurs à M composantes appartenant à l'ensemble convexe \mathcal{E}_1 , défini ci-dessous :

$$\mathcal{E}_1 = \{(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2) \mid (e_m, d_{1m}, d_{2m}) \in \mathcal{P}, m \in \{1 \dots M\}\}. \quad (3.13)$$

3.3.2 Formulation du problème CBP

Nous formulons le problème de détection et de localisation de particules comme la minimisation sous contrainte d'un critère quadratique basé sur (3.12) et (3.13) avec l'ajout d'une pénalisation l_1 . Cette approche est appelée dans la littérature *Continuous Basis Pursuit* (CBP) [EKANADHAM et collab. \[2011b\]](#). Dans notre cas il est exprimé comme suit :

$$\operatorname{argmin}_{\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2 \in \mathcal{E}_1} \|\mathbf{i} - (\mathbf{H}\mathbf{e} + \mathbf{H}_1\mathbf{d}_1 + \mathbf{H}_2\mathbf{d}_2)\|_2^2 + \lambda\|\mathbf{e}\|_1. \quad (3.14)$$

Il est à noter que cette régularisation implique une parcimonie structurée. En effet, quand le signal intensité \mathbf{e} est parcimonieux, alors les 2 vecteurs \mathbf{d}_1 et \mathbf{d}_2 le sont aussi à cause de la contrainte (3.13). Nous avons proposé de résoudre ce problème dans un premier temps avec la méthode du gradient proximal FISTA puis dans un second temps avec la méthode de l'active set de [LAWSON et HANSON \[1974\]](#) qui s'est avérée plus rapide, tout en ajoutant une étape supplémentaire de seuillage pour promouvoir la parcimonie des résultats. Nous exposerons en détail ces deux solveurs dans les sections suivantes.

3.4 FISTA CBP

Les solveurs typiques du problème (3.14) sont des méthodes proximales. Nous traitons ici la famille des méthodes *Iterative Shrinkage-Thresholding Algorithms* (ISTA) et plus particulièrement la version accélérée FISTA qui donne une meilleure vitesse de convergence. Cette famille de méthodes peut être vue comme une extension de l'algorithme du gradient classique aux fonctions convexes non différentiables. Elle est facilement implémentable et est adéquate aux problèmes de grande taille.

Posons f la partie quadratique de la fonction objectif :

$$f = \|\mathbf{i} - (\mathbf{H}\mathbf{e} + \mathbf{H}_1\mathbf{d}_1 + \mathbf{H}_2\mathbf{d}_2)\|_2^2 \quad (3.15)$$

et g la fonction de pénalité :

$$g = \lambda\|\mathbf{e}\|_1 + \mathbf{I}_{\mathcal{E}_1}, \quad (3.16)$$

avec $\mathbf{I}_{\mathcal{E}_1}$ la fonction indicatrice de l'ensemble \mathcal{E}_1 . Alors, la fonction objectif du problème CBP peut être écrite comme la somme de f et g . La méthode du gradient proximal FISTA s'écrit dans ce cas comme suit :

$$\begin{aligned} \tilde{\mathbf{v}}_{j+1} &= \tilde{\mathbf{u}}_j + \omega_j(\tilde{\mathbf{u}}_j - \tilde{\mathbf{u}}_{j-1}) \\ \tilde{\mathbf{u}}_{j+1} &= \operatorname{prox}_{\gamma_j g}(\tilde{\mathbf{v}}_j - \gamma_j \nabla f(\tilde{\mathbf{v}}_j)) \end{aligned} \quad (3.17)$$

pour chaque itération j , avec $\tilde{\mathbf{u}}_j \in \mathbb{R}^{3M}$, $\gamma_j \in \mathbb{R}$, $\operatorname{prox}_{\gamma_j g}$ l'opérateur proximal de la fonction $\gamma_j g$ défini comme suit [PARIKH et BOYD \[2014\]](#) :

$$\operatorname{prox}_{\gamma_j g}(\tilde{\mathbf{v}}) = \operatorname{argmin}_{\mathbf{y}} \left(g(\mathbf{y}) + \frac{1}{2\gamma_j} \|\mathbf{y} - \tilde{\mathbf{v}}\|_2^2 \right), \quad \mathbf{v} \in \mathbb{R}^{3M} \quad (3.18)$$

et $\omega_j \in [0, 1]$. Un simple choix de ω_j est donné par : $\omega_j = \frac{j}{j+3}$. Dans [BECK et TEBOULLE \[2009\]](#), les auteurs ont montré que FISTA converge à une vitesse de $O(\frac{1}{j^2})$.

3.4.1 Forme explicite de l'opérateur proximal

L'opérateur proximal peut être évalué très efficacement dans le cas présent. En effet : pour un point $\mathbf{y} = (\mathbf{a}^t, \mathbf{b}^t, \mathbf{c}^t)^t$, l'opérateur proximal $\operatorname{prox}_{\gamma g}$ est défini par :

$$\operatorname{argmin}_{(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2) \in \mathcal{E}_1} \lambda\|\mathbf{e}\|_1 + \frac{1}{2\gamma} (\|\mathbf{e} - \mathbf{a}\|_2^2 + \|\mathbf{d}_1 - \mathbf{b}\|_2^2 + \|\mathbf{d}_2 - \mathbf{c}\|_2^2).$$

Le problème est séparable et peut être résolu indépendamment pour chaque composante :

$$(e_m, d_{1m}, d_{2m}) = \text{proj}_{\mathcal{P}}(a_m - \lambda\gamma, b_m, c_m),$$

avec $\text{proj}_{\mathcal{P}}$ la projection d'un point de \mathbb{R}^3 sur le cône \mathcal{P} :

$$\text{proj}_{\mathcal{P}}(u, v, w) = \underset{(x, y, z) \in \mathcal{P}}{\text{argmin}} \quad (x - u)^2 + (y - v)^2 + (z - w)^2. \quad (3.19)$$

La résolution de (3.19) repose sur des théorèmes d'optimisation sous contraintes et de caractérisation de cône convexe. Ce problème admet un minimiseur unique car il s'agit d'un problème de projection sur un cône convexe. La Table 3.1 donne son expression explicite déduite des conditions d'optimalité BERTSEKAS [2016]. La démonstration est donnée en annexe en fin de ce document.

Cas	x^*	y^*	z^*
\mathcal{P}	u	v	w
\mathcal{P}^*	0	0	0
\mathcal{F}_1	$\frac{u+\alpha v}{1+\alpha^2}$	αx^*	w
\mathcal{F}_2	$\frac{u+\alpha w}{1+\alpha^2}$	v	αx^*
\mathcal{F}_3	$\frac{u-\alpha v}{1+\alpha^2}$	$-\alpha x^*$	w
\mathcal{F}_4	$\frac{u-\alpha w}{1+\alpha^2}$	v	$-\alpha x^*$
\mathcal{A}_1	$\frac{u+\alpha v+\alpha w}{1+2\alpha^2}$	αx^*	αx^*
\mathcal{A}_2	$\frac{u-\alpha v+\alpha w}{1+2\alpha^2}$	$-\alpha x^*$	αx^*
\mathcal{A}_3	$\frac{u-\alpha v-\alpha w}{1+2\alpha^2}$	$-\alpha x^*$	$-\alpha x^*$
\mathcal{A}_4	$\frac{u+\alpha v-\alpha w}{1+2\alpha^2}$	αx^*	$-\alpha x^*$

TABLEAU 3.1 – Solutions du problème de projection sur le cône \mathcal{P} . La première colonne correspond à l'ensemble auquel appartient (u, v, w) .

Les ensembles \mathcal{P}^* , \mathcal{F}_i et \mathcal{A}_i sont définis par :

$$\begin{aligned} \mathcal{P}^* &\triangleq \{(x, y, z) \mid |y + z| \leq -\frac{x}{\alpha}, |y - z| \leq -\frac{x}{\alpha}\} \\ \mathcal{F}_1 &\triangleq \{(x, y, z) \mid y \geq \alpha x, |z| \leq \alpha \frac{x+\alpha y}{1+\alpha^2}\} \\ \mathcal{F}_2 &\triangleq \{(x, y, z) \mid z \geq \alpha x, |y| \leq \alpha \frac{x+\alpha z}{1+\alpha^2}\} \\ \mathcal{F}_3 &\triangleq \{(x, y, z) \mid y \leq -\alpha x, |z| \leq \alpha \frac{x-\alpha y}{1+\alpha^2}\} \\ \mathcal{F}_4 &\triangleq \{(x, y, z) \mid z \leq -\alpha x, |y| \leq \alpha \frac{x-\alpha z}{1+\alpha^2}\} \\ \mathcal{A}_1 &\triangleq \{(x, y, z) \mid y + z \geq -\frac{x}{\alpha}, z \geq \alpha \frac{x+\alpha y}{1+\alpha^2}, y \geq \alpha \frac{x+\alpha z}{1+\alpha^2}\} \\ \mathcal{A}_2 &\triangleq \{(x, y, z) \mid z - y \geq -\frac{x}{\alpha}, z \geq \alpha \frac{x-\alpha y}{1+\alpha^2}, y \leq -\alpha \frac{x+\alpha z}{1+\alpha^2}\} \\ \mathcal{A}_3 &\triangleq \{(x, y, z) \mid y + z \leq \frac{x}{\alpha}, z \leq -\alpha \frac{x+\alpha y}{1+\alpha^2}, y \leq -\alpha \frac{x+\alpha z}{1+\alpha^2}\} \\ \mathcal{A}_4 &\triangleq \{(x, y, z) \mid z - y \leq \frac{x}{\alpha}, z \leq -\alpha \frac{x-\alpha y}{1+\alpha^2}, y \geq \alpha \frac{x-\alpha z}{1+\alpha^2}\}. \end{aligned}$$

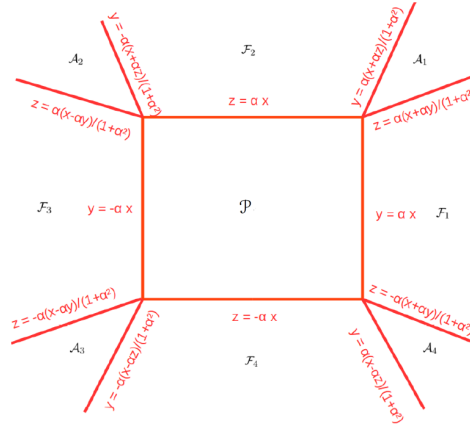


FIGURE 3.6 – Représentation graphique d'une tranche à $x = x_0 > 0$ du cône \mathcal{P} et des sous ensembles qui définissent, par morceaux, la solution dans le cas x positif.

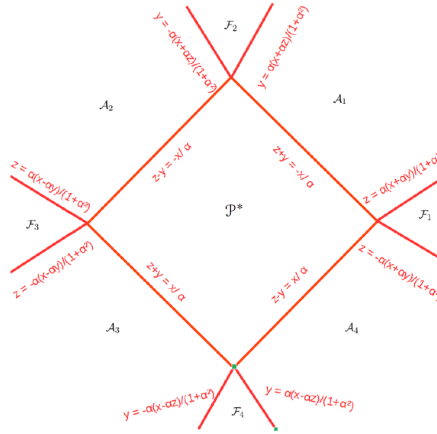


FIGURE 3.7 – Représentation graphique d'une tranche à $x = x_0 < 0$ des sous ensembles qui définissent, par morceaux, la solution dans le cas x négatif.

3.4.2 L'algorithme de FISTA CBP

Après avoir calculé l'opérateur proximal de la fonction de pénalité g , nous avons intégré son expression analytique à notre algorithme FISTA. Les variables d'entrée sont l'image acquise \mathbf{i} , les 3 dictionnaires issus de la discrétisation de Taylor d'ordre 1 : \mathbf{H} , \mathbf{H}_1 et \mathbf{H}_2 , le coefficient de parcimonie λ et ϵ un paramètre intervenant dans le critère de convergence. Les variables de sortie sont : le vecteur intensité \mathbf{e} , et les vecteurs déplacements \mathbf{d}_1 et \mathbf{d}_2 correspondants aux décalages horizontaux et verticaux. Chaque composante non nulle e_m de \mathbf{e} peut être associée à une particule détectée indexée par p . Son intensité \tilde{e}_p est égale à e_m et sa position sous-pixel peut être estimée à l'aide des équations ci-dessous déduites de (3.7) :

$$\mathbf{x}_{1p} = m_1 \Delta + \frac{d_{1m}}{e_m} \quad \text{and} \quad \mathbf{x}_{2p} = m_2 \Delta + \frac{d_{2m}}{e_m}. \quad (3.21)$$

On note $F = f + g$ la fonction objectif à minimiser. À chaque itération, après avoir obtenu le point d'itération courant \mathbf{s}_j , un point supplémentaire \mathbf{y}_{j+1} est choisi comme combinaison linéaire du point d'itération courant \mathbf{s}_j et le point d'itération précédent \mathbf{s}_{j-1} . Le point d'itération suivant \mathbf{s}_{j+1} est alors défini en 2 étapes : on applique d'abord la méthode de descente de gradient sur la partie quadratique f en \mathbf{y}_{j+1} puis l'opérateur proximal de γg sur le résultat obtenu. L'algorithme 2 décrit en détail le processus de reconstruction 2D avec FISTA :

Algorithm 2 FISTA

Input: $i, H, H_1, H_2, \lambda, \epsilon$
Output: $s = (e^t, \mathbf{d}_1^t, \mathbf{d}_2^t)^t$
 $\beta \leftarrow 0.5, \gamma \leftarrow 10$
while $|F(s_j) - F(s_{j-1})| \leq \epsilon$ **do**
 $w_j \leftarrow \frac{j}{j+3}$
 $\mathbf{y}_{j+1} \leftarrow s_j + w_j(s_j - s_{j-1})$
Repeat :
 $\mathbf{z} \leftarrow \text{prox}_{\gamma g}(\mathbf{y}_{j+1} - \gamma \nabla f(\mathbf{y}_{j+1}))$
Break if $f(\mathbf{z}) \leq \tilde{f}_\gamma(\mathbf{z}, \mathbf{y}_{j+1})$
Update $\gamma \leftarrow \beta \gamma$
Return $\gamma_{j+1} \leftarrow \gamma, s_{j+1} \leftarrow \mathbf{z}$
end while

Avec :

$$\begin{aligned} \tilde{f}_\gamma(\mathbf{z}, \mathbf{y}) &= f(\mathbf{y}) + \nabla f(\mathbf{y})^t (\mathbf{z} - \mathbf{y}) + \frac{1}{2\gamma} \|\mathbf{z} - \mathbf{y}\|_2^2 \\ F(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2) &= \|\mathbf{i} - (H\mathbf{e} + H_1\mathbf{d}_1 + H_2\mathbf{d}_2)\|_2^2 + \lambda \sum_{m=1}^M e_m \end{aligned}$$

En pratique, nous obtenons souvent un petit amas de détections non nulles dans le voisinage de chaque véritable particule. La largeur de cet amas varie en fonction de la tolérance ϵ . Pour garantir la convergence de l'algorithme, on a fixé ϵ à 10^{-12} . Cette valeur permet de réduire au maximum la taille de l'amas, ce qui permet de séparer les particules en chevauchement après extraction des maxima locaux. Cependant, le choix d'une telle valeur entraîne un temps de calcul très important et plus particulièrement dans le cas des pas très fins.

3.5 Active set CBP

3.5.1 Reparamétrisation

Afin d'accélérer la reconstruction 2D et pouvoir évaluer les résultats de notre approche sur une grande plage de Δ , nous avons opté pour un algorithme de reconstruction de second ordre de type active set. Nous avons ensuite reparamétrisé le problème CBP (3.12) afin de remplacer toutes les contraintes linéaires par des contraintes de positivité, ce qui permettrait d'appliquer l'algorithme 1 sur le problème induit. Cela est possible car \mathcal{E}_1 est un cône convexe et on peut déduire une description générative avec des coefficients positifs pour chaque sous-cône \mathcal{P} de \mathcal{E}_1 . Cette re-paramétrisation a été initialement suggérée par DUVAL et PEYRÉ [2017] dans le cas 1D, nous l'avons étendu au cas 2D. Dans notre cas, \mathcal{P} a une structure pyramidale à base carrée et peut être entièrement généré par 4 vecteurs dont les directions sont définies par les arêtes de la pyramide \mathcal{P} .

Soit $\alpha = \frac{\Delta}{2}$ et \tilde{P}_α la matrice dont les colonnes correspondent aux coordonnées cartésiennes du

vecteur directeur de chaque arête $\mathbf{x} = \begin{pmatrix} 1 \\ \alpha \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} 1 \\ -\alpha \end{pmatrix}$, $\mathbf{z} = \begin{pmatrix} 1 \\ \alpha \end{pmatrix}$, et $\mathbf{w} = \begin{pmatrix} 1 \\ -\alpha \end{pmatrix}$:

$$\tilde{P}_\alpha = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \alpha & -\alpha & \alpha & -\alpha \\ \alpha & \alpha & -\alpha & -\alpha \end{pmatrix}. \quad (3.22)$$

Alors, le sous-cône \mathcal{P} peut être décrit sous la forme générative suivante :

$$\mathcal{P} = \{(a, b, c) \in \mathbb{R}^3 \mid (a, b, c) = (v_1, v_2, v_3, v_4) \tilde{P}_\alpha^t, v_i \geq 0, i \in \{1 \dots 4\}\}. \quad (3.23)$$

Ainsi, \mathcal{E}_1 peut être décrit comme ci-dessous :

$$\mathcal{E}_1 = \{(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2)^t \in \mathbb{R}^{3M} \mid (\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2)^t = P_\alpha \mathbf{u}, \mathbf{u} \in \mathbb{R}_+^{4M}\}, \quad (3.24)$$

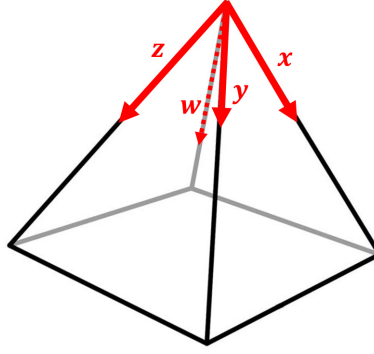


FIGURE 3.8 – \mathcal{P} est un cône pyramidal à base carrée, généré par la combinaison linéaire de quatre vecteurs \mathbf{x} , \mathbf{y} , \mathbf{z} et \mathbf{w} à coefficients positifs. Ces vecteurs sont les colonnes de la matrice $\tilde{\mathbf{P}}_\alpha$.

avec $\mathbf{P}_\alpha = \tilde{\mathbf{P}}_\alpha \otimes \mathbb{I}_M$, \otimes le produit de Kronecker et \mathbb{I}_M la matrice unitaire de taille M .

On a aussi :

$$\mathbf{H}\mathbf{e} + \mathbf{H}_1\mathbf{d}_1 + \mathbf{H}_2\mathbf{d}_2 = \mathbf{G}\mathbf{u}, \quad \text{avec } (\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2) \in \mathcal{E}_1 \quad \text{et} \quad \mathbf{G} = [\mathbf{H}\mathbf{H}_1\mathbf{H}_2]\mathbf{P}_\alpha. \quad (3.25)$$

La matrice \mathbf{G} est composée de 4 blocs dont chacun est la combinaison linéaire de \mathbf{H} , \mathbf{H}_1 et \mathbf{H}_2 . Elle contient $4M$ colonnes.

En raison de la première ligne de la matrice $\tilde{\mathbf{P}}_\alpha$, on a $\mathbf{1}^t \mathbf{e} = \mathbf{1}^t \mathbf{u}$ ce qui équivaut compte tenu de la positivité de \mathbf{e} et de \mathbf{u} , à l'égalité $\|\mathbf{e}\|_1 = \|\mathbf{u}\|_1$. Donc, nous pouvons remplacer le problème CBP avec des inégalités linéaires (3.14), par cette nouvelle formulation contenant uniquement des contraintes de positivité :

$$\underset{\mathbf{u} \in \mathbb{R}_+^{4M}}{\operatorname{argmin}} \|\mathbf{i} - \mathbf{G}\mathbf{u}\|_2^2 + \lambda \|\mathbf{u}\|_1. \quad (3.26)$$

3.5.2 Algorithme de l'active set CBP

Le problème CBP après reparamétrisation (3.26) a la même forme que celle de BP. Nous pourrions donc exploiter l'algorithme active set bp développé en haut, pour estimer les solutions de CBP. Le retour aux variables basiques \mathbf{e} , \mathbf{d}_1 et \mathbf{d}_2 s'effectue en multipliant la solution \mathbf{u} par \mathbf{P}_α . Enfin la position sous pixel est déduite via les équations (3.21). Le déroulement de l'active set CBP est résumé par l'algorithme 3 décrit ci-dessous :

Algorithm 3 Active set algorithm for Continuous Basis Pursuit

Input: \mathbf{i} , \mathbf{H} , \mathbf{H}_1 , \mathbf{H}_2 , λ

Output: \mathbf{e} , $\mathbf{d}_1, \mathbf{d}_2$

1: $\mathbf{G} \leftarrow [\mathbf{H}\mathbf{H}_1\mathbf{H}_2]\mathbf{P}_\alpha$

2: $\mathbf{u} = \text{bp}(\mathbf{i}, \mathbf{G}, \lambda)$ {Résoudre avec l'algorithme 1 de l'active set BP.}

3: $(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2) \leftarrow \mathbf{P}_\alpha \mathbf{u}$

3.6 Résultats numériques

L'objectif de cette section est d'effectuer une comparaison rapide des différentes méthodes convexes développées ici, afin de voir l'impact de la régularisation l_1 et de l'approximation de Taylor d'ordre 1 sur les résultats de la détection. Les images synthétiques, traitées ici, ont une taille de 32×32 et sont générées suivant le modèle additif (2.1). La densité de particules dans chaque image est contrôlée par le nombre de particules par pixel N_{ppp} . Dans cette simulation, elle est fixée à 0.1. Toutes les particules ont une intensité de 1 et les positions de leurs images sont distribuées

aléatoirement selon une distribution uniforme. L'écart type de la PSF adopté σ_{psf} est égal à 0.6, ce qui correspond à un motif d'image de taille 4×4 pixels. Un bruit gaussien a été ajouté, son écart-type est égal à 5% de l'intensité maximale de l'image d'une particule, ce qui correspond à un $SNR = \frac{\max(i)}{\sigma_{bruit}} = 20$. Les performances de détection sont quantifiées par trois métriques : Precision, Rappel et Fscore. La Precision représente le taux de fausses alarmes et le Rappel correspond au pourcentage de bonnes détections. Enfin, le Fscore est défini comme une moyenne harmonique des deux métriques précédentes et correspond à un certain compromis entre les fausses alarmes et les non détections. Nous exposerons plus en détails les conditions de simulations, ainsi que toutes les métriques employées dans le chapitre 5. Les méthodes testées sont l'active set NNLS, l'active set BP et l'active set CBP. Pour alléger les écritures, on se contentera d'appeler chacune de ces méthodes par le nom du problème de minimisation associé. Un seuillage d'intensité est appliquée sur la solution active set pour réduire l'effet du bruit, suivi d'une étape d'agrégation, par détection de maxima locaux, en vue d'éliminer les détections multiples. Les paramètres de réglages sont le niveau de seuillage, le paramètre de parcimonie λ (pour BP et CBP) et le paramètre régissant la finesse de la grille de discrétisation Δ .

