



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut Supérieur de l'Aéronautique et de l'Espace

Présentée et soutenue par :

Marie LASSERRE

le lundi 20 novembre 2017

Titre :

Estimation non-ambigüe de cibles grâce à une représentation parcimonieuse
Bayésienne d'un signal radar large bande

École doctorale et discipline ou spécialité :

ED MITT : Signal, image, acoustique et optimisation

Unité de recherche :

Équipe d'accueil ISAE-ONERA SCANR

Directeur(s) de Thèse :

Mme Stéphanie BIDON (directrice de thèse)

Jury :

Mme Sylvie MARCOS Directrice de Recherche L2S - Présidente

Mme Stéphanie BIDON Professeur ISAE-SUPAERO - Directrice de thèse

M. Guillaume GINOLHAC Professeur Université Savoie Mont Blanc - Rapporteur

M. Pascal LARZABAL Professeur IUT Cachan - Rapporteur

M. François LE CHEVALIER Professeur TU Delft

M. Jean-Philippe OVARLEZ Maître de recherche ONERA - Rapporteur

M. Philippe POULIGUEN Responsable domaine scientifique DGA/MRIS

Remerciements

Je remercie en premier lieu MM. Guillaume Ginolhac, Pascal Larzabal et Jean-Philippe Ovarlez pour avoir accepté de rapporter cette thèse, et également Mme Sylvie Marcos qui a présidé le jury.

Je remercie également très sincèrement ma directrice de thèse Stéphanie Bidon. Lors de ces quatre années, Stéphanie a toujours été présente pour répondre à mes interrogations, et exigeante quant au travail rendu. Je lui en suis très reconnaissante et je suis certaine que je n'aurais pas accompli ce travail sans la rigueur de son encadrement.

Je souhaite également remercier Olivier Besson, qui a encadré mes travaux de thèse lors des deux premières années. Je regrette que cet encadrement ait été interrompu car je suis convaincue que mes travaux auraient grandement bénéficié de son expertise.

J'exprime ma gratitude à la Direction Générale de l'Armement qui a financé entièrement mon contrat de thèse, et plus particulièrement M. Philippe Pouliguen qui a participé à mon jury de thèse.

Je tiens par ailleurs à remercier M. François Le Chevalier, qui a participé à mon jury de thèse, mais qui m'a surtout permis de collaborer avec l'équipe MS³ de l'université de Delft. Merci donc à M. Alexander Yarovoy de m'avoir accueillie dans son équipe lors de plusieurs séjours, M. Oleg Krasnov pour son expertise sur le radar PARSAX, et bien entendu M. Nikita Petrov pour nos fructueux échanges!

J'en viens maintenant à ceux qu'on ne remercie que le dernier jour alors que ce sont eux qui nous font sourire, tous les jours. Merci donc aux collègues/amis que j'ai côtoyés pendant ces 4 ans à l'ISAE. Les premiers rencontrés ce sont bien sûr les camarades de l'ENSICA, avec qui j'ai passé 18 mois dans le cocon du DAS : Joël, Romain, François, Guillaume, Matthieu, Yves, Laurent, Gilles, Valérie, Flavio (mon 1er co-bureau), Gaël, Arthur, Jacobo, Saïd et Souleymane. Non, je n'oublie pas Michèle! Quels bons moments nous avons passés accompagnés des écureuils, même si nous savions que le temps était compté... Mais c'est grâce au déménagement à Supaéro que j'ai pu rencontrer Nabil le combattant, Antoine le magicien, Pierrick l'aventurier et l'illustre Francesco. Grâce à vous, je garde de très bons souvenirs de ces années de thèse, et j'espère sincèrement continuer à alimenter ces souvenirs.

Mais non, je n'oublie pas l'unique M. Vincent Laquerbe! Tout avait pourtant si bien commencé, chacun dans son bureau, le calme plat dans les couloirs vides du DAS... De temps à autre, quelques coups de sang, l'orchidée qui vacille dans son pot, mais rien de bien méchant. Il a pourtant suffi d'un déménagement pour tout bouleverser, et c'est là que les Athéniens s'atteignirent! Pauvres Nabil et Antoine qui ont dû supporter nos éclats de voix, concours de patois, sifflements, chansons, bricolages et j'en passe. Oui je crois qu'on a un peu tout traversé en 3 ans, mais ensemble on va plus loin, non?

Pour finir, je remercie ma famille, et en particulier mes parents qui m'ont toujours accompagnée, sans jamais me diriger ni s'opposer à mes choix. Il n'a jamais été question de vanter les diplômes et les années d'études, pourtant à nous deux Matou on arrive presque à 20 ans ! C'est sans doute ce qui m'a fait garder les pieds sur terre et ne pas me prendre pour ce que je ne suis pas. Merci Papa et Maman, si j'en suis là aujourd'hui c'est évidemment grâce à vous.

Enfin, je remercie de tout coeur « Mon Vincent », avec qui j'ai déjà parcouru un long bout de chemin, sur lequel il m'a tour à tour poussée, tirée, relevée et toujours encouragée. J'espère continuer ensemble cette belle randonnée, qui se poursuit sur les sentiers d'Empeaux.

Acronymes

APES *Amplitude and Phase Estimation of a Sinusoid*
AR *Auto-Regressive*
BP *Basis Pursuit*
BPDN *Basis Pursuit De-Noising*
CDF *Cumulative Distribution Function*
CNR *Clutter-to-Noise Ratio*
CPI *Coherent Processing Interval*
DL1 *Développement limité d'ordre 1*
EM *Expectation-Maximization*
FIR *Finite Impulse Response*
IAA *Iterative Adaptive Algorithm*
iW-Capon *Iterative Wideband Capon*
MAP *Maximum A Posteriori*
MC *Monte-Carlo*
MCMC *Monte-Carlo Markov Chain*
MH *Metropolis-Hastings*
MiTI *Migrating Target Indicator*
ML *Maximum Likelihood*
MMSE *Minimum Mean Square Error*
MP *Matching Pursuit*
(n)MSE *(normalized) Mean Square Error*
OGSBI *Off-Grid Sparse Bayesian Inference*
OMP *Orthogonal Matching Pursuit*
PARSAX *Polarimetric Agile Radar in S- and X-band*
PCGS *Partially Collapsed Gibbs Sampler*
PRF *Pulse Repetition Frequency*
PSD *Power Spectral Density*
RVM *Relevance Vector Machine*
SAR *Synthetic Aperture Radar*
SINR *Signal-to-Interference-plus-Noise Ratio*
SNR *Signal-to-Noise Ratio*
SSR *Sparse Signal Representations*
W-APES *Wideband APES*
W-Capon *Wideband Capon*
W-IAA *Wideband IAA*
WN *White Noise*

Notations

Paramètres du signal

- c vitesse de la lumière dans le vide
- f_c fréquence porteuse
- B bande instantanée
- F_r fréquence de répétition des impulsions
- T_r période de répétition des impulsions
- T durée de l'impulsion
- δ_R résolution distance

Paramètres de traitement

- M Nombre d'impulsions considérées
- K Nombre de cases distances considérées
- μ Bande fractionnelle par sous-bande
- f_r Fréquence rapide normalisée
- f_D Fréquence Doppler normalisée
- d_a Distance d'ambiguïté
- v_a Vitesse d'ambiguïté
- n_{va} Facteur de dépliement (en vitesse)
- \mathbf{Y} Données exprimées dans le domaine fréquence-rapide / temps-long
- \mathbf{H} Dictionnaire définissant l'espace de SSR (taille $KM \times \bar{K}\bar{M}$)
- \mathbf{x} Vecteur parcimonieux à estimer (taille $\bar{K}\bar{M}$)
- \bar{M} Dimension du domaine vitesse discrétisé
- \bar{K} Dimension du domaine distance discrétisé
- N_t Nombre total de cibles
- \mathbf{u} Version vectorisée ligne-à-ligne de \mathbf{U}

Opérateurs mathématiques

- $\|\cdot\|_0$ Pseudo-norme ℓ_0 , abusivement appelée « norme ℓ_0 »
- $\|\cdot\|_1$ Norme ℓ_1

$\|\cdot\|_2$ Norme ℓ_2 , abusivement notée $\|\cdot\|$

⊙ Produit de Hadamart

$\pi(\boldsymbol{\theta})$ Distribution a priori de $\boldsymbol{\theta}$

$f(\boldsymbol{\theta}|\mathbf{y})$ Distribution a posteriori de $\boldsymbol{\theta}$

Table des matières

Introduction	1
1 État de l’art : estimation non-ambigüe de cibles par un radar large bande	5
1.1 Limites de l’estimation/détection de cibles avec un signal bande étroite	6
1.1.1 Modèle de signal bande étroite	6
1.1.2 Filtre adapté en bruit blanc	7
1.1.3 Ambiguïtés distance / vitesse	8
1.2 Utilisation d’un signal radar large bande	10
1.2.1 Modèle de signal large bande	10
1.2.2 Sommation cohérente large bande	11
1.2.3 Estimation spectrale pour la forme d’onde large bande	12
1.3 Les représentations parcimonieuses	17
1.3.1 Mise en forme du problème	17
1.3.2 Approches déterministes de représentation parcimonieuse	22
1.3.3 Approches Bayésiennes de représentation parcimonieuse	24
1.4 Une approche Bayésienne pour la représentation parcimonieuse d’un signal radar large bande	30
1.4.1 Présentation de l’algorithme	31
1.4.2 Exemple de résultats	34
1.4.3 Robustesse de l’algorithme	35
2 Algorithme Bayésien de SSR pour les cibles en chevauchement	39
2.1 Le problème des cibles en chevauchement	40
2.2 Algorithme Bayésien de SSR pour les cibles en chevauchement	43
2.2.1 Paramétrisation du dictionnaire	44

2.2.2	Modèle hiérarchique Bayésien	45
2.2.3	Schéma d'estimation	45
2.2.4	Estimation des paramètres de désalignement (ϵ^v, ϵ^r)	49
2.2.5	Remarque : prise en compte du désalignement via un dictionnaire linéarisé	53
2.3	Simulations numériques	55
2.3.1	Exemple de réalisation	55
2.3.2	Simulations de Monte-Carlo	57
2.4	Simulations sur données expérimentales	61
2.5	Conclusions	67
3	Algorithme Bayésien de SSR pour des cibles en chevauchement dans du fouillis diffus	69
3.1	Estimation de cibles dans du fouillis diffus	70
3.1.1	Caractéristiques du fouillis de sol	70
3.1.2	Algorithme Bayésien de SSR pour fouillis diffus	72
3.1.3	Exemple de résultats	76
3.2	Modification de l'algorithme robuste au désalignement pour prendre en compte le fouillis diffus	77
3.2.1	Modification du modèle hiérarchique Bayésien	77
3.2.2	Modification du schéma d'estimation	78
3.3	Simulations numériques	82
3.3.1	Exemples de réalisation	82
3.3.2	Performances en fonction du SINR des cibles	83
3.4	Simulations sur données expérimentales	86
3.5	Conclusions	88
4	Estimation de scènes radar à forte dynamique de puissance	91
4.1	Mise en évidence du problème	92

4.2	Nouvelle distribution a priori et modèle hiérarchique Bayésien associé	93
4.2.1	Paramétrisation du vecteur d'amplitude des cibles \mathbf{x}	94
4.2.2	Distribution a priori des autres paramètres	95
4.2.3	Distribution a priori des hyperparamètres	95
4.3	Nouveau schéma d'estimation	96
4.3.1	Échantillonnage du vecteur amplitude $\boldsymbol{\rho}$	97
4.3.2	Échantillonnage du vecteur phase $\boldsymbol{\phi}$	99
4.3.3	Échantillonnage des autres paramètres et hyperparamètres	99
4.3.4	Implémentation de l'algorithme	100
4.4	Exemples de réalisations de l'algorithme avec classes	101
4.4.1	Données synthétiques	102
4.4.2	Données semi-expérimentales	103
4.5	Simulations de Monte-Carlo sur données synthétiques	104
4.5.1	Intérêt des classes de puissance	106
4.5.2	Choix des classes de puissance	108
4.5.3	Robustesse du réglage des classes par rapport au niveau de bruit	109
4.6	Conclusions	109
	Conclusion	113
	A Étude de la distribution conditionnelle a posteriori du désalignement	117
A.1	Reformulation de la distribution conditionnelle a posteriori	117
A.2	Approximation de la posterior par une loi normale	119
A.2.1	Choix de loi de proposition pour l'échantillonnage de $\boldsymbol{\varepsilon}$	120
	B Calcul de $\ \tilde{\mathbf{h}}_i(\boldsymbol{\varepsilon}_i^v, \boldsymbol{\varepsilon}_i^r)\ ^2$	123
	C Densités de probabilité utilisées	125

Liste des publications	127
Bibliographie	129

Table des figures

1.1	Exemple de sortie de filtre adapté en bruit blanc pour un signal bande étroite contenant une seule cible	8
1.2	Exemple de traitement multi-PRF	9
1.3	Exemple de sommation cohérente d'un signal radar large bande	12
1.4	Estimation spectrale d'une cible lente (sommation cohérente large bande, W-APES, W-Capon, iW-Capon, W-IAA)	18
1.5	Estimation spectrale d'une cible rapide (sommation cohérente large bande, W-APES, W-Capon, iW-Capon, W-IAA)	19
1.6	Modèle hiérarchique Bayésien de l'algorithme non-robustifié	32
1.7	Schéma d'estimation de l'algorithme non-robustifié	35
1.8	Carte distance-vitesse estimées par l'algorithme non-robustifié et l'OMP ; scénario synthétique <i>on-grid</i> bruit blanc	36
1.9	Carte distance-vitesse estimées par l'algorithme non-robustifié et l'OMP ; scénario semi-expérimental PARSAX <i>on-grid</i> bruit blanc	36
1.10	Carte distance-vitesse estimées par l'algorithme sur scénarios <i>off-grid</i> , fouillis diffus, grande dynamique de puissance	38
2.1	Conséquence d'un désalignement en distance ou en vitesse pour l'algorithme non-robustifié	39
2.2	Modélisation du désalignement d'une cible en distance et en vitesse	44
2.3	Modèle hiérarchique Bayésien prenant en compte le désalignement par rapport à la grille d'analyse	46
2.4	Allure de la distribution conditionnelle a posteriori pour différentes valeurs de SNR	50
2.5	Schéma d'estimation prenant en compte le désalignement par rapport à la grille d'analyse	54
2.6	Cartes distance-vitesse estimées par les différents algorithmes ; scénario <i>off-grid</i> bruit blanc	56

2.7	Chaînes d'échantillonnage des paramètres estimés par l'algorithme prenant en compte le désalignement	57
2.8	Distributions a posteriori empiriques des paramètres estimés par l'algorithme prenant en compte le désalignement	58
2.9	Intérêt de la prise en compte du désalignement : performances en fonction de la valeur du désalignement	63
2.10	Comparaison entre les performances de l'algorithme robustifié et de OGSBI en fonction de la valeur du désalignement	64
2.11	Influence de la finesse de la grille d'analyse sur les performances de l'algorithme prenant en compte le désalignement	64
2.12	Sommation cohérente des données PARSAX considérées pour deux trains d'impulsions consécutifs	65
2.13	Cartes distance-vitesse estimées par les différents algorithmes ; scénario expérimental PARSAX <i>off-grid</i>	66
3.1	Modèle de fouillis composite de Chan	71
3.2	Densité spectrale de puissance et SINR loss pour un spectre Gaussien, exponentiel ou de type AR d'ordre 1	73
3.3	Modèle hiérarchique Bayésien prenant en compte la composante diffuse du fouillis	76
3.4	Cartes distance-vitesse estimées par les différents algorithmes ; scénario synthétique <i>on-grid</i> fouillis diffus	77
3.5	Modèle hiérarchique Bayésien prenant en compte le désalignement et le fouillis diffus	78
3.6	Cartes distance-vitesse estimées par les différents algorithmes ; scénario synthétique <i>off-grid</i> fouillis diffus	84
3.7	Distributions a posteriori empiriques des paramètres estimés par l'algorithme prenant en compte le désalignement et le fouillis diffus	85
3.8	Cartes distance-vitesse estimées par les différents algorithmes ; scénario synthétique <i>off-grid</i> fouillis diffus dépendant de la case distance	85
3.9	Intérêt de la prise en compte du fouillis diffus : performances en fonction du SINR des cibles	87
3.10	Cartes distance-vitesse estimées par les différents algorithmes ; scénario semi-expérimental PARSAX <i>off-grid</i>	89

4.1	Composante continue de l'amplitude avec un prior Bernoulli-Student-t hiérarchique sur l'amplitude complexe, pour différents réglages de la loi de mélange	94
4.2	Schéma hiérarchique Bayésien pour l'estimation de scènes à forte dynamique de puissance	96
4.3	Composante continue de la distribution conditionnelle a posteriori de l'amplitude pour le nouveau prior	98
4.4	Schéma d'estimation pour l'estimation de scènes à fortes dynamiques de puissance	101
4.5	Exemple de réalisation 1D de l'algorithme utilisant des classes de puissance; scénario synthétique <i>on-grid</i> bruit blanc	103
4.6	Exemple de réalisation 2D de l'algorithme utilisant des classes de puissance; scénario synthétique <i>off-grid</i> bruit blanc	104
4.7	Cartes distance-vitesse estimées par les algorithmes avec ou sans classes; scénario semi-expérimental PARSAX <i>on-grid</i> à forte dynamique	105
4.8	Intérêt de l'algorithme avec classes de puissances : performances en fonction du SNR des cibles faibles	107
4.9	Choix des classes de puissances : performances en fonction du SNR des cibles faibles	109
4.10	Robustesse du réglage des classes de puissances en fonction de la puissance de bruit estimée : performances en fonction du SNR des cibles faibles	110
4.11	Influence d'un prior non-informatif sur la puissance du bruit : performances en fonction du SNR des cibles faibles	110
A.1	Distribution conditionnelle a posteriori de $\varepsilon_{\bar{m}}$ pour différentes valeurs de ϕ_m et κ_m	118
A.2	Comparaison entre la posterior conditionnelle de $\varepsilon_{\bar{m}}$ et son approximation par une loi normale pour différentes valeurs de SNR	120
A.3	Échantillonnage de $\varepsilon_{\bar{m}}$ pour une loi de proposition uniforme ou normale pour différentes valeurs de SNR	121

Liste des tableaux

2.1	Caractéristiques des cibles utilisées dans le scénario synthétique 2D, repris plusieurs fois dans le manuscrit	56
2.2	Scénario utilisé pour les simulations de Monte-Carlo	59
2.3	Paramètres de traitement utilisés pour les simulations de Monte-Carlo	59
2.4	Paramètres des données PARSAX	61
2.5	Paramètres de traitement utilisés pour les données PARSAX	61
3.1	Distributions a priori utilisées dans le modèle hiérarchique Bayésien prenant en compte le désalignement et le fouillis diffus	78
4.1	Paramètres de traitement pour le scénario 1D	102
4.2	Performances d'estimation de l'algorithme avec classes ou sans classes pour différents réglages de la loi de mélange	106
A.1	Performances de l'échantillonnage de $\varepsilon_{\bar{m}}$ avec une loi de proposition uniforme ou normale suivant le SNR post-traitement : taux d'acceptation moyen τ et MSE sur $\hat{\varepsilon}_{\bar{m}}$	122

Introduction

Contexte de l'étude

Dans le domaine du radar, l'une des problématiques majeures reste l'estimation et la détection de cibles en environnement complexe. En effet, ce sujet a été traité dès les prémises de la technologie radar : le télémobiloscope inventé par Christian Hülsmeier au début du XX^{ème} siècle avait pour objectif d'empêcher les collisions entre bateaux sur le Rhin, même par temps de fort brouillard. L'idée est alors d'utiliser une onde électromagnétique et de détecter des objets en mouvement en analysant l'onde réfléchi. Les radars conventionnels utilisent une forme d'onde dite « à bande étroite », ce qui signifie que la bande de modulation correspond à moins de 10% de la fréquence porteuse. Les traitements radar sont intrinsèquement limités par les ambiguïtés distance (inversement proportionnelles à la fréquence de répétition des impulsions, PRF) et/ou les ambiguïtés vitesse (proportionnelles à la PRF). Par exemple, si on fait le choix d'une faible PRF, alors de nombreuses ambiguïtés vitesse apparaissent et compromettent à la fois l'estimation non-ambigüe de la vitesse des cibles situées en-dehors du fouillis, et la détection de cibles situées dans les vitesses aveugles (lobes de repliement du fouillis).

Les traitements « multi-PRF » sont couramment utilisés pour lever les ambiguïtés : il s'agit d'émettre plusieurs séries d'impulsions avec différentes PRF (non-multiples) et de combiner les sorties de traitement correspondant à chaque PRF pour distinguer la vraie cible de ses fantômes. L'inconvénient majeur de cette méthode est qu'elle limite le temps d'illumination de chaque série d'impulsion utilisant une PRF constante, ce qui peut contrarier la résolution Doppler. La forme d'onde large bande, plutôt utilisée dans le domaine de l'imagerie radar qu'en détection, offre une alternative aux traitements classiques multi-PRF. En effet, la résolution distance étant inversement proportionnelle à la largeur de bande, une forme d'onde large bande diminue la taille des cases distance. Les cibles en mouvement sont alors susceptibles de « migrer » d'une case distance à une autre lors du temps de traitement. Il apparaît donc un phénomène de couplage distance-vitesse qui peut être utilisé pour lever les ambiguïtés vitesse, tout en conservant un temps d'illumination convenable (à l'inverse d'un traitement multi-PRF).

Objectifs de la thèse et contributions

On étudie donc dans ces travaux, financés par la Direction Générale de l'Armement, la possibilité d'un traitement radar large bande à faible PRF permettant une estimation de la scène radar non-ambigüe en vitesse. On se place en outre dans le cadre des représentations parcimonieuses. L'objectif d'une telle représentation est d'exprimer les données dans une base adéquate, afin que chaque cible corresponde à un seul atome de cette base. On obtient ainsi une

représentation du signal d'intérêt sans lobes secondaires. Les algorithmes de représentation parcimonieuse développés dans ces travaux utilisent une approche Bayésienne : les paramètres inconnus sont considérés comme des variables aléatoires qui suivent une certaine distribution a priori. Une telle approche permet d'injecter une potentielle connaissance du comportement des variables inconnues, mais également, si le modèle est approprié, d'ajuster les paramètres de réglage de l'algorithme de manière plus « physique ».

Les travaux présentés dans ce manuscrit reposent sur un algorithme de représentations parcimonieuses développé dans [24]. Ce dernier permet d'estimer la distance et la vitesse non-ambiguë de cibles migrantes grâce à un algorithme MCMC (*Monte-Carlo Markov Chain*); le modèle hiérarchique Bayésien associé utilise une distribution a priori qui force la parcimonie du paramètre d'intérêt. Cet algorithme est néanmoins développé sous hypothèse de bruit blanc, et est limité au cas de cibles situées sur la grille d'analyse distance-vitesse associée au dictionnaire de représentations parcimonieuses. Afin de l'utiliser sur des données expérimentales, cet algorithme a donc été robustifié au cas de cibles en chevauchement sur la grille d'analyse, d'abord sous hypothèse de bruit blanc, puis en prenant en compte le fouillis (et une potentielle composante diffuse du fouillis). Pour cela, le modèle de signal a d'abord été modifié, puis les nouveaux paramètres ont été intégrés au modèle hiérarchique Bayésien utilisé dans [24]. L'algorithme MCMC utilisé pour estimer les paramètres d'intérêt a ensuite été rectifié en conséquence.

D'autre part, les différents algorithmes implémentés (prenant en compte ou pas le désalignement et le fouillis) reposent sur le réglage d'hyperparamètres qui ajustent la connaissance a priori du niveau de puissances des cibles présentes sur la scène radar. Dans le cas de scènes à forte dynamique de puissance, c'est-à-dire où cohabitent des cibles fortes et des cibles faibles, un tel réglage peut être difficile et peut conduire à la dégradation des performances d'estimation de la scène (en particulier l'estimation des cibles faibles). Un algorithme capable d'estimer de telles scènes a donc été proposé ; il repose sur une nouvelle distribution a priori qui utilise des « classes de puissance ».

Organisation du manuscrit

Le manuscrit comporte cinq chapitres décrits ci-après.

Le premier chapitre a pour objectif d'introduire les traitements classiques d'estimation de cibles, en bande étroite puis en large bande, afin de justifier l'utilisation d'une forme d'onde large bande pour la détection. On présente ensuite le principe des représentations parcimonieuses, qui sont utilisées dans ces travaux, ainsi que certains algorithmes déterministes et Bayésiens correspondants. Enfin, on décrit l'algorithme Bayésien de représentation parcimonieuse de [24] qui a servi de base aux travaux présentés dans ce manuscrit. On introduit en particulier les limites de cet algorithme, qui ont motivé les travaux de robustification entrepris.

Le deuxième chapitre est consacré à la robustification de l'algorithme de base de [24] au cas des cibles en chevauchement sur la grille d'analyse distance-vitesse. On montre qu'il

convient de prendre en compte ce potentiel désalignement sous peine de fortement dégrader les performances d'estimation. Bien que ce travail de robustification ait été entrepris en plusieurs étapes (1D/2D, avec/sans migration [3, 1, 4, 6]), on présente directement dans ce chapitre la prise en compte du désalignement en distance et en vitesse pour des cibles migrantes. On décrit donc la modification du modèle de signal via un modèle non-linéaire (par opposition au modèle linéaire communément utilisé dans la littérature) et l'intégration des nouveaux paramètres au modèle hiérarchique Bayésien. Une étude statistique basée sur des simulations de Monte-Carlo est utilisée pour quantifier l'intérêt du nouvel algorithme robustifié au désalignement des cibles par rapport à la version non-robustifiée, et par rapport à un algorithme de la littérature. Enfin, l'algorithme prenant en compte le désalignement est testé avec succès sur des données expérimentales enregistrées par le radar PARSAX de l'université de Delft.

Dans le troisième chapitre, on robustifie l'algorithme prenant en compte le désalignement des cibles à la présence de fouillis de sol, et en particulier à celle d'une composante diffuse. Pour cela, on fusionne le modèle hiérarchique Bayésien prenant en compte le désalignement des cibles (présenté dans le chapitre 2) à celui proposé dans [21] qui prend en compte le fouillis via un modèle auto-regressif de bruit. On obtient ainsi un algorithme capable d'estimer la localisation non-ambigüe de cibles migrantes en désalignement par rapport à la grille d'analyse distance-vitesse [7], y compris pour les cibles situées dans les vitesses aveugles. Une analyse statistique utilisant des simulations de Monte-Carlo montre l'intérêt de la prise en compte du fouillis par rapport au modèle bruit blanc. L'algorithme final est testé sur des données semi-expérimentales construites à partir de mesures de fouillis du radar PARSAX.

Enfin, le point de départ des travaux présentés dans le quatrième chapitre est le suivant : les algorithmes développés nécessitent le réglage d'hyperparamètres qui ajustent la connaissance a priori de la puissance des cibles présentes sur la scène radar. Ce réglage peut être facilité en estimant au préalable le niveau moyen de puissance sur la scène, mais il ne permet toutefois pas d'estimer correctement les scènes à très forte dynamique de puissance. On propose donc dans ce chapitre une nouvelle distribution a priori qui repose sur l'utilisation de classes de puissance. Le nouvel algorithme correspondant au nouveau modèle hiérarchique Bayésien et au schéma d'estimation associé est alors comparé à l'algorithme sans classes de puissance sur différents scénarios synthétiques où cohabitent des cibles fortes et faibles ; un réglage des classes de puissance est également proposé pour de tels scénarios [5], [2]. L'algorithme est finalement testé sur des données semi-expérimentales où des cibles synthétiques sont ajoutées à des mesures expérimentales de bruit thermique du PARSAX.

Enfin, le dernier chapitre résume les contributions des travaux présentés dans les précédents chapitres. Il propose ensuite de futures pistes d'étude afin de compléter ces travaux.

État de l'art : estimation non-ambigüe de cibles par un radar large bande

Les travaux présentés dans ce manuscrit s'inscrivent dans le cadre de la recherche d'un mode de détection non-ambigu de cibles dans un environnement complexe (fouillis diffus, cibles faibles en compétition avec des cibles fortes). Les radars de détection conventionnels utilisant une forme d'onde bande étroite sont limités par les ambiguïtés distance et vitesse. La forme d'onde large bande, classiquement utilisée en imagerie SAR, ouvre de nouvelles perspectives au problème de levée d'ambiguïté vitesse, tout en permettant un temps d'illumination de la scène plus long et une meilleure résolution distance.

Dans ce chapitre, on présente donc dans un premier temps le modèle de signal radar bande étroite et les problématiques associées, en particulier les ambiguïtés distance et vitesse. On introduit ensuite la forme d'onde large bande comme alternative, ainsi que des algorithmes d'estimation spectrale adaptés à cette forme d'onde spécifique. L'accent est ensuite mis sur les méthodes de représentation parcimonieuse (*Sparse Signal Representations*, SSR) qui permettent une représentation du signal d'intérêt dépourvue de lobes secondaires. La distinction est faite entre les algorithmes de résolution déterministes, et les approches Bayésiennes moins répandues. Enfin, un algorithme spécifique de représentation parcimonieuse d'un signal radar large bande [24] est présenté en détail ; c'est cet algorithme qui a servi de base au travail de recherche présenté dans la suite du manuscrit. Il est testé sur des données synthétiques et expérimentales, puis ses limites sont soulignées afin de motiver le travail de robustification entrepris dans ces travaux de thèse.