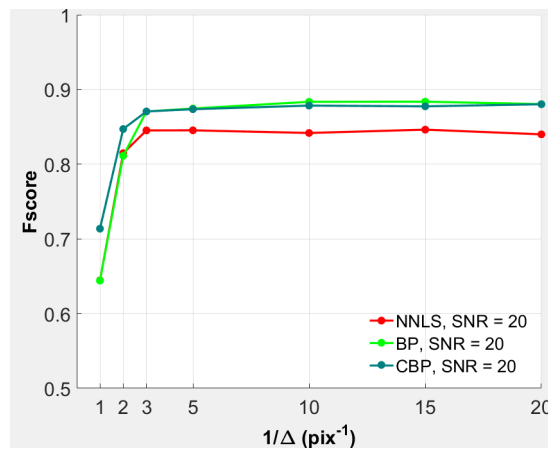


FIGURE 3.9 – Comparaison des méthodes NNLS, BP et CBP en termes de Fscore en fonction du pas de la grille Δ .

La figure 3.9 présente une comparaison du Fscore des 3 méthodes testées en fonction de la largeur de la taille de la grille Δ : NNLS qui constitue l'état de l'art des performances de détection en PIV et nos deux méthodes parcimonieuses qu'on a développées : BP et CBP. Les autres paramètres ont été réglés minutieusement de telle manière à avoir un Fscore maximal. Pour les 3 méthodes on a une saturation de performances à partir de $\Delta = \frac{1}{3}$ pixel. Le gain de détection est quasi nul à partir de cette valeur. Notre réglage est donc plus intéressant que celui suggéré par CHEMINET et collab. [2018], basé sur l'erreur de localisation, qui conduit à des pas de l'ordre de $\frac{1}{20}$. On peut aussi constater le gain en Fscore apporté par la régularisation l_1 dans le cas BP qui est de l'ordre de 4 %. Enfin, l'approche de dictionnaire d'interpolation combinée à une contrainte l_1 , qu'on a appelé CBP, s'est avérée moins pertinente car elle ne montre un intérêt qu'à des pas de grille très grossiers, incompatibles avec les taux de détection standards exigés en PIV. Une analyse plus approfondie de ces résultats en fonction de la N_{ppp} , en terme de Precision, de Rappel et de localisation sera traitée dans le chapitre 5 des simulations numériques.

3.7 Bilan

CHEMINET et collab. [2018] ont proposé dans leur article de reconstruire les centroïdes des images de particules en intégrant uniquement la contrainte de positivité et de résoudre le problème associé en s'appuyant sur l'active set NNLS. Combiné à une procédure d'extraction de maxima locaux, ce dernier a une structure gloutonne et présente un niveau de parcimonie hautement plus élevé que la méthode algébrique SMART. De plus, les performances de détection sont

excellentes, même à haute densité, en absence du bruit. Néanmoins, cette approche est fortement sensible au bruit et requiert une procédure supplémentaire de seuillage, difficilement réglable, en l'occurrence en présence de chevauchements. Pour surmonter ce problème, on a introduit un terme de régularisation l_1 . Ce dernier contrôle le niveau de parcimonie de la solution en améliorant le taux de fausses-détections. On a appelé cette approche BP. Effectivement, on a pu constater que cette dernière réduit sensiblement le nombre de fausses détections en comparaison avec la méthode de référence NNLS. Cela est d'autant plus vrai pour des solutions obtenues après seuillage des intensités faibles. On a pu également choisir un réglage du pas de discrétisation Δ plus favorable en mémoire et en temps de calcul, en gardant la même qualité de détection que celle obtenue par CHEMINET et collab. [2018] dans le cas d'un pas de $\frac{1}{20}$ pixel. Cependant, cette stratégie a le défaut de requérir un post-traitement afin de promouvoir la parcimonie du signal. En outre, la grille exigée est relativement fine. Ainsi on a opté pour la discrétisation de Taylor d'ordre 1 afin d'aboutir à une approximation plus précise du modèle continu même pour des pas de grille plus grossiers. Dans un contexte 2D, cette nouvelle discrétisation fait intervenir 2 dictionnaires dérivés supplémentaires associés à deux variables contraintes linéairement, ce qui introduit une difficulté algorithmique lors de l'inversion, par rapport au cas 1D évoqué par EKANADHAM et collab. [2011b]. Nous avons abordé ce problème dit CBP avec la méthode du gradient proximal FISTA. Cela a été possible après le calcul de l'expression analytique de l'opérateur proximal de la fonction correspondante aux contraintes du problème. Cependant, cet algorithme exige l'utilisation d'un paramètre de tolérance très petit pour garantir sa convergence, ce qui introduit un temps de calcul important, et plus particulièrement quand le pas de la grille utilisé est très fin. Ainsi, nous avons opté pour la méthode d'active set de LAWSON et HANSON [1974], une méthode plus rapide, tout en ajoutant une étape supplémentaire de seuillage pour promouvoir encore plus les résultats de détection. L'application de cette version d'active set a été rendu possible en transformant le problème CBP standard en un problème de minimisation avec uniquement des contraintes de positivité. Les évaluations sur des images de synthèse bruitées et à haute densité nous ont montré que l'approche CBP, à base d'interpolation de Taylor et de contrainte de parcimonie l_1 , conduit aux mêmes performances de détection que celles de BP qui utilise une approximation plus grossière. De plus, l'étape de post-processing est toujours nécessaire même en adoptant une approximation plus précise. Ceci a été contradictoire avec nos attentes, ce qui nous a poussé à remplacer le terme de pénalisation l_1 par une pénalisation l_0 afin de vérifier si cette conclusion est conservée si on change le type de parcimonie imposé. Le chapitre suivant met en exergue cela, et montre qualitativement l'intérêt de l'approche des dictionnaires d'interpolation quand elle est associée à la norme l_0 .

3.8 Références

- BECK, A. et M. TEBoulLE. 2009, «A fast iterative shrinkage-thresholding algorithm for linear inverse problems», *SIAM J. Imaging Sciences*, vol. 2, p. 183–202. [26](#), [31](#)
- BERTSEKAS, D. 2016, *Nonlinear Programming*, Athena scientific optimization and computation series, Athena Scientific, ISBN 9781886529052. [32](#)
- BIUCAS-DIAS, J. M. et M. A. T. FIGUEIREDO. 2007, «A new twist : Two-step iterative shrinkage/thresholding algorithms for image restoration», *IEEE Transactions on Image Processing*, vol. 16, n° 12, p. 2992–3004. [26](#)
- BLUMENSATH, T. et M. DAVIES. 2008, «Iterative thresholding for sparse approximations», *Journal of Fourier Analysis and Applications*, vol. 14, p. 629–654. [26](#)
- BREDIES, K. et D. LORENZ. 2007, «Linear convergence of iterative soft-thresholding», *J. Fourier Anal. Appl.*, vol. 14. [26](#)

- BRUCK, R. E. 1977, «On the weak convergence of an ergodic iteration for the solution of variational inequalities for monotone operators in hilbert space», *Journal of Mathematical Analysis and Applications*, vol. 61, n° 1, p. 159 – 164, ISSN 0022-247X. [26](#)
- BYRNE, C. 2008, *Applied Iterative Methods*, Ak Peters Series, Taylor & Francis, ISBN 9781568813424. [24](#)
- CANDÈS, E. et T. TAO. 2006, «Near-optimal signal recovery from random projections : Universal encoding strategies?», *IEEE Transactions on Information Theory*, vol. 52, p. 5406–5425. [25](#)
- CHEMINET, A., J. KRAWCZYNSKI et P. DRUAULT. 2018, «Particle image reconstruction for particle detection in particle tracking velocimetry», *Measurement Science and Technology*, vol. 29, n° 12, p. 125 202. [24](#), [28](#), [36](#), [37](#)
- CHEN, S. S., D. L. DONOHO et M. A. SAUNDERS. 1998, «Atomic decomposition by basis pursuit», *SIAM Journal on Scientific Computing*, vol. 20, n° 1, p. 33–61. [25](#), [26](#)
- COMBETTES, P. et V. WAJS. 2005, «Signal recovery by proximal forward-backward splitting», *SIAM J. Multiscale Model. Simul.*, vol. 4. [26](#)
- DAUBECHIES, I., M. DEFRISE et C. MOL. 2004, «An iterative thresholding algorithm for linear inverse problems with a sparsity constraints», *Communications on Pure and Applied Mathematics*, vol. 57. [26](#)
- DONOHO, D. 2006, «For most large underdetermined systems of linear equations the minimal l_1 -norm solution is also the sparsest solution», *Comm. Pure Appl. Math*, vol. 59. [25](#)
- DUVAL, V. et G. PEYRÉ. 2017, «Sparse spikes super-resolution on thin grids II : the continuous basis pursuit», *Inverse Problems*, vol. 33, n° 9, p. 095 008. [34](#)
- EFRON, B., T. HASTIE, I. JOHNSTONE et R. TIBSHIRANI. 2004, «Least angle regression” (with discussions)», *The Annals of Statistics*, vol. 32. [27](#)
- EKANADHAM, C., D. TRANCHINA et E. SIMONCELLI. 2011a, «Sparse decomposition of transformation-invariant signals with continuous basis pursuit», p. 4060 – 4063. [29](#)
- EKANADHAM, C., D. TRANCHINA et P. SIMONCELLI. 2011b, «Recovery of sparse translation-invariant signals with continuous basis pursuit», *IEEE Transactions on Signal Processing*, vol. 59, n° 10, p. 4735–4744. [31](#), [37](#)
- FIGUEIREDO, M., R. NOWAK et S. WRIGHT. 2008, «Gradient projection for sparse reconstruction : Application to compressed sensing and other inverse problems», *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, p. 586 – 597. [25](#)
- FIGUEIREDO, M. A. T. et R. D. NOWAK. 2003, «An em algorithm for wavelet-based image restoration», *IEEE Transactions on Image Processing*, vol. 12, n° 8, p. 906–916. [26](#)
- FIGUEIREDO, M. A. T., R. D. NOWAK et S. J. WRIGHT. 2007, «Gradient projection for sparse reconstruction : Application to compressed sensing and other inverse problems», *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, n° 4, p. 586–597. [26](#)
- KIM, S., K. KOH, M. LUSTIG, S. BOYD et D. GORINEVSKY. 2007, «An interior-point method for large-scale ℓ_1 -regularized least squares», *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, n° 4, p. 606–617. [26](#)
- LAWSON, C. L. et R. J. HANSON. 1974, *Solving Least Squares problems*, Prentice-Hall, Saddle River, NJ, USA, p. 149–199. [24](#), [27](#), [31](#), [37](#)

- NGUYEN, T. T., C. SOUSSEN, J. IDIER et E.-H. DJERMOUNE. 2017, «An optimized version of non-negative OMP», dans *XXVIe Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017*, Juan-les-Pins, France. [27](#)
- NOCEDAL, J. et S. WRIGHT. 1999, *Numerical Optimization*, Springer-Verlag, New-York. [27](#)
- OSBORNE, M., B. PRESNELL et B. TURLACH. 2000, «A new approach to variable selection in least squares problems», *IMA Journal of Numerical Analysis*, vol. 20, n° 3, p. 389–403. [27](#)
- PARIKH, N. et S. BOYD. 2014, «Proximal algorithms», *Found. Trends Optim.*, vol. 1, n° 3, p. 127–239. [25](#), [26](#), [31](#)
- PASSTY, G. B. 1979, «Ergodic convergence to a zero of the sum of monotone operators in hilbert space», *Journal of Mathematical Analysis and Applications*, vol. 72, n° 2, p. 383 – 390, ISSN 0022-247X. [26](#)
- TROPP, J. A. et S. J. WRIGHT. 2010, «Computational methods for sparse solution of linear inverse problems», *Proceedings of the IEEE*, vol. 98, n° 6, p. 948–958. [26](#), [27](#)
- WRIGHT, J., A. YANG, A. GANESH, S. SASTRY et L. YU. 2009, «Robust face recognition via sparse representation», *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, p. 210 – 227. [25](#)
- WRIGHT, S. J. 1997, *Primal-Dual Interior-Point Methods*, Society for Industrial and Applied Mathematics, USA, ISBN 089871382X. [26](#)
- ZHANG, Z., Y. XU, J. YANG, X. LI et D. ZHANG. 2015, «A survey of sparse representation : Algorithms and applications», *IEEE Access*, vol. 3, p. 490–530. [26](#)

Chapitre 4

Méthodes de reconstruction gloutonnes à grille

Sommaire

4.1 Dictionnaires d'interpolation et régularisation l_0	42
4.2 Continuous NNOMP (CNNOMP)	43
4.2.1 Aperçu de quelques méthodes gloutonnes usuelles	43
4.2.2 Retour sur le NNOMP	44
4.2.3 L'algorithme CNNOMP	46
4.2.4 Limitations de CNNOMP	47
4.3 KCNNOMP : Amélioration du CNNOMP	47
4.3.1 L'algorithme KCNNOMP	48
4.3.2 Intérêt de l'étape de raffinement	49
4.4 Bilan	50
4.5 Références	51

L'utilisation d'une approximation grossière avec la norme l_1 requiert l'utilisation d'une grille fine, de plus cela conduit à des clusters de détections autour de chaque particule, ce qui nécessite un post-processing minutieux et des réglages supplémentaires. Notre première idée a été d'augmenter l'ordre de l'approximation afin d'améliorer la précision du modèle discrétisé (3.3.1). On a également imposé le même type de pénalisation parcimonieuse pour garder l'aspect convexe du problème. Cependant, cette approche, dite CBP, n'apporte pas de réels avantages par rapport à la version basique BP. En effet, la capacité de détection des 2 approches est quasiment identique même si le dictionnaire est 4 fois plus large dans le cas de l'active set CBP. L'objet de ce chapitre est de vérifier si cette conclusion est aussi valable dans le cas d'une régularisation l_0 .

4.1 Dictionnaires d'interpolation et régularisation l_0 .

L'approximation de Taylor d'ordre 1 mène à l'expression matricielle suivante :

$$\mathbf{i} \simeq \mathbf{H}\mathbf{e} + \mathbf{H}_1\mathbf{d}_1 + \mathbf{H}_2\mathbf{d}_2, \quad (4.1)$$

avec \mathbf{e} , \mathbf{d}_1 et \mathbf{d}_2 des vecteurs contraints dans l'espace convexe \mathcal{E}_1 (3.13). Une approche parcimonieuse alternative consiste à intégrer une pénalisation de type l_0 afin d'augmenter le niveau de parcimonie des solutions recherchées. Cette approche peut être formulée comme suit :

$$\arg \min_{\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2 \in \mathcal{E}_0} \|\mathbf{i} - (\mathbf{H}\mathbf{e} + \mathbf{H}_1\mathbf{d}_1 + \mathbf{H}_2\mathbf{d}_2)\|_2^2 \quad (4.2)$$

$$\mathcal{E}_0 = \{(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2) \in \mathcal{E}_1 \mid \|\mathbf{e}\|_0 \leq K\}. \quad (4.3)$$

avec $\|\mathbf{e}\|_0$ la norme l_0 comptant le nombre de composantes non nulles de \mathbf{e} et K le nombre maximal de détections. Ce dernier paramètre est contrôlé par l'utilisateur. La gestion des contraintes linéaires avec une pénalité l_0 est problématique, car il devient difficile d'appliquer les méthodes gloutonnes non-négatives classiques, ainsi on a suivi la même procédure de reparamétrisation que celle adoptée dans la sous-section (3.5.1) du chapitre précédent, afin de se ramener à un problème de parcimonie avec uniquement des contraintes de positivité :

$$\arg \min_{\mathbf{u} \in \tilde{\mathcal{E}}_0} \|\mathbf{i} - \mathbf{G}\mathbf{u}\|_2^2 \quad (4.4)$$

$$\tilde{\mathcal{E}}_0 = \{\mathbf{u} \in \mathbb{R}_+^{4M}, \|\mathbb{1}_M \mathbb{1}_M \mathbb{1}_M \mathbb{1}_M \mathbf{u}\|_0 \leq K\}, \quad (4.5)$$

La matrice \mathbf{G} est redondante. Elle est composée de 4 blocs. Chaque bloc est la combinaison linéaire des 3 matrices \mathbf{H} , \mathbf{H}_1 et \mathbf{H}_2 construites via l'approximation de Taylor d'ordre 1 (3.25).

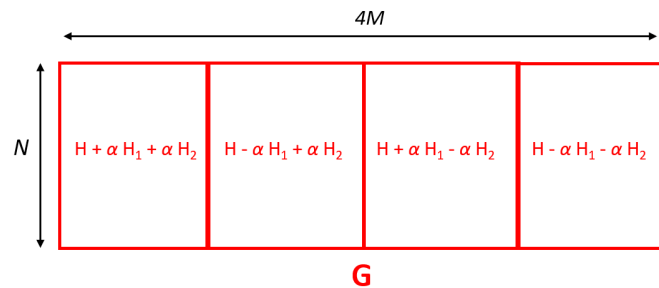


FIGURE 4.1 – La matrice \mathbf{G} est composée de 4 blocs. Chaque bloc est la combinaison linéaire de \mathbf{H} , \mathbf{H}_1 et \mathbf{H}_2 . Elle est de taille $N \times 4M$ avec M le nombre de nœuds de la grille et N le nombre de pixels de l'image acquise. α correspond à la moitié du pas de discrétisation.

Dans ce qui suit, on présentera un récapitulatif succinct des méthodes usuelles pour traiter les problèmes de parcimonie de type l_0 , avec et sans la contrainte de positivité. Nous décrirons plus en détail la méthode OMP non négative (NNOMP) et nous mettrons en exergue son intérêt en

pratique par rapport aux autres méthodes. Enfin, nous l'adapterons à la résolution du problème (4.4) à base de dictionnaires d'interpolation de Taylor.

4.2 Continuous NNOMP (CNNOMP)

4.2.1 Aperçu de quelques méthodes gloutonnes usuelles

Les algorithmes gloutons sont des algorithmes sous-optimaux destinés à traiter des problèmes de parcimonie de type l_0 . Ils sont initialisés par un vecteur nul (solution la plus parcimonieuse correspondante à un support vide) et changent progressivement la cardinalité de leur support à chaque itération en maximisant la descente du critère, d'où la dénomination "glouton". Ces algorithmes sont donc des algorithmes de descente car la norme résiduelle est décroissante à chaque itération. Dans cette section, on distingue deux catégories d'algorithmes gloutons : une première catégorie sans la contrainte de positivité et une seconde intégrant la contrainte de positivité.

Méthodes gloutonnes standards

Les algorithmes gloutons qui augmentent la cardinalité du support un par un à chaque itération comprennent le Matching pursuit (MP) MALLAT et ZHIFENG ZHANG [1993], l'Orthogonal matching pursuit (OMP) PATI et collab. [1993], et l'Orthogonal least squares (OLS) CHEN et collab. [1989]. Le principe du MP MALLAT et ZHIFENG ZHANG [1993] est le suivant : d'abord, on sélectionne l'atome contribuant le plus au résidu. Il s'agit de l'atome réalisant le maximum de corrélation avec ce dernier. Cette contribution est ensuite soustraite du résidu courant pour générer un nouveau résidu et sélectionner l'atome suivant. Il est à noter qu'un atome peut être sélectionné plusieurs fois durant cette procédure. Le critère d'arrêt est typiquement défini soit par le nombre maximal d'itérations, soit par la norme du résidu.

L'algorithme OMP PATI et collab. [1993] est une extension du MP. Sa différence par rapport à ce dernier réside dans la manière de mise à jour des coefficients des atomes sélectionnés. À chaque itération, l'OMP recalcule tous les coefficients en projetant le signal sur l'espace engendré par tous les atomes sélectionnés, contrairement à la méthode MP qui calcule seulement le coefficient du nouveau atome. La mise à jour des coefficients dans la technique OMP se fait via la méthode des moindres carrés non contraints (Unconstrained least squares, ULS). Cette projection orthogonale a l'avantage d'éviter la sélection d'un atome, préalablement choisi, dans les itérations suivantes. A noter qu'une itération MP est très rapide, mais le nombre d'itérations MP pour atteindre une valeur résiduelle donnée pourrait être sensiblement plus important que celui de l'OMP.

L'algorithme OLS CHEN et collab. [1989] met à jour les coefficients de la même manière que l'OMP mais diffère dans sa manière de sélectionner les atomes. En effet, l'OMP sélectionne l'atome ayant le produit scalaire le plus élevé avec le résidu courant (en valeur absolue) alors que l'OLS considère le produit scalaire maximal (en valeur absolue) entre la projection de l'atome sur le complément orthogonal du sous-espace sélectionné et le résidu courant. Les implémentations efficaces de l'OMP ou de l'OLS reposent sur l'idée que la solution du problème de moindres carrés peut être mise à jour rapidement en s'appuyant par exemple sur une factorisation matricielle de type Gram-Schmidt ou de Cholesky STURM et CHRISTENSEN [2012].

Au lieu de modifier la taille du support un par un, certains algorithmes produisent un support de cardinalité fixe à chaque itération, par exemple : CoSaMP NEEDELL et TROPP [2009], Subspace pursuit (SP) DAI et MILENKOVIC [2008] et Iterative hard thresholding (IHT) BLUMENSATH et DAVIES [2010]. CoSaMP et SP sont très similaires : à chaque itération de CoSaMP, 2K atomes sont sélectionnés pour augmenter le support puis les coefficients sont mis à jour en effectuant une projection orthogonale. Ces coefficients sont ensuite ordonnés en valeurs absolues afin de garder seulement les K atomes associés aux coefficients les plus élevés. De même, SP sélectionne K atomes pour augmenter le support et met à jour les coefficients en appliquant une projection orthogonale suivie d'un seuillage et d'une seconde projection orthogonale. Quant à l'IHT, le principe est de répéter un mouvement de descente suivi d'un seuillage BLUMENSATH et DAVIES [2008].

Par conséquent, l'IHT est très rapide à chaque itération, mais peut nécessiter un grand nombre d'itérations pour converger.

Méthodes gloutonnes non négatives

Les algorithmes gloutons destinés à la reconstruction de signaux positifs possèdent une structure proche des algorithmes gloutons standards. Certains algorithmes gloutons tels que MP ou IHT peuvent être adaptés facilement en algorithmes non négatifs en raison de la simplicité de leurs itérations. Cependant, ce n'est pas le cas des algorithmes gloutons orthogonaux tels que l'OMP et l'OLS. La prise en compte de la non-négativité au sein de ces deux algorithmes soulève des difficultés puisque la structure algorithmique est profondément modifiée.

La première extension non négative de l'OMP a été introduite par BRUCKSTEIN et collab. [2008]. Elle ne prend en compte que les produits scalaires positifs dans l'étape de sélection d'atome. En outre, la mise à jour des coefficients est effectuée en résolvant des sous problèmes de moindres carrés non négatifs (NNLS) au lieu de sous-problèmes ULS. À noter que NNLS n'a pas de solution analytique explicite contrairement à un problème ULS. Il faut donc privilégier une implémentation plus rapide de préférence récursive afin d'accélérer la résolution à chaque itération BRUCKSTEIN et collab. [2008], YAGHOABI et collab. [2015] et YAGHOABI et DAVIES [2015]. Une implémentation rapide du NNOMP a été proposée par YAGHOABI et collab. [2015]. Ce nouveau algorithme appelé Fast NNOMP (FNNOMP) combine les étapes de sélection d'atome et de mise à jour de coefficients en une seule étape de telle sorte que le vecteur résiduel soit orthogonal au sous-espace sélectionné à chaque itération. Ainsi, FNNOMP n'exige pas la résolution de sous-problèmes NNLS à chaque itération mais cela risque de donner des solutions différentes de celles du NNOMP.

L'extension non négative de l'OLS (NNOLS) a été introduite pour la première fois par YAGHOABI et DAVIES [2015]. Cette dernière requiert également la résolution de plusieurs sous problèmes NNLS. En effet, NNOLS nécessite de résoudre autant de sous-problèmes NNLS que de nombre d'atomes non sélectionnés à chaque itération. C'est un algorithme très coûteux. Une variante sous optimale a été proposée par les mêmes auteurs YAGHOABI et DAVIES [2015]. Elle est appelée Fast NNOLS (FNNOLS). Elle a l'avantage d'accélérer la reconstruction. Toutefois, comme dans le cas FNNOMP, elle risque de produire des solutions non négatives NGUYEN et collab. [2017].

Des extensions non négatives de CoSaMP, SP, et HTP ont été aussi introduites dans la littérature des méthodes gloutonnes par KIM et HALDAR [2016]. A l'instar de la méthode NNOMP de BRUCKSTEIN et collab. [2008], ces extensions (appelées respectivement NNCoSAMP, NNSP et NNHTP) ne prennent en compte que les produits scalaires positifs dans la sélection d'atomes et remplacent la projection orthogonale par la résolution de sous-problèmes NNLS. Il convient de noter que les extensions non négatives NNCoSAMP, NNSP et NNHTP n'arrivent pas à produire exactement une solution K-parcimonieuse à chaque itération, contrairement aux versions standards. Cela s'explique par le fait que les solutions NNLS diffèrent des solutions ULS. C'est l'un des inconvénients de l'utilisation de ces algorithmes en pratique. En outre ces derniers sont difficilement implémentables de manière récursive.

4.2.2 Retour sur le NNOMP

Dans ce qui suit, on s'intéressera uniquement à l'extension non négative de l'OMP (NNOMP) comme défini par BRUCKSTEIN et collab. [2008] pour résoudre (4.6) :

$$\operatorname{argmin} \|\mathbf{i} - \mathbf{H}\mathbf{e}\|_2^2 \quad \text{t.q.} \quad \mathbf{e} \in \mathbb{R}_+^M \quad \text{et} \quad \|\mathbf{e}\|_0 \leq K \quad (4.6)$$

Ce choix est motivé par plusieurs raisons relatives à sa mise en oeuvre. En effet, NNOMP a une structure algorithmique simple. Il est moins coûteux en terme de complexité algorithmique par rapport au NNOLS YAGHOABI et DAVIES [2015], et admet une implémentation récursive conduisant à des solutions exactes NGUYEN et collab. [2017] et NGUYEN et collab. [2019]. En outre, on possédait aussi les briques de codes de NGUYEN et collab. [2017], développées au sein du même

projet ANR, ce qui nous a encouragé à nous appuyer sur cette base algorithmique pour développer le nôtre.

À l'instar des 2 méthodes convexes NNLS et BP, le dictionnaire H a été construit via une discrétisation d'ordre 0. L'objectif est de trouver un vecteur K -parcimonieux non-négatif donnant la meilleure approximation au sens des moindres carrés de \mathbf{i} par $H\mathbf{e}$.

On garde les mêmes notations que celles employées dans le chapitre précédent : \mathbf{r} désigne le résidu, S le support de \mathbf{e} , \bar{S} le complément de S , H_S le sous-dictionnaire de H indexé par l'ensemble S et \mathbf{h}_j un vecteur colonne (atome) du dictionnaire H . Afin d'alléger les notations, nous appellerons le dictionnaire H par le même nom après normalisation de ses colonnes. Comme dans le cas BP, les variables d'entrée sont : l'image acquise \mathbf{i} , le dictionnaire H construit via une discrétisation d'ordre 0, et K un facteur géré par l'utilisateur permettant de modifier le niveau de parcimonie de la solution recherchée. Les variables de sortie sont les intensités des détections réunies dans le vecteur \mathbf{e} .

NNOMP a quasiment la même la structure que l'OMP [NGUYEN et collab. \[2017\]](#). Ce dernier commence d'un support vide $S = \emptyset$; à chaque itération, il sélectionne l'atome ayant la plus forte corrélation avec le résidu courant; il ajoute l'indice de cet atome au support S et calcule ensuite la projection de l'observation \mathbf{i} sur sur le cône induit par les colonnes sélectionnées. En raison de la contrainte de non-négativité, NNOMP présente deux différences essentielles avec OMP :

1. La règle de sélection implique un produit scalaire sans valeur absolue.
2. La projection orthogonale calculée à chaque itération dans l'OMP est remplacée par un sous-problème NNLS dont la solution analytique n'est pas explicite et qui nécessite souvent un processus de résolution itératif pour accélérer la reconstruction [NGUYEN et collab. \[2017\]](#).

Algorithm 4 NNOMP

Input: \mathbf{i} , H , K

Output: $\mathbf{e} = \text{NNOMP}(\mathbf{i}, H, K)$

```

1:  $S \leftarrow \emptyset$ ,  $\mathbf{r} \leftarrow \mathbf{i}$ 
2:  $k \leftarrow 0$ 
3: while  $k \leq K$  &  $\max_{j \in \bar{S}} \langle \mathbf{h}_j, \mathbf{r} \rangle \geq 0$  do
4:    $l_1 \leftarrow \arg \max_{j \in \bar{S}} \langle \mathbf{h}_j, \mathbf{r} \rangle$ 
5:    $S \leftarrow S \cup \{l_1\}$ 
6:    $\mathbf{z} \leftarrow \arg \min_{\mathbf{z} \geq 0} \|\mathbf{i} - H_S \mathbf{z}\|_2^2$ 
7:    $\mathbf{r} \leftarrow \mathbf{i} - H_S \mathbf{z}$ 
8:    $k \leftarrow k + 1$ 
9: end while
10:  $\mathbf{e}_S \leftarrow \mathbf{z}$ ,  $\mathbf{e}_{\bar{S}} \leftarrow 0$ 
    
```

L'estimation des positions des détections peuvent être déduites systématiquement à partir des indices des composantes non nulles du vecteur solution \mathbf{e} . En effet, chaque composante non nulle e_m de \mathbf{e} correspond à une détection d'une particule indexée p dont la position sous-pixel est estimée comme suit :

$$\mathbf{x}_p = \mathbf{m}\Delta \quad (4.7)$$

avec \mathbf{m} les coordonnées 2D du nœud de la grille correspondant à l'indice m . Afin d'augmenter le niveau de parcimonie et par conséquent la qualité de reconstruction, [NGUYEN et collab. \[2017\]](#) ont intégré une opération additionnelle de rétrécissement de support. Cette opération est ajoutée juste après l'étape 6 de l'algorithme 4 et consiste à supprimer du support courant S tous les atomes n'appartenant pas au support de la solution NNLS. On a décidé de l'omettre dans l'algorithme 4 en raison de son inefficacité dans l'algorithme CNNOMP.

4.2.3 L'algorithme CNNOMP

Afin de résoudre le problème inverse (4.4), issu de la combinaison des dictionnaires d'interpolation de Taylor avec la contrainte de pénalisation l_0 , nous avons opté pour la même stratégie algorithmique que celle développée par NGUYEN et collab. [2017] (NNOMP) mais en effectuant quelques légères modifications. Dans ce qui suit, nous appellerons cette nouvelle version CNNOMP.

Comme dans le cas de l'active set CBP, les variables d'entrées sont l'observation \mathbf{i} , la matrice redondante G construite à partir de H , et des matrices dérivées H_1 et H_2 , et le paramètre de parcimonie K . CNNOMP se différencie légèrement de la version basique NNOMP dans la façon dont les atomes sont ajoutés au dictionnaire. Une image de particule, dans un modèle discrétisé via l'approximation de Taylor d'ordre 1, est décrite par 4 atomes liés linéairement, de la matrice G . L'ajout d'atomes à chaque itération se fait donc par un bloc de 4 : le premier atome est choisi comme décrit par NGUYEN et collab. [2017] et les 3 autres sont les atomes qui lui sont liés (voir figure 4.2). Contrairement au schéma itératif proposé par NGUYEN et collab. [2017], on a privilégié une implémentation de LAWSON et HANSON [1974], pour résoudre les sous problèmes NNLS à chaque itération, en raison de sa stabilité numérique. Le passage de la variable \mathbf{u} aux variables $(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2)$ et puis à la position sous-pixel \mathbf{x} se fait en suivant les équations (3.21) et (3.24).

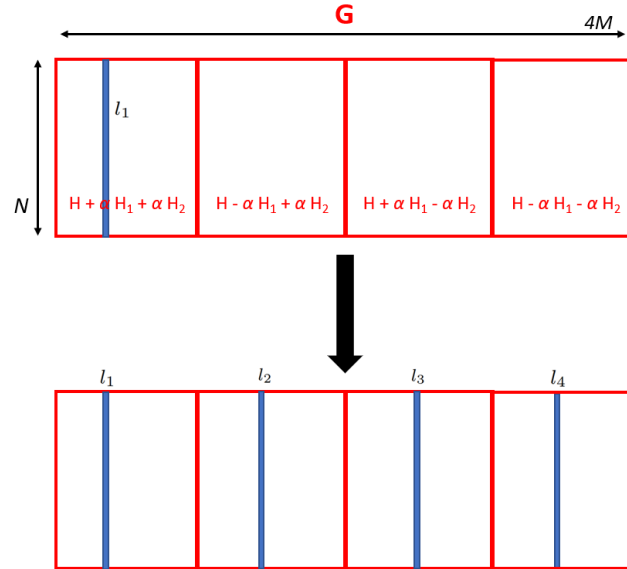


FIGURE 4.2 – Après avoir choisi l'atome \mathbf{g} de la matrice G , indexé l_1 , réalisant le maximum de corrélation avec le résidu, on ajoute les 3 autres colonnes : l_2 , l_3 et l_4 .

Algorithm 5 CNNOMP

Input: \mathbf{i}, G, K

Output: $(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2) = \text{CNNOMP}(\mathbf{i}, G, K)$

- 1: $S \leftarrow \emptyset, \mathbf{r} \leftarrow \mathbf{i}$
 - 2: $k \leftarrow 0$
 - 3: **while** $k \leq K$ & $\max_{j \in \bar{S}} \langle \mathbf{g}_j, \mathbf{r} \rangle \geq 0$ **do**
 - 4: $l_1 \leftarrow \underset{j \in \bar{S}}{\text{argmax}} \langle \mathbf{g}_j, \mathbf{r} \rangle$
 - 5: $S \leftarrow S \cup \{l_1, l_2, l_3, l_4\}$
 - 6: $\mathbf{z} \leftarrow \underset{\mathbf{z} \geq 0}{\text{argmin}} \|\mathbf{i} - G_S \mathbf{z}\|_2^2$
 - 7: $\mathbf{r} \leftarrow \mathbf{i} - G_S \mathbf{z}$
 - 8: $k \leftarrow k + 1$
 - 9: **end while**
 - 10: $\mathbf{u}_S \leftarrow \mathbf{z}, \mathbf{u}_{\bar{S}} \leftarrow 0$
 - 11: $(\mathbf{e}, \mathbf{d}_1, \mathbf{d}_2) \leftarrow P_\alpha \mathbf{u}$
-

Comme indiqué à la fin de la sous section précédente, l'opération de rétrécissement de support a été également testée dans ce contexte mais s'est avérée sans avantage notable sur la qualité des résultats. Ainsi, on a pas intégré cette étape dans l'algorithme 5.

4.2.4 Limitations de CNNOMP

Une illustration qualitative des premières itérations du CNNOMP est présentée dans la figure 4.3. Elle modélise la superposition de deux particules de même intensité et de même largeur de PSF ($\sigma_{psf} = 0.6$). Aucun bruit n'a été ajouté. Les positions réelles sont indiquées par des ronds bleus, tandis que les détections estimées sont désignées par des croix rouges. L'objectif de cette simulation est de visualiser le résultat issu de chaque itération afin d'appréhender le déroulement de l'algorithme CNNOMP et expliquer le comportement pathologique constaté lors de la résolutions des chevauchements. Chaque sous-figure $j \geq 2$ représente les détections obtenues à

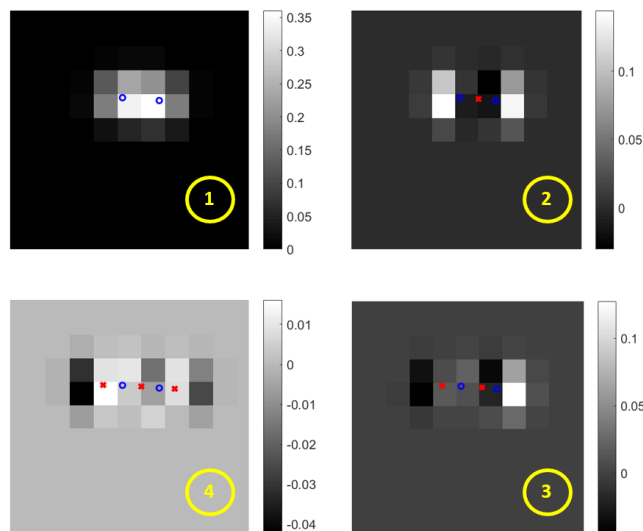


FIGURE 4.3 – Les 3 premières itérations de CNNOMP pour séparer deux particules de même intensité, qui se chevauchent et qui ont la même largeur de PSF. Les ronds bleus (o) correspondent aux particules réelles et les croix rouges (x) indiquent les détections estimées. Le pas de la grille Δ est égal à 1 pixel.

l'itération $j - 1$, ainsi que les images résiduelles r associées, qui diminuent en amplitude à chaque itération j . À la première itération, on voit une première détection au milieu des 2 particules. A l'issue des 2 itérations suivantes, on observe 2 autres détections aux centres des pixels, les plus brillants de l'image résiduelle. L'intensité maximale en valeur absolue de l'image résiduelle reste légèrement élevée. Elle est d'environ 3 pour cent par rapport à l'intensité maximale de l'image de la particule, ce qui peut donner lieu à d'autres détections, si on ne force pas l'arrêt du programme. On a également constaté, dans d'autres simulations qu'on a pas développé ici, que ce défaut demeure persistant quel que soit le pas de discrétisation utilisé, contrairement aux méthodes convexes : NNLS, BP et CBP qui ont des performances de détection qui s'améliorent quand le pas de discrétisation diminue.

4.3 KCNNOMP : Amélioration du CNNOMP

Comme montré plus haut, CNNOMP n'arrive pas à résoudre le chevauchement de particules, même via des grilles de discrétisation très fines. L'analyse précédente montre que la superposition des images de particules conduit systématiquement à une surdétection. C'est un défaut d'algorithme glouton purement incrémental. Afin de contourner ce problème, on s'est appuyé sur les travaux de [KNUDSON et collab. \[2014\]](#) sur le continuous OMP appliqués dans le cas de signaux électrophysiologies mono-dimensionnels. L'idée centrale de cette approche consiste à intégrer dans la boucle de CNNOMP une étape supplémentaire de minimisation sur les intensités et sur les

positions de particules afin de sortir de la grille de discrétisation figée et améliorer les estimations induites par l'étape de NNLS. C'est le KCNNOMP. Le déroulement de cet algorithme sera exposé en détail dans la sous section suivante.

4.3.1 L'algorithme KCNNOMP

Une modification de CNNOMP consiste à déplacer continuellement dans l'espace, à chaque itération, les positions des détections estimées par l'étape de NNLS et par conséquent affiner la valeur du résidu sur lequel on s'appuie pour le choix d'atomes. Les variables d'entrées sont l'image \mathbf{i} , le dictionnaire redondant de Taylor G , le facteur de parcimonie K et la fonction PSF h .

Les variables de sorties sont les mêmes que celles de CNNOMP. Les 3 premières lignes de la boucle principale demeurent inchangées et correspondent aux 3 étapes de base de CNNOMP :

1. Calcul du maximum de corrélation pour la sélection d'atomes
2. L'augmentation du support par un groupe de 4 atomes.
3. La minimisation d'un problème de moindres carrés sous contrainte de positivité sur le support courant S .

L'opérateur T a été introduit pour alléger les écritures. Il résume tout simplement les équations (3.21) et (3.24) permettant le passage aux variables de base \mathbf{e}_0 et son vecteur position associé \mathbf{x}_0 . Le résultat obtenu est utilisé pour initialiser un problème d'optimisation contraint non convexe afin d'améliorer les premières estimations. Ce dernier, indiqué ici par la fonction `locref` minimise l'erreur quadratique moyenne entre les images de particules observées et le modèle génératif (2.1), en cherchant l'emplacement et l'intensité optimale des particules (4.8) :

$$\min_{\mathbf{e} \geq 0, \mathbf{x}} \sum_{n=1}^N \left(\mathbf{i}(n) - \sum_{j=1}^k h(n - \mathbf{x}_j) \mathbf{e}_j \right)^2 \quad (4.8)$$

Ceci permet de diminuer le résidu et empêche les erreurs de modèle de générer des détections parasites. Cette idée est également présente dans l'algorithme IPR de WIENEKE [2013], mais ici toutes les détections ont la possibilité se déplacer simultanément.

Pour minimiser un critère très similaire à (4.8) des auteurs ont proposé BFGS DENOYELLE [2018] ou LBFGS CORNIC et collab. [2020]. Nous avons préféré une méthode de trust-region implémentée par la fonction Matlab `fmincon`, initialisée par les premières estimations de localisation et d'intensités de particules détectées (voir l'algorithme 7).

L'expression analytique du gradient de la fonction coût est fourni à `fmincon` afin d'accélérer sa convergence et optimiser son temps de fonctionnement. Les positions recherchées sont contraintes de rester à l'intérieur du domaine défini par l'image traitée et les intensités doivent être non négatives.

Algorithm 6 KCNNOMP

Input: \mathbf{i} , G , K , h

Output: \mathbf{e}, \mathbf{x}

- 1: $S \leftarrow \emptyset, \mathbf{r} \leftarrow \mathbf{i}$
 - 2: $k \leftarrow 0$
 - 3: **while** $k \leq K$ & $\max_{j \in S} \langle \mathbf{g}_j, \mathbf{r} \rangle \geq 0$ **do**
 - 4: $l_1 \leftarrow \operatorname{argmax}_{j \in \bar{S}} \langle \mathbf{g}_j, \mathbf{r} \rangle$
 - 5: $S \leftarrow S \cup \{l_1, l_2, l_3, l_4\}$
 - 6: $\mathbf{z}_0 \leftarrow \operatorname{argmin}_{\mathbf{z} \geq 0} \|\mathbf{i} - G_S \mathbf{z}\|_2^2$
 - 7: $(\mathbf{e}_0, \mathbf{x}_0) \leftarrow T(\mathbf{z}_0)$
 - 8: $(\mathbf{e}, \mathbf{x}) \leftarrow \operatorname{locref}(\mathbf{i}, h, \mathbf{e}_0, \mathbf{x}_0)$
 - 9: $r(n) \leftarrow i(n) - \sum_{j=1}^k h(n - \mathbf{x}_j) e_j \quad n \in [1, N]$
 - 10: $k \leftarrow k + 1$
 - 11: **end while**
-

Algorithm 7 Affinement de position

Input: i, h, e_0, x_0
Output: $(e, x) = \text{locref}(i, h, e_0, x_0)$

 1: $k \leftarrow \text{length}(e_0)$

 2: $\text{costfun}(e, x) = \sum_{n=1}^N (i(n) - \sum_{j=1}^k h(n - x_j) e_j)^2$

 3: $(e, x) \leftarrow \text{fmincon}(@\text{costfun}, e_0, x_0, 'trust - region')$
4.3.2 Intérêt de l'étape de raffinement

On garde les mêmes conditions de simulation que celles de la figure 4.3. Aucun bruit n'est ajouté. Les ronds bleus (o) correspondent aux centroïdes des images de particules et les croix rouges (x) correspondent aux positions des détections. La variation du niveau de grille du background dans chaque sous-figure reflète la variation d'intensité de l'image résiduelle r à chaque itération j . Les sous-figures 2 et 3 correspondent à la première itération et les sous-figures 4 et 5 correspondent à la deuxième itération. Lors de la première itération, la minimisation sur les positions et les intensités n'affecte pas l'image résiduelle, et la position de la détection demeure inchangée. Cependant à l'itération 2, les positions estimées sont nettement corrigées et la valeur du résidu descend drastiquement de 0.15 à 10^{-7} . Noter que la reconstruction sans correction des positions initialement estimées, aurait introduit une troisième détection à droite où il y avait un pic d'intensité dans la carte résiduelle (voir la sous-figure 4 de la figure 4.3).