Sommaire

1.1	Limites de l'estimation/détection de cibles avec un signal bande étroite	6
1.1.1	Modèle de signal bande étroite	6
1.1.2	Filtre adapté en bruit blanc	7
1.1.3	Ambiguïtés distance / vitesse	8
1.2	Utilisation d'un signal radar large bande	10
1.2.1	Modèle de signal large bande	10
1.2.2	Sommation cohérente large bande	11
1.2.3	Estimation spectrale pour la forme d'onde large bande	12
1.3	Les représentations parcimonieuses	17

1.3.1	Mise en forme du problème	17
1.3.2	Approches déterministes de représentation parcimonieuse	22
1.3.3	Approches Bayésiennes de représentation parcimonieuse	24
1.4	Une approche Bayésienne pour la représentation parcimonieuse d'un signal radar large bande	30
1.4.1	Présentation de l'algorithme	31
1.4.2	Exemple de résultats	34
1.4.3	Robustesse de l'algorithme	35

1.1 Limites de l'estimation/détection de cibles avec un signal bande étroite

Dans un premier temps, on se focalise sur la forme d'onde bande étroite utilisée par les radars conventionnels pour détecter et/ou estimer des cibles en environnement complexe. On considère qu'un signal radar est dit à bande étroite si le rapport entre la bande du signal B et la fréquence porteuse f_c est largement inférieur à 10 %. Ainsi, on présente dans cette section le modèle de signal bande étroite et le filtre adapté correspondant. On introduit alors la notion d'ambiguïtés distance et vitesse qui limitent les performances de détection de tels radars.

1.1.1 Modèle de signal bande étroite

Considérons un système radar qui émet M impulsions de durée T avec une fréquence de répétition (*Pulse Repetition Frequency* ou PRF) F_r . On utilise une forme d'onde de type *chirp* : chaque impulsion est modulée linéairement en fréquence autour de la fréquence porteuse f_c avec une bande B . Afin d'exprimer les données dans le domaine fréquence-rapide / temps-long, on applique le filtre adapté distance aux données, puis une transformée de Fourier sur le temps-court sur un segment sélectionné comportant K cases distances. On obtient alors le signal suivant correspondant à la m -ième impulsion ($m \in \{0, \dots, M - 1\}$) et la k -ième sous-bande ($k \in \{0, \dots, K - 1\}$), où ne se trouve qu'une seule cible d'amplitude complexe α_t , de retard initial τ_0 et se déplaçant à la vitesse v

$$\begin{aligned}
[\mathbf{Y}]_{k,m} &= \alpha_t \exp \left\{ j2\pi m \frac{2v}{c} f_c T_r \right\} \exp \left\{ -j2\pi k \frac{B}{K} \tau_0 \right\} \\
&= \alpha_t \exp \{ j2\pi m f_D \} \exp \{ -j2\pi k f_r \}.
\end{aligned} \tag{1.1}$$

On note $f_D = 2vf_c T_r / c$ la fréquence Doppler de la cible normalisée par rapport à $F_r = 1/T_r$, et $f_r = \tau_0 B / K$ la fréquence normalisée associée à la distance. Le signal utile dans le cas de plusieurs cibles s'écrit alors

$$\mathbf{Y} = \sum_{t=1}^{N_t} \alpha_t \mathbf{A}_t \tag{1.2}$$

où N_t désigne le nombre total de cibles et $\mathbf{A}_t = \mathbf{a}_r \mathbf{a}_D^T$ tels que

$$\mathbf{a}_r = \left[1 \quad \dots \quad \exp \{-j2\pi k f_r\} \quad \dots \quad \exp \{-j2\pi(K-1)f_r\} \right]^T \quad (1.3)$$

$$\mathbf{a}_D = \left[1 \quad \dots \quad \exp \{j2\pi m f_D\} \quad \dots \quad \exp \{j2\pi(M-1)f_D\} \right]^T. \quad (1.4)$$

En réalité, le signal reçu est pollué par du bruit \mathbf{N} ; il s'écrit alors

$$\mathbf{Y} = \sum_{t=1}^{N_t} \alpha_t \mathbf{A}_t + \mathbf{N}. \quad (1.5)$$

Le bruit \mathbf{N} contient tout d'abord le bruit thermique dû à l'agitation des électrons dans les appareils de mesure. Il est en général considéré blanc Gaussien, et sa puissance σ^2 est relativement bien connue. A ce bruit thermique s'ajoutent les échos indésirables dûs à l'environnement : les arbres, le sol, les bâtiments présents sur la scène radar. Ce « fouillis de sol » (*clutter*) peut grandement perturber l'estimation des paramètres des cibles présentes sur la scène.

Dans la suite du manuscrit, une version vectorisée ligne-à-ligne de l'équation (1.5) sera souvent utilisée

$$\mathbf{y} = \sum_{t=1}^{N_t} \alpha_t \mathbf{a}_t + \mathbf{n} \quad (1.6)$$

où \mathbf{y} , \mathbf{a}_t et \mathbf{n} correspondent à la vectorisation ligne-à-ligne de \mathbf{Y} , \mathbf{A}_t et \mathbf{N} respectivement. On cherche alors à estimer la distance, la vitesse et l'amplitude des cibles présentes sur la scène radar, qui sont associées aux paramètres f_D , f_r et α_t de l'équation (1.1).

1.1.2 Filtre adapté en bruit blanc

On définit le filtre adapté comme celui maximisant le rapport signal-à-bruit (*Signal-to-noise ratio*, SNR) en sortie de filtre dans le cas où une seule cible est présente dans la scène radar. On suppose dans ces travaux que le bruit \mathbf{n} suit une distribution Gaussienne centrée de matrice de covariance \mathbf{R} : $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$. La fonction vraisemblance $f(\mathbf{y}|\alpha)$ s'écrit alors

$$f(\mathbf{y}|\alpha) = \frac{1}{\pi^{KM} |\mathbf{R}|} \exp \left\{ -(\mathbf{y} - \alpha \mathbf{a}_t)^H \mathbf{R}^{-1} (\mathbf{y} - \alpha \mathbf{a}_t) \right\}. \quad (1.7)$$

Dans ce cas-là, l'amplitude de la cible en sortie de filtre adapté maximise la fonction vraisemblance [66], ce qui donne :

$$\hat{\alpha}_{\text{ML}} = \arg \min_{\alpha} (\mathbf{y} - \alpha \mathbf{a}_t)^H \mathbf{R}^{-1} (\mathbf{y} - \alpha \mathbf{a}_t). \quad (1.8)$$

Si on considère en outre que le bruit \mathbf{n} est un bruit blanc, c'est-à-dire $\mathbf{R} = \sigma^2 \mathbf{I}$, alors

$$\hat{\alpha}_{\text{ML}} = \arg \min_{\alpha} (\mathbf{y} - \alpha \mathbf{a}_t)^H (\mathbf{y} - \alpha \mathbf{a}_t).$$

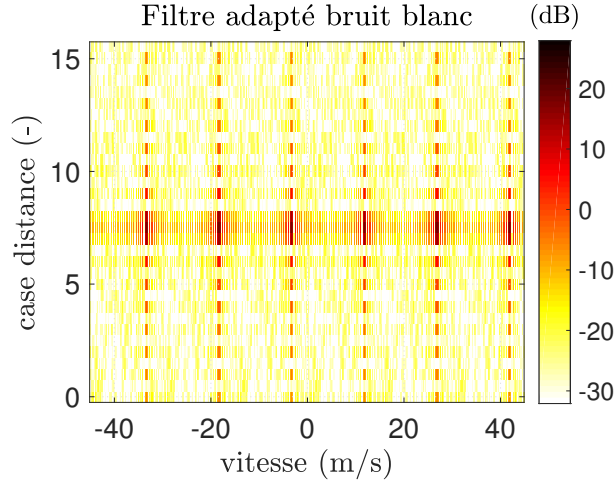


FIGURE 1.1 – Exemple de sortie de filtre adapté en bruit blanc pour un signal bande étroite contenant une seule cible à $v \approx -3$ m/s, $k_0 = 7.5$. $K = 16$, $M = 32$, $F_r = 1$ kHz, $n_{va} = 6$.

On obtient alors simplement le résultat suivant

$$\begin{aligned}\alpha_{\text{ML}} &= \frac{\mathbf{a}^H \mathbf{y}}{\mathbf{a}^H \mathbf{a}} \\ &= \frac{1}{KM} \mathbf{a}^H \mathbf{y}\end{aligned}$$

qui n'est autre que la transformée de Fourier 2D du signal reçu \mathbf{y} . Plus exactement, il s'agit d'après (1.3), (1.4) d'une transformée de Fourier sur le temps-long et d'une transformée de Fourier inverse sur le temps-court. Notons que si le bruit est coloré, il faut minimiser la fonction (1.8) pour implémenter le filtre adapté, ce qui nécessite en pratique l'estimation de la matrice de covariance du bruit \mathbf{R} .

On représente sur la Fig. 1.1 la sortie du filtre adapté bruit blanc pour un signal bande étroite ne contenant qu'une seule cible de vitesse non nulle. On reconnaît bien les lobes secondaires en distance et en vitesse, mais on voit surtout que le pic principal correspondant à la cible est répété sur l'axe des vitesses tous les $v_a = 15$ m/s. Ce phénomène dit d'ambiguïté vitesse est détaillé dans la section suivante.

1.1.3 Ambiguïtés distance / vitesse

Un système radar émettant des impulsions avec une PRF constante subit des ambiguïtés en distance et/ou en vitesse. En effet, la distance maximale mesurable sans ambiguïté vaut $d_a = (cT_r)/2$ où $T_r = 1/F_r$. Pour augmenter cette distance, il faut donc diminuer la PRF. Cependant, il existe également une plage d'ambiguïté en vitesse, limitée par $v_a = c/(2f_c T_r)$. Ainsi, si on utilise ici une faible PRF pour limiter les ambiguïtés distance, alors on augmente le nombre d'ambiguïtés vitesse. Ce phénomène est observé sur la Fig. 1.1 : la signature de la cible est répétée sur l'axe des vitesses avec une périodicité $v_a = 15$ m/s. Les pics qui ne

correspondent pas à la cible sont appelés « cibles fantômes ». Notons que cette périodicité est aussi valable pour le *clutter* : les éléments réfléchissants à vitesse nulle présents à chaque case distance sont répétés à v_a , $2v_a$, $-v_a$, $-2v_a$ et ainsi de suite, faisant apparaître les zones dites de « vitesses aveugles » où les lobes de repliement du *clutter* masquent les cibles présentes à cet endroit.

Dans le cadre de l'estimation ou de la détection des cibles, l'enjeu est donc de lever les ambiguïtés vitesse afin de discriminer la vraie cible de ses fantômes. Pour cela, certains radars adoptent un mode « multi-PRF » : ils émettent un train d'impulsions avec différentes valeurs de PRF, ce qui donne différentes valeurs de vitesse ambiguë v_a [95, chap. 3]. Ainsi, le pic principal n'est pas répété avec la même périodicité suivant la valeur de la PRF. En utilisant des PRF non multiples il est alors possible d'identifier la vraie cible comme correspondant au seul pic présent pour toutes les valeurs de PRF. Un exemple est présenté sur la Fig.1.2 avec les valeurs de PRF suivantes : 1.1 kHz, 1.4 kHz, 3 kHz. En comparant les trois cartes distance-vitesse, on identifie la vraie cible comme celle possédant une vitesse égale à -3 m/s.

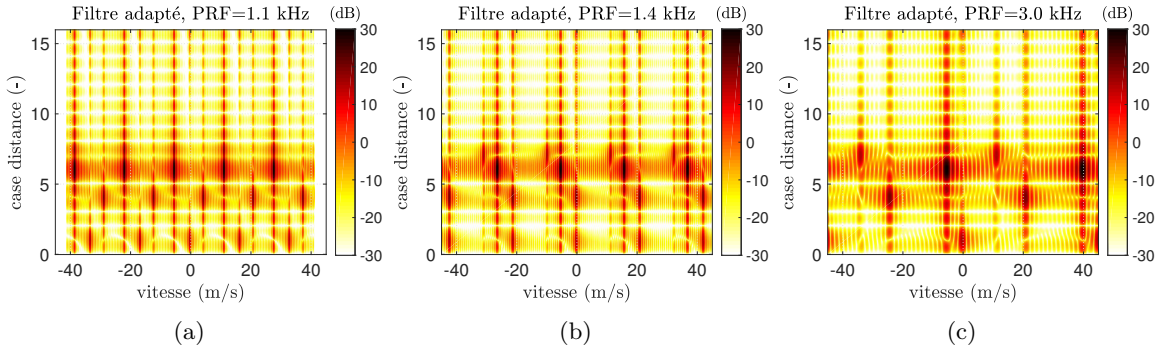


FIGURE 1.2 – Traitement multi-PRF d'un signal contenant une seule cible à $v = -3$ m/s, $k_0 = 7$. 1.2(a) $F_r = 1.1$ kHz, $v_a = 16.5$ m/s, 1.2(b) $F_r = 1.4$ kHz, $v_a = 21$ m/s, 1.2(c) $F_r = 3$ kHz, $v_a = 45$ m/s. $K = 16$, $M = 32$.

L'inconvénient d'un tel traitement multi-PRF est qu'il limite le temps d'illumination relatif à chaque rafale à PRF constante, ce qui peut diminuer la résolution Doppler. Il semble donc intéressant de concevoir un algorithme capable de lever les ambiguïtés vitesse en utilisant une unique PRF. L'utilisation d'un signal radar large bande à PRF constante peut être une solution à ce problème [71, 87, 44, 93]. En effet, on rappelle que la résolution d'un radar à compression d'impulsion est donnée par

$$\delta_R = \frac{c}{2B} \quad (1.9)$$

où $c = 3 \times 10^8$ m/s désigne la vitesse de l'onde électromagnétique. L'utilisation d'une large bande de modulation permet donc d'obtenir une meilleure résolution distance, c'est-à-dire que la taille des cases utilisées pour échantillonner l'espace distance est faible. En contre-partie, une cible se déplaçant à grande vitesse est alors susceptible de passer d'une case distance à la suivante lors du temps de traitement (*Coherent Processing Interval*, CPI). Ce phénomène dit de « migration distance » constitue une information supplémentaire qui peut être mise à profit

pour lever les ambiguïtés vitesse, par exemple en exploitant la pente de la signature spectrale 2D d'une cible [76], ou en utilisant les transformées de Keystone et Keystone-Wigner dans [88].

1.2 Utilisation d'un signal radar large bande

Dans le cadre de cette thèse, on utilise donc un signal radar large bande pour concevoir un mode non-ambigu d'estimation / détection de cibles. Une telle forme d'onde est classiquement utilisée en imagerie SAR [126, 109], mais peu dans le cadre de la détection de cibles [87, 44]. Pourtant, elle permet d'une part d'augmenter le temps d'illumination de la scène (par rapport à un traitement multi-PRF) et d'autre part d'améliorer la résolution distance.

Notons que diminuer la taille de la case distance grâce à la forme d'onde large bande suppose que l'on prenne en compte des problématiques dues à la haute résolution. Tout d'abord, la modélisation du fouillis (de mer ou de sol) par une distribution Gaussienne n'est plus appropriée [10, 43, 52]. En effet, à haute résolution la représentation temporelle du bruit possède de nombreux pics qui sont vus comme des cibles par les détecteurs classiques basés sur une hypothèse de Gaussianité, ce qui augmente considérablement la probabilité de fausse-alarme [52]. La distribution du bruit doit donc posséder une plus longue queue [43], comme c'est le cas par exemple avec un modèle composite K-distribué [112]. D'autre part, la modélisation d'une cible par un point réfléchissant peut être remise en question en haute-résolution. En effet, si la taille des cases d'analyse est faible alors une cible correspond à plusieurs points réfléchissants, qui peuvent être étalés en distance et en Doppler. Une cible est alors parfois modélisée comme la somme de plusieurs points réfléchissants [98], mais cette modélisation ne prend pas en compte le phénomène de dispersion dû à la haute résolution. Ce problème de dispersion n'est d'ailleurs pas spécifique à la haute résolution, et reste ouvert. Bien qu'ils constituent très certainement une limite, ces problèmes ne sont pas adressés dans ces travaux : on considère un bruit Gaussien, et les cibles sont modélisées comme la somme de plusieurs points réfléchissants, même si elles sont étendues en distance.

On présente dans cette section le modèle de signal large bande, ainsi que des algorithmes d'estimation spectrale adaptés à cette forme d'onde.

1.2.1 Modèle de signal large bande

Comme dans la section précédente, on considère un signal composé de M impulsions émises avec une PRF F_r , et correspondant à K sous-bandes. On les exprime dans le domaine fréquence-rapide / temps-long en utilisant le même traitement qu'en bande étroite (filtre adapté distance puis transformée de Fourier sur le temps-court). On obtient alors l'expression

suivante [45] :

$$\begin{aligned}\mathbf{Y}_{k,m} &= \alpha_t \exp \left\{ j2\pi m \frac{2v}{c} f_c T_r \right\} \exp \left\{ -j2\pi k \frac{B}{K} \tau_0 \right\} \exp \left\{ j2\pi km \frac{2v}{c} \frac{B}{K} T_r \right\} \\ &= \alpha_t \exp \{ j2\pi m f_D \} \exp \{ -j2\pi k f_r \} \exp \left\{ j2\pi km f_D \frac{B}{K f_c} \right\}.\end{aligned}\quad (1.10)$$

On reconnaît dans l'expression (1.10) deux exponentielles complexes associées aux fréquences f_D (temps long) et f_r (temps court) qui correspondent au signal bande étroite (1.1), ainsi qu'un terme de couplage distance / vitesse qui modélise le phénomène de migration distance. Il est alors possible d'exprimer le signal reçu comme une matrice \mathbf{Y} de taille $K \times M$

$$\mathbf{Y} = \sum_{t=1}^{N_t} \alpha_t \mathbf{A}_t + \mathbf{N}.\quad (1.11)$$

Contrairement au signal bande étroite, la signature de la t -ième cible s'exprime comme $\mathbf{A}_t = \mathbf{A}_{rD} \odot \{ \mathbf{a}_r \mathbf{a}_D^T \}$, \odot correspondant au produit de Hadamard, avec

$$[\mathbf{A}_{rD}]_{k,m} = \exp \left\{ j2\pi km f_D \frac{B}{K f_c} \right\}\quad (1.12)$$

$$[\mathbf{a}_r]_k = \exp \{ -j2\pi k f_r \}\quad (1.13)$$

$$[\mathbf{a}_D]_m = \exp \{ j2\pi m f_D \}.\quad (1.14)$$

1.2.2 Sommation cohérente large bande

Grâce à la définition de la signature de cible (1.10), il est possible de sommer de manière cohérente les M impulsions considérées. La sommation cohérente pour la fréquence rapide f_r et la fréquence Doppler f_D s'écrit alors [22]

$$S(f_r, f_D) = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} \mathbf{Y}_{k,m} \exp \{ -j2\pi m f_D (1 + \mu k) \} \exp \{ +j2\pi k f_r \}\quad (1.15)$$

où l'on définit le paramètre $\mu = (B/f_c)/K$ comme étant la bande fractionnelle pour chaque sous-bande. La sommation cohérente S peut être calculée sur un domaine fréquentiel arbitraire. En revanche, son implémentation rapide [22] est limitée à l'intervalle $[0, 1[\times[-n_{va}/2, n_{va}/2[$ où n_{va} désigne le facteur de dépliement de la carte, ce qui correspond à un intervalle de vitesse $[-n_{va}v_a/2, n_{va}v_a/2[$. Dans la suite du manuscrit, on utilise l'implémentation rapide de la sommation cohérente donc on se limite à ce domaine de définition.

On représente sur la Fig. 1.3 un exemple de sommation cohérente large bande pour un signal contenant 1 cible à vitesse nulle, $k_0 = \tau_0 B = 8$. On reconnaît sur cette représentation la forme caractéristique de la sommation cohérente large bande, à savoir un « papillon » dont les lobes secondaires en vitesse vérifient [22]

- une position $v_p = v \pm p v_a$
- un étalement en vitesse $\Delta v = p v_a B / f_c$

- un étalement en distance $\Delta k = pMB/f_c$
- un niveau relatif $\Delta S_{\text{dB}} = -20 \log_{10} \{pMB/f_c\}$ dB.

Il est intéressant de remarquer que ces grandeurs ne dépendent pas des caractéristiques de la cible, mais seulement de la bande fractionnelle B/f_c , du nombre d'impulsions M et de la vitesse ambiguë v_a .

En outre, on reconnaît localement la sommation cohérente bande étroite à savoir un sinus cardinal en deux dimensions (cf Fig. 1.3(b)). En effet, on remarque des lobes secondaires en distance et en vitesse atténués de -13.5 dB, puis -17.9 dB, et ainsi de suite. Dans la suite du manuscrit, le terme « lobes secondaires en vitesse » fait référence aux lobes secondaires séparés de v_a observables sur la Fig. 1.3(a).

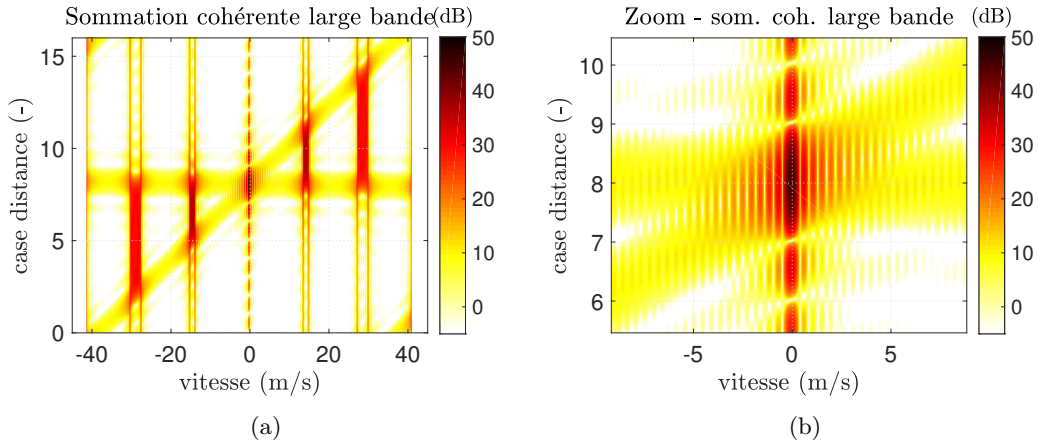


FIGURE 1.3 – Somme cohérente d'un signal radar large bande contenant une seule cible à $v = 0$ m/s, $k_0 = 8$. 1.3(a) Carte distance-vitesse entièrement dépliée, 1.3(b) Zoom autour du pic de la cible. $K = 16$, $M = 32$, $n_{\text{va}} = 6$, $v_a = 15$ m/s, $B/f_c=10\%$.

Ainsi, si on compare le traitement simple de sommation cohérente en large bande (Fig. 1.3) à celui en bande étroite (Fig. 1.1), on voit que grâce à sa forme particulière en papillon, la sommation cohérente large bande permet une identification plus directe de la « vraie » cible. De plus, le niveau des lobes secondaires en large bande est plus faible que celui des lobes de repliement en bande étroite. En revanche, il ne faut pas sous-estimer la complexité qu'engendre la forme d'onde large bande via la migration distance, par exemple en termes de non-stationnarités.

1.2.3 Estimation spectrale pour la forme d'onde large bande

Les méthodes d'estimation spectrale consistent à estimer la valeur du signal dans le domaine temps-rapide / fréquence-lente à chaque point de la carte distance-vitesse. En bande étroite, la méthode d'estimation spectrale la plus simple est donc une transformée de Fourier 2D, mais elle a l'inconvénient de produire de forts lobes secondaires. La méthode du périodogramme de Welch permet une estimation spectrale moins bruitée car elle utilise une fenêtre glissante avant de calculer la transformée de Fourier 2D du signal [114]. D'autres méthodes

ont ensuite été développées, qui améliorent considérablement la résolution spectrale et réduisent les lobes secondaires, par exemple les algorithmes Capon [34] et APES (*Amplitude and Phase Estimation of a Sinusoid*) [72]. Ces deux algorithmes calculent un filtre à réponse impulsionnelle finie (FIR, *Finite Impulse Response*) qui permet d'obtenir une estimation de l'amplitude au sens des moindres carrés ; l'originalité de ces méthodes réside dans l'utilisation de sous-matrices des données (en recouvrement) pour estimer la matrice de covariance de celles-ci. Contrairement à Capon, APES construit le filtre en soustrayant à la matrice de covariance des données la contribution des autres cibles. L'estimation du niveau d'amplitude des cibles est meilleure qu'avec Capon, mais ces dernières sont moins bien résolues.

Malgré la diminution du niveau des lobes secondaires en large bande par rapport aux lobes de repliement en bande étroite, un signal large bande peut mener à de nombreux problèmes d'estimation (dus à la forme de sa fonction d'ambiguïté) : cibles masquées aux vitesses aveugles et/ou dans les lobes secondaires d'une cible plus forte, levée d'ambiguïtés vitesse. Les algorithmes classiques bande étroite que nous venons de mentionner doivent être adaptés à la forme d'onde large bande afin de prendre en compte le phénomène de migration distance. Dans [44], des versions étendues des algorithmes APES (W-APES) et Capon (W-Capon, iW-Capon) sont présentées. On évoque également dans cette section une autre méthode d'estimation spectrale proche de la méthodologie de APES et Capon : l'algorithme IAA (*Iterative Adaptive Approach*) appliqué à un signal large bande (W-IAA) [87].

1.2.3.1 Algorithmes de référence étendus : W-APES, W-Capon, iW-Capon

Les algorithmes Capon et APES sont adaptés à la forme d'onde large bande dans [44], d'abord sous hypothèse de stationnarité (W-APES, W-Capon) puis sans cette hypothèse grâce à une méthode itérative (iW-Capon). Ces trois algorithmes sont décrits ci-après.

Dans un premier temps, $N_L N_M$ sous-matrices $\mathbf{Y}_{p,q}$ sont construites en appliquant une fenêtre glissante sur les données \mathbf{Y} ; elles sont vectorisées pour former les sous-vecteurs $\mathbf{y}_{p,q}$ que l'on peut écrire comme

$$\mathbf{y}_{p,q} = \alpha \mathbf{a}_{p,q} + \mathbf{e}_{p,q}.$$

On suppose que les sous-vecteurs de bruit $\mathbf{e}_{p,q}$ sont centrés, indépendants et qu'ils suivent une distribution complexe Gaussienne :

$$\begin{aligned} \mathbf{e}_{p,q} &\sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_{p,q}) \\ \mathcal{E} \{ \mathbf{e}_{p,q} \mathbf{e}_{r,s}^H \} &= \mathbf{Q}_{p,q} \delta_{p,r} \delta_{q,s}. \end{aligned}$$

Deux remarques sont alors adressées dans [44] : 1) l'hypothèse d'indépendance des sous-vecteurs $\mathbf{e}_{p,q}$ n'est pas correcte pour deux fenêtres adjacentes en recouvrement ; 2) il est fort probable que la matrice de covariance $\mathbf{Q}_{p,q}$ ne soit pas stationnaire à cause de la migration distance. Ici, par stationnaire on entend que les sous-matrices $\mathbf{Q}_{p,q}$ sont indépendantes des indices p, q : la matrice de covariance des sous-vecteurs de bruit est la même sur toute la carte, $\mathbf{Q}_{p,q} = \mathbf{Q}$.

Les algorithmes W-APES et W-Capon sont développés sous cette hypothèse de stationnarité, à l'image des algorithmes APES et Capon originaux. Il s'agit donc de calculer le filtre FIR qui donne l'estimateur du maximum de vraisemblance de l'amplitude α . W-APES construit ce filtre en utilisant la matrice de covariance du bruit \mathbf{Q} , alors que W-Capon utilise lui la matrice de corrélation des échantillons.

W-APES L'estimateur W-APES est défini comme l'estimateur du maximum de vraisemblance de l'amplitude α . Ne pouvant être calculé sous forme analytique [44], une procédure itérative est mise en place qui revient à maximiser la fonction log-vraisemblance $\Lambda(\{\mathbf{y}_{p,q}\}|\alpha, \mathbf{Q})$ par rapport à α puis par rapport à \mathbf{Q} :

$$\alpha = \frac{\sum_{p,q} \mathbf{a}_{p,q}^H \mathbf{Q}^{-1} \mathbf{y}_{p,q}}{\sum_{p,q} \mathbf{a}_{p,q}^H \mathbf{Q}^{-1} \mathbf{a}_{p,q}} \quad (1.16)$$

$$\mathbf{Q} = \frac{1}{N_M N_L} \sum_{p,q} (\mathbf{y}_{p,q} - \alpha \mathbf{a}_{p,q}) (\mathbf{y}_{p,q} - \alpha \mathbf{a}_{p,q})^H. \quad (1.17)$$

On rappelle que les vecteurs directionnels $\mathbf{a}_{p,q}$ prennent en compte le phénomène de migration distance ; leur expression est donnée dans [44]. Il est proposé d'initialiser l'algorithme par $\mathbf{Q} = \hat{\mathbf{R}} = \frac{1}{N_M N_L} \sum_{p,q} \mathbf{y}_{p,q} \mathbf{y}_{p,q}^H$ pour assurer la convergence.

W-Capon Une autre approche, plus simple du point de vue calculatoire, est décrite dans [44]. Elle consiste à remplacer la matrice de covariance du bruit \mathbf{Q} par la matrice de corrélation des échantillons $\hat{\mathbf{R}} = \frac{1}{N_M N_L} \sum_{p,q} \mathbf{y}_{p,q} \mathbf{y}_{p,q}^H$. On obtient alors une expression de l'estimateur W-Capon, cette fois sous forme analytique :

$$\hat{\alpha}^{\text{w-capon}} = \frac{\sum_{p,q} \mathbf{a}_{p,q}^H \hat{\mathbf{R}}^{-1} \mathbf{y}_{p,q}}{\sum_{p,q} \mathbf{a}_{p,q}^H \hat{\mathbf{R}}^{-1} \mathbf{a}_{p,q}}. \quad (1.18)$$

Ces deux algorithmes permettent donc d'estimer une scène comprenant des cibles migrantes, mais l'hypothèse de stationnarité dégrade considérablement leurs performances, comme illustré dans la section 1.2.3.3.

iW-Capon Un nouvel algorithme est présenté dans [44], qui passe outre l'hypothèse de stationnarité grâce à une procédure itérative de type CLEAN [108, 60]. L'idée principale est de remplacer les matrices de covariance du bruit $\mathbf{Q}_{p,q}$ par un estimé des matrices de covariance des données $\mathbf{R}_{p,q}$. L'algorithme iW-Capon comporte alors 2 étapes :

$$\hat{\alpha}^{\text{iw-capon}} = \frac{\sum_{p,q} \mathbf{a}_{p,q}^H \hat{\mathbf{R}}_{p,q}^{-1} \mathbf{y}_{p,q}}{\sum_{p,q} \mathbf{a}_{p,q}^H \hat{\mathbf{R}}_{p,q}^{-1} \mathbf{a}_{p,q}} \quad (1.19)$$

où $\hat{\mathbf{R}}_{p,q}$ est une matrice structurée, estimée de $\mathbf{R}_{p,q}$, que l'on calcule comme

$$\hat{\mathbf{R}}_{p,q} = \sum_{t=1}^{\hat{N}_t} |\hat{\alpha}_t|^2 \hat{\mathbf{a}}_{t,p,q} \hat{\mathbf{a}}_{t,p,q}^H + \hat{\mathbf{R}}_n \quad (1.20)$$

avec

\hat{N}_t le nombre estimé de cibles

$\hat{\alpha}_t$ l'amplitude complexe estimée de la t -ième cible

$\hat{\mathbf{a}}_{t,p,q}$ le sous-vecteur directionnel de la t -ième cible

$\hat{\mathbf{R}}_n$ la matrice de covariance correspondant à la composante stationnaire du bruit (indépendante des indices (p, q)).

Autrement dit, cet algorithme calcule un estimateur de type Capon de l'amplitude complexe $\hat{\alpha}^{iw-capon}$ en utilisant une estimée de la matrice de covariance des données $\hat{\mathbf{R}}_{p,q}$ qui comporte une composante non-stationnaire et une composante stationnaire. L'enjeu est alors d'estimer la composante stationnaire de $\hat{\mathbf{R}}_{p,q}$ grâce à une procédure itérative de type CLEAN. $\hat{\mathbf{R}}_n$ est d'abord estimée comme la matrice de covariance des données auxquelles on a enlevé la contribution de la cible la plus forte. Cette estimation est ensuite raffinée en enlevant successivement les contributions des autres cibles présentes dans le signal (de la plus forte à la moins forte). A noter que le nombre de cibles \hat{N}_t est supposé connu dans [44], ce qui constitue un critère d'arrêt simple. Néanmoins, il est possible de définir d'autres critères d'arrêt si le nombre de cibles n'est pas connu, par exemple un critère de blancheur.

1.2.3.2 Iterative Adaptive Approach (W-IAA)

Une autre méthode d'estimation spectrale adaptative a été présentée dans [86, 87] qui consiste à appliquer l'algorithme IAA décrit dans [120] à un signal radar large bande. On utilise l'acronyme « W-IAA » pour faire référence à cet algorithme. Celui-ci est similaire aux méthodes d'estimation spectrale de référence Capon et APES, mais se distingue des approches W-APES, W-Capon, iW-Capon puisqu'elle n'utilise pas de fenêtre glissante balayant les données pour former des sous-matrices.

L'idée développée dans [120, 87] consiste à exprimer les données à travers une grille d'analyse bi-dimensionnelle et à estimer l'amplitude des cibles potentiellement présentes à chaque case de cette grille. Plus précisément, les données vectorisées \mathbf{y} sont écrites sous la forme :

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (1.21)$$

où \mathbf{H} est une matrice de taille $K\bar{M} \times \bar{K}\bar{M}$ qui provient de la discrétisation de l'espace distance-vitesse ; elle contient les vecteurs directionnels adaptés à la forme d'onde large bande $\mathbf{a}_{\bar{i}}(f_{r_{\bar{i}}}, f_{D_{\bar{i}}})$ avec $\bar{i} \in \{0, \dots, \bar{K}\bar{M} - 1\}$, correspondant à chaque case de la grille d'analyse. Le vecteur \mathbf{x} de taille $\bar{K}\bar{M} \times 1$ contient l'information d'amplitude pour chaque case ; c'est le paramètre d'intérêt. Le vecteur de bruit \mathbf{n} est supposé blanc Gaussien de puissance σ^2 .

Dans un premier temps, il s'agit de construire la matrice diagonale \mathbf{P} de taille $\bar{K}\bar{M} \times \bar{K}\bar{M}$

qui contient la puissance présente à chaque case de la grille d'analyse : $P_{\bar{i},\bar{i}} = |x_{\bar{i}}|^2$. On utilise alors la matrice de covariance du bruit et des interférences $\mathbf{Q}_{\bar{i}} = \mathbf{R} - P_{\bar{i},\bar{i}} \mathbf{a}_{\bar{i}} \mathbf{a}_{\bar{i}}^H$ pour calculer la fonction coût des moindres carrés pondérés [120]

$$(\mathbf{y} - \mathbf{H}\mathbf{x})^H \mathbf{Q}_{\bar{i}}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \quad (1.22)$$

puis construire le filtre FIR donnant l'estimateur ML (*Maximum Likelihood*) de \mathbf{x} qui minimise cette fonction coût

$$\hat{\mathbf{x}}^{\text{w-iaa}} = \frac{\mathbf{a}_{\bar{i}}^H \mathbf{Q}_{\bar{i}}^{-1} \mathbf{y}}{\mathbf{a}_{\bar{i}}^H \mathbf{Q}_{\bar{i}}^{-1} \mathbf{a}_{\bar{i}}}.$$

On peut remarquer que cette approche est très similaire à celle de l'algorithme APES. Ensuite, en utilisant la définition de $\mathbf{Q}_{\bar{i}}$ il est possible de réécrire l'estimateur $\hat{\mathbf{x}}^{\text{w-iaa}}$ comme

$$\hat{\mathbf{x}}^{\text{w-iaa}} = \frac{\mathbf{a}_{\bar{i}}^H \hat{\mathbf{R}}^{-1} \mathbf{y}}{\mathbf{a}_{\bar{i}}^H \hat{\mathbf{R}}^{-1} \mathbf{a}_{\bar{i}}}. \quad (1.23)$$

où la matrice $\hat{\mathbf{R}}$ est estimée comme

$$\hat{\mathbf{R}} = \mathbf{H}\mathbf{P}\mathbf{H}^H. \quad (1.24)$$

Cette procédure itérative simple comporte donc 2 étapes : i) estimer la matrice de covariance $\hat{\mathbf{R}}$ (Eq. (1.24)), ii) estimer le vecteur d'amplitude $\hat{\mathbf{x}}^{\text{w-iaa}}$ (Eq. (1.23)). L'initialisation du vecteur \mathbf{x} se fait grâce à la sommation cohérente large bande (1.15). D'après [87], 15 itérations suffisent à la convergence de l'algorithme, ce qui en fait une méthode très rapide.

1.2.3.3 Résultats

Les algorithmes W-APES, W-Capon, iW-Capon [44] et W-IAA [87] sont maintenant comparés sur des données synthétiques afin d'illustrer leur comportement. Ils sont testés sur un scénario comprenant une cible de SNR égal à 45 dB, avec

$$\text{SNR} = KM \times \frac{\mathcal{E}\{|\alpha_t|^2\}}{\sigma^2}, \quad (1.25)$$

se déplaçant à vitesse lente ($v \approx 0.2v_a$, Fig. 1.4) puis à vitesse rapide ($v \approx 1.2v_a$, Fig. 1.5); la cible est toujours située sur la grille d'analyse spectrale.

Dans le cas d'une cible à vitesse faible, on voit que les quatre algorithmes atténuent bien les lobes secondaires observés sur la sommation cohérente (Fig. 1.4(a)). Notons tout de même qu'ils ne sont pas complètement éliminés par les algorithmes W-APES et W-Capon, en particulier autour de 17 m/s c'est-à-dire $v + v_a$.

En revanche, la différence entre les quatre algorithmes est beaucoup plus marquée lorsqu'il s'agit d'une cible rapide, c'est-à-dire dont la migration est plus importante. On voit effectivement que les algorithmes basés sur une hypothèse de stationnarité (W-APES et W-Capon,

Fig. 1.5(b) et 1.5(c) respectivement) ne parviennent pas à éliminer les lobes secondaires en vitesse. En revanche, la prise en compte de la non-stationnarité dans l'algorithme iW-Capon permet une estimation satisfaisante de la cible, en termes de localisation distance-vitesse et d'amplitude (Fig. 1.5(d)). En ce qui concerne l'algorithme W-IAA, c'est le vecteur $\hat{\mathbf{x}}^{\text{w-iaa}}$ que l'on représente sur la Fig. 1.5(e) et qui ne contient, en théorie, que l'amplitude des cibles. On voit qu'en fait certaines composantes du bruit sont vues comme des cibles de faible amplitude (en-dessous de 0 dB).

1.3 Les représentations parcimonieuses

Dans la section précédente, plusieurs algorithmes adaptés à la forme d'onde large bande ont été présentés. Néanmoins, ces algorithmes n'éliminent pas systématiquement les lobes secondaires, en particulier en vitesse, qui peuvent éventuellement masquer d'autres cibles. Les représentations parcimonieuses présentent un attrait particulier pour résoudre ce genre de problèmes puisqu'elles permettent la représentation du signal d'intérêt dépourvue de lobes secondaires. Précisons néanmoins que même s'ils n'apparaissent pas dans la représentation du signal, les lobes secondaires injectent de la corrélation entre les colonnes du dictionnaire, ce qui peut malgré tout perturber l'estimation de cibles localisées à leur endroit.

1.3.1 Mise en forme du problème

Afin d'introduire le modèle de signal utilisé en SSR, considérons le cas simple d'une scène radar ne contenant qu'une seule cible ; on néglige également le bruit \mathbf{n} . Le signal s'écrit alors

$$\begin{aligned} \mathbf{y} &= \alpha_t \mathbf{a}_t \\ &= \frac{1}{\sqrt{KM}} \mathbf{a}_t \times \sqrt{KM} \alpha_t \\ &\triangleq \mathbf{h}_i x_i \\ &\triangleq \mathbf{H} \mathbf{x} \end{aligned}$$

où \mathbf{x} est un vecteur de taille $\bar{K}\bar{M}$ qui représente l'amplitude complexe des cibles ; il possède ici un seul élément non-nul, $x_i = \sqrt{KM} \alpha_t$ (amplitude post-traitement de la cible). Ce vecteur peut donc être qualifié de « parcimonieux » : $\|\mathbf{x}\|_0 = 1 \ll \bar{K}\bar{M}$, où $\|\cdot\|_0$ désigne la norme ℓ_0 ¹, c'est-à-dire le nombre d'éléments non-nuls du vecteur considéré. Il permet bien une représentation sans lobes secondaires de la cible présente dans le signal, puisque celle-ci n'est associée qu'à un seul élément du vecteur parcimonieux \mathbf{x} .

La matrice \mathbf{H} de taille $KM \times \bar{K}\bar{M}$ est un dictionnaire qui définit l'espace de représentation du vecteur parcimonieux \mathbf{x} . Comme pour le modèle de l'IAA (1.21), il découle de la discrétisation de l'espace distance-vitesse, et \bar{K}, \bar{M} correspondent aux dimensions de reconstruction

1. En réalité, ℓ_0 n'est pas une norme à proprement parler mais on n'utilise cet abus de langage dans le reste du manuscrit.

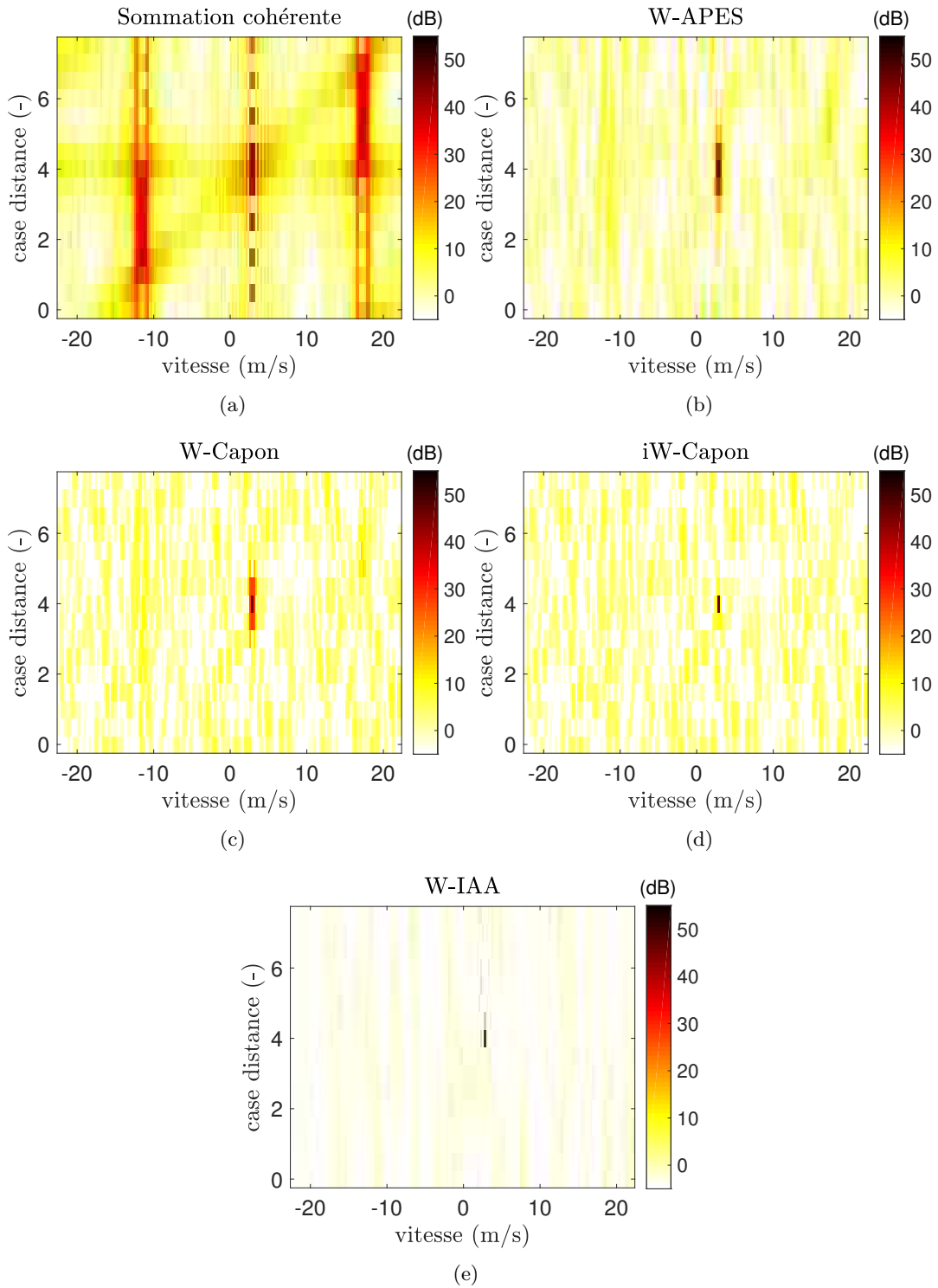


FIGURE 1.4 – Cartes distance-vitesse obtenues après sommation cohérente large bande 1.4(a) et grâce à différents algorithmes d’estimation spectrale adaptative : 1.4(b) W-APES, 1.4(c) W-Capon, 1.4(d) iW-Capon, 1.4(e) W-IAA. Ce scénario correspond à une cible à faible vitesse ($v = .2v_a$) à la 4ème case distance. $K = 8$, $M = 32$, $n_{va} = 3$, $v_a = 15$ m/s, $B/f_c=10\%$, $\sigma^2 = 1$.

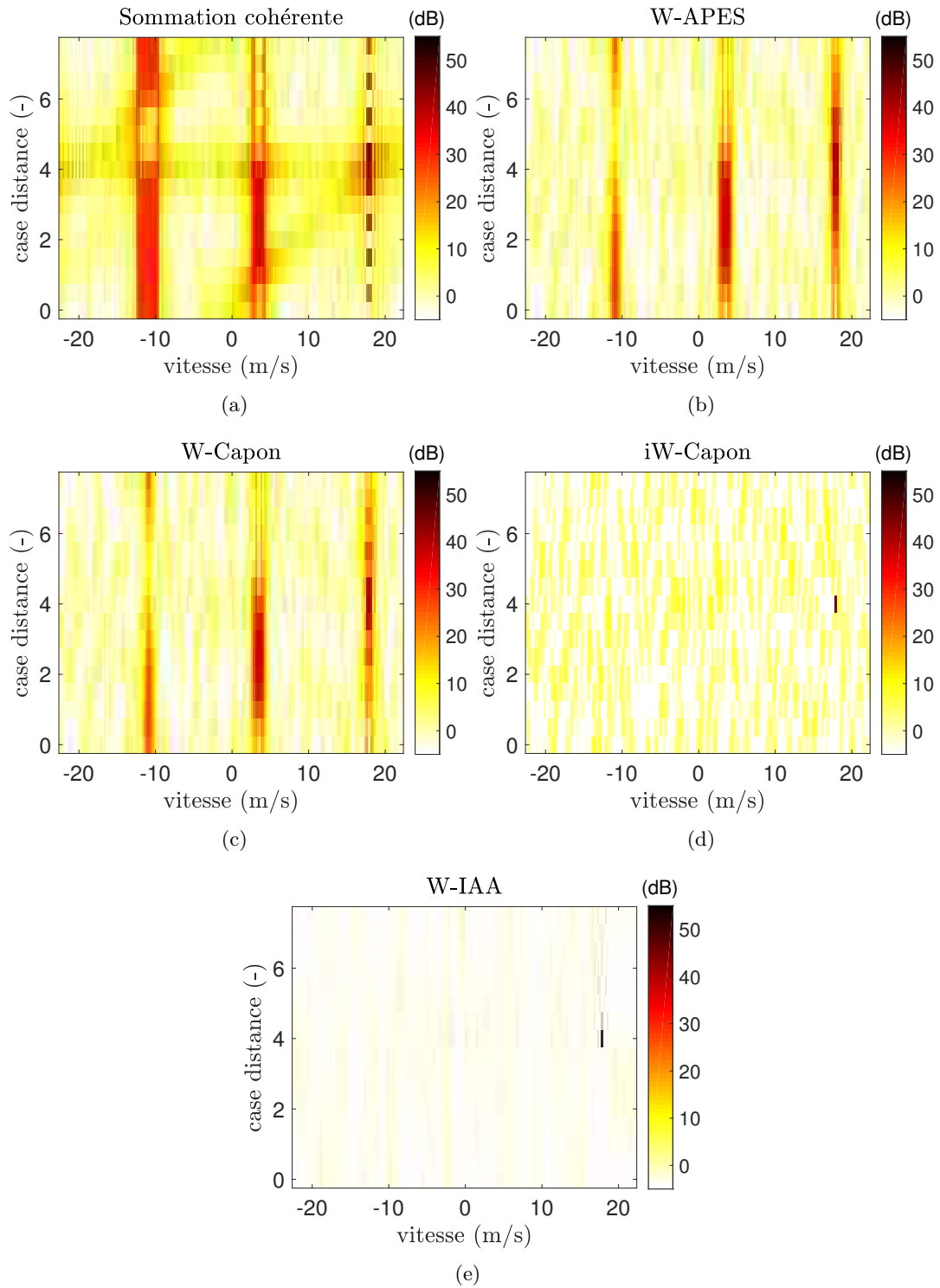


FIGURE 1.5 – Cartes distance-vitesse obtenues après sommation cohérente large bande 1.5(a) et grâce à différents algorithmes d'estimation spectrale adaptative : 1.5(b) W-APES, 1.5(c) W-Capon, 1.5(d) iW-Capon, 1.5(e) W-IAA. Ce scénario correspond à une cible à forte vitesse ($v = 1.2v_a$) à la 6ème case distance. $K = 8$, $M = 32$, $n_{va} = 3$, $v_a = 15$ m/s, $B/f_c = 10\%$, $\sigma^2 = 1$.

de la scène en distance et en vitesse respectivement. La \bar{i} -ième colonne du dictionnaire \mathbf{H} s'écrit

$$[\mathbf{h}_{\bar{i}}]_i = \frac{1}{\sqrt{KM}} \exp\{-j2\pi f_{r\bar{i}}k\} \exp\{j2\pi f_{d\bar{i}}(1 + \mu k)m\} \quad (1.26)$$

où $f_{r\bar{i}}$ et $f_{d\bar{i}}$ sont les fréquences rapide et Doppler normalisées pour la \bar{i} -ième case de la grille d'analyse. On utilise donc le même dictionnaire que pour W-IAA, au facteur de normalisation près $1/\sqrt{KM}$. Les indices k, m, i et $\bar{k}, \bar{m}, \bar{i}$ utilisés pour indiquer le dictionnaire \mathbf{H} sont liés par les relations :

$$\begin{aligned} i &= m + kM, & k &\in \{0, \dots, K-1\}, & m &\in \{0, \dots, M-1\}, \\ \bar{i} &= \bar{m} + \bar{k}\bar{M}, & \bar{k} &\in \{0, \dots, \bar{K}-1\}, & \bar{m} &\in \{0, \dots, \bar{M}-1\}. \end{aligned}$$

Grille d'analyse en distance La grille d'analyse en distance est divisée en \bar{K} cases avec éventuellement un facteur de *zero-padding* n_{zp}^r , de sorte que $\bar{K} = n_{zp}^r K$. La fréquence rapide normalisée pour une cible située à la \bar{k} -ième case distance s'écrit alors

$$f_{r\bar{i}} = \frac{\bar{k}}{\bar{K}}. \quad (1.27)$$

Grille d'analyse en vitesse Afin de lever les ambiguïtés, la carte distance-vitesse est dépliée avec un facteur n_{va} sur l'axe des vitesses; la grille d'analyse en vitesse s'étend donc de $-n_{va}v_a/2$ à $n_{va}v_a/2$. Cette grille est divisée en \bar{M} cases avec un éventuel facteur de *zero-padding* n_{zp}^v , de sorte que $\bar{M} = n_{zp}^v n_{va} M$. La fréquence Doppler normalisée s'écrit alors

$$f_{d\bar{i}} = \frac{n_{va}}{\bar{M}} \bar{m}' \quad (1.28)$$

où $\bar{m}' = \bar{m}$ si $\bar{m} \in \{0, \dots, \bar{M}/2 - 1\}$; $\bar{m}' = \bar{m} - \bar{M}$ si $\bar{m} \in \{\bar{M}/2, \dots, \bar{M} - 1\}$, en supposant \bar{M} pair.

Le raisonnement décrit ici peut être appliqué à une scène radar comprenant plusieurs cibles noyées dans du bruit. On écrit alors

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \quad (1.29)$$

Le vecteur parcimonieux \mathbf{x} possède ainsi N_t éléments non-nuls, pourvu que les cibles soient situées sur la grille d'analyse, qui correspondent à l'amplitude complexe post-traitement de chaque cible. Le vecteur parcimonieux \mathbf{x} constitue donc le paramètre d'intérêt que l'on cherche à estimer.

1.3.1.1 Résolution d'un problème linéaire

En considérant le modèle de signal (1.29), plusieurs méthodes peuvent être mises en œuvre pour estimer le vecteur parcimonieux \mathbf{x} . On considère dans un premier temps le cas $\mathbf{n} = \mathbf{0}$.

La solution des moindres carrés au problème $\mathbf{y} = \mathbf{H}\mathbf{x}$ s'écrit $\mathbf{x}_{LS} = \mathbf{H}^+\mathbf{y}$ où \cdot^+ désigne la pseudo-inverse, puisque le dictionnaire \mathbf{H} n'est pas inversible. En effet, il possède plus de colonnes que de lignes, ce qui signifie que le problème $\mathbf{y} = \mathbf{H}\mathbf{x}$ est mal posé : on ne peut trouver une unique solution à ce problème et il faut donc le contraindre.

Une première approche consisterait à chercher la solution avec le minimum d'énergie, ce qui revient à résoudre le problème

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_2 \quad \text{tel que} \quad \mathbf{H}\mathbf{x} = \mathbf{y} \quad (1.30)$$

où $\|\mathbf{x}\|_2 = \sum_{\bar{i}=0}^{\bar{K}\bar{M}-1} |x_{\bar{i}}|^2$. Avec une telle formulation, $\hat{\mathbf{x}}$ possède en fait de nombreux éléments non-nuls de faible amplitude, alors que l'on cherche à obtenir un vecteur avec peu d'éléments non-nuls mais d'amplitude (potentiellement) élevée. L'idéal serait donc de contraindre \mathbf{x} par sa norme ℓ_0 et de résoudre le problème

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{tel que} \quad \mathbf{H}\mathbf{x} = \mathbf{y}. \quad (1.31)$$

Malheureusement, ce problème est dit « NP-difficile », ce qui signifie qu'il faudrait tester toutes les combinaisons possibles de coefficients de \mathbf{x} afin de trouver celle donnant la norme ℓ_0 la plus faible.

1.3.1.2 Régularisation sous contrainte ℓ_1

Etant données les difficultés liées à l'utilisation de la norme ℓ_0 , le problème (1.31) est relaxé en utilisant son approximation convexe, à savoir la norme ℓ_1 [84, 33, 16] :

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{tel que} \quad \mathbf{H}\mathbf{x} = \mathbf{y}. \quad (1.32)$$

Cette mise en forme, appelée *Basis-Pursuit* (BP) a été présentée dans [38]. Si l'on considère de nouveau le cas $\mathbf{n} \neq \mathbf{0}$, cela revient à résoudre le problème d'optimisation convexe

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{tel que} \quad \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2 \leq \delta, \quad (1.33)$$

dénoté *Basis-Pursuit De-Noising* (BPDN) [38] ; δ représente un paramètre de régularisation à déterminer. Le problème (1.33) peut être mis sous forme non-contrainte en utilisant un multiplicateur de Lagrange λ adéquat :

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \lambda \|\mathbf{x}\|_1 + \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2. \quad (1.34)$$

Notons qu'augmenter la valeur du paramètre λ tend à favoriser une faible norme ℓ_1 quelle que soit la ressemblance du modèle avec les données, donc à trouver une solution nulle : $\hat{\mathbf{x}} = \mathbf{0}$. Au contraire, faire tendre λ vers 0 revient à favoriser la ressemblance aux données, peu importe la parcimonie du vecteur estimé $\hat{\mathbf{x}}$ (idem pour δ dans le problème (1.33)). Ce problème d'optimisation convexe peut être vu comme un problème de *Second-Order Cone (SOC) programming* et être résolu par des méthodes dites du point intérieur [31]. Malheureusement ces méthodes

deviennent lourdes calculatoirement lorsque les données sont de grandes dimensions.

De la même manière, il est possible de contraindre non pas la ressemblance avec les données via le terme $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2$, mais la parcimonie du vecteur \mathbf{x} via le terme $\|\mathbf{x}\|_1$. Cela revient à résoudre le problème équivalent dit du LASSO (*Least Absolute Shrinkage and Selection Operator*) [105]

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2 \quad \text{tel que } \|\mathbf{x}\|_1 \leq \zeta. \quad (1.35)$$

Les trois problèmes (1.33), (1.34) et (1.35) sont équivalents : ils donnent les mêmes solutions, mais par des moyens différents, pourvu que les paramètres de régularisation δ , λ et ζ soient choisis de manière adéquate (ce qui n'est pas trivial). Ces différentes formulations ont permis le développement de plusieurs algorithmes, certains utilisant une approche déterministe et d'autres une approche Bayésienne.

1.3.2 Approches déterministes de représentation parcimonieuse

Plusieurs algorithmes déterministes ont été développés pour trouver la solution aux problèmes (1.33), (1.34) et (1.35), parmi lesquels les algorithmes de seuillage et les algorithmes gloutons. Ces algorithmes ne donnent pas toujours la solution exacte aux problèmes (1.33), (1.34) et (1.35), mais une solution approchée.