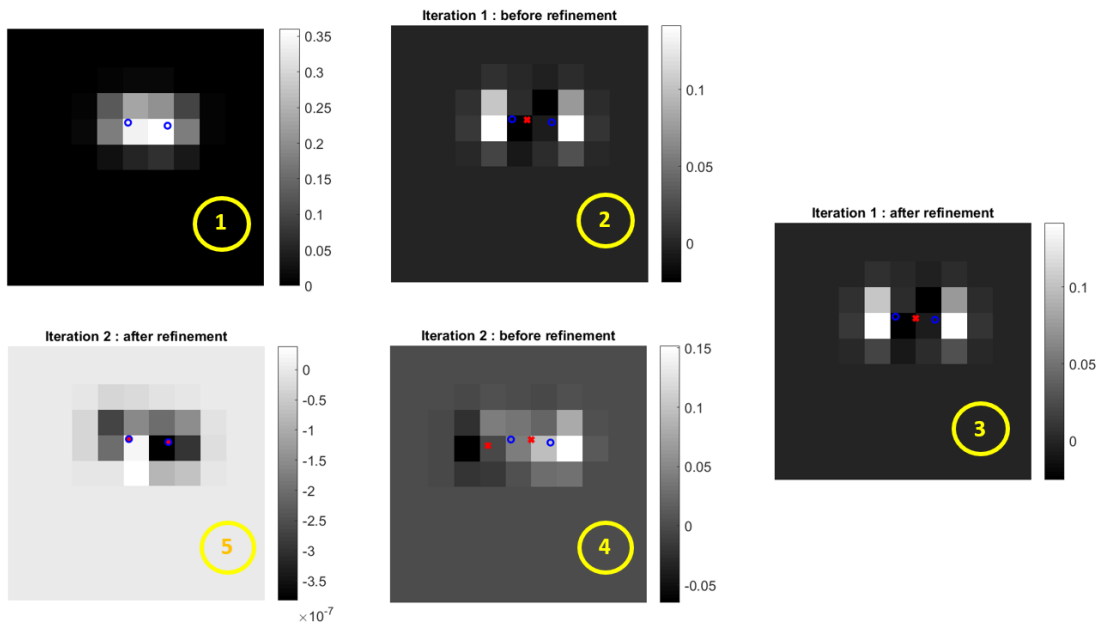


FIGURE 4.4 – Les deux sous-figures 2 et 3 correspondent à la première itération avant et après l'étape de minimisation sur les positions et les intensités et les deux sous-figures 4 et 5 correspondent à la deuxième itération avant et après minimisation sur les positions et les intensités. KCNNOMP parvient à séparer les deux particules qui se chevauchent. En effet, l'étape de minimisation non convexe corrige les positions estimées et réduit drastiquement la valeur absolue maximale de l'image résiduelle de 0,15 à 10^{-7} .

La figure 4.5 illustre bien le succès de KCNNOMP dans la résolution des chevauchements en absence du bruit. Elle correspond à une image de synthèse de taille 32×32 avec une densité de particule N_{ppp} élevée de l'ordre de 0.1 particules par pixel et un pas de discrétisation de 1 pixel. Les particules sont toujours indiquées par des ronds bleus et les détections par des croix rouges. Effectivement, toutes les particules sont quasiment détectées et très bien localisées par KCNNOMP contrairement à la version basique CNNOMP qui échoue chaque fois qu'il y a des interactions entre particules (voir la zone encadrée en jaune).

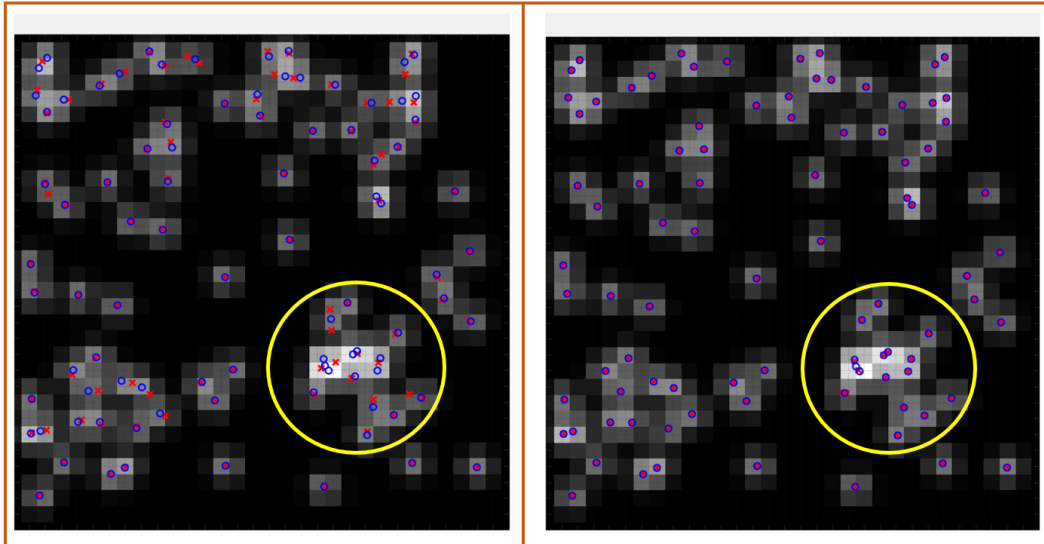


FIGURE 4.5 – Comparaison de résultats de reconstruction entre CNNOMP (à gauche) et KCNNOMP (à droite) sur une image de taille 32×32 en l’absence du bruit. Les ronds bleus correspondent aux vraies particules et les croix rouges correspondent aux détections observées. $\Delta = 1$ pixel et $N_{ppp} = 0.1$. Le cercle jaune entoure une zone à forte interactions de particules. L’algorithme KCNNOMP arrive à séparer les particules et à estimer correctement les positions contrairement à CNNOMP qui fonctionne seulement quand les recouvrements sont très partiels.

4.4 Bilan

Comme évoqué dans le chapitre précédent, l’adoption d’une grille fine avec une discrétisation d’ordre 0, comme dans le cas NNLS et BP, soulève plusieurs problèmes notamment :

1. La grande taille des dictionnaires mis en jeu.
2. La nécessité d’un post processing afin d’éliminer les détections multiples dues aux erreurs du modèle discrétisé et les détections parasites dûes au bruit.

Notre première approche a été d’augmenter l’ordre de l’approximation tout en imposant la non-négativité et la parcimonie comme à priori. L’approche CBP combinant une discrétisation d’ordre 1 et une parcimonie de type l_1 s’est avérée moins prometteuse puisqu’on a eu les mêmes défauts que les méthodes d’ordre 0. Ainsi, on a opté pour une régularisation de type l_0 pour vérifier que cela est dû simplement à la nature peu parcimonieuse de la norme l_1 et que l’apport de la nouvelle approximation du modèle serait significatif en changeant le type de parcimonie imposée. Pour ce fait, on s’est appuyé sur l’approche NNOMP de [NGUYEN et collab. \[2017\]](#), après re-paramétrisation du problème de minimisation initial, afin de développer notre algorithme CNNOMP. Ce dernier se différencie légèrement de la version basique NNOMP dans la façon dont les atomes sont ajoutés au dictionnaire. Dans notre cas, une image de particule est décrite par 4 atomes liés linéairement. L’ajout d’atomes à chaque itération se fait donc par un bloc de 4. Cette stratégie fonctionne bien avec des pas plus larges de l’ordre de 1 pixel mais échoue en présence de fortes interactions de particules dans l’image traitée.

Afin de mieux gérer ce problème, on s’est inspiré des travaux de [KNUDSON et collab. \[2014\]](#) dans le cas 1D, en incluant dans la boucle de CNNOMP une étape supplémentaire de minimisation sur les intensités et les positions des particules, juste après l’étape d’inversion NNLS. C’est le KCNNOMP. Noter que [KNUDSON et collab. \[2014\]](#) qualifient cette étape d’optionnelle et ne disent pas si les résultats qu’ils montrent sont obtenus ou non avec cette option.

Les premières évaluations sur des images non bruitées, à hautes densités, ont montré la pertinence de cette approche. En effet, l’étape de correction non convexe permet de chercher des positions plus adaptées aux observations en dehors de la grille, ce qui conduit à des résidus très précis et par conséquent un bon choix d’atomes dans les itérations suivantes.

Le chapitre suivant sera un chapitre de simulations et de commentaires de résultats, de toutes les méthodes qu'on a développés jusque là. On présentera plus en détail les conditions de simulation, le réglage de ces méthodes et leurs performances de détection et de localisation en fonction de la densité de particule, du rapport signal sur bruit et du temps de calcul.

4.5 Références

- BLUMENSATH, T. et M. DAVIES. 2008, «Iterative thresholding for sparse approximations», *Journal of Fourier Analysis and Applications*, vol. 14, p. 629–654. [43](#)
- BLUMENSATH, T. et M. E. DAVIES. 2010, «Normalized iterative hard thresholding : Guaranteed stability and performance», *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, n° 2, p. 298–309. [43](#)
- BRUCKSTEIN, A., M. ELAD et M. ZIBULEVSKY. 2008, «On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations», *Information Theory, IEEE Transactions on*, vol. 54, p. 4813 – 4820. [44](#)
- CHEN, S., S. A. BILLINGS et W. LUO. 1989, «Orthogonal least squares methods and their application to non-linear system identification», *International Journal of Control*, vol. 50, n° 5, p. 1873–1896. [43](#)
- CORNIC, P., B. LECLAIRE, F. CHAMPAGNAT, G. LE BESNERAIS, A. CHEMINET, C. ILLOUL et G. LOSFELD. 2020, «Double-frame tomographic PTV at high seeding densities», *Experiments in Fluids*, vol. 61, n° 2. [48](#)
- DAI, W. et O. MILENKOVIC. 2008, «Subspace pursuit for compressive sensing : Closing the gap between performance and complexity», *CoRR*, vol. abs/0803.0811. [43](#)
- DENOYELLE, Q. 2018, «Theoretical and Numerical Analysis of Super-Resolution Without Grid», , n° 2018PSLED030. [48](#)
- KIM, D. et J. HALDAR. 2016, «Greedy algorithms for nonnegativity-constrained simultaneous sparse recovery», *Signal Processing*, vol. 125. [44](#)
- KNUDSON, K. C., J. YATES, A. HUK et J. W. PILLOW. 2014, «Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit», dans *Advances in Neural Information Processing Systems 27*, édité par Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence et K. Q. Weinberger, Curran Associates, Inc., p. 1215–1223. [47](#), [50](#)
- LAWSON, C. L. et R. J. HANSON. 1974, *Solving Least Squares problems*, Prentice-Hall, Saddle River, NJ, USA, p. 149–199. [46](#)
- MALLAT, S. G. et ZHIFENG ZHANG. 1993, «Matching pursuits with time-frequency dictionaries», *IEEE Transactions on Signal Processing*, vol. 41, n° 12, p. 3397–3415. [43](#)
- NEEDELL, D. et J. TROPP. 2009, «Cosamp : Iterative signal recovery from incomplete and inaccurate samples», *Applied and Computational Harmonic Analysis*, vol. 26, n° 3, p. 301 – 321, ISSN 1063-5203. [43](#)
- NGUYEN, T. T., J. IDIER, C. SOUSSEN et E.-H. DJERMOUNE. 2019, «Non-negative orthogonal greedy algorithms», *IEEE Transactions on Signal Processing*, vol. 67, n° 21, p. 5643–5658. [44](#)
- NGUYEN, T. T., C. SOUSSEN, J. IDIER et E.-H. DJERMOUNE. 2017, «An optimized version of non-negative OMP», dans *XXVIe Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017*, Juan-les-Pins, France. [44](#), [45](#), [46](#), [50](#)

- PATI, Y. C., R. REZAIIFAR et P. S. KRISHNAPRASAD. 1993, «Orthogonal matching pursuit : recursive function approximation with applications to wavelet decomposition», dans *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, p. 40–44 vol.1. [43](#)
- STURM, B. et M. CHRISTENSEN. 2012, «Comparison of orthogonal matching pursuit implementations», p. 220–224. [43](#)
- WIENEKE, B. 2013, «Iterative reconstruction of volumetric particle distribution», *Measurement Science and Technology*, vol. 24, n° 2, p. 024 008. [48](#)
- YAGHOobi, M. et M. E. DAVIES. 2015, «Fast non-negative orthogonal least squares», dans *2015 23rd European Signal Processing Conference (EUSIPCO)*, p. 479–483. [44](#)
- YAGHOobi, M., D. WU et M. E. DAVIES. 2015, «Fast non-negative orthogonal matching pursuit», *IEEE Signal Processing Letters*, vol. 22, n° 9, p. 1229–1233. [44](#)

Chapitre 5

Résultats numériques

Sommaire

5.1 Conditions de simulation et métriques	54
5.1.1 Conditions de simulations numériques	54
5.1.2 Métriques de détection et de localisation	54
5.1.3 Intérêt du Fscore en PTV	56
5.1.4 Méthodes et paramètres de réglage	56
5.2 Méthodes convexes	57
5.2.1 Réglage de paramètres des méthodes convexes	57
5.2.2 Comparaison des méthodes convexes	58
5.3 Méthodes gloutonnes	59
5.3.1 Réglage de paramètres des méthodes gloutonnes	59
5.3.2 Comparaison des méthodes gloutonnes	60
5.4 Méthodes gloutonnes contre méthodes convexes	61
5.4.1 Comparaison de performances de détection	61
5.4.2 Comparaison de temps de calcul	62
5.5 Bilan	63
5.6 Références	63

Une série de tests de simulation a été conçue pour régler minutieusement les paramètres des méthodes proposées dans ce manuscrit. Ce réglage permet de comparer fidèlement ces méthodes par rapport à une variété de paramètres caractéristiques (SNR, N_{ppp} , Δ ...) et de sélectionner in fine la meilleure approche. Ces simulations permettent aussi de souligner l'impact du type de discrétisation et du type de contrainte de parcimonie imposés en terme de performance de détection et de temps de calcul. Ce chapitre est organisé comme suit : la première section décrit la procédure de génération d'images synthétiques et les métriques de performances de détection et de localisation adoptées. La deuxième section expose la stratégie de réglage utilisée pour comparer les méthodes convexes et souligne son intérêt dans le contexte PTV. Elle montre aussi l'inefficacité de la discrétisation de Taylor d'ordre 1 associée à une pénalisation l_1 (CBP) pour promouvoir les solutions de BP. La troisième section présente la même procédure de réglage appliquée à des méthodes gloutonnes non négatives qu'on a développées ici, afin de les comparer fidèlement. Elle met en exergue l'importance de l'interpolation de Taylor d'ordre 1 quand elle est associée à une contrainte de parcimonie l_0 , et plus particulièrement quand la résolution du problème induit se fait via KCNNOMP. Enfin, la dernière section présente une comparaison détaillée entre les meilleures méthodes de chaque catégorie : BP et KCNNOMP, en fonction de la densité de particules (N_{ppp}), du rapport signal sur bruit (SNR) et du temps de calcul.

5.1 Conditions de simulation et métriques

5.1.1 Conditions de simulations numériques

Les images PIV synthétiques traitées ici sont générées via le modèle additif décrit par l'équation (2.1). Toutes les particules ont une intensité de 1 et les projections de leurs images sont distribuées aléatoirement selon une distribution uniforme. La densité de particule par image est un facteur limitant majeur de la fiabilité de la PTV car une densité élevée empêche une identification et une localisation précise de particules. Ce paramètre est contrôlé par le nombre de particules par pixel (N_{ppp}) et varie entre 0.01 et 0.1. L'écart type de la PSF adopté σ_{psf} est égal à 0.6, ce qui correspond à un motif d'image de taille 4×4 pixels. L'équation (5.1) donne l'expression explicite unidimensionnelle de la fonction PSF utilisée dans nos simulations [LECORDIER et WESTERWEEL \[2004\]](#), [RAFFEL et collab. \[2007\]](#), [CORNIC et collab. \[2013\]](#) :

$$h(x) = \frac{1}{2} \left[\operatorname{erf} \left(\frac{x+0.5}{\sqrt{2}\sigma_{psf}} \right) - \operatorname{erf} \left(\frac{x-0.5}{\sqrt{2}\sigma_{psf}} \right) \right]. \quad (5.1)$$

Une distribution gaussienne à moyenne nulle représentant le bruit de la caméra, a été également ajoutée aux images. Son écart-type est d'environ 5 pour cent par rapport à l'intensité maximale des images de particules. Le rapport signal/bruit (SNR) est défini par $\text{SNR} = \frac{\max(i)}{\sigma_n}$ où σ_n est l'écart-type du bruit. Pour assurer la convergence statistique des métriques déployées, 30 images synthétiques de 32×32 pixels sont générées pour chaque densité N_{ppp} . La figure 5.1 présente quelques exemples d'images utilisés dans nos simulations.

5.1.2 Métriques de détection et de localisation

Ces métriques peuvent être divisées en deux catégories : des métriques de détection et des métriques de localisation.

Métriques de détection

Une détection est dite un vrai positif (TP) si elle se trouve au voisinage d'une particule et est la détection la plus proche de cette dernière. Le voisinage d'une particule est le disque centré autour du centroïde de son image \mathbf{x}_p avec un rayon de tolérance r inférieur ou égal à 0.5 pixel. Elle est désignée comme un faux positif (FP) si elle n'est pas située au voisinage d'une véritable particule.

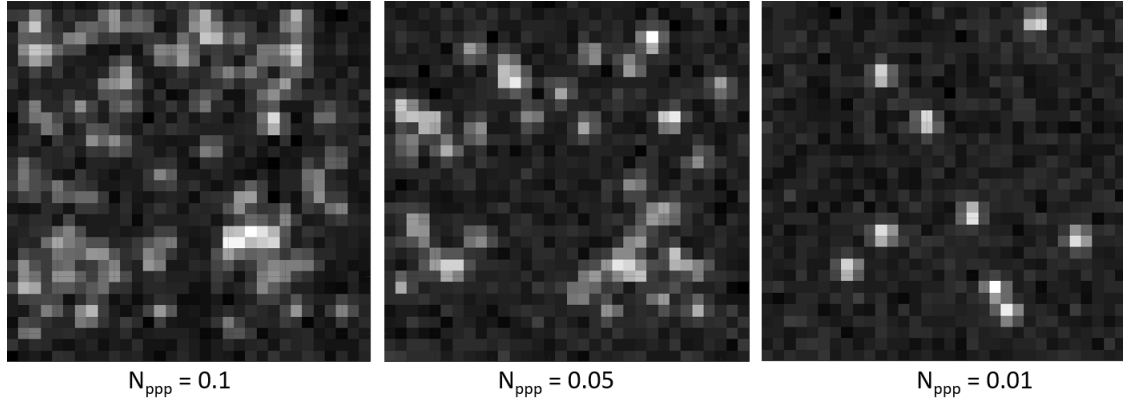


FIGURE 5.1 – Exemple d’images de synthèse utilisées dans nos simulations en présence d’un bruit modéré (SNR = 20). La densité des particules dans ces images varie de 0.01 à 0.1. $\sigma_{psf} = 0.6$ ce qui correspond à des motifs de taille d’environ 4×4 pixel.

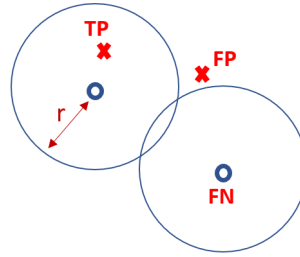


FIGURE 5.2 – Illustration de la notion de vrai positif (TP), faux positif (FP) et faux négatif (FN). Les ronds bleus correspondent aux véritables particules et les croix rouges correspondent aux détections.

Enfin, une particule est considérée comme un faux négatif (FN), si aucune détection n’est repérée dans son voisinage. La figure 5.2 illustre ces trois concepts de détection.

La "Precision" donne le pourcentage de vraies particules parmi toutes les particules détectées, et le "Rappel" est défini comme le nombre de vraies particules détectées divisé par le nombre total de particules [CORNIC et collab. \[2015\]](#) :

$$\text{Precision} = \frac{\#TP}{\#TP + \#FP} \quad \text{Rappel} = \frac{\#TP}{\#TP + \#FN}. \quad (5.2)$$

Ces mesures ont un comportement antagoniste. Typiquement la Precision augmente avec un paramètre de seuil, tandis que le Rappel diminue. Un moyen de combiner ces deux métriques est la moyenne harmonique Fscore :

$$\text{Fscore} = 2 \frac{\text{Precision} \times \text{Rappel}}{\text{Precision} + \text{Rappel}}. \quad (5.3)$$

Nous verrons plus loin son intérêt dans le cadre de la détection en PTV.

Métriques de localisation

La précision de la localisation est évaluée en utilisant la distance moyenne entre les positions des vraies détections TP $\tilde{\mathbf{x}}_p$ et les positions des particules associées \mathbf{x}_p :

$$\text{Erreur de localisation moyenne} = \frac{1}{\#TP} \sqrt{\sum_{p=1}^{\#TP} (\mathbf{x}_p - \tilde{\mathbf{x}}_p)^2}. \quad (5.4)$$

5.1.3 Intérêt du Fscore en PTV

Le Fscore est la moyenne harmonique de la Precision et du Rappel. Cet indice favorise légèrement la Precision par rapport au Rappel. Cela est intéressant en PTV où un pourcentage faible de faux positifs (FP) est important pour aboutir à une qualité d'appariement temporel élevée.