1.3.2.1 Algorithmes de seuillage

Les algorithmes de seuillage cherchent à résoudre le problème (1.34) (forme non contrainte), c'est-à-dire à minimiser la fonction coût

$$\lambda \|\mathbf{x}\|_1 + \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 = g(\mathbf{x}) + f(\mathbf{x})$$

que l'on décompose en deux termes : le terme en ℓ_1 , noté $g(\mathbf{x})$, et le terme en ℓ_2 , noté $f(\mathbf{x})$. La fonction f étant différentiable, on en calcule le gradient

$$\nabla f(\mathbf{x}) = \mathbf{H}^H(\mathbf{y} - \mathbf{H}\mathbf{x}).$$

Les algorithmes de seuillage reposent alors sur l'itération suivante (méthode de descente du gradient)

$$\hat{\mathbf{x}}^{(n+1)} = \mathcal{T} \left\{ \hat{\mathbf{x}}^{(n)} - t \nabla f \left(\hat{\mathbf{x}}^{(n)} \right) \right\} \quad (1.36)$$

$$= \mathcal{T} \left\{ \hat{\mathbf{x}}^{(n)} - t \mathbf{H}^H \left(\mathbf{y} - \mathbf{H} \hat{\mathbf{x}}^{(n)} \right) \right\}. \quad (1.37)$$

où t est une valeur de pas adéquate et \mathcal{T} une fonction de seuillage. La $n + 1$ -ième itération comprend donc 2 étapes : i) partir dans la direction inverse au gradient de f (le terme en ℓ_2 de la fonction coût) avec un pas t , ii) ajuster la valeur obtenue grâce à la fonction de seuillage $\mathcal{T}(\cdot)$ pour prendre en compte la parcimonie (via le terme ℓ_1 de la fonction coût). L'opération

de seuillage peut être « douce » (*Soft Thresholding Algorithm*) ou « dure » (*Hard Thresholding Algorithm*).

La fonction de seuillage dur utilisée dans l'algorithme IHT (*Iterative Hard Thresholding*) de [27, 28] s'écrit

$$\mathcal{T}\{z_i\} = \begin{cases} 0 & \text{si } |z_i| \leq \sqrt{\lambda} \\ z_i & \text{sinon} \end{cases} \quad (1.38)$$

Autrement dit, elle consiste à annuler les éléments de $\hat{\mathbf{x}}^{(n)} - t\mathbf{H}^H(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^{(n)})$ inférieurs à un certain seuil, ici $\sqrt{\lambda}$. Une autre fonction de seuillage dur est présentée dans [27], qui ne conserve que les N_t éléments d'amplitude la plus forte (« N_t -sparse algorithm »). L'algorithme IHT garantit une certaine erreur maximale et est robuste au bruit [28], ce qui en fait une solution intéressante grâce à ses performances et sa simplicité.

Une méthode de seuillage doux est proposée dans [42] puis reprise dans [16, 17]. La fonction de seuillage doux s'écrit

$$\mathcal{T}\{z_i\} = \begin{cases} 0 & \text{si } |z_i| \leq \alpha \\ (|z_i| - \alpha) \operatorname{sgn}(z_i) & \text{sinon} \end{cases} \quad (1.39)$$

où $\alpha = \lambda t$ désigne le seuil ; le réglage de ce seuil est donc un réglage du pas t . Contrairement au seuillage fort, les valeurs dépassant le seuil pré-défini ne sont pas conservées mais atténuées. L'algorithme ISTA de [42] ne convergeant qu'en $1/n$, n étant l'indice d'itération, des versions rapides ont été implémentées : les algorithmes FISTA [16] et NESTA [17] qui convergent en $1/n^2$.

1.3.2.2 Algorithmes gloutons

Une autre approche pour la résolution de problèmes de SSR a rencontré beaucoup de succès : les *greedy algorithms*, ou algorithmes gloutons. Ces derniers ont pour but de réduire la charge calculatoire en identifiant en premier lieu le support du paramètre d'intérêt \mathbf{x} . En effet, si l'on désigne par Γ le support de \mathbf{x} alors on peut re-écrire le problème (1.29) comme

$$\mathbf{y} = \mathbf{H}_\Gamma \mathbf{x}_\Gamma + \mathbf{n} \quad (1.40)$$

où \mathbf{H}_Γ désigne le dictionnaire \mathbf{H} privé des colonnes correspondant à des éléments de \mathbf{x} n'appartenant pas au support Γ ; de même pour \mathbf{x}_Γ . Si la parcimonie est vérifiée, alors la matrice \mathbf{H}_Γ possède beaucoup plus de lignes que de colonnes, donc le problème est sur-déterminé et on peut le résoudre par les moindres carrés comme

$$\hat{\mathbf{x}}_\Gamma = \left(\mathbf{H}_\Gamma^H \mathbf{H}_\Gamma \right)^{-1} \mathbf{H}_\Gamma^H \mathbf{y} \quad (1.41)$$

Une des premières méthodes utilisant cette approche a été décrite dans [60, 108] : c'est

l'algorithme CLEAN. Il consiste à rechercher dans le spectre du signal la valeur maximale, assimilée à la cible la plus forte, puis à retirer du signal la signature de cette cible. La procédure est répétée de manière itérative jusqu'à ce qu'un certain critère d'arrêt soit atteint (nombre de cibles, blancheur du signal résiduel). Cette approche peut devenir inefficace si le premier maximum ne correspond pas à une cible, mais à une interférence constructive entre lobes secondaires. Notons également qu'elle est susceptible de mener à de nombreuses fausses alarmes, c'est-à-dire d'estimer des cibles là où il n'y en a pas [2].

Plus récemment, les algorithmes de *Matching Pursuit* (MP) ont été développés et largement utilisés par la communauté. En premier lieu, Mallat et al. ont présenté l'algorithme MP dans [75]. Cet algorithme consiste à construire, à partir du dictionnaire \mathbf{H} , la base orthogonale qui maximise la projection des résidus. Pour cela, une procédure itérative est mise en place, qui identifie à chaque itération la colonne de \mathbf{H} qui maximise la projection des résidus sur celle-ci, puis met à jour les résidus

$$\begin{aligned}\bar{i}^{(n)} &= \arg \max_{i=0, \dots, KM-1} |\mathbf{h}_i^H \mathbf{r}^{(n)}| \\ \mathbf{r}^{(n+1)} &= \mathbf{r}^{(n)} - \frac{\mathbf{h}_{\bar{i}^{(n)}}^H \mathbf{r}^{(n)}}{\|\mathbf{h}_{\bar{i}^{(n)}}\|^2} \mathbf{h}_{\bar{i}^{(n)}}.\end{aligned}$$

L'algorithme *Orthogonal Matching Pursuit* (OMP) a plus tard été présenté dans [107], et a rencontré beaucoup de succès grâce à ses garanties de convergence. Le but est toujours de maximiser la projection du signal \mathbf{y} sur une base adéquate, mais cette fois-ci le paramètre d'intérêt $\hat{\mathbf{x}}_{\Gamma}^{(n)}$ est recalculé à chaque itération via (1.41) en utilisant la base $\mathbf{H}_{\Gamma}^{(n)}$ mise à jour. D'autres algorithmes ont été développés afin d'améliorer les performances de l'OMP. La principale amélioration est de prendre en compte plusieurs coefficients du vecteur \mathbf{x} à chaque itération (contre un seul pour l'OMP); elle est utilisée par exemple dans StOMP (*Stagewise OMP*) [49] qui converge en 10 itérations seulement, mais également dans ROMP (*Regularized OMP*) [80] qui généralise les performances de l'OMP afin d'estimer un signal compressible quelconque à partir de données bruitées. Enfin, CoSaMP (*Compressive Sampling Matching Pursuit*) [79] utilise aussi cette modification et y ajoute une opération de troncature (*pruning*) qui permet de généraliser les résultats à tout dictionnaire respectant les *Restrictive Isometry Properties* (RIP) [30].

1.3.3 Approches Bayésiennes de représentation parcimonieuse

Après avoir abordé les approches déterministes de représentations parcimonieuses, on présente dans cette section les approches Bayésiennes : le paramètre d'intérêt \mathbf{x} est considéré comme une variable aléatoire suivant une certaine distribution a priori (*prior distribution*) $\pi(\mathbf{x})$. L'intérêt d'une telle formulation est d'avoir la possibilité d'injecter une potentielle connaissance du comportement des variables inconnues, mais également de pouvoir ajuster les paramètres de réglage de manière plus « physique ». En effet, si on prend l'exemple de l'algorithme OMP [107] et que l'on ne connaît pas le nombre de cibles présentes sur la scène radar, alors on doit ajuster le critère d'arrêt en fonction des résidus attendus, ce qui n'est

pas trivial (en comparaison avec un réglage par rapport au niveau de bruit par exemple, plus « physique »).

La distribution conditionnelle $f(\mathbf{y}|\mathbf{x})$ renseigne quant à l'information relative à \mathbf{x} présente dans les données \mathbf{y} : elle permet de calculer la distribution a posteriori $f(\mathbf{x}|\mathbf{y})$ grâce à la règle de Bayes

$$f(\mathbf{x}|\mathbf{y}) \propto f(\mathbf{y}|\mathbf{x})\pi(\mathbf{x}). \quad (1.42)$$

On s'intéresse tout particulièrement à la distribution a posteriori $f(\mathbf{x}|\mathbf{y})$ puisqu'elle décrit le comportement du paramètre d'intérêt \mathbf{x} conditionnellement aux données \mathbf{y} , et sera utilisée pour calculer un estimateur de \mathbf{x} .

Dans cette section, nous ferons d'abord le lien entre les approches Bayésiennes et les approches déterministes de représentations parcimonieuses. Ensuite, nous détaillerons comment la parcimonie est forcée via les approches Bayésiennes, le pendant de la régularisation par la norme ℓ_1 dans les approches déterministes. Enfin, nous décrirons certains algorithmes Bayésiens de résolution.

1.3.3.1 Lien avec les approches déterministes

Le paramètre d'intérêt \mathbf{x} peut être estimé via l'estimateur MAP (*Maximum A Posteriori*) comme suit

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \max_{\mathbf{x}} f(\mathbf{x}|\mathbf{y}). \quad (1.43)$$

Comparé à l'estimateur du maximum de vraisemblance, l'estimateur MAP incorpore dans son problème d'optimisation l'information additionnelle dont peut disposer l'utilisateur sur le paramètre à estimer (via sa distribution a priori).

Si l'on considère le bruit \mathbf{n} Gaussien de moyenne nulle et de matrice de covariance $\mathbf{R} \in \mathbb{C}^{KM \times KM}$, alors étant donné le modèle de signal (1.29), $f(\mathbf{y}|\mathbf{x}) \sim \mathcal{CN}(\mathbf{H}\mathbf{x}, \mathbf{R})$. La distribution a posteriori de \mathbf{x} s'écrit alors

$$\begin{aligned} f(\mathbf{x}|\mathbf{y}) &\propto \frac{1}{\pi^{KM} |\mathbf{R}|} \exp \left\{ -(\mathbf{y} - \mathbf{H}\mathbf{x})^H \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \right\} \pi(\mathbf{x}) \\ \log f(\mathbf{x}|\mathbf{y}) &= \text{cste} - (\mathbf{y} - \mathbf{H}\mathbf{x})^H \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) + \log \pi(\mathbf{x}). \end{aligned}$$

En se limitant au cas du bruit blanc Gaussien, $\mathbf{R} = \sigma^2 \mathbf{I}_{KM}$, l'estimateur MAP s'écrit alors

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 - \sigma^2 \log \pi(\mathbf{x}). \quad (1.44)$$

On reconnaît dans cette expression un problème d'optimisation similaire à la forme non-contrainte (1.34) : il s'agit de minimiser la norme $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2$, sous une certaine contrainte $-\sigma^2 \log \pi(\mathbf{x})$. Ainsi, une distribution a priori $\pi(\mathbf{x})$ adéquate permet de forcer la parcimonie du vecteur \mathbf{x} . Notons que la puissance du bruit σ^2 n'est pas forcément connue, donc il pourra être nécessaire de lui assigner également une distribution a priori ; cette approche sera détaillée

par la suite.

1.3.3.2 Distributions forçant la parcimonie

Comme nous l'avons vu précédemment, le choix de la distribution a priori du vecteur d'intérêt \mathbf{x} est crucial puisqu'il doit, entre autres, assurer son caractère parcimonieux. Ainsi, la plupart des éléments du vecteur \mathbf{x} doit pouvoir prendre une valeur proche de 0 (voire exactement nulle), mais quelques éléments prennent eux une valeur potentiellement élevée. Donc si elle est continue, la distribution a priori $\pi(\mathbf{x})$ doit être piquée autour de 0, mais avec une queue lourde, ce qui revient à avoir un moment d'ordre 4 (kurtosis) positif.

Prior hiérarchiques Une stratégie largement adoptée est d'utiliser un schéma hiérarchique pour définir la distribution a priori $\pi(\mathbf{x})$:

$$\pi(\mathbf{x}) = \int \pi(\mathbf{x}|\boldsymbol{\sigma}_x^2)\pi(\boldsymbol{\sigma}_x^2)d\boldsymbol{\sigma}_x^2. \quad (1.45)$$

En supposant en outre que les éléments de \mathbf{x} sont indépendants et identiquement distribués (i. i. d.), alors

$$\pi(\mathbf{x}) = \prod_{\bar{i}=0}^{\bar{K}\bar{M}-1} \pi(x_{\bar{i}}), \quad (1.46)$$

$$\pi(\mathbf{x}|\boldsymbol{\sigma}_x^2) = \prod_{\bar{i}=0}^{\bar{K}\bar{M}-1} \pi(x_{\bar{i}}|\sigma_{x_{\bar{i}}}^2). \quad (1.47)$$

Cette formulation hiérarchique permet le développement simplifié des algorithmes d'estimation du paramètre d'intérêt \mathbf{x} .

En particulier, la distribution de Laplace $\pi(x_{\bar{i}}) = \frac{1}{2b}e^{-|x_{\bar{i}}|/b}$ est très populaire étant donnée sa forme piquée autour de 0 qui favorise la parcimonie du vecteur \mathbf{x} . De plus, si on écrit l'estimateur MAP pour une telle distribution, on obtient

$$\begin{aligned} \hat{\mathbf{x}}_{\text{MAP}} &= \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 - \sigma^2 \log \pi(\mathbf{x}) \\ \hat{\mathbf{x}}_{\text{MAP}} &= \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 - \sigma^2 \log \prod_{\bar{i}=0}^{\bar{K}\bar{M}-1} \pi(x_{\bar{i}}) \\ &= \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \sigma^2/b \sum_{\bar{i}=0}^{\bar{K}\bar{M}-1} |x_{\bar{i}}| \\ &= \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \sigma^2/b \|\mathbf{x}\|_1 \end{aligned}$$

et on reconnaît le problème d'optimisation (1.34). Autrement dit, résoudre le problème déterministe de minimisation ℓ_2 sous contrainte ℓ_1 revient à calculer l'estimateur MAP pour une distribution a priori de type Laplace. En fait, si on considère $\mathbf{x} \in \mathbb{R}^{\bar{K}\bar{M} \times 1}$ on peut obtenir une

distribution de Laplace en utilisant une loi $f(x_{\bar{i}}|\sigma_{x_{\bar{i}}}^2)$ Gaussienne centrée de variance $\sigma_{x_{\bar{i}}}^2$ avec une loi de mélange $\pi(\sigma_{x_{\bar{i}}}^2)$ exponentielle de paramètre $1/(2b^2)$ [57] :

$$\begin{aligned}
\pi(x_{\bar{i}}) &= \int_{-\infty}^{\infty} f(x_{\bar{i}}|\sigma_{x_{\bar{i}}}^2)\pi(\sigma_{x_{\bar{i}}}^2)d\sigma_{x_{\bar{i}}}^2 \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_{x_{\bar{i}}}^2}} \exp\left\{-\frac{1}{2\sigma_{x_{\bar{i}}}^2}|x_{\bar{i}}|^2\right\} \times \frac{1}{2b^2} \exp\left\{-\sigma_{x_{\bar{i}}}^2/(2b^2)\right\} \mathbb{I}_{[0,+\infty)}(\sigma_{x_{\bar{i}}}^2)d\sigma_{x_{\bar{i}}}^2 \\
&= \frac{1}{b^2\sqrt{2\pi}} \int_0^{\infty} \exp\left\{-\frac{1}{2}\left(x_{\bar{i}}^2\gamma^{-2} + b^{-2}\gamma^2\right)\right\} d\gamma \\
&= \frac{1}{2b} e^{-|x_{\bar{i}}|/b}.
\end{aligned}$$

grâce à un changement de variable et en utilisant le calcul d'intégrale donné dans [11]. Si $\mathbf{x} \in \mathbb{C}^{\bar{K}\bar{M}\times 1}$, alors en utilisant une loi $f(x_{\bar{i}}|\sigma_{x_{\bar{i}}}^2)$ complexe Gaussienne centrée de variance $\sigma_{x_{\bar{i}}}^2$ et une loi de mélange $\pi(\sigma_{x_{\bar{i}}}^2)$ exponentielle on obtient une distribution de Laplace pour la partie réelle et la partie imaginaire de $x_{\bar{i}}$.

D'autre part, la loi de Student-t est également intéressante pour forcer la parcimonie car elle est symétrique, comme la loi Gaussienne, mais possède une plus lourde queue (c'est-à-dire qu'elle a davantage tendance à favoriser les valeurs éloignées de sa moyenne). Un schéma hiérarchique est présenté dans [106] pour l'obtenir, utilisé ensuite dans l'algorithme *Relevance Vector Machine* (RVM). Ce prior hiérarchique, et le schéma d'estimation associé, ont été très largement repris dans la littérature, par exemple [13, 65]. Il s'agit donc d'utiliser une distribution Gaussienne centrée $f(x_{\bar{i}}|\sigma_{x_{\bar{i}}}^2)$ de variance $\sigma_{x_{\bar{i}}}^2$ et une loi de mélange $\pi(\sigma_{x_{\bar{i}}}^2)$ inverse-Gamma de paramètres (a, b) :

$$\begin{aligned}
\pi(x_{\bar{i}}) &= \int_{-\infty}^{\infty} f(x_{\bar{i}}|\sigma_{x_{\bar{i}}}^2)\pi(\sigma_{x_{\bar{i}}}^2)d\sigma_{x_{\bar{i}}}^2 \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_{x_{\bar{i}}}^2}} \exp\left\{-\frac{1}{2\sigma_{x_{\bar{i}}}^2}|x_{\bar{i}}|^2\right\} \times \frac{b^a}{\Gamma(a)} \sigma_{x_{\bar{i}}}^{-2(a+1)} e^{-b\sigma_{x_{\bar{i}}}^{-2}} \mathbb{I}_{[0,+\infty)}(\sigma_{x_{\bar{i}}}^2)d\sigma_{x_{\bar{i}}}^2 \\
&= \int_0^{\infty} \frac{1}{\sqrt{2\pi}} \frac{b^a}{\Gamma(a)} \sigma_{x_{\bar{i}}}^{-2(1/2+a+1)} e^{-\sigma_{x_{\bar{i}}}^{-2}(|x_{\bar{i}}|^2/2+b)} d\sigma_{x_{\bar{i}}}^2 \\
&= \frac{1}{\sqrt{2\pi b}} \frac{\Gamma(a+1/2)}{\Gamma(a)} \left(1 + \frac{|x_{\bar{i}}|^2}{2b}\right)^{-(a+1/2)}
\end{aligned}$$

qui est bien une loi de Student-t [106]. Notons que le même raisonnement est applicable si $\mathbf{x} \in \mathbb{C}^{\bar{K}\bar{M}\times 1}$.

D'autres distributions forçant la parcimonie peuvent également être utilisées grâce à un schéma hiérarchique. Par exemple, une loi de Laplace $f(x_{\bar{i}}|\sigma_{x_{\bar{i}}}^2)$ et une loi de mélange $\pi(\sigma_{x_{\bar{i}}}^2)$ inverse-Gamma [123] donnent une distribution de Pareto généralisée très piquée en 0 comme la loi de Laplace, mais possédant une queue semblable à la loi de Student-t [12, 14].

Distributions mixtes Une autre manière de forcer la parcimonie du paramètre d'intérêt \mathbf{x} est d'utiliser une distribution a priori dite « mixte » [68, 24]. Une telle distribution permet d'assigner à certains éléments du vecteur \mathbf{x} une valeur exactement nulle, avec une certaine probabilité $1 - w$; les autres coefficients suivent une distribution continue avec une probabilité w . La distribution a priori de $\pi(\mathbf{x})$ peut alors s'écrire

$$\begin{aligned}\pi(\mathbf{x}) &= \prod_{i=0}^{\bar{K}\bar{M}-1} \pi(x_{\bar{i}}) \\ &= \prod_{i=0}^{\bar{K}\bar{M}-1} (1 - w)\delta(x_{\bar{i}}) + wf_c(x_{\bar{i}})\end{aligned}$$

et fait partie de la classe des distributions *spike and slab* [78, 121]. L'avantage d'une telle distribution est qu'elle permet de découpler le réglage de la parcimonie (liée à la probabilité $1 - w$) et celui du niveau d'amplitude des cibles (liée à la distribution $f_c(x_{\bar{i}})$) contrairement aux distributions présentées précédemment qui doivent à la fois être piquées autour de 0 (réglage de la parcimonie) et posséder une longue queue (valeur non-nulles potentiellement élevées).

Notons que la distribution $f_c(x_{\bar{i}})$ peut être obtenue grâce à un schéma hiérarchique, par exemple en suivant le schéma RVM pour une distribution de Student-t [121, 24, 29]. D'autre part, la probabilité w qu'une cible soit présente à la case d'analyse concernée est souvent inconnue, même si l'on pressent qu'elle doit approcher $N_t/(\bar{K}\bar{M})$. Un étage peut donc être ajouté au schéma hiérarchique : la probabilité w est considérée comme une variable aléatoire avec une certaine distribution a priori. Notons que c'est une distribution mixte qui a été choisie dans l'algorithme détaillé dans la Section 1.4 et dans les versions robustifiées des chapitres 2 et 3.

1.3.3.3 Algorithmes de résolution Bayésienne

Comme nous l'avons mentionné précédemment, pour estimer le paramètre d'intérêt \mathbf{x} il faut étudier la distribution a posteriori $f(\mathbf{x}|\mathbf{y})$. Jusqu'à présent, nous nous sommes focalisés sur l'estimateur MAP, c'est-à-dire sur la recherche de la valeur de \mathbf{x} maximisant cette distribution. Il est toutefois possible d'estimer le paramètre d'intérêt \mathbf{x} d'une autre manière, par exemple en calculant l'estimateur MMSE (*Minimum Mean Square Error*) qui minimise l'erreur quadratique moyenne :

$$\hat{\mathbf{x}}_{\text{MMSE}} = \int \mathbf{x}f(\mathbf{x}|\mathbf{y})d\mathbf{x}, \quad (1.48)$$

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \max_{\mathbf{x}} f(\mathbf{x}|\mathbf{y}). \quad (1.49)$$

L'estimateur MMSE correspond donc à la moyenne a posteriori du vecteur \mathbf{x} , alors que l'estimateur MAP donne le mode principal de cette distribution. Malheureusement, la distribution a posteriori $f(\mathbf{x}|\mathbf{y})$ ne peut pas toujours être intégrée ou maximisée analytiquement en forme directe. Des méthodes d'approximation sont alors mises en place pour calculer les estimateurs ;

on distingue alors les méthodes d’approximation stochastiques (utilisées dans ces travaux) et les méthodes d’approximation déterministes [26, chap. 10].

Méthodes d’approximation stochastiques Dans l’hypothèse de moyens de calculs infinis, ces méthodes donneraient un résultat exact, mais, ne disposant que d’une ressource finie, on obtient une approximation du résultat. Parmi les méthodes d’approximation stochastiques, les méthodes de Monte-Carlo par chaînes de Markov (*Monte-Carlo Markov Chain*, MCMC) sont très largement utilisées et appliquées dans de nombreux domaines [89]. Ici, il s’agit de tirer $N_r + N_{bi}$ échantillons $\mathbf{x}^{(n)}$ suivant la distribution a posteriori $f(\mathbf{x}|\mathbf{y})$ afin de calculer les estimateurs MMSE et MAP comme suit

$$\hat{\mathbf{x}}_{\text{MMSE}} = \frac{1}{N_r} \sum_{n=1}^{N_r} \mathbf{x}^{(n+N_{bi})} \quad (1.50)$$

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \max_{\mathbf{x}^{(N_{bi} < n < N_r)}} f(\mathbf{x}^{(n)}|\mathbf{y}) \quad (1.51)$$

où N_{bi} désigne la période de chauffe de l’algorithme et N_r le nombre d’échantillons utilisés pour calculer les estimateurs. Plus précisément, on peut utiliser un échantillonneur de Gibbs lorsque le tirage suivant la loi a posteriori $f(\mathbf{x}|\mathbf{y})$ n’est pas trivial : on génère alors de manière itérative des échantillons $\mathbf{x}^{(n)}$ suivant leur distribution conditionnelle a posteriori $f(\mathbf{x}^{(n)}|\mathbf{y}, \sigma_x^2)$. Après une période de chauffe N_{bi} , ces échantillons suivent leur distribution a posteriori, et on peut alors approcher les estimateurs grâce à N_r échantillons comme précédemment. Cette méthode est utilisée dans l’algorithme détaillé dans la section suivante, et dans les travaux présentés dans la suite du manuscrit.

D’autre part, lorsque la loi conditionnelle a posteriori utilisée dans l’échantillonneur de Gibbs n’est pas une distribution connue, une étape de Metropolis-Hastings (MH) peut être utilisée² : il s’agit de tirer un échantillon suivant une loi de proposition, puis d’accepter ou rejeter cet échantillon avec une certaine probabilité d’acceptation [90]. On parle alors d’échantillonneur de Gibbs hybride ; cet algorithme sera détaillé dans le chapitre 2. Mentionnons également le *partially collapsed Gibbs sampler* (PCGS) [110, 83, 48] qui permet d’accélérer la convergence d’un échantillonneur de Gibbs « classique » en marginalisant une ou plusieurs distributions conditionnelles a posteriori par rapport à une autre variable supplémentaire du schéma hiérarchique Bayésien.

Néanmoins, l’inconvénient principal des méthodes MCMC réside dans leur charge calculatoire puisqu’il faut répéter le tirage de variables aléatoires $N_r + N_{bi}$ fois.

Méthodes d’approximation déterministes Ces méthodes utilisent une approximation analytique de la distribution a posteriori pour calculer les estimateurs, par exemple en supposant que cette distribution peut se mettre sous une certaine forme factorisée. Puisqu’elles utilisent une approximation de la distribution a posteriori, ces méthodes ne peuvent pas donner un résultat exact [26] mais elles possèdent en contre-partie un coût calculatoire réduit par

2. Notons que l’algorithme MH peut également être utilisé indépendamment d’un échantillonneur de Gibbs.

rapport aux méthodes MCMC.

L'algorithme *Expectation-Maximization* (EM) rencontre un succès considérable comme moyen de calcul des estimateurs ML ou MAP, par exemple dans [57, 106] et dans l'algorithme OGSBI de [119] utilisé dans ce manuscrit à titre de comparaison. Considérons en premier lieu le cas de l'estimateur du maximum de vraisemblance où le but est de maximiser la fonction log-vraisemblance : $\hat{\mathbf{x}}_{ML} = \arg \max \log f(\mathbf{y}|\mathbf{x})$. L'idée de l'algorithme EM est d'introduire des « données cachées », notées \mathbf{z} , afin de former un set de données complètes $\{\mathbf{y}, \mathbf{z}\}$ et de maximiser la log-vraisemblance des données complètes $f(\mathbf{y}, \mathbf{z}|\mathbf{x})$ (et pas celle des données « incomplètes » $f(\mathbf{y}|\mathbf{x})$). Cette maximisation se fait alors en deux étapes

E-step on approxime $f(\mathbf{y}, \mathbf{z}|\mathbf{x})$ par la fonction $Q(\mathbf{x}, \mathbf{x}^{old}) = \mathcal{E}_{\mathbf{z}|\mathbf{y}, \mathbf{x}^{old}} \{\log f(\mathbf{y}, \mathbf{z}|\mathbf{x})\}$

M-step on maximise la fonction $Q(\mathbf{x}, \mathbf{x}^{old})$: $\mathbf{x}^{new} = \arg \max_{\mathbf{x}} Q(\mathbf{x}, \mathbf{x}^{old})$.

L'algorithme EM garantit alors la croissance de la log-vraisemblance des données incomplètes, à moins qu'elle ait déjà atteint un maximum local [26]. Afin d'utiliser l'algorithme EM, il convient donc de choisir les données manquantes \mathbf{z} . La méthode dite de « Type I » [57] utilise l'hyperparamètre σ_x^2 comme donnée manquante et permet de calculer

$$\begin{aligned}\hat{\mathbf{x}}_{ML} &= \arg \max_{\mathbf{x}} Q(\mathbf{x}, \mathbf{x}^{old}) \\ \hat{\mathbf{x}}_{MAP} &= \arg \max_{\mathbf{x}} Q(\mathbf{x}, \mathbf{x}^{old}) + \log \pi(\mathbf{x}).\end{aligned}$$

La méthode dite de « Type II » [106, 119] utilise par contre le vecteur \mathbf{x} comme variable cachée pour obtenir $\hat{\sigma}_x^2$; une approximation est ensuite utilisée pour calculer $\hat{\mathbf{x}}_{MAP}$ comme $\hat{\mathbf{x}}_{MAP} = \arg \max_{\mathbf{x}} f(\mathbf{y}|\mathbf{x}, \hat{\sigma}_x^2)$.