La figure ci-dessous 5.3 montre les performances de détection de la méthode convexe NNLS en fonction du seuil s . Nous fixons la densité de particules N_{ppp} à 0.1 et le SNR à 20.

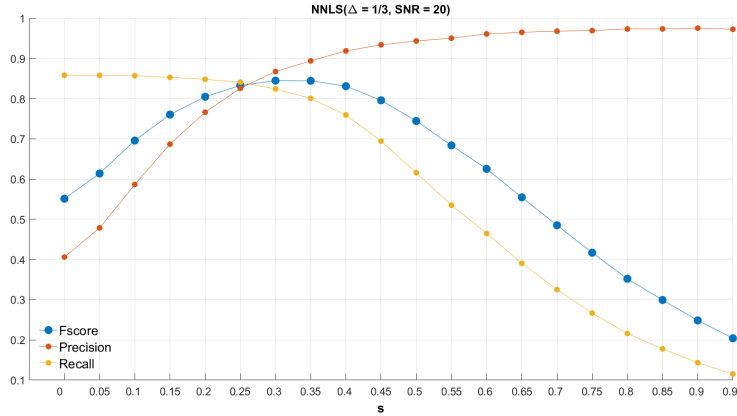


FIGURE 5.3 – Courbes de Precision, de Rappel et de Fscore en fonction du paramètre seuil s à la sortie de l'algorithme NNLS. A l'optimum du Fscore, autour de ≈ 0.3 , on a Precision \geq Rappel.

Une augmentation du niveau du seuil s améliore la métrique Precision mais détériore le Rappel. Nous observons également qu'il existe un seuil particulier s_0 où la Precision (P), le Rappel (R) et le Fscore (F) se croisent : $F(s_0) = P(s_0) = R(s_0)$. À ce seuil s_0 , la valeur absolue de la pente de la courbe Precision est supérieure à celle du Rappel. On peut prouver que dans ce cas, le seuil correspondant à un Fscore optimal s_{opt} est supérieur à s_0 , et donc :

$$P(s_{opt}) > R(s_{opt}). \quad (5.5)$$

En effet, au seuil s_0 et seulement à ce seuil, les dérivées du Fscore, de la Precision et du Rappel sont liées par l'expression simple suivante :

$$F'(s_0) = \frac{R'(s_0) + P'(s_0)}{2}. \quad (5.6)$$

Puisque $P'(s_0) \geq |R'(s_0)|$, F augmente localement à s_0 . Ainsi, s_{opt} est supérieur à s_0 et par conséquent $P(s_{opt}) > R(s_{opt})$.

Nous pensons que ce comportement est typique au problème de détection que nous traitons ici. En effet, la valeur maximale de la Precision atteinte dans ce type de problème sera généralement proche de 1, car les particules peuvent toujours être détectées en utilisant un seuil important. Inversement, le rappel maximal sera généralement inférieur à 1, car les particules se chevauchent. Lorsque les 2 courbes de Precision et de Rappel se croisent, la valeur de la Precision a déjà baissé bien plus que le Rappel n'a augmenté, parce que l'écart par rapport à la valeur maximale est plus important pour la métrique Precision. Ainsi, la pente de l'indice Precision à son intersection avec l'indice Rappel est généralement plus importante que la pente absolue de Rappel. En utilisant le même raisonnement que dans le cas de la figure 5.3, la Precision au point où le Fscore atteint son maximum sera plus élevée que le Rappel associé. Nous pensons donc qu'un point de fonctionnement au maximum du Fscore est un compromis intéressant pour la PTV qui a besoin de détections fiables pour l'association temporelle.

5.1.4 Méthodes et paramètres de réglage

Les méthodes testées ici sont les méthodes obtenues par une discrétisation d'ordre 0 : NNLS, BP et NNOMP, et les méthodes obtenues par interpolation de Taylor d'ordre 1 : CNNOMP, CNNOMP

et sa version modifiée KCNNOMP. NNLS, BP et CBP s'appuient sur des solveurs convexes. Cependant, ces algorithmes donnent de petits clusters de détections au voisinage de chaque particule, et sont sensibles au bruit. Afin d'atténuer ce défaut, une étape d'extraction de maxima locaux, suivi d'un seuillage est alors effectuée. Par conséquent, chaque méthode convexe a trois paramètres à régler : Δ contrôle l'espacement de la grille, λ contrôle le niveau de parcimonie de la solution recherchée, et s contrôle l'impact du bruit sur la qualité de détection. Noter que pour les méthodes NNLS et BP, les coordonnées de chaque maxima sont prises comme estimations des positions de particules. Cela limite la précision de localisation en cas de discrétisation grossière. Afin de réduire les erreurs de localisation, ces maxima sont introduits pour initialiser la fonction d'affinement de localisation (voir l'algorithme 7). NNOMP, CNNOMP et KCNNOMP s'appuient sur des solveurs gloutons. Chacune de ces 3 méthodes a deux paramètres à régler : Δ contrôle la finesse de la grille et K est un paramètre de parcimonie. Le tableau 5.1 ci-dessous est un récapitulatif de toutes les méthodes développées ici, ainsi que leurs paramètres associés.

Catégorie	Méthodes	Ordre de l'approximation	Paramètres
l_1	NNLS	0	s, Δ
	BP	0	s, λ, Δ
	CBP	1	
l_0	NNOMP	0	K, Δ
	CNNOMP	1	
	KCNNOMP	1	

TABLEAU 5.1 – Les paramètres des méthodes convexes et gloutonnes.

5.2 Méthodes convexes

5.2.1 Réglage de paramètres des méthodes convexes

La qualité de reconstruction 2D avec les méthodes convexes dépend essentiellement de trois paramètres : Δ , λ et s . L'un des paramètres clés est le pas de discrétisation Δ qui entraîne un compromis entre les besoins en mémoire/calcul et les performances de détection. Un réglage adéquat de Δ dépend essentiellement du niveau de SNR et de la densité d'ensemencement N_{ppp} dans les images traitées. Une difficulté se pose lors de la recherche d'une métrique de réglage afin de choisir les paramètres optimaux permettant une qualité de reconstruction élevée. L'adoption de la moyenne des erreurs de localisation, comme métrique de sélection, a conduit **CHEMINET et col-lab. [2018]** à sélectionner un pas Δ de l'ordre de $\frac{1}{20}$ pixel, ce qui entraîne en pratique un encombrement mémoire et un temps de calcul importants. Il existe quelques travaux en tomoPIV qui se sont intéressés au réglage d'un paramètre similaire à savoir le rapport voxel sur pixel $\frac{v}{p}$. Par exemple, **CHAMPAGNAT et collab. [2014]** ont proposé un rapport $\frac{v}{p}$ égal à $\frac{1}{2}$. Dans **THOMAS et collab. [2014]**, les auteurs ont préconisé des $\frac{v}{p} \simeq 1$, voire légèrement inférieurs en fonction de l'algorithme de reconstruction. La métrique utilisée est le facteur Q qui prédit une qualité de corrélation dans une logique de PIV classique. Cette métrique n'est pas pertinente dans notre contexte de PTV. Ici, nous avons adopté une politique de réglage alternative basée sur la métrique Fscore.

Afin de sélectionner les paramètres optimaux au sens du Fscore, nous avons traité une série d'images, comme expliqué dans la sous section 5.1.1. Ces images comprennent une densité de particules élevée $N_{ppp} = 0, 1$. La première simulation était dans un cas sans bruit et la seconde était en présence d'un bruit modéré (SNR = 20). Pour chaque pas Δ variant de 1 pixel à $\frac{1}{20}$, nous avons optimisé le paramètre de parcimonie λ et le seuil d'intensité s pour obtenir une valeur de Fscore maximale. Cette procédure a été appliquée au NNLS, BP et au CBP. Les résultats de performances de détection, dont la moyenne a été calculée sur 30 images, sont présentées dans la figure 5.4 :

Lorsque Δ diminue, le Fscore augmente, quelle que soit la méthode utilisée. Pour la méthode NNLS, le Fscore est presque maximal à $\Delta = \frac{1}{10}$ pixel en absence du bruit. Utiliser un pas de discrétisation plus fin que cette valeur nécessiterait plus de mémoire et de calcul avec un gain négligeable

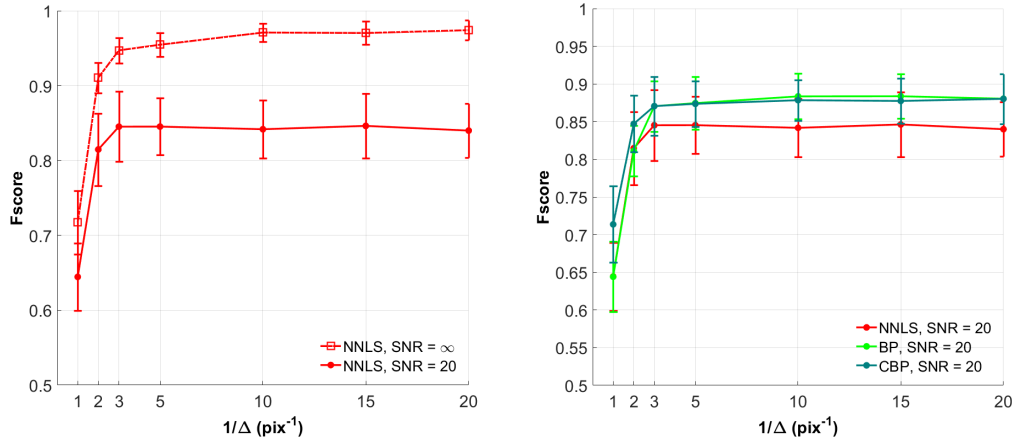


FIGURE 5.4 – Évolution du Fscore des méthodes convexes en fonction de Δ . Pour un pas Δ variant de 1 à $\frac{1}{20}$ pixel, nous avons optimisé le paramètre de parcimonie λ et le seuil s pour obtenir un Fscore maximal. Lorsque Δ diminue, Fscore augmente quelle que soit la méthode testée. Le Fscore est presque maximal à $\Delta = \frac{1}{10}$ pour NNLS en absence du bruit et à $\Delta = \frac{1}{3}$ pour toutes les méthodes convexes en présence d'un bruit modéré (SNR = 20).

en détection. Pour NNLS, BP et CBP avec un SNR de = 20, les performances de détections ne varient quasiment plus à partir de $\Delta = \frac{1}{3}$ pixel.

5.2.2 Comparaison des méthodes convexes

Comparaison des performances de détection

La figure 5.5 ci-dessous compare les mesures de performances de détection des méthodes convexes, obtenues avec un réglage optimal en termes de Fscore, en fonction de la densité de particules N_{ppp} . Comme prévu, on observe la diminution de la Precision et du Rappel au fur et à mesure que la densité d'ensemencement N_{ppp} croît car l'occurrence de superposition des images de particules augmente dans les clichés acquis. Nous pouvons également observer que la méthode NNLS est très sensible au bruit vu que les valeurs de ces métriques de détection sont très inférieures à celles obtenues dans le cas bruité. Notons aussi que les courbes de Precision et de Rappel, tracés ici, sont cohérentes avec celles de la figure 9 de CHEMINET et collab. [2018], En effet, le comportement de son algorithme NNLS avec $\Delta = \frac{1}{20}$ est très similaire au notre avec $\Delta = \frac{1}{3}$. Cette figure montre également l'intérêt de la régularisation l_1 en présence du bruit : la courbe de rappel BP est quasi identique à celle de NNLS mais sa Precision est légèrement améliorée. Cela montre que BP parvient à éliminer plus de fausses détections produites par le bruit. CBP utilise le même pas de discrétisation que BP, son dictionnaire est 4 fois plus large que celui de BP, mais donne des performances de détection similaires à ce dernier.

Comparaison des performances de localisation

Dans le cas d'un bruit modéré, NNLS et BP atteignent presque leur Fscore maximal quand $\Delta = \frac{1}{3}$ pixel. Toutefois, ce pas demeure relativement large et risque d'entraîner des erreurs de localisation plus importantes. L'objectif de ce paragraphe est d'évaluer les erreurs de localisation entraînées par $\Delta = \frac{1}{3}$ et de montrer l'impact de l'étape de minimisation non convexe sur la précision de localisation. La figure 5.6 représente l'erreur de localisation moyenne entre les détections estimées et les particules associées en utilisant la méthode NNLS et BP dans le même cas bruité traité précédemment. Cet indice a été calculé pour NNLS et BP avant et après raffinement, en utilisant le pas de discrétisation optimal. Nous avons également ajouté le cas $\Delta = \frac{1}{20}$ proposé par CHEMINET et collab. [2018] et nous avons eu des résultats de localisation du même ordre que ceux de la figure 8b de son article CHEMINET et collab. [2018], correspondant à un pas de $\frac{1}{20}$ pixel : l'er-

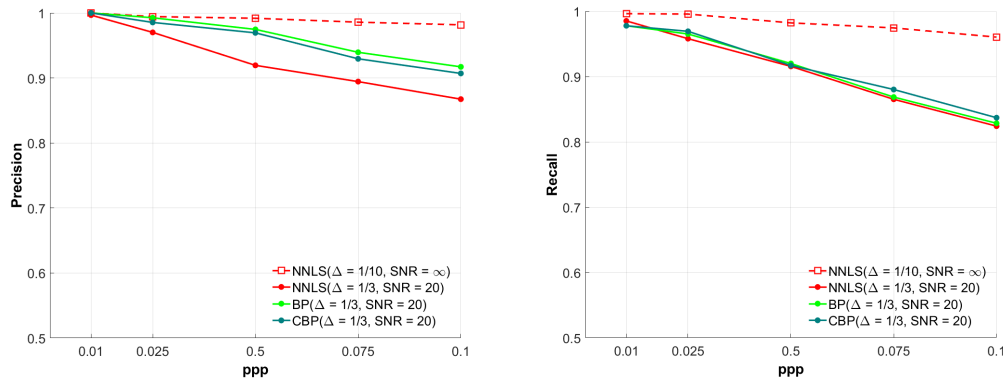


FIGURE 5.5 – Comparaison des performances de détection entre les méthodes NNLS (rouge), BP (vert) et CBP (bleu) en fonction de la densité des particules. Les courbes continues correspondent à un SNR de 20 et les courbes en pointillées correspondent à un SNR infini. Les barres d'erreur ont été omises pour éviter de surcharger les figures. Leur demi-largeur est d'environ 0.04. Les performances de détection de NNLS se dégradent nettement en présence du bruit. La régularisation l_1 (BP) améliore la précision de NNLS en éliminant les détections parasites générées par le bruit. CBP conduit à la même performance de détection que BP.

leur moyenne de localisation de BP et de NNLS avec un pas de grille $\Delta = \frac{1}{3}$ pixel après raffinement, varie de 0.1 à 0.2 pour une densité d'ensemencement $N_{\text{ppp}} \geq 0.025$. Idem pour NNLS avec un pas très fin de l'ordre de $\frac{1}{20}$ pixel.

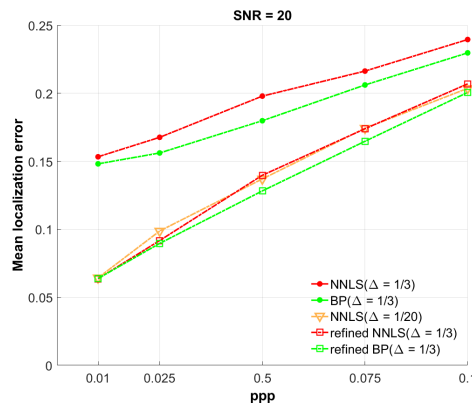


FIGURE 5.6 – Comparaison de l'erreur de localisation moyenne, avant et après raffinement de localisation, en fonction de la densité d'ensemencement. Lorsque le pas de discrétisation est large, les erreurs de localisation sont importantes. Cependant, le processus de raffinement améliore la précision de localisation et conduit à des erreurs comparables à celles obtenues avec $\Delta = \frac{1}{20}$ pixel.

Nous pouvons donc conclure que l'utilisation d'une grille fine n'est pas essentielle pour une localisation précise si un affinement de localisation est effectué après détection car une reconstruction avec un pas de $\Delta = \frac{1}{3}$ entraîne les mêmes erreurs de localisations, que celles de $\Delta = \frac{1}{20}$.

5.3 Méthodes gloutonnes

5.3.1 Réglage de paramètres des méthodes gloutonnes

Dans cette sous-section, nous suivons la même procédure de réglage que celle utilisée pour les méthodes convexes afin de comparer fidèlement les 3 méthodes gloutonnes présentées dans ce manuscrit : NNOMP, CNNOMP et KCNNOMP. Le paramètre régissant la finesse de la grille Δ et le coefficient de parcimonie K ont été réglés de telle façon à assurer un Fscore maximal. La figure 5.7 montre les courbes des 3 méthodes l_0 en fonction de Δ . La densité d'ensemencement est égal à

$N_{\text{ppp}} = 0.1$. Chaque mesure est obtenue en utilisant le coefficient de parcimonie optimal en termes de Fscore. Nous n'avons pas intégré à cette figure les autres valeurs correspondant à Δ inférieur à $\frac{1}{5}$ car le bruit prédomine sur l'erreur de modèle.

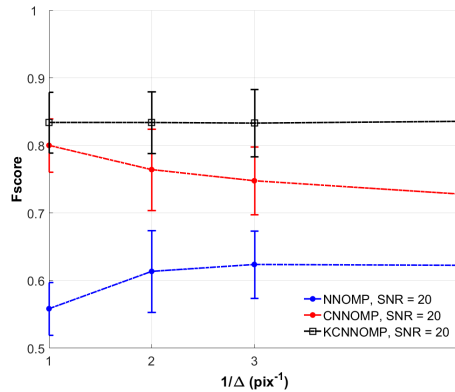


FIGURE 5.7 – Évolution de la moyenne et de l'écart-type du Fscore des méthodes gloutonnes en fonction de Δ . Chaque méthode est utilisée avec son paramètre de parcimonie optimal. La sensibilité globale du Fscore en fonction de Δ est négligeable pour KCNNOMP et également faible pour les deux autres méthodes NNOMP et CNNOMP. Lorsque Δ diminue, on observe des comportements antagonistes des 3 méthodes : la performance de détection de NNOMP se détériore, celle de CNNOMP s'améliore et celle de KCNNOMP demeure inchangée. Pour garantir un meilleur compromis entre une qualité élevée de détection et un faible temps de calcul, nous avons choisi $\Delta = \frac{1}{3}$ pixel pour NNOMP et $\Delta = 1$ pixel pour CNNOMP et KCNNOMP.

Notons d'abord que le Fscore de KCNNOMP n'est pas sensible à Δ : la variation de la moyenne du Fscore est négligeable par rapport à son écart-type. Ce comportement est une conséquence directe du raffinement hors grille pour la réduction du résidu de l'image. Il est donc clair que pour KCNNOMP, la sélection d'un pas de grille grossier de l'ordre de 1 pixel est suffisante. On peut aussi constater que la sensibilité du Fscore de NNOMP et CNNOMP en fonction de Δ est plus faible, en comparaison avec les méthodes convexes. Pour NNOMP, à mesure que Δ diminue, la moyenne du Fscore augmente légèrement et commence à stagner à partir de $\Delta = 1/3$ pixel. Inversement, pour CNNOMP, lorsque Δ diminue, la moyenne de Fscore baisse légèrement, donc $\Delta = 1$ pixel est le meilleur choix.

On conclut donc que le réglage optimal serait de fixer Δ à $\frac{1}{3}$ pixel pour NNOMP, et 1 pixel pour CNNOMP et KCNNOMP.

5.3.2 Comparaison des méthodes gloutonnes

Afin de comparer fidèlement les méthodes gloutonnes, nous avons tracé les courbes de Precision et de Rappel en fonction de densité d'ensemencement de particules, en utilisant les paramètres optimaux. La figure 5.8 présente une comparaison entre NNOMP NGUYEN et collab. [2017] et nos 2 algorithmes gloutons continus : CNNOMP et KCNNOMP. En cas de faibles densités ($N_{\text{ppp}} = 1\%$), il n'y a pratiquement pas de chevauchements et la qualité de détection est très élevée pour toutes les méthodes sauf la méthode NNOMP perturbée par le bruit. Lorsque la probabilité de chevauchement augmente, nous constatons une nette diminution des valeurs de Precision et de Rappel, et plus particulièrement pour NNOMP. CNNOMP semble être plus robuste aux chevauchements et sa version modifiée KCNNOMP se comporte encore mieux.

Nous pouvons également remarquer que dans des simulations sans bruit, KCNNOMP parvient à séparer presque toutes les particules qui se chevauchent, grâce à l'étape de raffinement. Ces performances de détection sont identiques à celles obtenues avec NNLS (voir figure 5.5) avec un pas de grille Δ de $\frac{1}{10}$ pixel. C'est un excellent résultat de KCNNOMP car on aboutit à une qualité de reconstruction identique à celle de NNLS avec une grille 10 fois plus grossière.

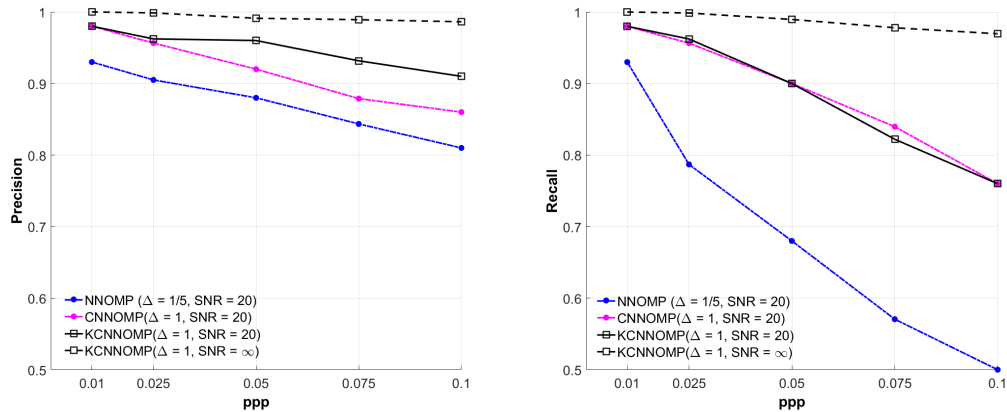


FIGURE 5.8 – Comparaison des performances de détection des méthodes gloutonnes en fonction de densité de particules. Les courbes continues correspondent à un SNR de 20 et les courbes pointillées correspondent à un SNR infini. La performance de détection de NNOMP se dégrade considérablement à mesure que la densité d'ensemencement augmente. CNNOMP améliore NNOMP et est relativement robuste aux chevauchements. Sa version modifiée KCNNOMP est plus performante, en particulier en absence du bruit.