On mentionne également la méthode d'approximation variationnelle Bayésienne (VB) qui a été utilisée dans [23] pour donner une version accélérée de l'algorithme de base présenté dans la section suivante. Il s'agit alors de choisir une distribution variationnelle qui minimise la divergence de Kullback-Leibler, par exemple en utilisant une forme factorisée de cette distribution [26]. La méthode VB peut être vue comme une extension de l'algorithme EM : ce dernier calcule l'estimateur MAP comme la valeur la plus probable de chaque paramètre, alors que VB calcule (une approximation de) la distribution a posteriori complète des paramètres et variables cachées pour obtenir les estimateurs désirés.

1.4 Une approche Bayésienne pour la représentation parcimonieuse d'un signal radar large bande

Dans cette section, l'algorithme développé dans [24] est détaillé puisqu'il a servi de base au travail de recherche présenté dans la suite du manuscrit. Cet algorithme permet d'obtenir une représentation parcimonieuse d'une scène radar large bande grâce à une approche Bayésienne. Il possède néanmoins certaines limitations présentées en fin de section, qui ont motivé la robustification progressive de l'algorithme.

1.4.1 Présentation de l'algorithme

On rappelle que l'on considère un signal composé de M impulsions émises avec une faible PRF F_r . Ce signal, correspondant à K cases distances, est exprimé dans le domaine fréquence-rapide / temps-long.

1.4.1.1 Mise en forme du problème

Le modèle de signal considéré initialement est celui décrit par (1.5), (1.12), (1.13), (1.14). Le vecteur bruit \mathbf{N} est considéré blanc Gaussien de puissance σ^2 . Le signal est ensuite interprété via les représentations parcimonieuses et le modèle de signal (1.29) introduit précédemment est adopté :

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}.$$

Afin d'estimer le paramètre d'intérêt \mathbf{x} , il faut choisir un dictionnaire \mathbf{H} adéquat. Nous avons vu précédemment qu'un dictionnaire basé sur la signature de cible peut être adopté ; il est décrit par les équations (1.26), (1.27), (1.28).

Approximation du dictionnaire Dans [24], la signature de la cible est interprétée différemment et un dictionnaire « approché » est adopté. En effet, la signature de la cible $[\mathbf{A}]_{k,m}$ est réécrite comme suit :

$$[\mathbf{A}]_{k,m} = \exp \left\{ j2\pi \left(-\tau_0 \frac{B}{K} k + \frac{2vf_c}{c} T_{r,k} m \right) \right\} \quad (1.52)$$

avec

$$T_{r,k} = (1 + \mu k) T_r.$$

Cette formulation permet de re-interpréter la signature de la cible : pour une sous-bande donnée, la signature est une sinusoïde 1D de fréquence Doppler non-normalisée $2vf_c/c$ mais avec une fréquence d'échantillonnage $T_{r,k}$ qui dépend proportionnellement de l'indice de sous-bande [24]. Cela illustre le phénomène de repliement en vitesse, qui augmente avec l'indice de sous-bande.

L'idée est alors d'utiliser une version tronquée du théorème de reconstruction de Shannon [96] pour exprimer la signature $[\mathbf{A}]_{k,m}$: \mathbf{A} peut être interpolée à partir d'une cisoïde 2D « non-repliée » $\bar{\mathbf{A}}$ de fréquence $(-\tau_0, 2vf_c/c)$ et de périodes d'échantillonnage $(B/K, \bar{T}_r)$

$$[\mathbf{A}]_{k,m} = \sum_{\bar{m}=0}^{\bar{M}-1} [\bar{\mathbf{A}}]_{k,\bar{m}} \operatorname{sinc} \left\{ \pi \left(m \frac{T_{r,k}}{\bar{T}_r} - \bar{m} \right) \right\} \quad (1.53)$$

où \bar{T}_r respecte le théorème de Shannon. \bar{M} désigne le nombre d'impulsions virtuelles considérées ; il est choisi de manière à ce que la formule (1.53) reste bien une formule d'interpolation, et pas d'extrapolation. Il est le pendant de la discrétisation de l'espace Doppler détaillé dans (1.28). En remplaçant l'expression de $[\mathbf{A}]_{k,m}$ (1.53) dans le modèle de signal (1.5), et après

vectorisation, on obtient le dictionnaire approché

$$\mathbf{H} = \mathbf{T}\mathbf{F}^{-1}$$

avec \mathbf{T} la matrice de mesure diagonale par bloc qui reprend la formule (1.53), et \mathbf{F} la matrice de passage du domaine fréquence-courte / temps-long au domaine temps-rapide / fréquence-lente. Ces matrices sont explicitées dans [24].

L'algorithme présenté dans [24] et détaillé par la suite peut être utilisé de manière équivalente avec le « vrai » dictionnaire ou bien sa version approchée.

1.4.1.2 Modèle hiérarchique Bayésien

Le paramètre d'intérêt \mathbf{x} , qui contient l'information d'amplitude des cibles, est estimé via le paradigme Bayésien. Pour cela, un modèle hiérarchique Bayésien est mis en place : les paramètres inconnus sont vus comme des variables aléatoires suivant une certaine distribution a priori. Ces distributions peuvent dépendre d'hyperparamètres inconnus auxquels on assigne une distribution a priori, et ainsi de suite. Le modèle hiérarchique utilisé dans [24] est représenté graphiquement dans la Fig. 1.6 ; il est détaillé par la suite.

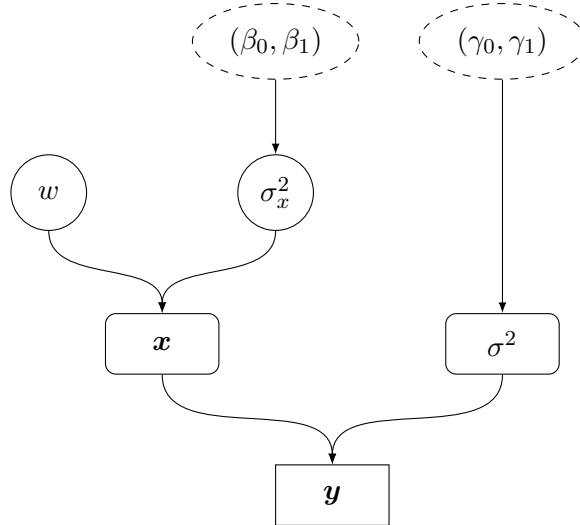


FIGURE 1.6 – Modèle hiérarchique Bayésien utilisé dans l'algorithme [24]. Les paramètres dans les cercles en pointillés doivent être réglés par l'utilisateur.

Vraisemblance On rappelle que le bruit \mathbf{n} est considéré blanc Gaussien de moyenne nulle et de variance σ^2 . Ainsi, les données \mathbf{y} sont également distribuées suivant une Gaussienne de variance σ^2 mais de moyenne non-nulle $\mathbf{H}\mathbf{x}$: $\mathbf{y} \sim \mathcal{CN}(\mathbf{H}\mathbf{x}, \sigma^2 \mathbf{I}_{KM})$. La fonction de vraisemblance s'écrit alors

$$f(\mathbf{y}|\mathbf{x}, \sigma^2) \propto \frac{1}{\sigma^{2KM}} \exp \left\{ -\sigma^{-2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 \right\}. \quad (1.54)$$

Les paramètres σ^2 et \mathbf{x} étant inconnus, ils sont considérés comme des variables aléatoires et un étage est ajouté au modèle hiérarchique afin de leur attribuer une distribution a priori.

Paramètres

Puissance du bruit La distribution a priori assignée au paramètre σ^2 permet de simplifier les calculs : une distribution inverse-Gamma (conjuguée de la Gaussienne) est choisie, $\sigma^2 \sim \mathcal{IG}(\gamma_0, \gamma_1)$. Elle s'écrit :

$$\pi(\sigma^2 | \gamma_0, \gamma_1) \propto \frac{e^{-\gamma_1/\sigma^2}}{(\sigma^2)^{\gamma_0+1}} \mathbb{I}_{[0,+\infty)}(\sigma^2) \quad (1.55)$$

Notons cependant que cette distribution peut respecter dans une certaine mesure le sens physique puisque ses paramètres (γ_0, γ_1) permettent d'ajuster la moyenne et la variance de la distribution en fonction de la connaissance que l'utilisateur a du niveau de bruit thermique.

Vecteur d'amplitude Comme expliqué précédemment, la distribution a priori assignée au vecteur \mathbf{x} doit forcer la parcimonie de ce dernier. Dans [24], le choix d'une distribution mixte est fait. Les éléments du vecteur \mathbf{x} sont considérés i.i.d. et un prior de type Bernoulli-Gaussien est assigné à chacun d'eux : $x_{\bar{i}} \sim \text{BerCN}(w, 0, \sigma_x^2)$. La distribution a priori de $x_{\bar{i}}$ s'écrit donc

$$\pi(x_{\bar{i}} | w, \sigma_x^2) = (1 - w)\delta(|x_{\bar{i}}|) + w \frac{1}{\pi\sigma_x^2} \exp\left\{-\frac{|x_{\bar{i}}|^2}{\sigma_x^2}\right\}. \quad (1.56)$$

Avec cette distribution, les éléments $x_{\bar{i}}$ sont nuls avec une probabilité $(1 - w)$; ils suivent une distribution Gaussienne centrée de puissance σ_x^2 avec une probabilité w . Ainsi, comme mentionné précédemment, le réglage de la parcimonie et celui de la puissance des cibles sont découplés. Les hyperparamètres w et σ_x^2 sont inconnus, donc un dernier étage est ajouté au modèle hiérarchique Bayésien.

Hyperparamètres

Probabilité Puisque le niveau de parcimonie de la scène radar n'est a priori pas connu, une distribution a priori uniforme est choisie pour la probabilité w : $w \sim \mathcal{U}_{[0,1]}$.

Puissance des cibles Comme pour la puissance du bruit σ^2 , une distribution a priori inverse-Gamma est choisie pour la puissance des cibles σ_x^2 : $\sigma_x^2 | \beta_0, \beta_1 \sim \mathcal{IG}(\beta_0, \beta_1)$.

Pour finir, l'utilisateur doit régler les paramètres (γ_0, γ_1) et (β_0, β_1) afin d'ajuster l'allure des distributions a priori de la puissance de bruit thermique σ^2 et de la puissance des cibles σ_x^2 respectivement, suivant la connaissance qu'il a de ces paramètres.

1.4.1.3 Schéma d'estimation

Le schéma hiérarchique Bayésien présenté dans la section précédente est utilisé pour estimer le paramètre d'intérêt \mathbf{x} . Deux estimateurs sont calculés dans [24] : l'estimateur MMSE et l'estimateur MAP, dont on a donné l'expression dans (1.48), (1.49). Ces deux estimateurs dépendent de la distribution a posteriori $f(\mathbf{x}|\mathbf{y})$ calculée dans [24] comme

$$f(\mathbf{x}|\mathbf{y}) \propto \frac{B(1 + \|\mathbf{x}\|_0, 1 + \bar{K}\bar{M} - \|\mathbf{x}\|_0)\Gamma(\beta_0 + \|\mathbf{x}\|_0)}{(\beta_1 + \|\mathbf{x}\|_2^2)^{\beta_0 + \|\mathbf{x}\|_0} [\|\mathbf{y} - \mathbf{H}_i\mathbf{x}\|_2^2 + \gamma_1]^{\bar{K}\bar{M} + \gamma_0}} \prod_{\bar{i}/x_{\bar{i}}=0} \delta(|x_{\bar{i}}|) \quad (1.57)$$

où $B(\cdot)$ désigne la fonction Beta. Cette distribution ne peut être intégrée ou maximisée analytiquement en forme directe, donc les estimateurs théoriques sont approchés grâce à une méthode de type MCMC. Plus précisément, un échantillonneur de Gibbs est implémenté. Il consiste à générer de manière itérative des échantillons $(\sigma^{2(n)}, \mathbf{x}^{(n)}, w^{(n)}, \sigma_x^{2(n)})$ qui suivent chacun leur distribution conditionnelle a posteriori respective. Après une période de chauffe N_{bi} , ces échantillons suivent leur distribution a posteriori, et on peut alors approcher les estimateurs MMSE et MAP grâce à N_r échantillons en les calculant comme décrit dans (1.50), (1.51). Les distributions conditionnelles a posteriori de chacun des paramètres sont obtenues via la distribution jointe a posteriori

$$f(\sigma^2, \mathbf{x}, w, \sigma_x^2|\mathbf{y}) \propto f(\mathbf{y}|\mathbf{x}, \sigma^2)\pi(\mathbf{x}|w, \sigma_x^2)\pi(w)\pi(\sigma_x^2)\pi(\sigma^2). \quad (1.58)$$

Elles sont résumées dans la Fig. 1.7 ; les calculs sont détaillés dans [24]. Ce schéma d'estimation permet donc d'obtenir deux estimés du vecteur d'intérêt \mathbf{x} , mais également l'estimateur MMSE des autres paramètres σ^2 , w et σ_x^2 qui peuvent être des indicateurs de la qualité de l'estimation.

1.4.2 Exemple de résultats

L'algorithme présenté dans [24] est ensuite testé sur des données synthétiques et semi-expérimentales. Il est comparé à un autre algorithme de représentation parcimonieuse : l'algorithme OMP [85], souvent utilisé comme référence dans le cadre des SSR, et qui peut être utilisé avec le même dictionnaire parcimonieux \mathbf{H} .

1.4.2.1 Données synthétiques

Dans un premier temps, des données synthétiques large bande sont générées ; elles contiennent 4 cibles dans du bruit blanc de puissance σ^2 . Les cibles, classées de la plus proche à la plus éloignée, ont un SNR post-traitement (1.25) de 20, 15, 30 et 15 dB respectivement. La quatrième cible est cachée dans le lobe secondaire de la première ; la deuxième cible est elle située à v_a , ce qui poserait problème en présence de clutter puisqu'elle serait masquée par les lobes secondaires de ce dernier.

Require: $\mathbf{y}, \mathbf{H}, (\gamma_0, \gamma_1), (\beta_0, \beta_1)$
Ensure: $\hat{\mathbf{x}}_{\text{MMSE}}, \hat{\sigma}_{\text{MMSE}}^2, \hat{w}_{\text{MMSE}}, \hat{\sigma}_x^2$

```

{Initialisation}
 $w^{(0)} \sim \mathcal{U}_{[0,1]}$ 
 $\sigma_x^2{}^{(0)} \sim \mathcal{IG}(\beta_0, \beta_1)$ 
 $\mathbf{x}^{(0)} \sim \prod_{i=0}^{\bar{K}\bar{M}-1} \mathcal{BerCN}(w^{(0)}, 0, \sigma_x^2{}^{(0)})$ 
{Itérations}
for  $n = 1$  à  $N_{bi} + N_r$  do
   $\sigma_x^2{}^{(n)} \sim \mathcal{IG}(KM + \gamma_0, \|\mathbf{y} - \mathbf{H}\mathbf{x}^{(n-1)}\|_2^2 + \gamma_1)$  /*step 1*/
   $w^{(n)} \sim \mathcal{Be}(1 + \|\mathbf{x}^{(n-1)}\|_0, 1 + \bar{K}\bar{M} - \|\mathbf{x}^{(n-1)}\|_0)$  /*step 2*/
   $\sigma_x^2{}^{(n)} \sim \mathcal{IG}(\beta_0 + \|\mathbf{x}^{(n-1)}\|_0, \beta_1 + \|\mathbf{x}^{(n-1)}\|_2^2)$  /*step 3*/
  for  $\bar{i} = 0$  à  $\bar{K}\bar{M} - 1$  do
     $x_{\bar{i}}^{(n)} \sim \mathcal{BerCN}(w_{\bar{i}}, \mu_{\bar{i}}, \eta_{\bar{i}}^2)$  /*step 4*/
  end for
end for
{Estimateurs}
 $\hat{\theta}_{\text{MMSE}} = \frac{1}{N_r} \sum_{n=1}^{N_r} \theta^{(n+N_{bi})}$ 

```

FIGURE 1.7 – Schéma d’estimation de l’algorithme de [24]. Les paramètres $(w_{\bar{i}}, \mu_{\bar{i}}, \eta_{\bar{i}}^2)$ sont détaillés dans [24]; ils dépendent de l’intégralité des autres paramètres mis à jour.

On peut voir sur la Fig. 1.8 que la scène est correctement estimée par les deux algorithmes : la localisation des cibles en distance et en vitesse est adéquate, et leur amplitude est correctement estimée. En particulier, il n’y a pas de fausses alarmes dans les lobes secondaires en vitesse. Notons toutefois que le critère d’arrêt utilisé avec l’algorithme OMP est le plus avantageux, c’est-à-dire que le nombre de cibles est connu. Si on utilise un critère d’arrêt sur les résidus, alors de nombreuses fausses alarmes apparaissent sur la carte.

1.4.2.2 Données expérimentales

Dans un second temps, on compare ces deux algorithmes sur des données semi-expérimentales. Des cibles synthétiques (les mêmes que dans la section précédente) sont ajoutées à du bruit thermique mesuré par le radar PARSAX de l’université TU Delft [70]. Ainsi, on peut tester les algorithmes sur un scénario « réaliste » tout en ayant des cibles coopératives. La Fig. 1.9 montre que les deux algorithmes estiment correctement la scène radar en termes de localisation distance/vitesse et d’amplitude.

1.4.3 Robustesse de l’algorithme

D’après ces résultats, l’algorithme proposé dans [24] semble convaincant. Il convient néanmoins de souligner certains points :

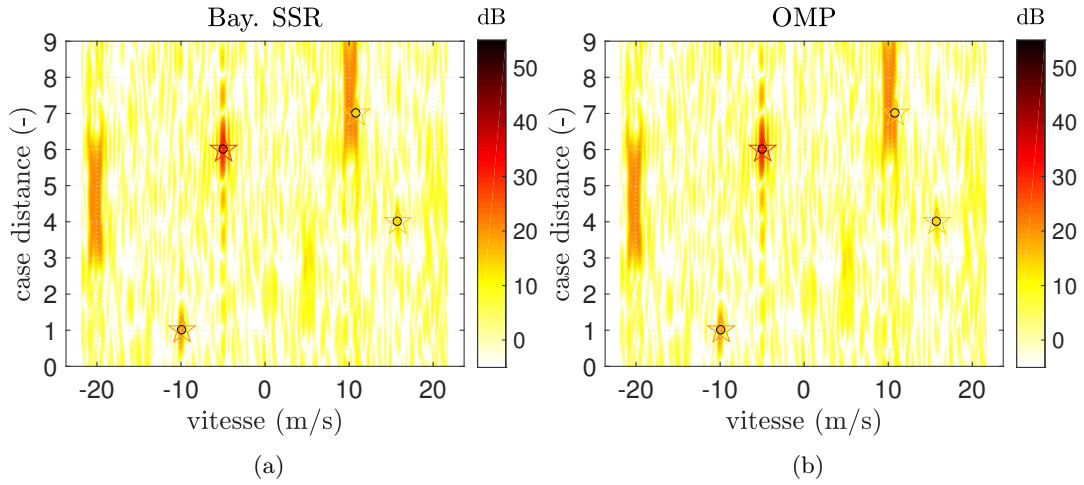


FIGURE 1.8 – Carte distance-vitesse estimée par deux algorithmes de représentation parcimonieuse. Le scénario consiste en 4 cibles de SNR=15, 20, 30, 15 dB (de la plus proche à la plus éloignée) dans du bruit blanc de puissance σ^2 . Le fond représente la sommation cohérente large bande, les étoiles la vraie scène et les ronds la scène estimée. $K = 10$, $M = 32$, $n_{va} = 3$, $v_a = 15$ m/s, $B/f_c=10\%$, $\sigma^2 = 1$. 1.8(a) algorithme de [24] 1.8(b) algorithme OMP de [85] avec le nombre de cibles connu.

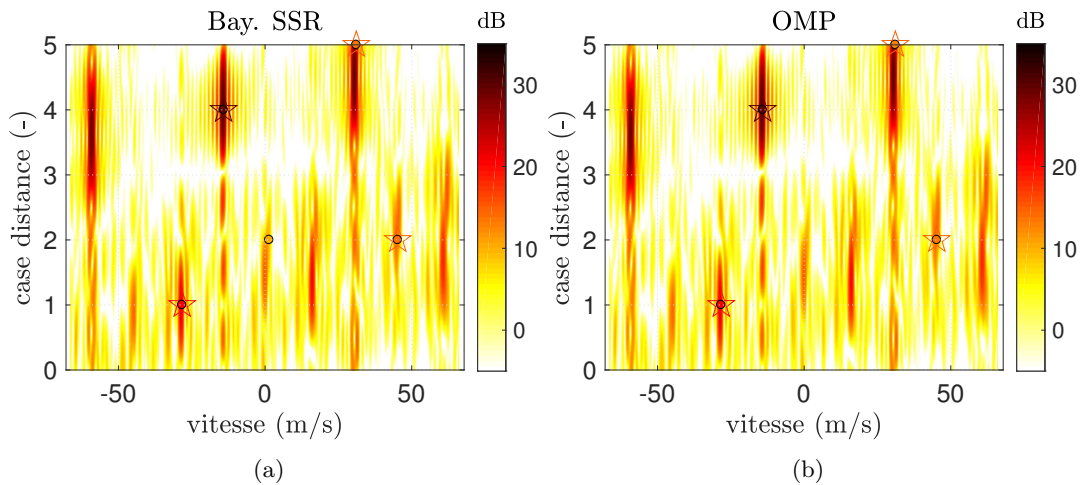


FIGURE 1.9 – Carte distance-vitesse estimée par deux algorithmes de représentation parcimonieuse. Le scénario consiste en 4 cibles de SNR=15, 20, 30, 15 dB (de la plus proche à la plus éloignée) dans du bruit thermique mesuré par le radar PARSAX. Les données sont normalisées de sorte que $\sigma^2 = 1$. Le fond représente la sommation cohérente large bande, les étoiles la vraie scène et les ronds la scène estimée. $K = 6$, $M = 32$, $n_{va} = 3$, $v_a = 45$ m/s, $B = 100$ MHz, $f_c = 3.315$ GHz. 1.8(a) algorithme de [24] 1.8(b) algorithme OMP de [85] avec le nombre de cibles connu.

1. Les cibles synthétiques ajoutées au bruit blanc sont délibérément situées sur la grille

d'analyse utilisée pour construire le dictionnaire \mathbf{H} . Autrement dit,

$$\exists \bar{i} \in \{0, \dots, \bar{K}\bar{M} - 1\} | f_r = f_{r\bar{i}}, f_D = f_{D\bar{i}}$$

où (f_r, f_D) désignent les fréquences rapide et Doppler normalisées de la cible en question.

2. Le bruit \mathbf{n} est considéré blanc Gaussien.
3. La dynamique de puissance des cibles est réduite (SNR minimum de 15 dB, maximum de 30 dB).

Si on teste l'algorithme sur des scénarios ne respectant pas chacune de ces conditions, alors on obtient les résultats présentés dans la Fig. 1.10. On voit sur la Fig. 1.10(a) que si les cibles sont en chevauchement entre deux cases d'analyse, alors l'algorithme de [24] ne parvient pas à estimer correctement la scène. Plus précisément : l'énergie de la troisième cible (en chevauchement à la fois sur la grille d'analyse en distance et sur celle en vitesse) est répartie sur les cases distance adjacentes. Cela conduit à des fausses-alarmes dans les lobes secondaires en distance et en vitesse. La quatrième cible (cachée dans le lobe secondaire) est masquée par la troisième cible et n'est pas vue par l'algorithme. D'autre part, on voit sur la Fig. 1.10(b) que si le bruit est coloré, de nombreuses fausses alarmes apparaissent dans les lobes secondaires du clutter, ici à $\pm v_a$. Par conséquent, il est difficile de dire si la deuxième cible, située à v_a , est correctement estimée par l'algorithme. Pour finir, la Fig. 1.10(c) montre que l'algorithme de [24] est mis en difficulté par un scénario à forte dynamique de puissance. En effet, la quatrième cible de SNR=12 dB masquée par une cible forte à 45 dB n'est pas estimée.

Ces observations ont motivé le travail de robustification de l'algorithme entrepris dans ce manuscrit. Dans un premier temps, l'algorithme a été robustifié au désalignement par rapport à la grille d'analyse afin d'estimer correctement des cibles potentiellement en chevauchement. Le cas du bruit coloré ayant été abordé en amont dans [21], l'algorithme correspondant a été utilisé pour produire un algorithme robuste à la fois aux cibles en chevauchement et au bruit coloré. Enfin, le cas d'une scène radar à forte dynamique de puissance a été traité à part, puis l'algorithme obtenu a été robustifié au problème des cibles en chevauchement.

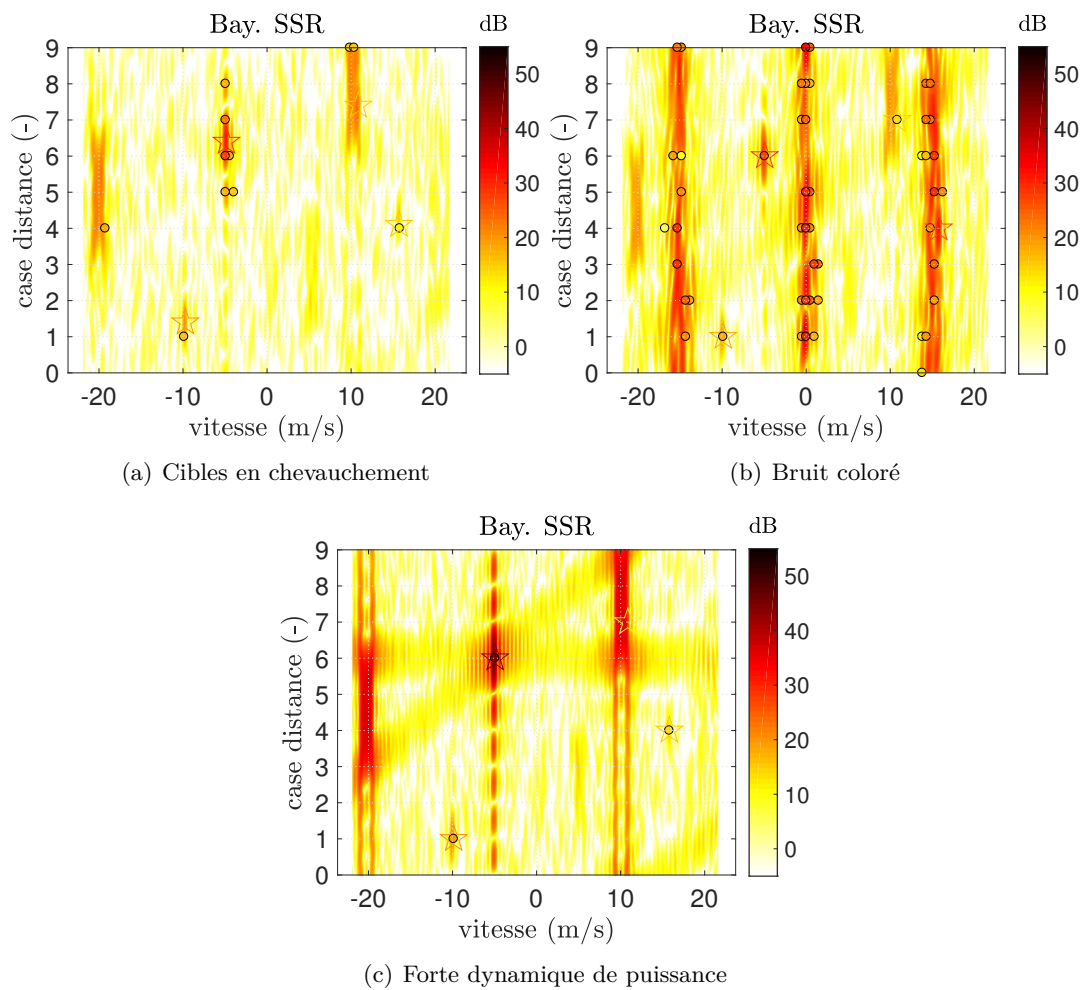


FIGURE 1.10 – Carte distance-vitesse estimée par l’algorithme de [24] sur différents scénarios. Les 4 cibles ont un SNR=15, 20, 45, 12 dB (de la plus proche à la plus éloignée). Le fond représente la sommation cohérente large bande, les étoiles la vraie scène et les cercles la scène estimée. 1.10(a) Cibles en chevauchement, 1.10(b) Bruit coloré, 1.10(c) Forte dynamique de puissance.

Algorithme Bayésien de SSR pour les cibles en chevauchement

Dans ce chapitre, on présente les travaux de robustification de l'algorithme « de référence » de [24] au cas de cibles en chevauchement. Comme mentionné dans le chapitre précédent, l'algorithme de référence est originellement limité au cas de cibles situées sur la grille d'analyse utilisée pour construire le dictionnaire \mathbf{H} . Néanmoins, il existe toujours un certain désalignement entre les cibles et cette grille d'analyse dans les scénarios réels. Lorsque ce désalignement n'est pas pris en compte dans l'algorithme de référence, la levée d'ambiguïté vitesse n'a pas lieu, comme en témoigne la Fig. 2.1. Le scénario consiste en une cible à vitesse nulle chevauchant deux cases d'analyse adjacentes en distance (Fig. 2.1(a)) ou en vitesse (Fig. 2.1(b)). On voit qu'un désalignement en distance a pour conséquence de multiples fausses alarmes dans les lobes secondaires en distance, soit $v = 0$ m/s et $k = 0, \dots, K - 1$. Un désalignement en vitesse provoque par contre majoritairement des fausses-alarms dans les lobes secondaires en vitesse, soit $v = \pm v_a$. Il est donc nécessaire de prendre en compte ce désalignement afin d'améliorer les performances de l'algorithme de référence sur des scénarios plus réalistes.

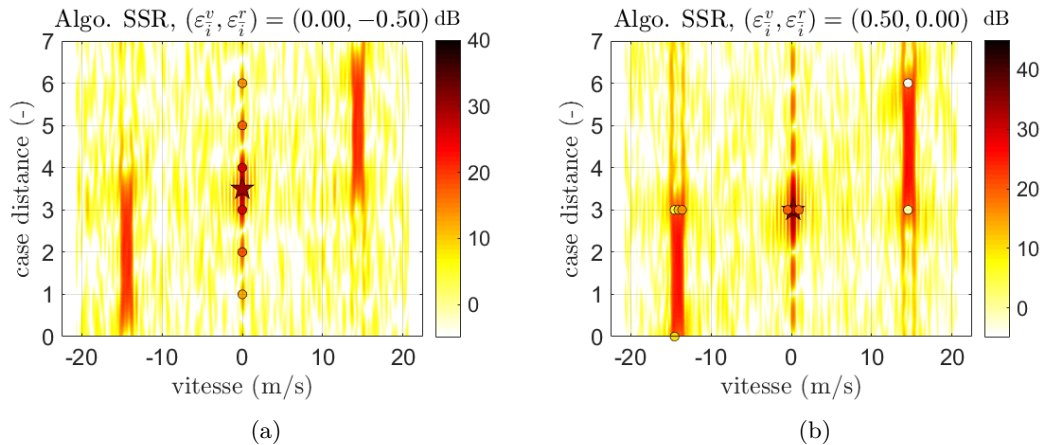


FIGURE 2.1 – Conséquence d'un désalignement par rapport à la grille d'analyse en distance 2.1(a) ou en vitesse 2.1(a) sur l'estimation par l'algorithme non-robustifié de [24]. Le fond représente la sommation cohérente large bande, les étoiles la vraie scène et les ronds la scène estimée par l'algorithme de référence [24]. $K = 8$, $M = 32$, $n_{va} = 3$, $v_a = 15$ m/s, $B/f_c = 10\%$, $\sigma^2 = 1$.

Ce travail de robustification a été mené en plusieurs étapes. Dans un premier temps, la

migration distance n'a pas été prise en compte. On s'est alors concentré sur le cas unidimensionnel d'une seule case distance, $K = 1$, (il s'agit alors de traiter le désalignement par rapport à la grille d'analyse en vitesse [3], [1]) puis le cas bidimensionnel, $K > 1$ [4]. La migration des cibles et leur désalignement en distance et en vitesse a enfin été pris en compte dans une dernière étape [6]. Ces différentes étapes ne sont pas entièrement détaillées ici afin de ne pas alourdir le propos. On présente donc directement la prise en compte du désalignement en distance et vitesse pour des cibles migrantes. L'annexe A apporte tout de même des détails soutenant l'approche adoptée, en se limitant au cas unidimensionnel.

Sommaire

2.1	Le problème des cibles en chevauchement	40
2.2	Algorithme Bayésien de SSR pour les cibles en chevauchement	43
2.2.1	Paramétrisation du dictionnaire	44
2.2.2	Modèle hiérarchique Bayésien	45
2.2.3	Schéma d'estimation	45
2.2.4	Estimation des paramètres de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$	49
2.2.5	Remarque : prise en compte du désalignement via un dictionnaire linéarisé	53
2.3	Simulations numériques	55
2.3.1	Exemple de réalisation	55
2.3.2	Simulations de Monte-Carlo	57
2.4	Simulations sur données expérimentales	61
2.5	Conclusions	67

2.1 Le problème des cibles en chevauchement

Comme expliqué dans le chapitre précédent, les données \mathbf{y} sont exprimées via les représentations parcimonieuses grâce au modèle

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \quad (2.1)$$

Il s'agit alors d'estimer le vecteur parcimonieux \mathbf{x} qui contient l'information d'amplitude des cibles présentes dans la scène radar. Le dictionnaire \mathbf{H} repose sur une grille d'analyse contenant $\bar{K} \times \bar{M}$ éléments. On rappelle que la \bar{i} -ième colonne du dictionnaire \mathbf{H} s'écrit

$$[\mathbf{h}_{\bar{i}}]_i = \frac{1}{\sqrt{\bar{K}\bar{M}}} \exp\{-j2\pi f_{r\bar{i}}k\} \exp\{j2\pi f_{d\bar{i}}(1 + \mu k)m\}$$

où $f_{r\bar{i}} = \bar{k}/\bar{K}$ et $f_{d\bar{i}} = n_{va}\bar{m}'/\bar{M}$ sont les fréquences rapide et Doppler normalisées pour la \bar{i} -ième case de la grille d'analyse (\bar{k} -ième case distance, \bar{m} -ième case vitesse).

Présentation et quantification du problème Les cibles présentes sur la scène radar ne sont pas nécessairement situées sur la grille d'analyse distance-vitesse, autrement dit, la

fréquence rapide et/ou la fréquence Doppler des cibles ne correspondent pas aux fréquences (f_{r_i}, f_{d_i}) du dictionnaire. Ce problème de désalignement par rapport à la grille d’analyse a d’abord été observé par Malioutov [74], puis quantifié à plusieurs reprises [59, 92, 122, 50]. D’après ces travaux, il est primordial de prendre en compte ce désalignement sous peine de fortement dégrader les performances des algorithmes utilisés.

Raffinement de la grille L’approche la plus intuitive pour résoudre le problème dit de *grid mismatch* consisterait à essayer de diminuer le désalignement en raffinant la grille [74, 50, 51]. Le modèle adopté pour le signal est le suivant :

$$\mathbf{y} = \mathbf{H}(\tilde{\mathbf{f}}_r, \tilde{\mathbf{f}}_d)\mathbf{x} + \mathbf{n} \quad (2.2)$$

où $(\tilde{\mathbf{f}}_r, \tilde{\mathbf{f}}_d)$ correspondent à la grille d’analyse sur-échantillonnée (le désalignement n’est donc pas explicité dans le modèle). L’approche proposée dans [74], qui est une des premières études sur le sujet, consiste alors à faire une première estimation avec une grille d’analyse comportant $\bar{K} \times \bar{M}$ points puis à raffiner localement et de manière adaptative cette grille jusqu’à ce qu’elle soit suffisamment fine. L’approche est différente dans [50] : la grille est sur-échantillonnée, et les “vraies” fréquences sont estimées directement à partir des points de la grille $(\tilde{\mathbf{f}}_r, \tilde{\mathbf{f}}_d)$ via une équation quadratique. Néanmoins, lorsque la grille est très fine, le dictionnaire \mathbf{H} gagne en cohérence, ce qui peut mettre à mal les performances d’algorithmes classiques de SSR. En effet, si la cohérence entre deux colonnes de \mathbf{H} est trop importante, alors la distinction entre deux cibles localisées aux fréquences correspondant à ces colonnes est compromise [51]. L’algorithme mis en œuvre dans [51] consiste alors à exclure de l’analyse les fréquences appartenant à la “bande de cohérence” de l’indice considéré (*Band Exclusion*). A noter que dans [51], l’erreur due au désalignement apparaît explicitement dans le modèle : elle est contenue dans le bruit \mathbf{n} .

Estimation conjointe du vecteur d’amplitude et du désalignement Le modèle (2.2) a donc été utilisé pour éviter le désalignement en raffinant la grille, mais il a également été repris dans le but d’estimer le désalignement. En particulier, le cas du dictionnaire de Fourier 1D a été étudié dans [61] :

$$\mathbf{H} \triangleq \mathbf{F} = \begin{bmatrix} \mathbf{f}_0 & \dots & \mathbf{f}_{\bar{M}-1} \end{bmatrix}$$

$$[\mathbf{f}_{\bar{m}}]_m = \frac{1}{\sqrt{\bar{M}}} e^{j2\pi m f_{d\bar{m}}}.$$

\mathbf{F} est donc paramétré par la grille $\mathbf{f}_d = [f_{d0} \dots f_{d\bar{M}-1}]$; cette grille n’est pas fixée, mais est considérée comme déterministe et estimée conjointement à \mathbf{x} . Le problème est alors résolu en utilisant une approche Bayésienne et un algorithme *Expectation-Maximization* (EM). Un autre modèle propose d’expliciter les perturbations dues au *grid mismatch* via une matrice \mathbf{E} que l’on ajoute au dictionnaire non-perturbé \mathbf{H} ; le modèle est alors le suivant :

$$\mathbf{y} = (\mathbf{H} + \mathbf{E})\mathbf{x} + \mathbf{n} \quad (2.3)$$

ou de manière équivalente en utilisant le dictionnaire de Fourier

$$\mathbf{y} = (\mathbf{F} + \mathbf{E})\mathbf{x} + \mathbf{n}. \quad (2.4)$$

C'est le modèle adopté pour quantifier l'influence d'une perturbation dans [92] et [59] (cas de l'algorithme de *Basis Pursuit*), et dans [122] lorsque la perturbation est due au *grid mismatch* seulement. Ce modèle de perturbation sur le dictionnaire est également utilisé dans [125, 103]. L'algorithme *WSS - TLS* (*Weighted Structured Sparse Total Least Squares*) de [125] repose sur la formulation et la résolution d'un problème de moindres carrés totaux sous contrainte l_1 . Dans [103], l'algorithme OMP est modifié pour prendre en compte les perturbations dues au *grid mismatch* : il s'agit alors de forcer la perturbation sur les colonnes de \mathbf{H} dans une certaine limite puis de résoudre le problème résultant en gardant une approche de type OMP.

Un cas particulier du modèle (2.3) est très souvent utilisé : c'est le cas où \mathbf{E} peut se mettre sous la forme $\mathbf{G}\mathbf{D}_\varepsilon$ où \mathbf{G} est une matrice quelconque et \mathbf{D}_ε une matrice diagonale [58, 64, 63, 119, 62, 124, 100, 101]. En fait, ce modèle est utilisé dans le cadre du dictionnaire de Fourier, et résulte d'un développement limité d'ordre 1 autour de la fréquence $\tilde{f}_{d\bar{m}}$:

$$\mathbf{f}(f_{d_n}) = \mathbf{f}(\tilde{f}_{d\bar{m}}) + \mathbf{g}(\tilde{f}_{d\bar{m}}) (f_{d_n} - \tilde{f}_{d\bar{m}}) \quad (2.5)$$

avec $\mathbf{g}(\tilde{f}_{d\bar{m}}) = \frac{d}{d\tilde{f}_{d\bar{m}}} \mathbf{f}(\tilde{f}_{d\bar{m}})$. Si l'on considère l'intégralité des fréquences, on obtient le modèle suivant :

$$\mathbf{y} = (\mathbf{F} + \mathbf{G}\mathbf{D}_\varepsilon)\mathbf{x} + \mathbf{n}$$

avec $\mathbf{F} = [\mathbf{f}(\tilde{f}_{d_0}) \ \dots \ \mathbf{f}(\tilde{f}_{d_{M-1}})]$, $\mathbf{G} = [\mathbf{g}(\tilde{f}_{d_0}) \ \dots \ \mathbf{g}(\tilde{f}_{d_{M-1}})]$ et $\mathbf{D}_\varepsilon = \text{diag}(\boldsymbol{\varepsilon})$ où le vecteur $\boldsymbol{\varepsilon}$ représente le désalignement. Les matrices \mathbf{F} et \mathbf{G} sont alors entièrement connues et il s'agit d'estimer \mathbf{x} et $\boldsymbol{\varepsilon}$. Les méthodes d'estimation utilisées avec ce modèle sont alors très variées. Dans [64, 58], le problème est repris sous la forme du WSS-TLS de [125] appliqué au problème particulier de la localisation de source et résolu via un algorithme de descente alternée (*alternating-descent algorithm*) [64] ou OMP [58]. C'est toujours le cas de la localisation de sources qui est traité par l'algorithme OGSBI (*Off-Grid Sparse Bayesian Inference*) de [119], mais sous une perspective Bayésienne cette fois, et résolu via un algorithme EM. L'approche utilisée dans [62] se distingue des autres par le fait de considérer \mathbf{F} et \mathbf{G} comme deux dictionnaires, et il faut alors estimer deux vecteurs parcimonieux \mathbf{x} et $\mathbf{z} \triangleq \mathbf{D}_\varepsilon \mathbf{x}$, ce qui est fait via un algorithme variationnel Bayésien. A noter que dans [100, 101], \mathbf{x} et $\boldsymbol{\varepsilon}$ sont estimés conjointement via des algorithmes classiques (OMP, LASSO et FISTA) augmentés ; il est mentionné qu'un problème de dédoublement de cible peut apparaître s'il y a ambiguïté sur la position de la cible par rapport à la grille (case \bar{m} et désalignement de +0.5 ou bien case $\bar{m} + 1$ et désalignement de -0.5 ?) ; la solution proposée consiste alors à faire fusionner *a posteriori* les deux cibles obtenues (formule spécifiée dans [100, 101]). Enfin, une approche parcimonieuse « par bloc » (*block sparse estimator*) est associée au modèle de signal (2.1) dans [63] et dans [124] après reformulation du problème. Dans [124], le développement limité (2.5) est utilisé pour reformuler l'expression de la matrice de covariance des données $\mathbf{R} = \mathcal{E} \{ \mathbf{y}\mathbf{y}^H \}$, et le problème est résolu via un algorithme EM.

Continuité des fréquences Pour finir, certains travaux récents s'affranchissent du problème de désalignement en se passant de grille d'analyse : les fréquences sont alors estimées sur le continuum $[0, 1)$. De nombreux travaux [102, 19, 32] utilisent pour cela la norme atomique définie comme [37]

$$\|\mathbf{x}^*\|_{\mathcal{A}} = \inf \left\{ \sum_{\mathbf{a} \in \mathcal{A}} c_{\mathbf{a}} : \mathbf{x}^* = \sum_{\mathbf{a} \in \mathcal{A}} c_{\mathbf{a}} \mathbf{a}, c_{\mathbf{a}} \geq 0, \forall \mathbf{a} \in \mathcal{A} \right\}. \quad (2.6)$$

La forme duale de la norme atomique permet d'écrire le problème de minimisation

$$\min_{\mathbf{x}^*} \|\mathbf{x}^*\|_{\mathcal{A}} \quad \text{tel que} \quad \mathbf{y} = \mathbf{x}^* \quad (2.7)$$

sous forme d'un problème de type *Semidefinite program* [102] que l'on peut facilement résoudre en utilisant des routines existantes, par exemple SeDuMi ou SDPT3. Dans le cas d'un bruit additif blanc Gaussien, le problème de minimisation ℓ_2 sous contrainte ℓ_1 (1.34) est remplacé par le suivant [19]

$$\min_{\mathbf{x}^*} \frac{1}{2} \|\mathbf{y} - \mathbf{x}^*\|_2^2 + \tau \|\mathbf{x}^*\|_{\mathcal{A}}. \quad (2.8)$$

Il s'agit alors de choisir le paramètre de régularisation adéquat τ pour résoudre ce problème via *Semidefinite programming* [19]. Le problème de minimisation (2.8) a depuis été étendu, entre autres, au cas de données incomplètes [117] et à la localisation de sources pour les réseaux premiers entre eux (*co-prime arrays*) [99]. Notons que [40] puis [73, 118] ont utilisé la norme atomique sur des données bi-dimensionnelles, notamment dans le cas de vecteurs d'observation multiples (*Multiple Measurement Vectors*, MMV) [73, 118].

Une autre approche consiste à expliciter la parcimonie du vecteur d'intérêt \mathbf{x} à travers une matrice structurée de rang faible construite à partir des données \mathbf{y} [39, 73]. Cette approche à l'avantage de ne pas utiliser de discrétisation de l'espace distance-vitesse, et évite ainsi le problème de désalignement par rapport à la grille d'analyse. L'algorithme EMaC (*Enhancement Matrix Completion*) présenté dans [39] propose de ré-organiser la matrice des données en une matrice de Hankel par bloc de rang faible, puis de minimiser sa norme nucléaire (la somme de ses valeurs singulières). Comme dans le cas de la norme atomique, ce problème convexe peut être mis sous forme de SDP pour être résolu par des routines connues de type SeDuMi. Les travaux présentés dans [73] combinent les approches de minimisation de la norme atomique et de la norme nucléaire.

2.2 Algorithme Bayésien de SSR pour les cibles en chevauchement

Comme expliqué dans le chapitre 1, l'objectif est de robustifier l'algorithme présenté dans [24] au cas des cibles en chevauchement sur la grille d'analyse. Le modèle hiérarchique Bayésien de [24] est conçu de telle manière que les paramètres ont un sens physique bien identifié, ce qui facilite le réglage des hyperparamètres. En outre, il offre une grande flexibilité de modélisation

que l'on utilise ici pour traiter le cas des cibles en chevauchement.

2.2.1 Paramétrisation du dictionnaire

Un modèle dit « paramétrique » est adopté dans ces travaux¹; il consiste à paramétrer directement le dictionnaire \mathbf{H} pour faire apparaître une variable de désalignement. Dans le cas d'une analyse bi-dimensionnelle distance-vitesse, les vecteurs $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$ sont introduits pour modéliser le désalignement par rapport à la grille d'analyse en vitesse et en distance respectivement; le modèle de signal s'écrit alors

$$\mathbf{y} = \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)\mathbf{x} + \mathbf{n}. \quad (2.9)$$

La paramétrisation du dictionnaire \mathbf{H} consiste à re-écrire les fréquences normalisées $(f_{d\bar{i}}, f_{r\bar{i}})$ afin de faire apparaître les variables de désalignement [6]

$$f_{d\bar{i}} = \frac{\bar{m}' + \varepsilon_i^v}{\bar{M}} n_{va} \quad (2.10)$$

$$f_{r\bar{i}} = \frac{\bar{k} + \varepsilon_i^r}{\bar{K}}, \quad (2.11)$$

avec $(\varepsilon_i^v, \varepsilon_i^r) \in [-0.5, 0.5) \times [-0.5, 0.5)$. La Fig. 2.2 illustre la modélisation du désalignement pour une cible (croix rouge) située à la \bar{i} -ième case d'analyse; ici on a $\varepsilon_i^v = 0.4$ et $\varepsilon_i^r = -0.3$.

On rappelle que l'expression de $f_{d\bar{i}}$ de (2.10), même dans le cas d'une cible alignée sur la grille d'analyse, n'est pas habituelle. Tout d'abord, la prise en compte du facteur de repliement permet d'estimer sans ambiguïté des cibles situées au-delà de la première ambiguïté vitesse. Ensuite, l'indice signé \bar{m}' permet de prendre en compte une vitesse positive ou négative sur l'intervalle étendu $[-n_{va}v_a/2, n_{va}v_a/2)$.

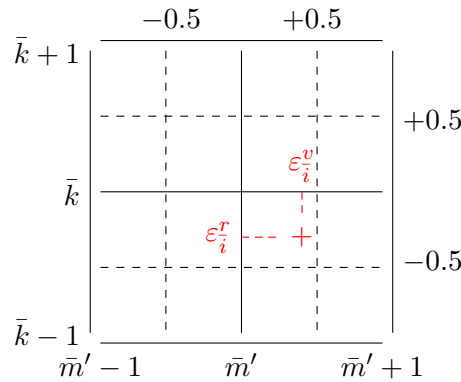


FIGURE 2.2 – Modélisation du désalignement d'une cible située à la \bar{i} -ième case d'analyse (croix rouge). ε_i^v et ε_i^r représentent le désalignement par rapport à la grille d'analyse en vitesse et en distance respectivement.

1. Ce choix est justifié dans [1] et discuté en 2.2.5.

2.2.2 Modèle hiérarchique Bayésien

Par rapport à l'algorithme de [24], il s'agit maintenant d'estimer à la fois le vecteur \mathbf{x} qui contient l'information d'amplitude des cibles, mais également les vecteurs de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$ qui permettent de corriger leur position en distance et en vitesse. Le schéma hiérarchique Bayésien est alors modifié en conséquence.

Vraisemblance Le bruit \mathbf{n} est toujours considéré blanc Gaussien de moyenne nulle et de variance σ^2 . Les données \mathbf{y} sont donc distribuées suivant une Gaussienne de variance σ^2 mais de moyenne non-nulle $\mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)\mathbf{x}$:

$$f(\mathbf{y}|\mathbf{x}, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \sigma^2) \propto \frac{1}{\sigma^{2KM}} \exp \left\{ -\sigma^{-2} \|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)\mathbf{x}\|_2^2 \right\}. \quad (2.12)$$

Les paramètres $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$, \mathbf{x} et σ^2 étant inconnus, une distribution a priori doit leur être attribuée. Les distributions a priori des paramètres \mathbf{x} et σ^2 , et celles des hyperparamètres w et σ_x^2 sont inchangées par rapport au modèle non-robustifié de [24]. La seule modification du schéma hiérarchique Bayésien de [24] consiste à y intégrer les paramètres de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$.

Paramètres de désalignement Le choix est fait de conditionner la distribution a priori des vecteurs de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$ au vecteur d'amplitude \mathbf{x} . L'idée est la suivante : si aucune cible n'est présente à la case d'analyse, alors la valeur correspondante de désalignement est nulle. On considère donc les éléments des vecteurs de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$ i.i.d. et on impose une distribution a priori jointe à chacun des éléments $(\varepsilon_i^v, \varepsilon_i^r)$, conditionnée par la valeur de x_i . Dans le cas où une cible est bien présente à la case d'analyse considérée ($x_i \neq 0$), une distribution a priori non-informative uniforme est choisie. On a donc :

$$\pi(\varepsilon_i^v, \varepsilon_i^r | x_i = 0) = \delta(\varepsilon_i^v) \delta(\varepsilon_i^r) \quad (2.13)$$

$$\pi(\varepsilon_i^v, \varepsilon_i^r | x_i \neq 0) = \mathbb{I}_{[-0.5, 0.5]}(\varepsilon_i^v) \mathbb{I}_{[-0.5, 0.5]}(\varepsilon_i^r). \quad (2.14)$$

Le modèle hiérarchique résultant est représenté graphiquement dans la Fig. 2.3. Etant donné que la distribution a priori des vecteurs $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$ est non-informative, l'utilisateur n'a pas besoin de régler de paramètres supplémentaires par rapport au modèle non-robustifié de [24].

2.2.3 Schéma d'estimation

Le schéma hiérarchique Bayésien présenté dans la section 2.2.2 est utilisé pour estimer les paramètres d'intérêt \mathbf{x} , $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$. Plus précisément, on calcule l'estimateur MMSE de ces

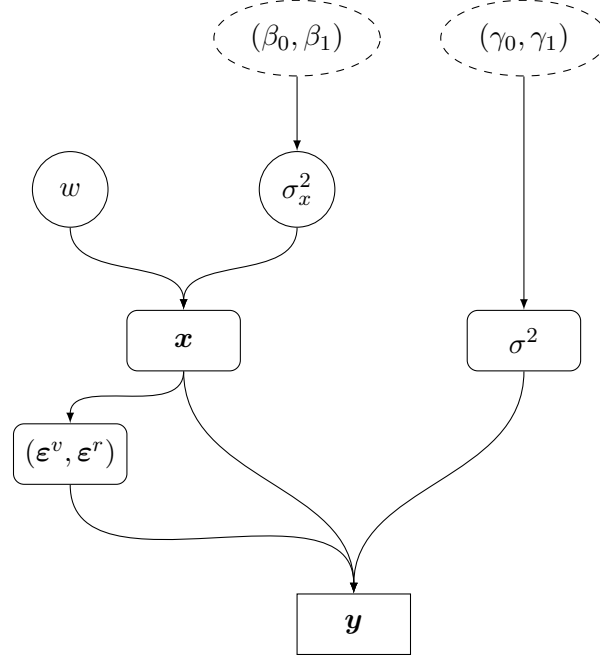


FIGURE 2.3 – Modèle hiérarchique Bayésien prenant en compte le désalignement par rapport à la grille d’analyse. Les paramètres dans les cercles en pointillés doivent être réglés par l’utilisateur ; il n’y a pas de réglage supplémentaire par rapport au modèle non-robustifié de [24].

paramètres

$$\hat{\mathbf{x}}_{\text{MMSE}} = \int \mathbf{x} f(\mathbf{x}|\mathbf{y}) d\mathbf{x}, \quad (2.15)$$

$$(\hat{\epsilon}^v, \hat{\epsilon}^r)_{\text{MMSE}} = \int (\epsilon^v, \epsilon^r) f(\epsilon^v, \epsilon^r|\mathbf{y}) d\epsilon^v d\epsilon^r. \quad (2.16)$$

Comme dans le cas non-robustifié, on ne peut pas calculer de manière analytique sous forme directe ces estimateurs. Un échantillonneur de Gibbs est alors implémenté. Son principe est le suivant : des échantillons $(\sigma^{2(n)}, \epsilon^{v(n)}, \epsilon^{r(n)}, \mathbf{x}^{(n)}, w^{(n)}, \sigma_x^{2(n)})$ sont générés de manière itérative suivant leur distribution conditionnelle a posteriori respective ; après une période de chauffe N_{bi} , ces échantillons suivent leur distribution a posteriori respective [90]. Les estimateurs (2.15), (2.16) sont alors calculés comme

$$\begin{aligned} \hat{\mathbf{x}}_{\text{MMSE}} &= \frac{1}{N_r} \sum_{n=1}^{N_r} \mathbf{x}^{(n+N_{bi})} \\ \hat{\epsilon}^v_{\text{MMSE}} &= \frac{1}{N_r} \sum_{n=1}^{N_r} \epsilon^{v(n+N_{bi})} \\ \hat{\epsilon}^r_{\text{MMSE}} &= \frac{1}{N_r} \sum_{n=1}^{N_r} \epsilon^{r(n+N_{bi})}. \end{aligned}$$

Afin d'implémenter l'échantillonneur de Gibbs, il faut calculer de manière analytique les distributions conditionnelles a posteriori des paramètres inconnus. Elles sont obtenues grâce à la distribution jointe a posteriori

$$f(\sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, w, \sigma_x^2 | \mathbf{y}) \propto f(\mathbf{y} | \mathbf{x}, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \sigma^2) \pi(\mathbf{x} | w, \sigma_x^2) \pi(w) \pi(\sigma_x^2) \pi(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \pi(\sigma^2). \quad (2.17)$$

On détaille par la suite le calcul des distributions conditionnelles a posteriori des paramètres et des hyperparamètres.

2.2.3.1 Distributions conditionnelles a posteriori des paramètres \mathbf{x} et σ^2

Puissance du bruit La puissance du bruit σ^2 est tirée suivant sa distribution conditionnelle a posteriori

$$\begin{aligned} f(\sigma^2 | \mathbf{y}, \mathbf{x}, w, \sigma_x^2, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) &\propto f(\mathbf{y} | \mathbf{x}, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \sigma^2) \pi(\sigma^2) \\ &\propto \sigma^{-2KM} \exp \left\{ -\sigma^{-2} \|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x}\|^2 \right\} \\ &\quad \times \exp \left\{ -\gamma_1 \sigma^{-2} \right\} \sigma^{-2(\gamma_0+1)} \mathbb{I}_{[0,+\infty)}(\sigma^2) \\ &\propto \sigma^{-2(KM+\gamma_0+1)} \exp \left\{ -\sigma^{-2} \left(\|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x}\|^2 + \gamma_1 \right) \right\} \mathbb{I}_{[0,+\infty)}(\sigma^2). \end{aligned}$$

On reconnaît une distribution inverse-Gamma $\mathcal{IG}(\gamma_0 + KM, \gamma_1 + \|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x}\|^2)$. Par rapport au cas non-robustifié de [24], le tirage de la puissance de bruit σ^2 dépend des paramètres de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$ via le dictionnaire \mathbf{H} .