5.4 Méthodes gloutonnes contre méthodes convexes

5.4.1 Comparaison de performances de détection

Comme indiqué dans les sous-sections précédentes, pour un SNR de 20, BP avec un pas de grille de $\frac{1}{3}$ pixel est la meilleure approche convexe et KCNNOMP avec un pas de 1 pixel est la meilleure approche gloutonne. Une étude comparative s'impose alors afin de trancher sur la meilleure approche à retenir. Pour cela, on a tracé plusieurs courbes de Precision et de Rappel en fonction du niveau de bruit SNR et de la densité de particules N_{ppp} .

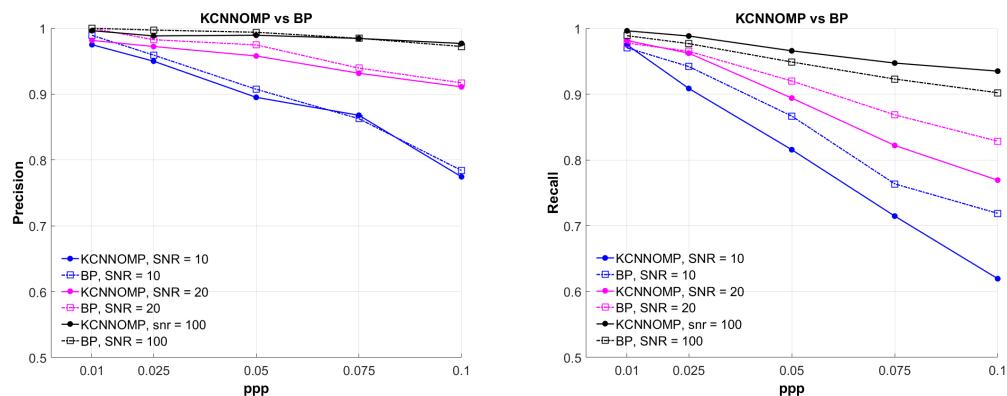


FIGURE 5.9 – Comparaison entre BP (courbe continue) et KCNNOMP (courbe en pointillée) en fonction de la densité des particules. Chaque couleur est associée à un niveau de SNR (noir : 100, rose : 20 et bleu : 10). Pour les SNR bas, BP est plus performant en détection que KCNNOMP et vice-versa pour les SNR élevés.

La figure 5.9 illustre cette comparaison : pour des valeurs de SNR élevées, les valeurs de Precision des deux algorithmes sont presque identiques mais les valeurs du Rappel de KCNNOMP sont légèrement plus élevées. Cela prouve que pour des SNR élevés, l'utilisation de la méthode KCNNOMP est plus pertinente que BP. À noter que dans ce cas, $\lambda_{opt} = 0$ pour BP, ce qui correspond implicitement à une approche NNLS. Toutefois, lorsque le SNR est faible, la tendance est inversée. Les courbes de Precision et de Rappel de BP sont supérieures à celles de KCNNOMP. Cet écart est plus visible sur les courbes de rappel, en particulier pour les densités élevées. Cela prouve qu'une approche convexe avec une pénalisation de type l_1 est plus robuste au bruit.

5.4.2 Comparaison de temps de calcul

Nous avons évalué le temps de calcul des 2 meilleures méthodes BP et KCNNOMP sur la même série d'images de taille 32×32 en utilisant les paramètres optimaux. Nous avons également ajouté la méthode NNLS avec un pas de grille $\Delta = \frac{1}{20}$ pixel qui correspond à l'état de l'art des méthodes à grille 2D en PTV [CHEMINET et collab. \[2018\]](#). L'optimiseur non convexe de raffinement de localisation, utilisé à plusieurs reprises dans KCNNOMP et une seule fois après BP, a été réglé différemment. Le paramètre de tolérance de l'optimiseur a été fixé à 10^{-8} pour le post-traitement BP et à 10^{-14} pour KCNNOMP. Cette valeur est inférieure pour le raffinement associé au BP car nous estimons que le chevauchement a été essentiellement traité à l'étape d'inversion.

La sous-figure de gauche [5.10a](#) présente le temps de calcul résultant en secondes, en fonction de la densité des particules N_{ppp} . Nous pouvons observer que la méthode BP est très rapide même pour des densités N_{ppp} élevés (≤ 1 sec). Le raffinement de localisation augmente légèrement son temps d'exécution, mais le temps de calcul global demeure faible en comparaison avec les autres méthodes. A l'instar de NNLS, le temps d'exécution de BP commence par une valeur élevée mais augmente très lentement contrairement à la méthode KCNNOMP qui commence par une valeur très faible et augmente rapidement avec l'augmentation de la densité de particules. Nous pouvons également constater que la quasi-totalité du temps d'exécution de KCNNOMP est due à l'étape de minimisation non convexe (voir la sous-figure [5.10b](#)) : pour une densité N_{ppp} de 1%, le rapport entre le temps de minimisation non convexe et le temps d'exécution total est de 82% et pour une densité N_{ppp} de 10% le rapport dépasse 99%.

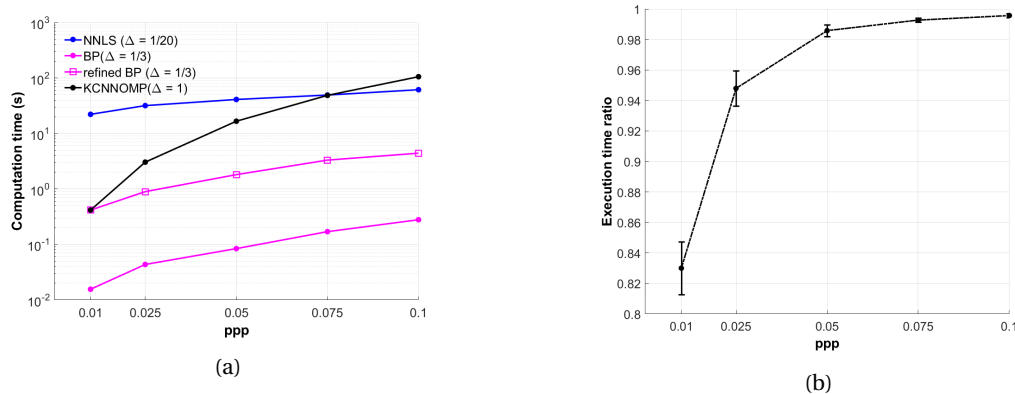


FIGURE 5.10 – À gauche : comparaison du temps de calcul entre BP, BP avec une étape postérieure de raffinement de position et de KCNNOMP en utilisant les paramètres optimaux. NNLS avec $\Delta = \frac{1}{20}$, adopté par [CHEMINET et collab. \[2018\]](#), a également été tracé ici. À droite : le rapport entre le temps de minimisation non convexe et le temps d'exécution global.

Quand il s'agit de traiter des images plus large, le temps de calcul de notre méthode gloutonne KCNNOMP augmente rapidement. Pour réduire son temps d'exécution, plusieurs pistes sont envisageables : on peut par exemple suivre la procédure de [CHEMINET et collab. \[2018\]](#) consistant à segmenter l'image originale en blocs de pixels séparées, pouvant être traités séquentiellement. Ce procédé permet une réduction drastique du coût de calcul sans détérioration significative des performances de détection [CHEMINET et collab. \[2018\]](#). Une telle technique sera également efficace si elle est associée à des méthodes convexes. À noter que, pour les méthodes convexes BP et CBP, une autre façon d'accomplir la réduction de dimensionnalité est d'utiliser les techniques *safe screening* [FERCOQ et collab. \[2015\]](#) [HERZET et DRÉMEAU \[2018\]](#). Ces dernières ne sont pas des heuristiques. Elles permettent d'identifier par de simples tests, avant ou pendant l'optimisation, la plupart de nœuds ne comportant pas de particules. Leur complexité est négligeable par rapport à l'effort de calcul typique nécessaire pour résoudre le problème de reconstruction 2D [HERZET et collab. \[2019\]](#).

5.5 Bilan

Ce chapitre décrit en détail les conditions de nos simulations numériques, la politique de réglage adoptée et les principaux résultats obtenus.

Dans la première section, on a expliqué le processus de simulation, les métriques de performances de détection et de localisation utilisées, ainsi que leurs principales propriétés dans un problème de détection en PTV, notamment la caractérisation du point correspondant à une valeur de Fscore maximale. Nous avons ensuite proposé une politique de réglage unique de tous les paramètres (pas de discrétisation, seuils et paramètres de régularisation) pour toutes les méthodes développées : convexes et gloutonnes. Cette politique nous a conduit à fixer le pas de grille à $\frac{1}{3}$ pixel pour NNLS, BP, CBP et NNOMP et à 1 pixel pour CNNOMP et KCNNOMP. On a également montré que la meilleure approche convexe assurant un bon compromis entre temps d'exécution et performances de détection est BP et que la meilleure approche l_0 est KCNNOMP. Enfin, on a identifié 2 domaines d'emploi intéressants de ces 2 méthodes : à fort SNR, KCNNOMP atteint des performances de détections identiques à celles de NNLS avec un dictionnaire 25 fois plus petit. A faible SNR, la régularisation l_1 fournit les meilleures performances avec des grilles assez grossières (de l'ordre de $\frac{1}{3}$ pixel), ce résultat nuance les conclusions de CHEMINET et collab. [2018] qui préconisent des pas de discrétisation de l'ordre de $\frac{1}{20}$ pixel. Quant au temps d'exécution, on a constaté que la méthode BP est la plus rapide dans le traitement des images à hautes densités, même en ajoutant une étape postérieure de raffinement.

5.6 Références

- CHAMPAGNAT, F., P. CORNIC, A. CHEMINET, B. LECLAIRE, G. LE BESNERAIS et A. PLYER. 2014, «Tomographic PIV : Particles versus blobs», *Measurement Science and Technology*, vol. 25, n° 8, p. 084 002. [57](#)
- CHEMINET, A., J. KRAWCZYNSKI et P. DRUAULT. 2018, «Particle image reconstruction for particle detection in particle tracking velocimetry», *Measurement Science and Technology*, vol. 29, n° 12, p. 125 202. [v](#), [57](#), [58](#), [62](#), [63](#)
- CORNIC, P., F. CHAMPAGNAT, A. CHEMINET, B. LECLAIRE et G. LE BESNERAIS. 2015, «Fast and efficient particle reconstruction on a 3D grid using sparsity», *Experiments in Fluids*, vol. 56, n° 3, 62, ISSN 0723-4864. [55](#)
- CORNIC, P., F. CHAMPAGNAT, A. CHEMINET, B. LECLAIRE et G. LEBESNERAIS. 2013, «Computationally efficient sparse algorithms for tomographic piv reconstruction», . [54](#)
- FERCOQ, O., A. GRAMFORT et J. SALMON. 2015, «Mind the duality gap : safer rules for the lasso», *Proceedings of the 32nd International Conference on Machine Learning*, vol. 37, p. 333–342. [62](#)
- HERZET, C., C. DORFFER et A. DRÉMEAU. 2019, «Gather and conquer : Region-based strategies to accelerate safe screening tests», *IEEE Transactions on Signal Processing*, vol. 67, n° 12, p. 3300–3315. [62](#)
- HERZET, C. et A. DRÉMEAU. 2018, «Joint screening tests for Lasso», dans *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, p. 4084–4088. [62](#)
- LECORDIER, B. et J. WESTERWEEL. 2004, «The europiv synthetic image generator (sig)», . [54](#)
- NGUYEN, T. T., C. SOUSSEN, J. IDIER et E.-H. DJERMOUNE. 2017, «An optimized version of non-negative OMP», dans *XXVIe Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017*, Juan-les-Pins, France. [60](#)
- RAFFEL, M., C. WILLERT, S. WERELEY et J. KOMPENHANS. 2007, *Particle Image Velocimetry - A Practical Guide (2nd Edition)*, ISBN 978-3-540-72307-3. [54](#)

THOMAS, L., B. TREMBLAIS et L. DAVID. 2014, «Optimization of the volume reconstruction for classical tomo-piv algorithms (mart, bimart and smart) : synthetic and experimental studies», *Measurement Science and Technology*, vol. 25, n° 3, p. 035 303. [57](#)

Chapitre 6

Conclusions et Perspectives

Sommaire

6.1 Résumé de contributions	66
6.1.1 Rappel succinct de la problématique de thèse	66
6.1.2 Contributions sur l'aspect convexe	66
6.1.3 Contributions sur l'aspect glouton	67
6.1.4 Contributions sur le réglage et la comparaison des méthodes	68
6.1.5 Sélection de la meilleure méthode	68
6.2 Perspectives	68
6.2.1 Réduction de la dimensionnalité et validation sur des images réelles	68
6.2.2 Méthodes gloutonnes à sélection multiple	68
6.2.3 Vers une reconstruction sans grille	69
6.3 Références	69

Dans ce dernier chapitre, nous rappellerons brièvement la problématique de cette thèse, et nous résumerons nos principales contributions ainsi que les perspectives possibles de nos travaux.

6.1 Résumé de contributions

6.1.1 Rappel succinct de la problématique de thèse

La Vélocimétrie par Image de Particule (PIV) est une technique expérimentale de mesure de champs de vitesses. La procédure d'estimation consiste à découper 2 images successives en fenêtres d'interrogation et de les corrélérer afin d'estimer le déplacement local d'une zone de fluide et par conséquent le vecteur vitesse associé. Toutefois, cette opération agit comme un filtre passe bas qui lisse les gradients, ce qui conduit à une résolution spatiale réduite.

La Vélocimétrie par Suivi de Particules (PTV) contourne la limitation de l'effet moyen de la PIV en suivant individuellement l'évolution temporelle de la position de chaque particule, fournissant ainsi à chaque instant et en tout point de l'écoulement comprenant une particule, une mesure de vitesse bien résolue.

FENG et collab. [2011] ont montré numériquement que les sources d'erreur en PTV, lors de l'utilisation de caméras scientifiques à grande vitesse, est principalement lié à l'incertitude sur la position des particules. Ainsi, une bonne détection/localisation est une condition sine qua non pour assurer une mesure précise en PTV. Parmi les techniques usuelles pour détecter et localiser les particules, on peut mentionner celles basées sur une procédure de seuillage dynamique à savoir la DTB et l'EDT. D'autres techniques à base de corrélation sont également utilisées notamment la méthode (PMCM) et sa version améliorée (CCM) comprenant 2 filtres mis en série. Ces techniques ont l'avantage d'être rapides et facilement implémentables. Cependant, leurs performances de détection se détériorent fortement quand l'occurrence de chevauchement augmentent dans l'image.

Afin de tenir compte du phénomène de superposition, un modèle additif de formation d'image, a été adopté par plusieurs auteurs. Ce dernier suppose que chaque pixel illuminé de l'image acquise, est le résultat de contribution d'un nombre fini de particules ponctuelles, favorisant ainsi des reconstructions parcimonieuses. La détection/localisation de particules avec ce modèle est souvent faite en 2 étapes : primo, on discrétise le modèle sur une grille figée de nœuds, via une approximation de Taylor d'ordre 0 pour chaque pixel de l'image. Cela aboutit à une relation matricielle linéaire entre les observations et une distribution d'intensité de particules. Deuxio, on inverse le système linéaire en intégrant la contrainte de positivité par exemple par moindres carrés sous contrainte de positivité, ou en intégrant la parcimonie afin d'extraire une solution unique.

Cette approche à grille sépare les particules seulement quand la grille adoptée est extrêmement fine. CHEMINET et collab. [2018] ont montré que leur algorithme NNLS, basé sur un optimiseur de type active set et combiné à une extraction de maxima locaux, requiert au minimum un pas de vingtième de pixel pour gérer les chevauchements, ce qui engendre un encombrement mémoire et un temps de calcul très important. De plus, leur technique est fortement sensible au bruit de mesure, et requiert une procédure supplémentaire de seuillage, difficilement réglable, en l'occurrence en présence de chevauchements.

6.1.2 Contributions sur l'aspect convexe

Afin d'améliorer la robustesse de la méthode NNLS vis à vis du bruit et augmenter le niveau de parcimonie des solutions, on a introduit un terme de régularisation l_1 au problème inverse formulé par CHEMINET et collab. [2018], et on a résolu le problème induit en modifiant la fonction $lsqnonneg$, une variante de l'active set, afin de l'adapter à notre problème BP. Effectivement, on a pu constater que cette dernière réduit sensiblement le nombre de fausses détections en comparaison avec NNLS. Le résultat est encore plus pertinent après filtrage des détections à faibles intensités.

On a pu également choisir un réglage du pas de discrétisation Δ plus optimal en mémoire et en temps de calcul, en gardant la même qualité de détection que celle obtenue par [CHEMINET et collab. \[2018\]](#) dans le cas d'un pas de $\frac{1}{20}$ pixel. Cela a été obtenu en choisissant la plus grande valeur de Δ garantissant une valeur de Fscore maximale, en présence d'un bruit modéré. Dans notre cas Δ_{opt} est égal à $\frac{1}{3}$ pixel. Ce choix de Δ à base de métriques de détection, ne se fait pas au détriment de la précision de localisation, car en affinant les positions estimées, on aboutit aux mêmes erreurs de localisation obtenues dans le cas de grilles très fines. Néanmoins, la stratégie BP a le défaut de requérir un post-processing afin de promouvoir la parcimonie du signal. En outre, la grille exigée est relativement fine.

Ainsi, on a poussé l'ordre l'approximation de Taylor à l'ordre 1 lors de la procédure de discrétisation, afin d'aboutir à une approximation plus précise du modèle continu, même pour des pas de grille plus grossiers. Dans un contexte 2D, cette nouvelle discrétisation fait intervenir 2 dictionnaires dérivés supplémentaires associés à deux variables contraintes linéairement. On a baptisé ce problème CBP.

On l'a abordé avec la méthode du gradient proximal FISTA. Cela a été possible après le calcul de l'expression analytique de l'opérateur proximal de la fonction modélisant les contraintes du problème. Cependant, cet algorithme exige l'utilisation d'un paramètre de tolérance très petit pour garantir une meilleure convergence, ce qui introduit un temps de calcul important et notamment quand les pas sont fins.

Afin d'accélérer la résolution, on a utilisé la méthode de l'active set précédemment appliquée au problème BP. Cela a été possible en reparamétrisant le problème standard de CBP en un problème avec uniquement des contraintes de positivité. On a également introduit une étape de supplémentaire de seuillage pour promouvoir la qualité de détection.

Les évaluations sur des images de synthèse bruitées et à haute densité nous ont montré que l'approche CBP, à base d'interpolation de Taylor d'ordre 1 et de contrainte de parcimonie l_1 , simulé avec ses paramètres optimaux, conduit aux mêmes performances de détection que celles de BP. De plus, l'étape de post-processing est toujours nécessaire pour promouvoir la détection. Ainsi, on a choisi BP comme étant la meilleure méthode convexe, assurant le meilleur compromis performance de détection par rapport aux besoins en mémoire et en coût de calcul.

6.1.3 Contributions sur l'aspect glouton

Comme évoqué précédemment, l'approche CBP est moins prometteuse puisqu'on a eu les mêmes défauts que les méthodes d'ordre 0. Cela nous a poussé à explorer les résultats d'une régularisation l_0 .

Pour ce fait, On a gardé la même reparamétrisation développée pour le cas CBP, et on a résolu le problème induit en s'appuyant sur la méthode NNOMP de [NGUYEN et collab. \[2017\]](#). Notre algorithme, qu'on a appelé CNNOMP, se différencie légèrement de NNOMP dans la façon d'augmenter le support, puisque à chaque itération l'ajout d'atomes se fait par un bloc de 4. Cette stratégie fonctionne bien avec des pas plus larges de l'ordre de 1 pixel mais échoue en présence de fortes interactions de particules dans l'image traitée. Nous montrons en effet qu'en cas de recouvrement l'algorithme glouton a tendance à placer une particule à peu près au milieu des deux vraies particules avec une intensité plus élevée que l'intensité nominale. Cette particule ne sera jamais remise en cause et laissera dans le résidu deux particules fantômes de plus faible intensité. En fonction du réglage de l'algorithme cela se traduira par deux ou trois détections mal placées.

Afin de mieux gérer les chevauchements, on a ajouté une étape supplémentaire de minimisation non convexe, juste après l'étape d'inversion NNLS. C'est le KCNNOMP. L'objectif de cette optimisation est de remettre en cause simultanément la localisation des particules pré-détectées pour leur permettre de rejoindre une meilleure position.

Nos évaluations sur des images à hautes densités sans bruit, ont montré que la reconstruction dans ce cas est quasi parfaite. En effet, l'étape de correction non convexe permet de chercher des positions plus optimales en dehors de la grille, ce qui conduit à des résidus précis et par conséquent un bon choix d'atomes dans les itérations suivantes.

On a également montré grâce à notre politique de réglage basée sur la maximisation de Fscore, qu'en présence d'un bruit modéré et une densité élevée, on pourrait adopter des pas de l'ordre de 1 pixel sans dégrader la qualité de reconstruction.

6.1.4 Contributions sur le réglage et la comparaison des méthodes

Afin de comparer équitablement toutes les méthodes gloutonnes et convexes exposées dans ce manuscrit, on a opté pour un réglage de paramètres à base de la métrique de détection Fscore [AIT TILAT et collab. \[2019\]](#). L'idée principale de cette stratégie est de faire varier sur une large plage tous les paramètres régissant la qualité de détection, en s'appuyant sur des images de synthèse à forte densité de particules, afin de sélectionner la combinaison conduisant à un meilleur compromis entre qualité de détection et besoins en mémoire et temps de calcul. Pour chaque valeur du pas de discrétisation Δ , on sélectionne les paramètres de seuillage et de parcimonie correspondants à la valeur la plus élevée du Fscore. Ensuite, on choisit le pas à partir duquel la courbe Fscore commence à se stabiliser. Un argument qualitatif est fourni pour justifier qu'un réglage maximisant le Fscore correspond à un point de fonctionnement qui privilégie la Precision par rapport au Rappel. Or un faible taux de faux positifs est propice à l'appariement temporel des détections en PTV. Ce réglage via la métrique Fscore s'avère donc particulièrement pertinent pour la PTV.