Vecteur d'amplitude Le vecteur d'amplitude \mathbf{x} est tiré élément par élément ; il s'agit donc d'exprimer la distribution conditionnelle a posteriori des éléments $x_{\bar{i}}$. On calcule

$$\begin{aligned} f(x_{\bar{i}} | \mathbf{y}, \sigma^2, w, \sigma_x^2, \varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r, \mathbf{x}_{-\bar{i}}) &\propto f(\mathbf{y} | \mathbf{x}, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \sigma^2) \pi(x_{\bar{i}} | w, \sigma_x^2) \\ &\propto \exp \left\{ -\sigma^{-2} \|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x}\|^2 \right\} \\ &\quad \times \left[(1-w) \delta(|x_{\bar{i}}|) + \frac{w}{\pi \sigma_x^2} \exp \left\{ -\frac{|x_{\bar{i}}|^2}{\sigma_x^2} \right\} \right]. \end{aligned}$$

En suivant l'approche proposée dans [47, 24], on introduit le vecteur $\mathbf{e}_{\bar{i}} = \mathbf{y} - \sum_{\bar{j} \neq \bar{i}} \mathbf{h}_{\bar{j}} x_{\bar{j}}$ pour calculer

$$\begin{aligned} \|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x}\|^2 &= \|\mathbf{e}_{\bar{i}} - \mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r) x_{\bar{i}}\|^2 \\ &= \mathbf{e}_{\bar{i}}^H \mathbf{e}_{\bar{i}} - 2 \operatorname{Re} \left\{ x_{\bar{i}}^* \mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}} \right\} + \|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2 |x_{\bar{i}}|^2. \end{aligned}$$

La distribution conditionnelle a posteriori s'écrit alors

$$\begin{aligned}
f(x_{\bar{i}}|\mathbf{y}, \sigma^2, w, \sigma_x^2, \varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r, \mathbf{x}_{-\bar{i}}) &\propto \exp \left\{ -\sigma^{-2} \left[|x_{\bar{i}}|^2 \|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2 - 2 \operatorname{Re} \left\{ x_{\bar{i}}^* \mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}} \right\} \right] \right\} \\
&\times \left[(1-w) \delta(|x_{\bar{i}}|) + \frac{w}{\pi \sigma_x^2} \exp \left\{ -\frac{|x_{\bar{i}}|^2}{\sigma_x^2} \right\} \right] \\
&\propto (1-w_{\bar{i}}) \delta(|x_{\bar{i}}|) + \frac{w_{\bar{i}}}{\pi \eta_{\bar{i}}^2} \exp \left\{ -\frac{|x_{\bar{i}} - \mu_{\bar{i}}|^2}{\eta_{\bar{i}}^2} \right\}
\end{aligned}$$

avec

$$\begin{aligned}
\eta_{\bar{i}}^2 &= \left(\sigma^{-2} + \sigma_x^{-2} \|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2 \right)^{-1} \\
\mu_{\bar{i}} &= \sigma^{-2} \eta_{\bar{i}}^2 \mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}} \\
w_{\bar{i}} &= \frac{w}{(1-w) \sigma_x^2 \eta_{\bar{i}}^{-2} \exp \left\{ -\eta_{\bar{i}}^{-2} |\mu_{\bar{i}}|^2 \right\} + w}.
\end{aligned}$$

On reconnaît là une distribution Bernoulli-Gaussienne de paramètres $(w_{\bar{i}}, \mu_{\bar{i}}, \eta_{\bar{i}}^2)$. Notons que $\|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2 = 1$ donc $\eta_{\bar{i}}^2 = (\sigma^{-2} + \sigma_x^{-2})^{-1}$.

2.2.3.2 Distributions conditionnelles a posteriori des hyperparamètres w et σ_x^2

Taux d'occupation Étant donné la fonction de vraisemblance, la distribution conditionnelle a posteriori du taux d'occupation w s'écrit :

$$\begin{aligned}
f(w|\mathbf{y}, \sigma^2, \mathbf{x}, \sigma_x^2, \varepsilon^v, \varepsilon^r) &\propto f(\mathbf{y}|\mathbf{x}, \varepsilon^v, \varepsilon^r, \sigma^2) \pi(\mathbf{x}|w, \sigma_x^2) \pi(w) \\
&\propto \prod_{\bar{i}=0}^{\bar{K}\bar{M}-1} \left[(1-w) \delta(|x_{\bar{i}}|) + \frac{w}{\pi \sigma_x^2} \exp \left\{ -\frac{|x_{\bar{i}}|^2}{\sigma_x^2} \right\} \right] \times \mathbb{I}_{[0,1]}(w) \\
&\propto (1-w)^{n_0} w^{n_1} \mathbb{I}_{[0,1]}(w)
\end{aligned}$$

où $n_1 = \|\mathbf{x}\|_0$ et $n_0 = \bar{K}\bar{M} - \|\mathbf{x}\|_0$. On reconnaît une distribution Beta $\mathcal{B}e(1+n_1, 1+n_0)$, donc le tirage du taux d'occupation w est inchangé par rapport au cas non-robustifié de [24].

Puissance des cibles Comme dans le cas de la puissance de bruit σ^2 , on a choisi une distribution a priori inverse-Gamma pour la puissance des cibles σ_x^2 , qui est conjuguée de la

Gaussienne. On obtient alors des calculs similaires, et on a

$$\begin{aligned}
f(\sigma_x^2 | \mathbf{y}, \sigma^2, \mathbf{x}, w, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) &\propto f(\mathbf{y} | \mathbf{x}, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \sigma^2) \pi(\mathbf{x} | w, \sigma_x^2) \pi(\sigma_x^2) \\
&\propto \prod_{\bar{i}=0}^{\bar{K}\bar{M}-1} \left[(1-w) \delta(|x_{\bar{i}}|) + \frac{w}{\pi \sigma_x^2} \exp \left\{ -\frac{|x_{\bar{i}}|^2}{\sigma_x^2} \right\} \right] \\
&\quad \times \exp \left\{ -\beta_1 \sigma_x^{-2} \right\} \sigma_x^{-2(\beta_0+1)} \mathbb{I}_{[0,+\infty)}(\sigma_x^2) \\
&\propto \sigma_x^{-2(n_1+\beta_0+1)} \exp \left\{ -\sigma_x^{-2} (\|\mathbf{x}\|_2^2 + \beta_1) \right\} \mathbb{I}_{[0,+\infty)}(\sigma_x^2).
\end{aligned}$$

On reconnaît à nouveau une distribution inverse-Gamma $\mathcal{IG}(\beta_0 + n_1, \beta_1 + \|\mathbf{x}\|_2^2)$, également inchangée par rapport au cas non-robustifié.

2.2.4 Estimation des paramètres de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$

Le désalignement des cibles par rapport à la grille d'analyse a donc été pris en compte via la paramétrisation directe du dictionnaire \mathbf{H} par deux vecteurs de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$. Il s'agit alors d'estimer ces paramètres d'intérêt, ici en utilisant un échantillonneur de Gibbs modifié. Pour cela, on cherche à calculer la distribution jointe conditionnelle a posteriori $f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \mathbf{y}, \sigma^2, \mathbf{x})$, puis à tirer des échantillons suivant cette loi. Ce travail a été présenté dans diverses publications : dans le cas unidimensionnel [3], [1], dans le cas bidimensionnel sans migration [4] puis dans le cas bidimensionnel avec migration [6], ce dernier étant présenté ci-après.

2.2.4.1 Calcul de la distribution conditionnelle a posteriori

Les paramètres de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$ sont également tirés élément par élément ; on cherche donc à calculer la distribution conditionnelle a posteriori de $(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)$. La distribution a priori de $(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)$ étant conditionnée par $x_{\bar{i}}$, on fait une disjonction des cas :

$$\begin{aligned}
f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \mathbf{y}, \sigma^2, \mathbf{x}; x_{\bar{i}} = 0) &\propto f(\mathbf{y} | \mathbf{x}, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \sigma^2) \pi(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | x_{\bar{i}} = 0) \\
&= \delta(\varepsilon_{\bar{i}}^v) \delta(\varepsilon_{\bar{i}}^r).
\end{aligned}$$

D'autre part, si $x_{\bar{i}} \neq 0$:

$$\begin{aligned}
f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \mathbf{y}, \sigma^2, \mathbf{x}, \varepsilon_{-\bar{i}}^v, \varepsilon_{-\bar{i}}^r; x_{\bar{i}} \neq 0) &\propto f(\mathbf{y} | \mathbf{x}, \varepsilon^v, \varepsilon^r, \sigma^2) \pi(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | x_{\bar{i}} \neq 0) \\
&\propto \exp \left\{ -\sigma^{-2} \left[|x_{\bar{i}}|^2 \|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2 - 2 \operatorname{Re} \left\{ x_{\bar{i}}^* \mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}} \right\} \right] \right\} \\
&\quad \times \mathbb{I}_{[-0.5, 0.5]}(\varepsilon_{\bar{i}}^v) \mathbb{I}_{[-0.5, 0.5]}(\varepsilon_{\bar{i}}^r) \\
&\propto \exp \left\{ 2\sigma^{-2} \operatorname{Re} \left\{ x_{\bar{i}}^* \mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}} \right\} \right\} \mathbb{I}_{[-0.5, 0.5]}(\varepsilon_{\bar{i}}^v) \mathbb{I}_{[-0.5, 0.5]}(\varepsilon_{\bar{i}}^r) \\
&\propto \exp \left\{ \operatorname{Re} \left\{ \sum_{i=0}^{KM-1} 2\sigma^{-2} x_{\bar{i}}^* [\mathbf{h}_{\bar{i}}^H(0, 0)]_i [\mathbf{e}_{\bar{i}}]_i \right. \right. \\
&\quad \left. \left. \exp \left\{ j2\pi \left(\frac{\varepsilon_{\bar{i}}^r}{K} k - \frac{\varepsilon_{\bar{i}}^v}{M} n_{va} (1 + \mu k) m \right) \right\} \right\} \right\} \\
&\quad \times \mathbb{I}_{[-.5, .5]}(\varepsilon_{\bar{i}}^v) \mathbb{I}_{[-.5, .5]}(\varepsilon_{\bar{i}}^r).
\end{aligned}$$

Finalement on obtient une expression concise de la distribution conditionnelle a posteriori de $(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)$

$$\begin{aligned}
f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \mathbf{y}, \sigma^2, \mathbf{x}, \varepsilon_{-\bar{i}}^v, \varepsilon_{-\bar{i}}^r; x_{\bar{i}} \neq 0) &\propto \exp \left\{ \sum_{i=0}^{KM-1} \kappa_i \cos \left(2\pi \left[\frac{\varepsilon_{\bar{i}}^v}{M} n_{va} (1 + \mu k) m - k \frac{\varepsilon_{\bar{i}}^r}{K} \right] - \phi_i \right) \right\} \\
&\quad \times \mathbb{I}_{[-.5, .5]}(\varepsilon_{\bar{i}}^v) \mathbb{I}_{[-.5, .5]}(\varepsilon_{\bar{i}}^r). \tag{2.18}
\end{aligned}$$

avec $\kappa_i = |[\mathbf{z}_{\bar{i}}]_i|$, $\phi_i = \angle[\mathbf{z}_{\bar{i}}]_i$, et $\mathbf{z}_{\bar{i}} = 2\sigma^{-2} x_{\bar{i}}^* [\mathbf{h}_{\bar{i}}^*(0, 0) \odot \mathbf{e}_{\bar{i}}]$. La Fig. 2.4 représente l'allure de la distribution (2.18) pour deux valeurs de SNR. On voit que pour un SNR faible la distribution $f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \mathbf{y}, \sigma^2, \mathbf{x}, \varepsilon_{-\bar{i}}^v, \varepsilon_{-\bar{i}}^r; x_{\bar{i}} \neq 0)$ est plutôt étalée, alors que pour des valeurs plutôt élevées (SNR=25 dB) elle est très piquée.

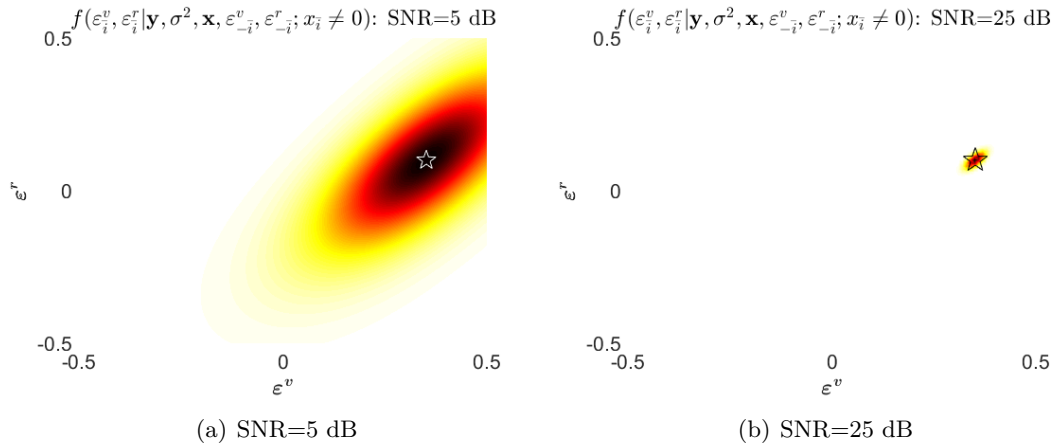


FIGURE 2.4 – Allure de la distribution conditionnelle a posteriori $f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \mathbf{y}, \sigma^2, \mathbf{x}, \varepsilon_{-\bar{i}}^v, \varepsilon_{-\bar{i}}^r; x_{\bar{i}} \neq 0)$ pour différentes valeurs de SNR : 2.4(a) SNR=5 dB, 2.4(b) SNR=25 dB.

2.2.4.2 Modification de l'échantillonneur de Gibbs

La distribution (2.18) n'appartient à aucune classe de distributions connues. Une étape spécifique doit alors être ajoutée à l'échantillonneur de Gibbs pour tirer les échantillons $(\varepsilon_i^v, \varepsilon_i^r)$. Différentes méthodes ont été testées pour échantillonner correctement ces paramètres ; elles sont résumées ci-après. De plus, une étude approfondie de la distribution (2.18) est présentée dans l'annexe A dans le cas unidimensionnel où $K = 1$. Cette étude préliminaire a permis les développements suivants concernant l'échantillonnage des paramètres de désalignement.

Inversion de CDF Tout d'abord, on teste la méthode classique d'inversion de CDF (*Cumulative Distribution Function*), en se limitant au cas unidimensionnel ($K = 1$). Cette méthode repose sur le calcul de la CDF conditionnelle a posteriori de $\varepsilon_{\bar{m}}$, qui est effectué empiriquement à chaque étape de l'échantillonneur de Gibbs à partir de la distribution conditionnelle a posteriori $f(\varepsilon_{\bar{m}}^{(n-1)} | \mathbf{y}, \mathbf{x}^{(n)}, \sigma^{2(n)}, \boldsymbol{\varepsilon}_{-\bar{m}}^{(n)})$. On rappelle les étapes de la méthode d'inversion de CDF [46], appliquée à l'échantillonnage de $\varepsilon_{\bar{m}}$:

1. On calcule la CDF de $\varepsilon_{\bar{m}}^{(n-1)}$.
2. On tire un échantillon uniformément distribué sur l'intervalle $[0, 1]$.
3. On trouve l'antécédent de cet échantillon suivant la CDF calculée à la première étape : c'est $\varepsilon_{\bar{m}}^{(n)}$.

Le calcul de la CDF se fait empiriquement : on calcule en réalité la somme cumulative de la distribution conditionnelle a posteriori avec un certain pas, dont le choix est primordial. Par souci d'efficacité calculatoire, le pas est limité à 10^{-3} .

Metropolis-Hastings Une étape de Metropolis-Hastings (MH) peut être ajoutée à l'échantillonneur de Gibbs afin de tirer conjointement $(\varepsilon_i^v, \varepsilon_i^r)$. L'intérêt de l'algorithme de MH est qu'il peut être intégré dans l'échantillonneur de Gibbs simplement. En effet, l'échantillonneur de Gibbs hybride ainsi formé est assuré de converger même en ne faisant qu'un seul tirage de MH à chaque itération de l'échantillonneur de Gibbs [90]. L'algorithme de MH fait partie de la classe des méthodes d'acceptation-rejet ; sa formulation générale est la suivante [90] :

- Candidat : $c \sim q(c|x^{(n)})$
- Taux d'acceptation : $\rho(x^{(n)}, c) = \min \left\{ 1, \frac{f(c) q(x^{(n)}|c)}{f(x^{(n)}) q(c|x^{(n)})} \right\}$
- Nouvel échantillon : $x^{(n+1)} = \begin{cases} c & \text{avec probabilité } \rho(x^{(n)}, c) \\ x^{(n)} & \text{avec probabilité } 1 - \rho(x^{(n)}, c). \end{cases}$

où $q(c|x^{(n)})$ est la loi de proposition et $f(x^{(n)})$ la loi cible. Le choix de la loi de proposition est primordial car plus elle est proche de la loi cible, meilleur est le taux d'acceptation $\rho(x^{(n)}, c)$. Un meilleur taux d'acceptation permet ensuite une convergence plus rapide.

Dans notre cas, le tirage de MH permet d'obtenir un échantillon $(\varepsilon_i^v, \varepsilon_i^r)$. Un candidat (c^v, c^r) est tiré suivant une loi de proposition $q(c^v, c^r | \varepsilon_i^v, \varepsilon_i^r)$. Il est accepté ou rejeté avec

le taux d'acceptation

$$\rho\left((\varepsilon_i^{v(n)}, \varepsilon_i^{r(n)}), (c^v, c^r)\right) = \min \left\{ 1, \frac{f(c^v, c^r | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r)}{f(\varepsilon_i^{v(n)}, \varepsilon_i^{r(n)} | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r)} \frac{q(\varepsilon_i^{v(n)}, \varepsilon_i^{r(n)} | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r)}{q(c^v, c^r | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r)} \right\} \quad (2.19)$$

qui dépend de la ressemblance entre la loi cible $f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r; x_{\bar{i}} \neq 0)$ et la loi de proposition $q(c^v, c^r | \varepsilon_i^{v(n)}, \varepsilon_i^{r(n)})$.

La Fig. 2.4, ainsi que les développements résumés dans l'annexe A, suggèrent qu'une loi de proposition adaptée puisse dépendre du SNR de la cible située à la \bar{i} -ième case d'analyse :

- une loi de proposition uniforme pour un SNR faible : $\varepsilon_i^v, \varepsilon_i^r \sim_{q_1} \mathcal{U}_{[-.5, .5] \times [-.5, .5]}$.
- une loi de proposition Gaussienne pour un SNR fort : $\varepsilon_i^v, \varepsilon_i^r \sim_{q_2} \mathcal{N}(\boldsymbol{\mu}_\varepsilon, \boldsymbol{\Sigma}_\varepsilon)$.

Ainsi, on n'utilise pas une, mais deux lois de proposition de formes différentes suivant le SNR estimé, q_1 ou q_2 , alors que la forme générique de l'algorithme MH n'utilise qu'une seule loi de proposition, dont les paramètres dépendent de l'itération précédente. A cause de ce changement de loi de proposition, les propriétés de convergence ne sont plus théoriquement assurées, mais nous avons constaté que les performances d'estimation sont malgré tout très acceptables.

Les moments d'ordre 1 et 2 de la loi de proposition Gaussienne q_2 changent à chaque itération et doivent donc être estimés au préalable. Tout d'abord, la moyenne est estimée comme le mode de la distribution $f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r; x_{\bar{i}} \neq 0)$. Pour cela, (2.18) est reformulée :

$$f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r; x_{\bar{i}} \neq 0) \propto \exp \left\{ \text{Re} \left\{ \sum_{i=1}^{KM-1} [\mathbf{z}_{\bar{i}}]_i \exp \{j2\pi (f_{rci}k - f_{dci}(1 + \mu k)m)\} \right\} \right\} \mathbb{I}_{[-.5, .5]}(\varepsilon_i^v) \mathbb{I}_{[-.5, .5]}(\varepsilon_i^r) \quad (2.20)$$

avec $f_{rci} = \varepsilon_i^r / \bar{K}$ et $f_{dci} = n_{va} \varepsilon_i^v / \bar{M}$. On reconnaît là la sommation cohérente adaptée à la forme d'onde large bande du vecteur $\mathbf{z}_{\bar{i}}$ [77, p.522]. Ainsi, pour estimer $\boldsymbol{\mu}_\varepsilon$ il suffit d'estimer la localisation du maximum de cette sommation cohérente restreinte au domaine $(\varepsilon_i^v, \varepsilon_i^r) \in [-.5, .5] \times [-.5, .5]$. Ensuite, on utilise un développement limité d'ordre 2 de la fonction (2.18) autour de son maximum, c'est-à-dire autour de $\boldsymbol{\mu}_\varepsilon$. La distribution $f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r; x_{\bar{i}} \neq 0)$ est alors approchée par

$$\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \sigma^2, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r; x_{\bar{i}} \neq 0 \sim \mathcal{N}(\boldsymbol{\mu}_\varepsilon, \boldsymbol{\Sigma}_\varepsilon). \quad (2.21)$$

Cette démarche est donc assez lourde calculatoirement puisqu'elle nécessite le calcul d'une sommation cohérente à chaque itération, avant le tirage des candidats.

Remarque. *Finally, the method of inversion of CDF gives similar performances to those of MH in terms of estimation of target amplitude, but a more important error on the estimation of the misalignment². The method of inversion of CDF does not improve the*

2. Les deux méthodes ont été comparées dans le cas unidimensionnel seulement (une seule case distance).

performances de l'analyse par rapport à l'algorithme de MH malgré un coût calculatoire plus important, on décide de ne pas l'utiliser.

Tirage direct Malgré ses garanties de convergence, l'échantillonneur de Gibbs hybride peut être ralenti si la distribution cible dans l'étape de MH est très piquée, comme dans le cas d'un SNR élevé. En effet, si la loi de proposition, elle-même très piquée, ne colle pas correctement à la loi cible alors l'acceptation de l'échantillon est rare. En conséquence, la chaîne d'échantillonnage se retrouve « bloquée » sur une valeur, et peut mettre un certain temps à changer de valeur via l'acceptation d'un nouvel échantillon. Une autre méthode d'échantillonnage peut alors être adoptée : les échantillons $(\varepsilon_i^v, \varepsilon_i^r)$ sont directement tirés suivant leur distribution $f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \sigma^2, \mathbf{x}; x_i \neq 0)$ calculée sur le domaine $[-0.5, 0.5] \times [-0.5, 0.5]$ arbitrairement discrétisé. Cette méthode n'est pas satisfaisante d'un point de vue algorithmique, mais elle a l'avantage de produire un nouvel échantillon à chaque itération.

La Fig. 2.5 résume l'échantillonneur de Gibbs utilisé pour estimer les paramètres d'intérêt \mathbf{x} , $(\varepsilon^v, \varepsilon^r)$. On rappelle que cet algorithme fournit en outre l'estimateur MMSE des autres paramètres et hyperparamètres, en particulier de la puissance de bruit σ^2 qui pourrait être utilisé pour mettre en place un seuil de détection. Il est important de souligner que l'échantillonneur de Gibbs possède des garanties de convergence qui ne dépendent pas de son initialisation [90]. En revanche, cette convergence peut être plus ou moins rapide, et c'est bien ce que nous avons constaté.

2.2.5 Remarque : prise en compte du désalignement via un dictionnaire linéarisé

Comme mentionné dans la partie 2.1, le désalignement des cibles par rapport à la grille d'analyse est souvent pris en compte via un modèle linéaire (2.1). Cette approche a été testée en se limitant au cas unidimensionnel ($K = 1$) ; la méthodologie et les résultats sont présentés dans [1] et résumés ici. On applique un développement limité d'ordre 1 (DL1) autour de $\varepsilon = \mathbf{0}$ à la matrice de Fourier \mathbf{F} pour qu'elle ne soit plus directement paramétrée par ε . On notera $\mathbf{G}(\varepsilon)$ la matrice de Fourier après DL1 :

$$\mathbf{G}(\varepsilon) = \mathbf{F}(\mathbf{0}) + \left. \frac{\partial \mathbf{F}}{\partial \varepsilon} \right|_{\varepsilon=\mathbf{0}} \text{diag}(\varepsilon - \mathbf{0}). \quad (2.22)$$

Il s'agit alors de calculer la quantité

$$\left. \frac{\partial \mathbf{F}}{\partial \varepsilon} \right|_{\varepsilon=\mathbf{0}} = \begin{bmatrix} \left. \frac{\partial \mathbf{f}_0}{\partial \varepsilon_0} \right|_{\varepsilon_0=0} & \cdots & \left. \frac{\partial \mathbf{f}_{M-1}}{\partial \varepsilon_{M-1}} \right|_{\varepsilon_{M-1}=0} \end{bmatrix}.$$

En notant $\mathbf{D} = \text{diag} \left(j \frac{2\pi}{M} \left[0 \quad \dots \quad M-1 \right] \right)$ et $\mathbf{D}_\varepsilon = \text{diag}(\varepsilon)$ on obtient

$$\mathbf{G}(\varepsilon) = \mathbf{F}(\mathbf{0}) + \mathbf{D}\mathbf{F}(\mathbf{0})\mathbf{D}_\varepsilon. \quad (2.23)$$

Require: \mathbf{y} , (γ_0, γ_1) , (β_0, β_1)
Ensure: $\hat{\mathbf{x}}_{\text{MMSE}}$, $(\hat{\boldsymbol{\varepsilon}}^v, \hat{\boldsymbol{\varepsilon}}^r)_{\text{MMSE}}$, $\hat{\sigma}_{\text{MMSE}}^2$, \hat{w}_{MMSE} , $\hat{\sigma}_x^2_{\text{MMSE}}$

```

{Initialisation}
 $w^{(0)} \sim \mathcal{U}_{[0,1]}$ 
 $\sigma_x^2{}^{(0)} \sim \mathcal{IG}(\beta_0, \beta_1)$ 
 $\mathbf{x}^{(0)} \sim \prod_{i=0}^{\bar{K}\bar{M}-1} \text{BerCN}(w^{(0)}, 0, \sigma_x^2{}^{(0)})$ 
 $(\boldsymbol{\varepsilon}^v{}^{(0)}, \boldsymbol{\varepsilon}^r{}^{(0)}) \sim \mathcal{U}_{[-0.5,0.5] \times [-0.5,0.5]}$ 
{Itérations}
for  $n = 1$  à  $N_{bi} + N_r$  do
   $\sigma^2{}^{(n)} \sim \mathcal{IG}(KM + \gamma_0, \|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v{}^{(n-1)}, \boldsymbol{\varepsilon}^r{}^{(n-1)})\mathbf{x}^{(n-1)}\|_2^2 + \gamma_1)$  /*step 1*/
   $w^{(n)} \sim \mathcal{Be}(1 + n_1, 1 + n_0)$  /*step 2*/
   $\sigma_x^2{}^{(n)} \sim \mathcal{IG}(\beta_0 + n_1, \beta_1 + \|\mathbf{x}^{(n-1)}\|_2^2)$  /*step 3*/
  for  $\bar{i} = 0$  à  $\bar{K}\bar{M} - 1$  do
     $x_{\bar{i}}^{(n)} \sim \text{BerCN}(w_{\bar{i}}, \mu_{\bar{i}}, \eta_{\bar{i}}^2)$  /*step 4*/
     $(\varepsilon_{\bar{i}}^v{}^{(n)}, \varepsilon_{\bar{i}}^r{}^{(n)})$  via une étape de MH ou tirage direct. /*step 5*/
  end for
end for
{Estimateurs}
 $\hat{\theta}_{\text{MMSE}} = \frac{1}{N_r} \sum_{n=1}^{N_r} \theta^{(n+N_{bi})}$ 

```

FIGURE 2.5 – Schéma d'estimation prenant en compte le désalignement par rapport à la grille d'analyse; $n_1 = \|\mathbf{x}^{(n-1)}\|_0$, $n_0 = \bar{K}\bar{M} - \|\mathbf{x}^{(n-1)}\|_0$.

Le nouveau modèle s'écrit alors

$$\mathbf{y} = (\mathbf{F}(\mathbf{0}) + \mathbf{D}\mathbf{F}(\mathbf{0})\mathbf{D}_\varepsilon)\mathbf{x} + \mathbf{n} \triangleq \mathbf{G}(\boldsymbol{\varepsilon})\mathbf{x} + \mathbf{n}. \quad (2.24)$$

Ensuite, le même modèle hiérarchique Bayésien est utilisé pour calculer les estimateurs MMSE (2.15), (2.16). L'échantillonnage des paramètres \mathbf{x} , σ^2 , w , σ_x^2 est inchangé, si ce n'est que la matrice $\mathbf{F}(\boldsymbol{\varepsilon})$ est remplacée par la matrice $\mathbf{G}(\boldsymbol{\varepsilon})$. En revanche, l'échantillonnage des paramètres de désalignement est modifié : la distribution conditionnelle a posteriori de $\varepsilon_{\bar{m}}$ s'écrit simplement

$$f(\varepsilon_{\bar{m}} | \mathbf{y}, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{m}}, w, \sigma^2, \sigma_x^2) \propto \exp \left\{ -\frac{(\varepsilon_{\bar{m}} - \mu_{\varepsilon_{\bar{m}}})^2}{2\sigma_{\varepsilon_{\bar{m}}}^2} \right\} f(\varepsilon_{\bar{m}} | x_{\bar{m}}) \quad (2.25)$$

avec

$$\sigma_{\varepsilon_{\bar{m}}}^2 = \frac{\sigma^2}{2\|\mathbf{f}_{\bar{m}}\|_2^2 |x_{\bar{m}}|^2}$$

$$\mu_{\varepsilon_{\bar{m}}} = 2\mathcal{Re} \left\{ x_{\bar{m}}^* \mathbf{f}_{\bar{m}}^H \tilde{\boldsymbol{\varepsilon}}_{\bar{m}} \right\} \frac{\sigma_{\varepsilon_{\bar{m}}}^2}{\sigma^2}.$$

où $\mathbf{f}_{\bar{m}}$ désigne la \bar{m} -ième colonne de $\mathbf{DF}(\mathbf{0})$. Par conséquent,

$$f(\varepsilon_{\bar{m}}|\mathbf{y}, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{m}}, w, \sigma^2, \sigma_x^2) \propto \delta(\varepsilon_{\bar{m}}) \quad \text{si } x_{\bar{m}} = 0 \quad (2.26)$$

$$\varepsilon_{\bar{m}}|\mathbf{y}, \mathbf{x}, \boldsymbol{\varepsilon}_{-\bar{m}}, w, \sigma^2, \sigma_x^2 \sim \mathcal{N}_{[-0.5, +0.5]}(\mu_{\varepsilon_{\bar{m}}}, \sigma_{\varepsilon_{\bar{m}}}^2) \quad \text{si } x_{\bar{m}} \neq 0 \quad (2.27)$$

où $\mathcal{N}_{[-0.5, +0.5]}(\mu_{\varepsilon_{\bar{m}}}, \sigma_{\varepsilon_{\bar{m}}}^2)$ désigne une gaussienne réelle tronquée sur l'intervalle $[-0.5, +0.5]$ de moyenne $\mu_{\varepsilon_{\bar{m}}}$ et variance $\sigma_{\varepsilon_{\bar{m}}}^2$. L'échantillonnage de $\varepsilon_{\bar{m}}$ est donc simplifié par rapport à une paramétrisation directe du dictionnaire, puisqu'il existe dans la littérature des méthodes efficaces pour tirer des échantillons suivant une loi Gaussienne tronquée, par exemple [41].