6.1.5 Sélection de la meilleure méthode

On a montré que la meilleure approche convexe assurant un bon compromis entre temps d'exécution et performances de détection est BP et que la meilleure approche l_0 est KCNNOMP.

On a aussi identifié 2 domaines d'emploi intéressants de ces deux méthodes : à fort SNR, KCNNOMP atteint des performances de détections identiques à celles de BP avec seulement un dictionnaire 25 fois plus petit que celui de BP. A faible SNR, la régularisation l_1 fournit les meilleures performances avec des pas de grilles de l'ordre de $\frac{1}{3}$ pixel.

Quant au temps d'exécution, on a constaté que la méthode BP est la plus rapide dans le traitement des images à hautes densités, même en ajoutant une étape postérieure de raffinement.

6.2 Perspectives

6.2.1 Réduction de la dimensionnalité et validation sur des images réelles

Le temps de calcul de KCNNOMP explose quand on traite des images plus larges. Pour réduire son temps d'exécution, on peut par exemple suivre la procédure de [CHEMINET et collab. \[2018\]](#) consistant à segmenter l'image originale en agrégats de pixels séparés, et puis les traiter séquentiellement. Ce procédé réduit drastiquement le temps de calcul sans détérioration significative des performances de détection. Une telle technique sera également efficace si elle est appliquée à la méthode convexe BP. Une autre façon de réduire la dimensionnalité de BP est d'utiliser les techniques de *safe screening* [FERCOQ et collab. \[2015\]](#), [HERZET et DRÉMEAU \[2018\]](#), [HERZET et collab. \[2019\]](#). Ces dernières permettent d'identifier par des tests simples, avant ou pendant l'optimisation, la plupart de nœuds ne comportant pas de particules.

Tous les algorithmes 2D exposés dans ce manuscrit, peuvent être étendus en 3D. Ainsi, la réduction de complexité est indispensable pour gérer le problème 3D. Dans un premier temps, on se contentera de données synthétiques afin de re-régler les paramètres de seuil et de discrétisation, avant de traiter des images réalistes.

6.2.2 Méthodes gloutonnes à sélection multiple

On a vu que les méthodes gloutonnes échouaient en cas de recouvrement à cause de la stratégie de modification unique à chaque étape. Une solution à ce problème pourrait être apportée par des techniques implémentant des modifications multiples. [MHENNI et collab. \[2020\]](#) ont proposé

une telle technique. Cette procédure étend les méthodes gloutonnes en proposant plusieurs modifications simultanées. Les modifications multiples sont proposées via un algorithme d'homotopie. La restriction à une unique modification de cette nouvelle classe d'algorithmes correspond à OMP. En effet dans une méthode d'homotopie la première famille de solutions à un unique atome correspond à la maximisation de la corrélation entre signal et dictionnaire, comme pour OMP. Cette technique pourrait constituer une alternative économique à la méthode KCNNOMP (dont la phase de raffinement non convexe est très coûteuse). On pourrait commencer à l'implémenter dans le cas du dictionnaire de Taylor d'ordre 0 et l'étendre au cas du dictionnaire de Taylor d'ordre 1.

6.2.3 Vers une reconstruction sans grille

Comme évoqué dans le chapitre des méthodes convexes, l'adoption d'une grille fine soulève plusieurs problèmes notamment la taille des dictionnaires mis en jeu qui augmente de façon phénoménale avec la taille des images traitées.

Les résultats sur le KCNNOMP posent aussi la question sur le degrés de pertinence des méthodes à grille puisque la séparation n'aurait pas été possible sans sortir de la grille. De ce fait, une perspective intéressante serait d'explorer des approche hors-grille, c'est à dire des approches ne faisant pas intervenir une grille prédéterminée comme ce qui a été proposé par [DENOYELLE et collab. \[2019\]](#). Ces dernières se font généralement sur 3 phases :

1. Étape de corrélation pour introduire de nouvelles particules.
2. Optimisation sur les intensités.
3. Optimisation sur les intensités et les positions afin de diminuer plus le critère.

La plupart de ces briques ont été déjà développées dans nos travaux antérieurs : la phase 2 peut être faite avec l'algorithme Fista ou l'active set qu'on a déjà développés, en outre la recherche de solutions sera plus rapide car elle se fera sur un espace plus réduit. La phase 3 correspond à la même étape de minimisation non convexe utilisée dans le KCNNOMP. Seule la première étape est à creuser car elle fait intervenir un dictionnaire continu.

6.3 Références

- AIT TILAT, S., F. CHAMPAGNAT et C. HERZET. 2019, «A new sparsity based particle image reconstruction approach for particle detection», dans *Proceedings of the 13th International Symposium on Particle Image Velocimetry*. [68](#)
- CHEMINET, A., J. KRAWCZYNSKI et P. DRUAULT. 2018, «Particle image reconstruction for particle detection in particle tracking velocimetry», *Measurement Science and Technology*, vol. 29, n° 12, p. 125202. [66](#), [67](#), [68](#)
- DENOYELLE, Q., V. DUVAL, G. PEYRÉ et E. SOUBIES. 2019, «The sliding frank–wolfe algorithm and its application to super-resolution microscopy», *Inverse Problems*, vol. 36, n° 1, p. 014001. [69](#)
- FENG, Y., J. GOREE et B. LIU. 2011, «Errors in particle tracking velocimetry with high-speed cameras», *The Review of scientific instruments*, vol. 82, p. 053707. [66](#)
- FERCOQ, O., A. GRAMFORT et J. SALMON. 2015, «Mind the duality gap : safer rules for the lasso», *Proceedings of the 32nd International Conference on Machine Learning*, vol. 37, p. 333–342. [68](#)
- HERZET, C., C. DORFFER et A. DRÉMEAU. 2019, «Gather and conquer : Region-based strategies to accelerate safe screening tests», *IEEE Transactions on Signal Processing*, vol. 67, n° 12, p. 3300–3315. [68](#)
- HERZET, C. et A. DRÉMEAU. 2018, «Joint screening tests for Lasso», dans *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, p. 4084–4088. [68](#)

- MHENNI, R. B., S. BOURGUIGNON et J. IDIER. 2020, «A greedy sparse approximation algorithm based on l_1 -norm selection rules», dans *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, p. 5390–5394. [68](#)
- NGUYEN, T. T., C. SOUSSEN, J. IDIER et E.-H. DJERMOUNE. 2017, «An optimized version of non-negative OMP», dans *XXVIe Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017*, Juan-les-Pins, France. [67](#)

Annexe A

Annexes

A.1 Définitions

Pour $\alpha > 0$, on considère le cône pyramidal défini par

$$\mathcal{P} \triangleq \{(x, y, z) \in \mathbb{R}^3 \mid |y| \leq \alpha x, |z| \leq \alpha x\}. \quad (\text{A.1})$$

Noter que $x \geq 0$ est implicitement vérifiée.

A.2 Problème à résoudre

Le problème à résoudre est celui de la projection orthogonale d'un point de \mathbb{R}^3 sur \mathcal{C} . Ce problème se pose par exemple dans la résolution de problèmes quadratiques à pénalisation l_1 sous contrainte d'appartenance à un hypercône. Une approche de type ISTA/FISTA se ramène in fine au problème de projection étudié.

Formellement, le but est de calculer, pour tout $(a, b, c) \in \mathbb{R}^3$,

$$\text{proj}_{\mathcal{P}}(a, b, c) \triangleq \arg \min_{(x, y, z) \in \mathcal{P}} (x - a)^2 + (y - b)^2 + (z - c)^2. \quad (\text{A.2})$$

A.3 Résultats complets

On note $(x^*, y^*, z^*) = \text{proj}_{\mathcal{P}}(a, b, c)$. L'expression explicite de (x^*, y^*, z^*) est donnée en Table A.1. La solution est linéaire par morceaux, les dix morceaux sont dix sous espaces qui partitionnent \mathbb{R}^3 .

Cas	x^*	y^*	z^*
\mathcal{P}	a	b	c
\mathcal{P}^*	0	0	0
\mathcal{F}_1	$\frac{a+\alpha b}{1+\alpha^2}$	αx^*	c
\mathcal{F}_2	$\frac{a+\alpha c}{1+\alpha^2}$	b	αx^*
\mathcal{F}_3	$\frac{a-\alpha b}{1+\alpha^2}$	$-\alpha x^*$	c
\mathcal{F}_4	$\frac{a-\alpha c}{1+\alpha^2}$	b	$-\alpha x^*$
\mathcal{A}_1	$\frac{a+\alpha b+\alpha c}{1+2\alpha^2}$	αx^*	αx^*
\mathcal{A}_2	$\frac{a-\alpha b+\alpha c}{1+2\alpha^2}$	$-\alpha x^*$	αx^*
\mathcal{A}_3	$\frac{a-\alpha b-\alpha c}{1+2\alpha^2}$	$-\alpha x^*$	$-\alpha x^*$
\mathcal{A}_4	$\frac{a+\alpha b-\alpha c}{1+2\alpha^2}$	αx^*	$-\alpha x^*$

(A.3)

TABLEAU A.1 – Expression de $\text{proj}_{\mathcal{P}}(a, b, c)$ en fonctions des dix cas.

Les ensembles \mathcal{P}^* , \mathcal{F}_i et \mathcal{A}_i sont définis par :

$$\mathcal{P}^* \triangleq \left\{ (a, b, c), |b+c| \leq -\frac{a}{\alpha}, |b-c| \leq -\frac{a}{\alpha} \right\} \quad (\text{A.4})$$

$$\mathcal{F}_1 \triangleq \left\{ (a, b, c) \in \mathbb{R}^3 | b \geq \alpha a, |c| \leq \alpha \frac{a+\alpha b}{1+\alpha^2} \right\} \quad (\text{A.5})$$

$$\mathcal{F}_2 \triangleq \left\{ (a, b, c) \in \mathbb{R}^3 | c \geq \alpha a, |b| \leq \alpha \frac{a+\alpha c}{1+\alpha^2} \right\} \quad (\text{A.6})$$

$$\mathcal{F}_3 \triangleq \left\{ (a, b, c) \in \mathbb{R}^3 | b \leq -\alpha a, |c| \leq \alpha \frac{a-\alpha b}{1+\alpha^2} \right\} \quad (\text{A.7})$$

$$\mathcal{F}_4 \triangleq \left\{ (a, b, c) \in \mathbb{R}^3 | c \leq -\alpha a, |b| \leq \alpha \frac{a-\alpha c}{1+\alpha^2} \right\} \quad (\text{A.8})$$

$$\mathcal{A}_1 \triangleq \left\{ (a, b, c) | b+c \geq -\frac{a}{\alpha}, c \geq \alpha \frac{a+\alpha b}{1+\alpha^2}, b \geq \alpha \frac{a+\alpha c}{1+\alpha^2} \right\} \quad (\text{A.9})$$

$$\mathcal{A}_2 \triangleq \left\{ (a, b, c) | c-b \geq -\frac{a}{\alpha}, c \geq \alpha \frac{a-\alpha b}{1+\alpha^2}, b \leq -\alpha \frac{a+\alpha c}{1+\alpha^2} \right\} \quad (\text{A.10})$$

$$\mathcal{A}_3 \triangleq \left\{ (a, b, c) | b+c \leq \frac{a}{\alpha}, c \leq -\alpha \frac{a+\alpha b}{1+\alpha^2}, b \leq -\alpha \frac{a+\alpha c}{1+\alpha^2} \right\} \quad (\text{A.11})$$

$$\mathcal{A}_4 \triangleq \left\{ (a, b, c) | c-b \leq \frac{a}{\alpha}, c \leq -\alpha \frac{a+\alpha b}{1+\alpha^2}, b \geq \alpha \frac{a-\alpha c}{1+\alpha^2} \right\} \quad (\text{A.12})$$

Une représentation plane de ces espaces est donnée dans les figures ci dessous. La figure 1 représente le plan $x = cste \geq 0$ et la figure 2 représente le plan $x = cste \leq 0$

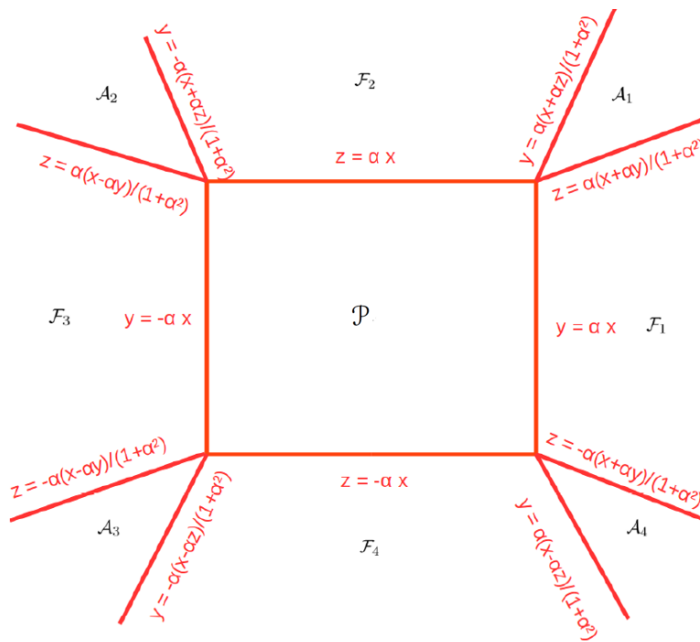


FIGURE A.1 – Représentation graphique d'une tranche à $x = x_0 > 0$ du cône \mathcal{P} et des sous ensembles qui définissent, par morceaux, la solution dans le cas x positif

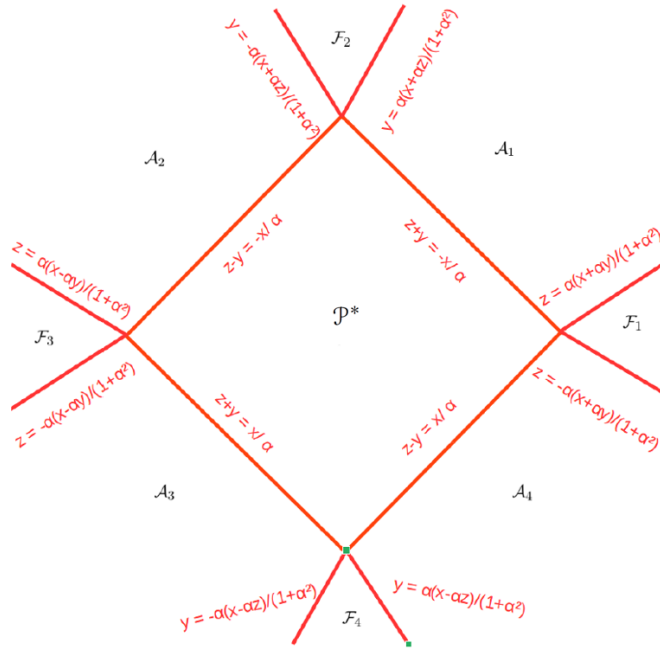


FIGURE A.2 – Représentation graphique d'une tranche à $x = x_0 < 0$ du cône polaire \mathcal{P}^* des sous ensembles qui définissent, par morceaux, la solution dans le cas x négatif

A.4 Démonstration

Résoudre le problème (2) équivaut à trouver la plus courte distance séparant le point (a, b, c) du \mathcal{P} . On distingue 4 cas :

1. l'ensemble des points (a, b, c) qui appartiennent à \mathcal{P} ,
2. l'ensemble des points (a, b, c) qui se projettent sur l'une des faces de \mathcal{P} ,
3. l'ensemble des points (a, b, c) qui se projettent sur l'une des arêtes de \mathcal{P} ,
4. l'ensemble des points (a, b, c) qui se projettent sur le sommet de \mathcal{P} .

A.4.1 Équations d'optimalité

Expression

Pour montrer ce résultat, il suffit de vérifier que le point (x^*, y^*, z^*) satisfait les conditions d'optimalité d'ordre 1 définies ci-dessous [BERTSEKAS \[2016\]](#) :

$$\begin{cases} (x^*, y^*, z^*) \in \mathcal{P} \\ \partial f_x(x^*, y^*, z^*)(x - x^*) + \partial f_y(x^*, y^*, z^*)(y - y^*) + \partial f_z(x^*, y^*, z^*)(z - z^*) \geq 0 \quad \forall (x, y, z) \in \mathcal{P} \end{cases} \quad (\text{A.13})$$

pour

$$f(x, y, z) = (x - a)^2 + (y - b)^2 + (z - c)^2. \quad (\text{A.14})$$

C'est à dire, vérifier que :

$$\begin{cases} (x^*, y^*, z^*) \in \mathcal{P} \\ (x^* - a)(x - x^*) + (y^* - b)(y - y^*) + (z^* - c)(z - z^*) \geq 0 \quad \forall (x, y, z) \in \mathcal{P} \end{cases} \quad (\text{A.15})$$

Symétrie

Soit (x^*, y^*, z^*) une solution du problème de minimisation (2). C'est à dire :

$$proj_{\mathcal{P}}(a, b, c) = (x^*, y^*, z^*)$$

A partir des équations d'optimalité, on déduit que :

$$proj_{\mathcal{P}}(a, c, b) = (x^*, z^*, y^*) \quad (\text{A.16})$$

$$proj_{\mathcal{P}}(a, -b, c) = (x^*, -y^*, z^*) \quad (\text{A.17})$$

$$proj_{\mathcal{P}}(a, -c, b) = (x^*, -z^*, y^*) \quad (\text{A.18})$$

$$proj_{\mathcal{P}}(a, -b, -c) = (x^*, -y^*, -z^*) \quad (\text{A.19})$$

$$proj_{\mathcal{P}}(a, b, -c) = (x^*, y^*, -z^*) \quad (\text{A.20})$$

Proof: On se contente de montrer l'équation (A.16) car le raisonnement est similaire pour les autres cas.

Soit (a, b, c) un point de \mathbb{R}^3 . On pose :

$$\begin{aligned} (x_1^*, y_1^*, z_1^*) &= proj_{\mathcal{P}}(a, c, b) \\ (x^*, y^*, z^*) &= proj_{\mathcal{P}}(a, b, c) \end{aligned}$$

Alors les 2 solutions (x_1^*, y_1^*, z_1^*) et (x^*, y^*, z^*) vérifient les conditions d'optimalités ci dessous :

$$\begin{cases} (x_1^*, y_1^*, z_1^*) \in \mathcal{P} \\ (x_1^* - a)(x - x_1^*) + (y_1^* - c)(y - y^*) + (z_1^* - b)(z - z_1^*) \geq 0 \quad \forall (x, y, z) \in \mathcal{P} \end{cases} \quad (\text{A.21})$$

et

$$\begin{cases} (x^*, y^*, z^*) \in \mathcal{P} \\ (x^* - a)(x - x^*) + (y^* - b)(y - y^*) + (z^* - c)(z - z^*) \geq 0 \quad \forall (x, y, z) \in \mathcal{P} \end{cases} \quad (\text{A.22})$$

L'équation (A.21) est équivalente à :

$$\begin{cases} (x_1^*, z_1^*, y_1^*) \in \mathcal{P} \\ (x_1^* - a)(x - x_1^*) + (z_1^* - b)(y - z_1^*) + (y_1^* - c)(z - y_1^*) \geq 0 \quad \forall (x, y, z) \in \mathcal{P} \end{cases} \quad (\text{A.23})$$

car y et z ont un rôle symétrique.

Les conditions (A.22) et (A.23) permettent d'identifier la relation entre ces 2 solutions :

$$\begin{cases} x_1^* = x^* \\ y_1^* = z^* \\ z_1^* = y^* \end{cases}$$

■

A.4.2 Cas des points appartenant au cône \mathcal{P}

Soit (a, b, c) un point du cône \mathcal{P} . Le point (a, b, c) vérifie bien les conditions d'optimalité car en posant :

$$\begin{cases} x^* = a \\ y^* = b \\ z^* = c \end{cases}$$

on a :

$$(x^* - a)(x - x^*) + (y^* - b)(y - y^*) + (z^* - c)(z - z^*) = 0 \quad \forall (x, y, z) \in \mathcal{P}.$$

A.4.3 Cas d'une projection sur une face

Ce cas correspond aux cas \mathcal{F}_i . On fait la démonstration dans le cas \mathcal{F}_1 et on généralisera au cas des autres faces. On rappelle que \mathcal{F}_1 est défini par :

$$\mathcal{F}_1 \triangleq \left\{ (a, b, c) \in \mathbb{R}^3 \mid b \geq \alpha a, |c| \leq \alpha \frac{a + \alpha b}{1 + \alpha^2} \right\}. \quad (\text{A.24})$$

On veut montrer que pour tout point (a, b, c) de \mathcal{F}_1 , le point

$$\begin{aligned} x^* &= \frac{a + \alpha b}{1 + \alpha^2} \\ y^* &= \alpha x^* \\ z^* &= c \end{aligned} \quad (\text{A.25})$$

satisfait les conditions d'optimalité (A.15).

Montrons d'abord que $(x^*, y^*, z^*) \in \mathcal{P}$

On a :

- $x^* \geq 0$ car $|c| \leq \alpha x^*$
- $|y^*| \leq \alpha x^*$ car $y^* = \alpha x^*$
- $|z^*| \leq \alpha x^*$ car $z^* = c$ et $|c| \leq \alpha x^*$

donc (x^*, y^*, z^*) appartient à \mathcal{P} .

Passons maintenant à l'inégalité, on s'appuie d'abord sur la propriété suivante

Propriété 1 La projection orthogonale de tout point (a, b, c) sur le plan $y = \alpha x$ s'écrit

$$\begin{aligned} x^* &= \frac{a + \alpha b}{1 + \alpha^2} \\ y^* &= \alpha x^* \\ z^* &= c \end{aligned} \quad (\text{A.26})$$

Proof: On remplace y par αx dans la définition de f (A.14), le problème devient non contraint sur les deux variables x et z :

$$f(x, z) = (x - a)^2 + (\alpha x - b)^2 + (z - c)^2 \quad (\text{A.27})$$

On annule le gradient de f pour tirer l'expression de x^* et z^*

$$\begin{cases} \partial f_x(x^*, z^*) = 0 \\ \partial f_z(x^*, z^*) = 0 \end{cases} \quad (\text{A.28})$$

Cela est équivalent à :

$$(x^* - a) + \alpha(\alpha x^* - b) = 0 \quad \text{et} \quad z^* - c = 0 \quad (\text{A.29})$$

Conclusion :

$$\begin{cases} x^* = \frac{a + \alpha b}{1 + \alpha^2} \\ y^* = \alpha x^* \text{ car on est sur la face } y = \alpha x \\ z^* = c \end{cases} \quad (\text{A.30})$$

■

Cette propriété implique que

Propriété 2

$$(x^* - a)x^* + (y^* - b)y^* + (z^* - c)z^* = 0 \quad (\text{A.31})$$

Proof: On a :

$$\begin{aligned}
 (x^* - a)x^* + (y^* - b)y^* + (z^* - c)z^* &= x^*((x^* - a) + \alpha(\alpha x^* - b)) \quad \text{car } z^* = c \text{ et } y^* = \alpha x^* \\
 &= x^*(x^*(1 + \alpha^2) - (a + \alpha b)) \\
 &= x^*((a + \alpha b) - (a + \alpha b)) \quad \text{car } x^* = \frac{a + \alpha b}{1 + \alpha^2} \\
 &= 0
 \end{aligned} \tag{A.32}$$

Donc la condition d'optimalité (A.15) se simplifie en :

$$(x^* - a)x + (y^* - b)y + (z^* - c)z \geq 0 \quad \forall (x, y, z) \in \mathcal{P} \tag{A.33}$$

Soit (x, y, z) un point du cône \mathcal{P} et (a, b, c) un point de \mathcal{F}_1 , on a :

$$\begin{cases} x^* - a = \frac{\alpha(b - \alpha a)}{1 + \alpha^2} \\ y^* - b = -\frac{(b - \alpha a)}{1 + \alpha^2} \end{cases} \tag{A.34}$$

Donc,

$$(x^* - a) = -\alpha(y^* - b). \tag{A.35}$$

Cela implique que :

$$\begin{aligned}
 (x^* - a)x + (y^* - b)y + (z^* - c)z &= -\alpha x(y^* - b) + y(y^* - b) \\
 &= (y - \alpha x)(y^* - b)
 \end{aligned} \tag{A.36}$$

Or $y - \alpha x \leq 0$ puisque (x, y, z) est dans \mathcal{P} et $y^* - b \leq 0$ car $b \geq \alpha a$.

Donc

$$(x^* - a)x + (y^* - b)y + (z^* - c)z \geq 0 \tag{A.37}$$

■

Généralisation aux autres facettes

Commençons par le cas \mathcal{F}_2

Soit (a, b, c) un point de \mathcal{F}_2 et (x^*, y^*, z^*) sa projection sur le cône \mathcal{P} .

Il est clair que le point (a, c, b) appartient à \mathcal{F}_1

Soit (x_1^*, y_1^*, z_1^*) la projection du point (a, c, b) sur le cône \mathcal{P}

alors :

$$\begin{aligned}
 x_1^* &= \frac{a + \alpha c}{1 + \alpha^2} \\
 y_1^* &= \alpha x_1^* \\
 z_1^* &= b
 \end{aligned}$$

Or d'après l'équation (A.16) :

$$(x_1^*, y_1^*, z_1^*) = (x^*, z^*, y^*) \tag{A.38}$$

D'où :

$$\begin{aligned}
 x^* &= \frac{a + \alpha c}{1 + \alpha^2} \\
 y^* &= b \\
 z^* &= \alpha x^*
 \end{aligned}$$

On peut donc déduire les solutions du cas \mathcal{F}_2 à partir du cas \mathcal{F}_1 en remplaçant b par c et en permutant les deux dernières composantes.

Passons maintenant au cas \mathcal{F}_3
 Soit (a, b, c) un point de \mathcal{F}_3 et (x^*, y^*, z^*) sa projection sur le cône \mathcal{P} .
 Le point $(a, -b, c)$ appartient à \mathcal{F}_1
 Soit (x_1^*, y_1^*, z_1^*) la projection du point $(a, -b, c)$ sur le cône \mathcal{P}
 alors :

$$\begin{aligned} x_1^* &= \frac{a-\alpha b}{1+\alpha^2} \\ y_1^* &= \alpha x_1^* \\ z_1^* &= c \end{aligned}$$

Or d'après l'équation (A.17) :

$$(x_1^*, y_1^*, z_1^*) = (x^*, -y^*, z^*) \quad (\text{A.39})$$

D'où :

$$\begin{aligned} x^* &= \frac{a-\alpha b}{1+\alpha^2} \\ y^* &= -\alpha x^* \\ z^* &= c \end{aligned}$$

On peut donc déduire les solutions du cas \mathcal{F}_3 à partir du cas \mathcal{F}_1 en remplaçant b par $-b$ et en remplaçant la deuxième composante par son opposée.

De façon similaire, on peut également déduire l'expression de la projection dans le cas \mathcal{F}_4 à partir du cas \mathcal{F}_2 en remplaçant c par $-c$ et en remplaçant la troisième composante par son opposée.

A.4.4 Cas d'une projection sur une arête

Soit l'espace \mathcal{A}_1 défini par :

$$\mathcal{A}_1 \triangleq \left\{ (a, b, c) \mid b + c \geq -\frac{a}{\alpha}, c \geq \alpha \frac{a + \alpha b}{1 + \alpha^2}, b \geq \alpha \frac{a + \alpha c}{1 + \alpha^2} \right\} \quad (\text{A.40})$$

On veut montrer que pour tout point de \mathcal{A}_1 , le point

$$\begin{aligned} x^* &= \frac{a + \alpha b + \alpha c}{1 + 2\alpha^2} \\ y^* &= \alpha x^* \\ z^* &= \alpha x^* \end{aligned} \quad (\text{A.41})$$

satisfait les les conditions d'optimalités définies dans (A.15).

Montrons d'abord que $(x^*, y^*, z^*) \in \mathcal{P}$

On a :

- $x^* \geq 0$ car $b + c \geq -\frac{a}{\alpha}$
- $|y^*| \leq \alpha x^*$ car $y^* = \alpha x^*$
- $|z^*| \leq \alpha x^*$ car $z^* = \alpha x^*$

donc (x^*, y^*, z^*) appartient à \mathcal{P}

Passons maintenant à l'inégalité : On s'appuie d'abord sur la propriété suivante :

Propriété 3 *La projection orthogonale de tout point (a, b, c) sur l'arête définie par $(y = \alpha x$ et $z = \alpha x)$ s'écrit :*

$$\begin{aligned} x^* &= \frac{a + \alpha b + \alpha c}{1 + 2\alpha^2} \\ y^* &= \alpha x^* \\ z^* &= \alpha x^* \end{aligned} \quad (\text{A.42})$$

Proof: On remplace y par αx et z par αx dans la définition de f (A.14), le problème devient non contraint sur la variable x

$$f(x) = (x - a)^2 + (\alpha x - b)^2 + (\alpha x - c)^2 \quad (\text{A.43})$$

On annule la dérivée première de f pour tirer l'expression de x^* ■

$$\frac{df}{dx}(x^*) = 0 \quad (\text{A.44})$$

Cela est équivalent à :

$$(x^* - a) + \alpha(\alpha x^* - b) + \alpha(\alpha x^* - c) = 0 \quad (\text{A.45})$$

Conclusion :

$$\begin{cases} x^* = \frac{a + \alpha b + \alpha c}{1 + 2\alpha^2} \\ y^* = \alpha x^* \\ z^* = \alpha x^* \end{cases} \quad (\text{A.46})$$

Cette propriété implique que

Propriété 4

$$(x^* - a)x^* + (y^* - b)y^* + (z^* - c)z^* = 0 \quad (\text{A.47})$$

Proof: On a :

$$\begin{aligned} (x^* - a)x^* + (y^* - b)y^* + (z^* - c)z^* &= x^*((x^* - a) + \alpha(\alpha x^* - b) + \alpha(\alpha x^* - c)) \quad \text{car } z^* = y^* = \alpha x^* \\ &= x^*(x^*(1 + 2\alpha^2) - (a + \alpha b + \alpha c)) \\ &= x^*((a + \alpha b + \alpha c) - (a + \alpha b + \alpha c)) \quad \text{car } x^* = \frac{a + \alpha b + \alpha c}{1 + 2\alpha^2} \\ &= 0 \end{aligned} \quad (\text{A.48})$$

Donc la condition d'optimalité (A.15) se simplifie en

$$(x^* - a)x + (y^* - b)y + (z^* - c)z \geq 0 \quad \forall (x, y, z) \in \mathcal{P} \quad (\text{A.49})$$

Soit (x, y, z) un point du cône \mathcal{P} et (a, b, c) un point de \mathcal{A}_1

On a :

$$\begin{cases} x^* - a = -\alpha \frac{(\alpha x - b) + (\alpha x - c)}{1 + 2\alpha^2} \\ y^* - b = -\frac{(\alpha x - b) + (\alpha^2(c - b))}{1 + 2\alpha^2} \\ z^* - c = -\frac{(\alpha x - c) + (\alpha^2(b - c))}{1 + 2\alpha^2} \end{cases} \quad (\text{A.50})$$

Donc

$$(x^* - a) = -\alpha((y^* - b) + (z^* - c)) \quad (\text{A.51})$$

Cela implique que :

$$\begin{aligned} (x^* - a)x + (y^* - b)y + (z^* - c)z &= -\alpha x(y^* - b) - \alpha x(z^* - c) + y(y^* - b) + z(z^* - c) \\ &= (y - \alpha x)(y^* - b) + (z - \alpha x)(z^* - c) \end{aligned} \quad (\text{A.52})$$

Or, d'après (A.40) :

$$\begin{cases} c \geq \alpha \frac{a + \alpha b}{1 + \alpha^2} \\ b \geq \alpha \frac{a + \alpha c}{1 + \alpha^2} \end{cases} \quad (\text{A.53})$$

donc

$$\begin{cases} -((a\alpha - c) + \alpha^2(b - c)) \leq 0 \\ -((a\alpha - b) + \alpha^2(c - b)) \leq 0 \end{cases} \quad (\text{A.54})$$

Ainsi

$$\begin{cases} y^* - b \leq 0 \\ z^* - c \leq 0 \end{cases} \quad (\text{A.55})$$

On a également : $y - \alpha x \leq 0$ et $z - \alpha x \leq 0$ puisque (x, y, z) est dans \mathcal{P}

Par conséquent :

$$(x^* - a)x + (y^* - b)y + (z^* - c)z \geq 0 \quad (\text{A.56})$$

Généralisation aux autres arrêtes

Commençons par le cas \mathcal{A}_2

Soit (a, b, c) un point de \mathcal{A}_2 et (x^*, y^*, z^*) sa projection sur le cône \mathcal{P} .

Il est clair que le point $(a, -b, c)$ appartient à \mathcal{A}_1

Soit (x_1^*, y_1^*, z_1^*) la projection du point $(a, -b, c)$ sur le cône \mathcal{P}

alors :

$$\begin{aligned} x_1^* &= \frac{a - \alpha b + \alpha c}{1 + 2\alpha^2} \\ y_1^* &= \alpha x_1^* \\ z_1^* &= \alpha x_1^* \end{aligned}$$

Or d'après l'équation (8) :

$$(x_1^*, y_1^*, z_1^*) = (x^*, -y^*, z^*) \quad (\text{A.57})$$

D'où :

$$\begin{aligned} x^* &= \frac{a - \alpha b + \alpha c}{1 + 2\alpha^2} \\ y^* &= -\alpha x^* \\ z^* &= \alpha x^* \end{aligned}$$

On peut donc déduire les solutions du cas \mathcal{A}_2 à partir du cas \mathcal{A}_1 en remplaçant b par $-b$ et en remplaçant la deuxième composante par son opposée.

Passons maintenant au cas \mathcal{A}_3 :

Soit (a, b, c) un point de \mathcal{A}_3 et (x^*, y^*, z^*) sa projection sur le cône \mathcal{P} .

Il est clair que le point $(a, -b, -c)$ appartient à \mathcal{A}_1

Soit (x_1^*, y_1^*, z_1^*) la projection du point $(a, -b, -c)$ sur le cône \mathcal{P}

alors :

$$\begin{aligned} x_1^* &= \frac{a - \alpha b - \alpha c}{1 + 2\alpha^2} \\ y_1^* &= \alpha x_1^* \\ z_1^* &= \alpha x_1^* \end{aligned}$$

Or d'après l'équation (8) :

$$(x_1^*, y_1^*, z_1^*) = (x^*, -y^*, -z^*) \quad (\text{A.58})$$

D'où :

$$\begin{aligned} x^* &= \frac{a - \alpha b - \alpha c}{1 + 2\alpha^2} \\ y^* &= -\alpha x^* \\ z^* &= -\alpha x^* \end{aligned}$$

On peut donc déduire les solutions dans le cas \mathcal{A}_3 à partir du cas \mathcal{A}_1 en remplaçant b par $-b$ et c par $-c$ et en remplaçant les 2 dernière composantes par leurs opposées.

On peut aussi déduire l'expression de la projection dans le cas \mathcal{A}_4 à partir du cas \mathcal{A}_1 en remplaçant c par $-c$ et en remplaçant la dernière composante par son opposée.

A.4.5 Cas de la projection sur le sommet du cône

Soit l'espace défini par \mathcal{P}^* :

$$\mathcal{P}^* = \left\{ (a, b, c), |b + c| \leq -\frac{a}{\alpha}, |b - c| \leq -\frac{a}{\alpha} \right\} \quad (\text{A.59})$$

On veut montrer que pour tout point de \mathcal{P}^* , le point $(0, 0, 0)$ satisfait les conditions d'optimalités définies dans (A.15).

Soit (x, y, z) un point du cône \mathcal{P} et (a, b, c) un point de \mathcal{P}^* . En remplaçant (x^*, y^*, z^*) par $(0, 0, 0)$ dans (A.15), la condition d'optimalité se simplifie en

$$ax + by + cz \leq 0. \quad (\text{A.60})$$

On sait d'après [BERTSEKAS, 2016, Prop B.13 Appendice B] que l'ensemble de points (a, b, c) qui vérifient cette propriété est appelé « cône polaire » de \mathcal{P} .

Donc pour prouver notre résultat, il suffit de montrer que \mathcal{P}^* est le cône polaire de \mathcal{P} .

Détermination du cône polaire

Le théorème de Minkowski-Weyl permet de caractériser le cône polaire de tout cône à partir d'une description générative finie de ce dernier à savoir que si un cône est décrit par

$$\left\{ (x, y, z) \mid (x, y, z)^t = \sum_{j=1}^r \beta_j \mathbf{u}_j, \quad \beta_j \geq 0, \quad j = 1, \dots, r \right\} \quad (\text{A.61})$$

pour des vecteurs $\mathbf{u}_1, \dots, \mathbf{u}_r$ donnés (appelés vecteurs générateurs) alors le cône polaire est déterminé par

$$\{ \mathbf{a} \in \mathbb{R}^3 \mid \mathbf{a}^t \mathbf{u}_j \leq 0, \quad j = 1, \dots, r \}. \quad (\text{A.62})$$

Il nous faut donc trouver une description générative du cône \mathcal{P} , c'est à dire caractériser les vecteurs générateurs du cône \mathcal{P} .

Dans notre cas \mathcal{P} est généré par les 4 vecteurs ci-dessous (ce sont les vecteurs directeurs des arêtes du cône) :

$$\begin{cases} \mathbf{u}_1 = (1, \alpha, \alpha)^t \\ \mathbf{u}_2 = (1, -\alpha, \alpha)^t \\ \mathbf{u}_3 = (1, \alpha, -\alpha)^t \\ \mathbf{u}_4 = (1, -\alpha, -\alpha)^t \end{cases} \quad (\text{A.63})$$

car $\forall (x, y, z) \in \mathcal{P}$ on peut toujours trouver un quadruplet $(\beta_1, \beta_2, \beta_3, \beta_4)$ positif tel que :

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ \alpha & -\alpha & \alpha & -\alpha \\ \alpha & \alpha & -\alpha & -\alpha \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (\text{A.64})$$

en posant :

$$\begin{cases} \beta_1 = \frac{x}{4} \left(1 + \frac{y}{\alpha x}\right) \left(1 + \frac{z}{\alpha x}\right) \\ \beta_2 = \frac{x}{4} \left(1 - \frac{y}{\alpha x}\right) \left(1 + \frac{z}{\alpha x}\right) \\ \beta_3 = \frac{x}{4} \left(1 + \frac{y}{\alpha x}\right) \left(1 - \frac{z}{\alpha x}\right) \\ \beta_4 = \frac{x}{4} \left(1 - \frac{y}{\alpha x}\right) \left(1 - \frac{z}{\alpha x}\right) \end{cases} \quad (\text{A.65})$$

Noter que la positivité des β_j est due au fait que x est positif et y et z sont inférieurs à αx .
En conclusion, le cône polaire de \mathcal{P} est l'espace défini par :

$$\begin{cases} a + \alpha b + \alpha c \leq 0 \\ a - \alpha b + \alpha c \leq 0 \\ a + \alpha b - \alpha c \leq 0 \\ a - \alpha b - \alpha c \leq 0 \end{cases} \quad (\text{A.66})$$

avec $(a, b, c) \in \mathbb{R}^3$, ou, de manière concise, l'espace défini par (A.59).

A.5 Références

BERTSEKAS, D. 2016, *Nonlinear Programming*, Athena scientific optimization and computation series, Athena Scientific, ISBN 9781886529052. [III](#), [X](#)

Titre: Détection et localisation de particules dans des images PIV via des approches parcimonieuses à grille.

Mots clés: PIV, détection/localisation, discrétisation, dictionnaire d'interpolation, parcimonie.

Résumé: Cette thèse trouve son application dans le domaine de la vélocimétrie par image de particules (PIV). L'objectif principal est la détection/localisation 2D de particules en contexte de fortes densités et donc de forts chevauchements. On considère des approches qui discrétisent la relation entre les caractéristiques des particules et l'image observée et résolvent le problème inverse associé avec des techniques de parcimonie. Dans ce contexte l'algorithme NNLS a montré une très bonne capacité à résoudre les chevauchements avec un a priori de positivité, pourvu que l'on adopte des pas de discrétisation de l'ordre de $\frac{1}{20}$ pixel. Cependant, cela entraîne un dictionnaire très volumineux et par conséquent un coût de calcul et de mémoire très importants. Pour réduire la taille du dictionnaire, nous avons considéré une discrétisation alternative basée sur l'approximation de Taylor d'ordre 1, à laquelle on a associé dans un premier temps une contrainte de parcimonie l_1 et dans un second temps une contrainte de parcimonie l_0 . Dans le cas l_1 , nous avons résolu le problème d'optimisation induit (CBP) en utilisant une méthode proximale, puis une méthode d'active set. L'implémentation de cette dernière a été grandement simplifiée après reparamétrisation de CBP en un problème avec des contraintes de positivité. Cependant, cette approche n'a pas de gain significatif sur la qualité de détection et de localisation ou sur le temps de calcul, en présence de fortes interactions de particules, par rapport à une méthode de régularisation basique BP utilisant une approximation grossière. Nous avons également utilisé la nouvelle reparamétrisation pour construire l'algorithme CNNOMP basé sur des structures gloutonnes de type OMP non négatif. En adjoignant à chaque itération de cet algorithme une minimisation non linéaire du résidu, on obtient le KCNNOMP qui a une capacité de résolution de chevauchements à fort SNR avec des pas grossiers de discrétisation. Nous avons proposé et démontré l'intérêt d'une stratégie de réglage unique pour tous les types d'algorithmes étudiés et leurs paramètres (pas de discrétisation, paramètres de régularisation) par maximisation du Fscore. Cette stratégie nous a permis d'identifier deux domaines d'emplois : à fort SNR, KCNNOMP atteint des performances de détections identiques à celles de NNLS avec un dictionnaire 25 fois plus petit. A faible SNR, la régularisation l_1 fournit les meilleures performances avec des grilles assez grossières dont le pas est de l'ordre de $\frac{1}{3}$ pixel.

Title: Particle detection and localization in PIV images with grid-based sparsity techniques.

Keywords: PIV, detection and localization, discretization, interpolation dictionary, sparsity.

Abstract: This thesis finds its application in the field of particle image velocimetry (PIV). The main goal is to identify and localize particles in high density conditions. The approaches considered here, discretise the relationship between particle characteristics and the observed image, and solve the associated inverse problem with sparsity techniques. In this context, NNLS algorithm has shown a very good ability to handle overlapping with positivity prior, provided that the adopted grid size step is equal to $\frac{1}{20}$ pixel. However, this results in a very large dictionary which raises memory and computational issues. To reduce the size of the dictionary, we've considered an alternative discretization based on first order Taylor approximation, to which we've associated an l_1 sparse constraint and an l_0 sparse constraint. In the case of l_1 sparsity, we've solved the related optimization problem (CBP) using a proximal method and then with an active set method. The implementation of the latter has been greatly simplified after the reparameterization of CBP into an optimization problem with positivity constraints. However, this approach does not yield a significant gain in detection and localization quality or in computation time, in the presence of high particle interactions, compared to the basic regularized method BP which uses a coarse approximation. We've also used the new reparameterization to build the CNNOMP algorithm, which is based on greedy structures similar to non-negative OMP method. By adding to each iteration of this algorithm a non-linear minimization of the residual, we obtain the KCNNOMP algorithm. The latter has a great capability to handle particle overlapping in high SNR conditions, with coarse discretization steps. We've also proposed and demonstrated the interest of a unique setting strategy for all types of algorithms studied here and their parameters (discretization steps, regularization parameters), by maximizing the Fscore metric. This strategy allowed us to identify two areas of use : when the SNR is high, KCNNOMP achieves identical detection performance to that of NNLS with a dictionary 25 smaller. When the SNR is low, the l_1 regularisation provides the best performance with coarse grids whose grid step size is around $\frac{1}{3}$ pixel.