Néanmoins, la méthode reposant sur un DL1 offre de moins bonnes performances que celle utilisant une paramétrisation directe du dictionnaire. En particulier, on compare dans [1] ces deux approches et on montre que les performances de l'approche reposant sur un DL1 dépendent fortement de la valeur du désalignement $\varepsilon_{\bar{m}}$; l'écart avec la méthode de paramétrisation directe est d'autant plus grande que le SNR augmente (voir Fig. 2.10).

2.3 Simulations numériques

On présente dans cette section les résultats obtenus sur données synthétiques par l'algorithme robuste au désalignement par rapport à la grille d'analyse. Les données synthétiques sont générées en ajoutant des cibles à un bruit blanc Gaussien suivant le modèle (1.6).

2.3.1 Exemple de réalisation

On considère dans cette section un scénario à 4 cibles, dont les caractéristiques sont détaillées dans le Tableau 2.1. En particulier, la cible 4 est située dans le lobe secondaire de la cible 3, et la cible 2 est située à $+v_a$. On représente sur la Fig. 2.6 les cartes distance-vitesse estimées par l'algorithme non-robustifié de [24] (dénommé « Algo. SSR : WN+sans mismatch ») et par la version prenant en compte le désalignement de [1, 4, 6] présentée dans ce chapitre (« Algo. SSR : WN+mismatch »). Comme attendu, l'algorithme non-robustifié ne permet pas d'estimer correctement la scène : la cible 1 n'est presque pas vue (≈ 0 dB), l'énergie de la cible 3 est répartie sur les cases distance 6 et 7 (et dans les lobes secondaires en vitesse). Enfin, les picots estimés autour de la cible 4 peuvent correspondre à une répartition de l'énergie de la cible, ou bien aux lobes secondaires de la cible 3. Par contre, l'algorithme prenant en compte le désalignement estime correctement la scène radar, en levant les ambiguïtés vitesse : les lobes secondaires des cibles 1, 3 et 4 ne sont pas vus comme des cibles, et la cible 2 à $+v_a$ est correctement estimée.

On représente sur la Fig. 2.7 les chaînes d'échantillonnage de l'algorithme prenant en compte le désalignement ; on ne représente l'échantillonnage de \mathbf{x} , $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$ que pour la deuxième cible. Ces chaînes convergent rapidement, y compris les chaînes d'échantillonnage des paramètres de désalignement. Les échantillons tirés sont également utilisés pour construire les

TABLE 2.1 – Caractéristiques des cibles utilisées dans cet exemple de réalisation (ce scénario est repris plusieurs fois dans le manuscrit). Les cibles sont numérotées de la plus proche à la plus éloignée.

	case distance	vitesse (m/s)	ε_i^v	ε_i^r
# 1	1	-9.23	0.3	0.4
# 2	4	14.81	-0.4	0.1
# 3	6	-4.59	0.2	0.4
# 4	7	10.41	0.2	0.4

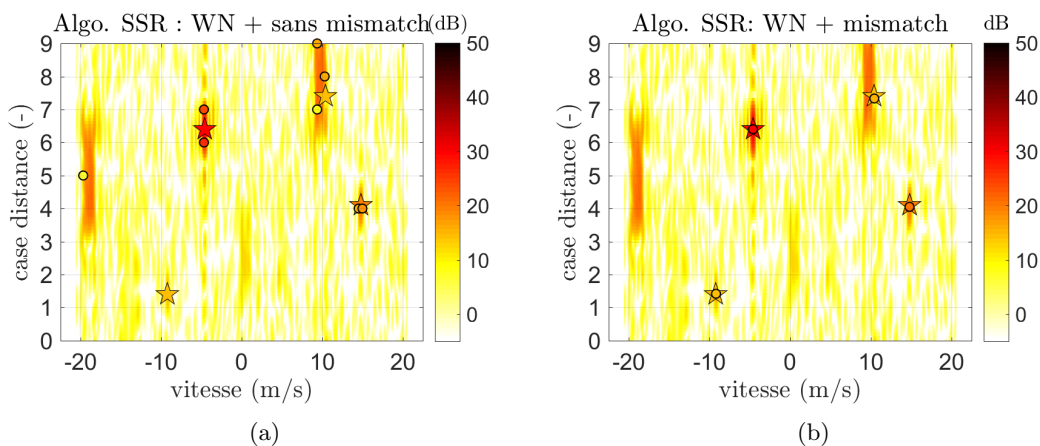


FIGURE 2.6 – Cartes distance-vitesse estimées par les différents algorithmes 2.6(a) Algorithme Bayésien de SSR non robustifié [24]; 2.6(b) Algorithme Bayésien de SSR prenant en compte le désalignement. Le fond représente la sommation cohérente large bande, les étoiles la vraie scène et les ronds la scène estimée. $K = 10$, $M = 32$, $n_{va} = 3$, $v_a = 15$ m/s, $B/f_c=10\%$, $\sigma^2 = 1$.

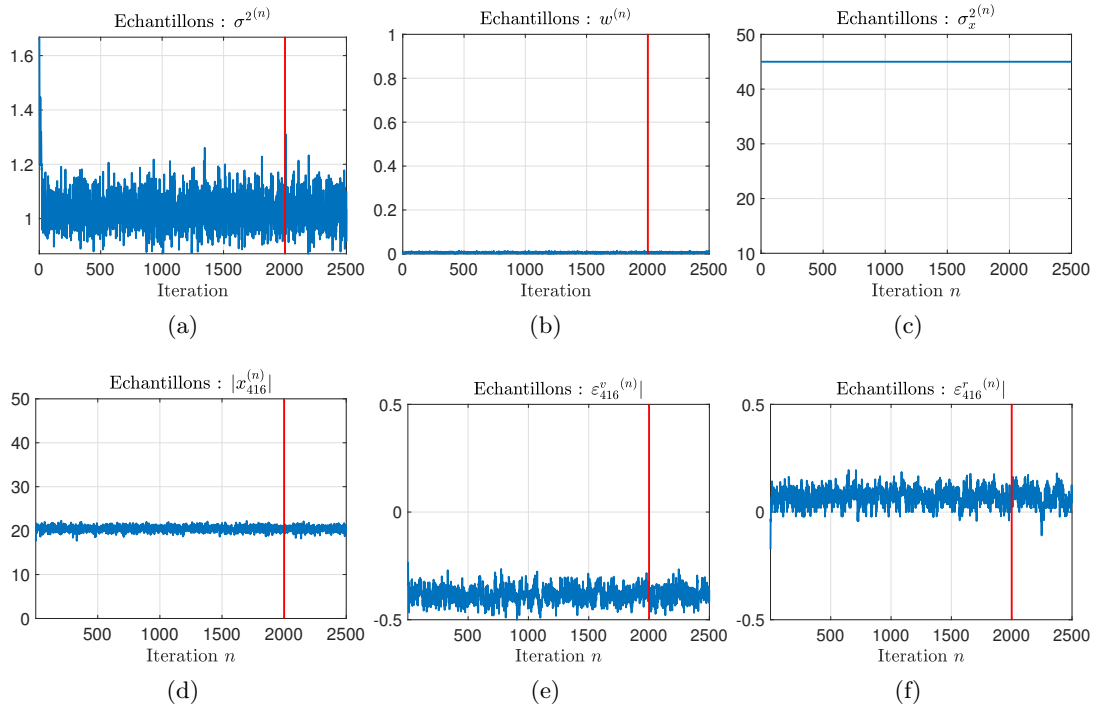


FIGURE 2.7 – Chaînes d'échantillonnage des paramètres σ^2 , w et σ_x^2 , et de \mathbf{x} et $(\varepsilon^v, \varepsilon^r)$ pour la cible 2. La ligne verticale rouge représente la fin de la période de chauffe choisie.

distributions a posteriori empiriques des différents paramètres ; elles sont représentées sur la Fig. 2.8. Il est intéressant de voir que ces distributions sont piquées (moments d'ordre 4 positifs et tous supérieurs à 2.9), même dans le cas d'une distribution a priori non-informative (paramètre w par exemple). En outre, les estimateurs MMSE calculés possèdent effectivement une faible erreur quadratique moyenne.

2.3.2 Simulations de Monte-Carlo

On procède ensuite à des simulations de Monte-Carlo (MC) afin d'approfondir la comparaison entre l'algorithme prenant en compte le désalignement et la version non-robustifiée de [24]. Les critères adoptés sont l'erreur quadratique moyenne normalisée (*normalized Mean Square Error*, nMSE) de la scène estimée $\mathbf{H}(\hat{\varepsilon}^v, \hat{\varepsilon}^r)\hat{\mathbf{x}}$, celle du vecteur d'amplitude $\hat{\mathbf{x}}$ et celle

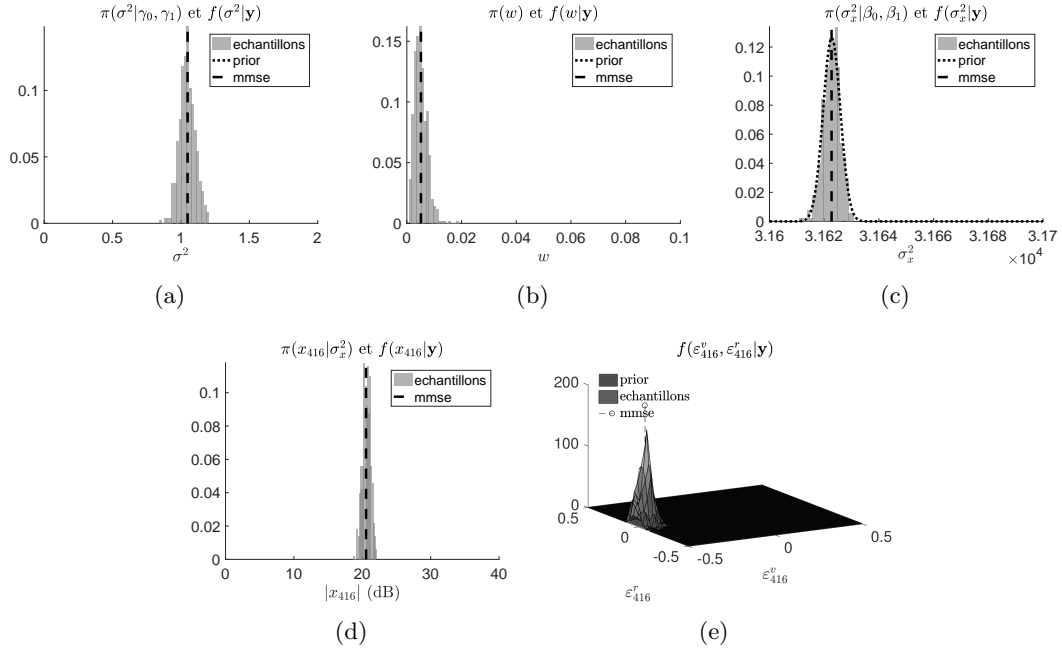


FIGURE 2.8 – Distributions a posteriori empiriques construites à partir des échantillons tirés des paramètres σ^2 , w et σ_x^2 , ainsi que ceux de l'élément des vecteurs \mathbf{x} et (ϵ^v, ϵ^r) correspondant à la cible 2. La ligne verticale pointillée représente l'estimateur MMSE de chaque paramètre ; la courbe pointillée la distribution a priori.

de chaque élément $\hat{x}_{\bar{i}}$:

$$\text{nMSE}(\mathbf{H}(\hat{\epsilon}^v, \hat{\epsilon}^r)\hat{\mathbf{x}}) = \frac{1}{N_{mc}} \sum_{n=1}^{N_{mc}} \frac{\|\mathbf{H}(\hat{\epsilon}^v(n), \hat{\epsilon}^r(n))\hat{\mathbf{x}}^{(n)} - \mathbf{H}(\epsilon^v, \epsilon^r)\mathbf{x}\|_2^2}{\|\mathbf{H}(\epsilon^v, \epsilon^r)\mathbf{x}\|_2^2} \quad (2.28)$$

$$\text{nMSE}(\hat{\mathbf{x}}) = \frac{1}{N_{mc}} \sum_{n=1}^{N_{mc}} \frac{\|\hat{\mathbf{x}}^{(n)} - \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} \quad (2.29)$$

$$\text{nMSE}(\hat{x}_{\bar{i}}) = \frac{1}{N_{mc}} \sum_{n=1}^{N_{mc}} \frac{|\hat{x}_{\bar{i}}^{(n)} - x_{\bar{i}}|^2}{|x_{\bar{i}}|^2}. \quad (2.30)$$

2.3.2.1 Intérêt de la robustification au désalignement

Dans un premier temps, on évalue l'intérêt de la robustification au désalignement en comparant les performances de l'algorithme robustifié à celles de l'algorithme non robustifié, en fonction de la valeur du désalignement en distance ou en vitesse. Le scénario consiste en une seule cible avec un SNR de 20, 30 ou 40 dB en désalignement par rapport à la grille d'analyse en distance ou en vitesse : $\epsilon_{\bar{i}}^v = 0$ et $\epsilon_{\bar{i}}^r \neq 0$, ou bien $\epsilon_{\bar{i}}^v \neq 0$ et $\epsilon_{\bar{i}}^r = 0$. La cible possède une vitesse nulle et se trouve à la 4^{ème} case distance. Les paramètres de scénario et de traitement sont résumés dans les Tableaux 2.2 et 2.3 respectivement. En particulier, la grille d'analyse n'est pas sur-échantillonnée : $n_{zp}^r = n_{zp}^d = 1$ (on montre dans la section 2.3.2.3 que cela

n'améliore pas les performances). La période de chauffe de l'échantillonneur de Gibbs N_{bi} et le nombre de réalisations N_r sont fixés empiriquement : les chaînes d'échantillonnage doivent avoir convergé avant N_{bi} , et N_r doit permettre d'avoir assez d'échantillons pour le calcul des estimateurs MMSE. La distribution a priori de la puissance du bruit σ^2 est non-informative : nous avons observé dans [2] qu'un tel choix ne détériore pas de manière significative les performances (l'algorithme testé utilise un modèle hiérarchique Bayésien quelque peu différent et sera présenté dans le chapitre 4). Par contre, la distribution a priori de la puissance des cibles σ_x^2 est informative et il est préférable que sa moyenne corresponde à la puissance maximale observée sur la carte (obtenue par exemple par sommation cohérente)³. Pour l'évaluation des performances sur données synthétiques, l'algorithme est initialisé grâce aux vraies valeurs des paramètres. Ce réglage permet de diminuer le temps de chauffe et donc la charge calculatoire (déjà élevée puisqu'on procède à $N_{mc} = 200$ simulations de Monte-Carlo). Pour traiter des données expérimentales (ou semi-expérimentales), on peut par exemple initialiser l'algorithme en procédant au préalable à une sommation cohérente large bande. Il faut alors ajuster la période de chauffe N_{bi} pour que les chaînes d'échantillonnage ait bien convergé.

TABLE 2.2 – Scénario pour les simulations de MC.

Nombre de cases distance	$K = 8$
Nombre d'impulsions	$M = 32$ $10 \log_{10}(KM) = 24.1$ dB.
Fréquence porteuse	$f_c = 10$ GHz
Largeur de bande	$B = 1$ GHz
Vitesse ambiguë	$v_a = 15$ m/s
Puissance du bruit	$\sigma^2 = 1$

TABLE 2.3 – Paramètres de traitement utilisés pour les simulations de Monte-Carlo

Facteur de dépliement	$n_{va} = 3$
Grille d'analyse	$\tilde{K} = K, \tilde{M} = n_{va}M$
Nombre d'itérations	$N_{bi} = 100, N_r = 1000$
Initialisation	Vraies valeurs
Prior de σ^2	$(\gamma_0, \gamma_1) = (0, 0)$
Prior de σ_x^2	$(m_{\sigma_x^2}, \sqrt{var_{\sigma_x^2}}) = (40, 5)$ dB

D'après la Fig.2.9, l'algorithme prenant en compte le désalignement surpasse la version non-robustifiée à partir d'une valeur de désalignement très faible ($\approx .03$ à 20 dB), quelle que soit la métrique observée. On remarque également que les performances de l'algorithme non-robustifié ne dépendent pas de la valeur du SNR de la cible pour un désalignement supérieur à 0.1, mais elles se dégradent quand le désalignement augmente. En revanche, les performances de l'algorithme robustifié ne dépendent pas de la valeur du désalignement : l'estimation de la scène radar est aussi juste pour un fort que pour un faible désalignement,

3. Le chapitre 4 se focalise plus précisément sur les conséquences de ce réglage et propose une approche alternative.

et elles sont les mêmes pour un désalignement en distance ou en vitesse. Pour finir, on voit que les performances de l’algorithme prenant en compte le désalignement s’améliorent lorsque le SNR de la cible augmente ; on perd environ 10 dB d’erreur lorsqu’on augmente le SNR de 10 dB.

2.3.2.2 Comparaison avec un algorithme de la littérature

L’algorithme prenant en compte le désalignement détaillé ici a également été comparé à un autre algorithme de la littérature, l’algorithme OGSBI [119] souvent pris comme référence. OGSBI repose sur un modèle linéaire pour prendre en compte le désalignement, comme détaillé en 2.2.5. Ensuite, une distribution a priori de type Laplacien est utilisée pour forcer la parcimonie du vecteur \mathbf{x} , et l’algorithme de résolution est un algorithme EM [119]. Cette comparaison, limitée au cas unidimensionnel ($K = 1$), est détaillée dans [1] et reproduite en partie sur la Fig. 2.10. Elle montre que les performances de l’algorithme OGSBI ne dépendent pas de la valeur du désalignement, mais qu’elles sont bien inférieures à celles de l’algorithme proposé. En particulier, on voit dans [1] que la scène estimée par OGSBI comporte de nombreuses fausses alarmes, et est donc bien moins « parcimonieuse ». Cette différence est probablement due au choix d’une distribution a priori Laplacienne, qui impose qu’un seul hyperparamètre règle à la fois le niveau de parcimonie et la puissance des cibles. Ce n’est pas le cas avec une distribution Bernoulli-Gaussienne, puisque celle-ci découple ce réglage via sa structure mixte et les hyperparamètres w et σ_x^2 . En outre, l’algorithme EM donne une solution sous-optimale comparée à celle de l’algorithme MCMC utilisé ici : il peut converger vers un maximum local, et pas global [90, chap. 5].

2.3.2.3 Performance en fonction de la finesse de la grille d’analyse

Le problème des cibles en chevauchement est souvent contourné en raffinant la grille d’analyse [74, 50, 51]. On teste donc les performances de l’algorithme proposé (prenant en compte le désalignement via une paramétrisation directe) en fonction de la finesse de la grille. A cause de l’augmentation rapide du coût calculatoire, on se limite à une analyse unidimensionnelle ($K = 1$). Le scénario correspond à une seule cible (Fig. 2.11(a)) dont la fréquence f_{d_i} est uniformément distribuée sur l’intervalle $[0, 1/M]$, et possédant un SNR de 10, 20 et 40 dB, ou deux cibles (Fig. 2.11(b)) de même SNR égal à 10, 20 et 40 dB. On compare donc la MSE normalisée de la scène reconstruite $\mathbf{F}(\hat{\boldsymbol{\varepsilon}})\hat{\mathbf{x}}$ estimée par l’algorithme prenant en compte le désalignement et sa version non-robustifiée de [24] pour une grille d’analyse en vitesse possédant \bar{M} éléments tels que $\bar{M} = nM$, $n = 1, \dots, 4$.

D’après la Fig. 2.11, les performances de l’algorithme prenant en compte le désalignement (ronds) ne dépendent pas de la finesse de la grille d’analyse. Autrement dit, il n’y a pas d’intérêt à sur-échantillonner la grille d’analyse du dictionnaire quand on utilise l’algorithme robustifié au désalignement. En revanche, l’algorithme non-robustifié de [24] améliore ses performances quand on raffine la grille d’analyse, et ce dès lors que $\bar{M} = 2M$, jusqu’à atteindre les performances de l’algorithme prenant en compte le désalignement pour un scénario avec

une seule cible (Fig. 2.11(a)0). Par contre, pour un scénario à plusieurs cibles (2 cibles pour la Fig. 2.11(b)) les performances de l’algorithme non-robustifié ne dépassent pas celles de l’algorithme robustifié, en particulier lorsqu’on augmente le SNR.

2.4 Simulations sur données expérimentales

On teste enfin l’algorithme proposé prenant en compte le désalignement sur des données expérimentales enregistrées par le radar PARSAX de TU Delft (Pays-Bas) le 16/11/2010 [70]; on le compare alors à la version non-robustifiée de [24]. Lors de l’enregistrement de ce jeu de données, le radar pointait en direction de l’autoroute entre Rotterdam et La Haye un mardi après-midi; on s’attend donc à observer des voitures s’approchant et s’éloignant du radar. Les paramètres pour ce jeu de données PARSAX sont résumés dans le Tableau 2.4; les paramètres de traitement dans le Tableau 2.5. Comme on ne connaît pas la réalité terrain, on calcule la sommation cohérente pour deux trains d’impulsions de longueur M consécutifs; elles sont représentées sur la Fig. 2.12. Grâce à ces valeurs de sommation cohérente, on peut calculer le *Migrating Target Indicator* (MiTI) de la scène. Les vraies cibles sont alors identifiées comme ayant une forme caractéristique de « doublet » avec deux pics à la même vitesse. Par convention, les cibles à vitesse positives s’approchent du radar; en comparant « à l’œil nu » les deux sommations cohérentes (Fig. 2.12(a) et 2.12(b)) on est également capable de distinguer les vraies cibles de leurs lobes secondaires. Sur la Fig. 2.12 on indique la position de ces vraies cibles en les entourant de noir.

TABLE 2.4 – Paramètres des données PARSAX

Nombre de cases distance	$K = 16$
Nombre d’impulsions	$M = 64$
	$10 \log_{10}(KM) = 30 \text{ dB.}$
Fréquence porteuse	$f_c = 3.315 \text{ GHz}$
Largeur de bande	$B = 100 \text{ MHz}$
Vitesse ambiguë	$v_a = 45 \text{ m/s}$
Puissance du bruit	$\sigma^2 = 1$

TABLE 2.5 – Paramètres de traitement pour les données PARSAX

Facteur de dépliement	$n_{va} = 2$
Grille d’analyse	$K = K, M = n_{va}M$
Nombre d’itérations	$N_{bi} = 1000, N_r = 1000$
Initialisation	Max de sommation coh.
Prior de σ^2	$(\gamma_0, \gamma_1) = (0, 0)$
Prior de σ_x^2	$(m_{\sigma_x^2}, \sqrt{\text{var}_{\sigma_x^2}}) = (40, 5)\text{dB}$

On représente sur la Fig. 2.13 les cartes distance-vitesse estimées par l’algorithme prenant en compte le désalignement (Fig. 2.13(b)) et la version non-robustifiée de [24] (Fig. 2.13(a))

à partir du jeu de données PARSAX. On voit que la version non-robustifiée ne permet pas de lever les ambiguïtés : au niveau de la case distance 1368, la vraie cible et ses lobes secondaires sont vus comme des cibles ; il en est de même pour la case distance 1372. En revanche, l'algorithme prenant en compte le désalignement lève effectivement les ambiguïtés vitesse : seules les vraies cibles sont estimées par l'algorithme, et pas leurs lobes secondaires. De plus, de nombreuses cibles sont estimées autour de $v = 0$ m/s : il s'agit du fouillis de sol. Le fouillis ne correspond ici qu'à des discrets à vitesse nulle, mais il peut également posséder une composante diffuse ; dans ce cas-là, l'estimation sous hypothèse de bruit blanc peut être fortement perturbée⁴.

4. La notion de fouillis de sol et sa prise en compte sont détaillées dans le chapitre suivant.

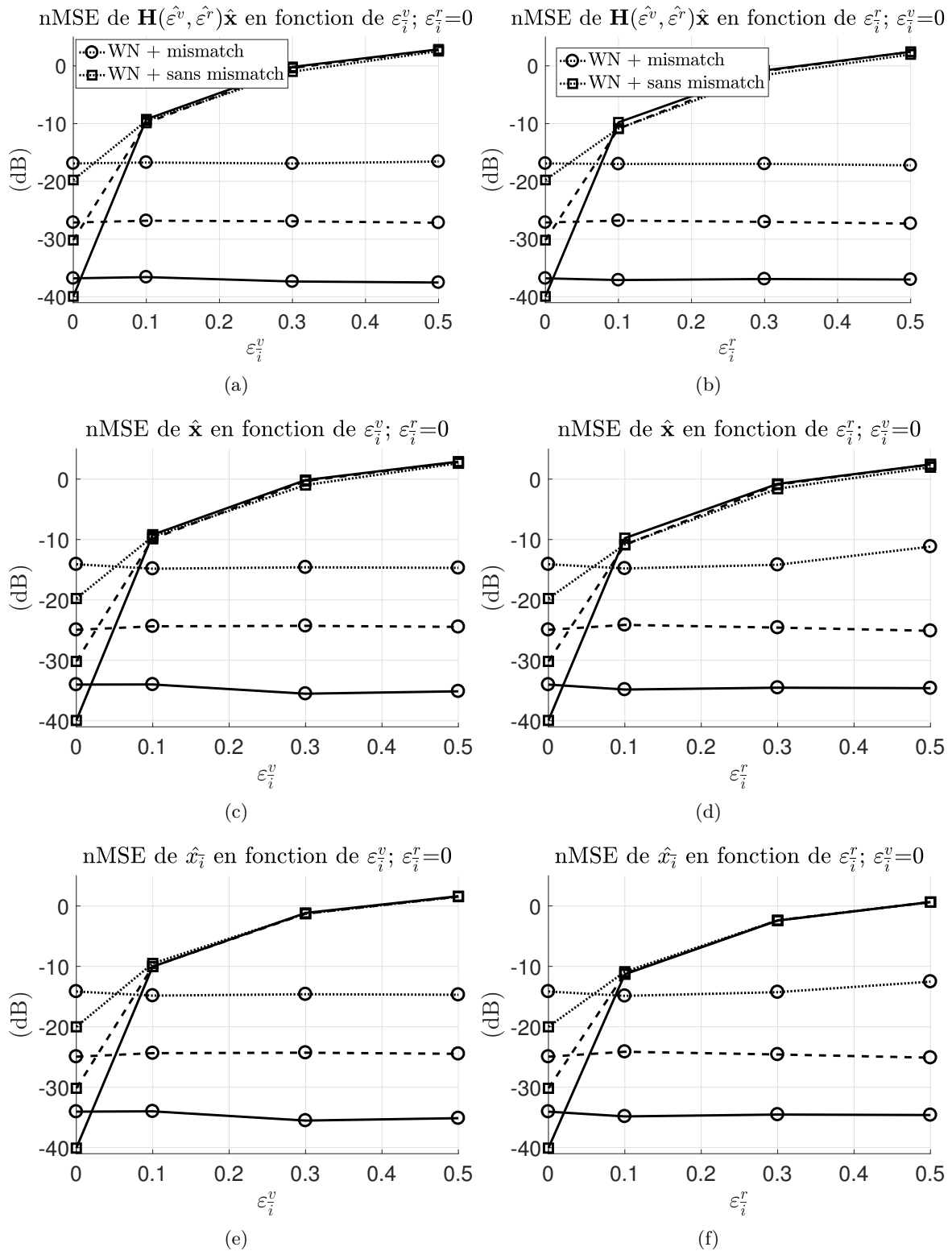


FIGURE 2.9 – nMSE de la scène reconstruite $\mathbf{H}(\hat{\varepsilon}^v, \hat{\varepsilon}^r)\hat{\mathbf{x}}$ (2.9(a),2.9(b)), du vecteur d'amplitude $\hat{\mathbf{x}}$ (2.9(c),2.9(d)) et de l'élément de \mathbf{x} où se situe la cible \hat{x}_i (2.9(e),2.9(f)) en fonction du désalignement en vitesse (colonne de gauche) ou en distance (colonne de droite). Le scénario consiste en une cible avec un SNR de 20 dB (pointillés), 30 dB (tirets) ou 40 dB (trait plein).

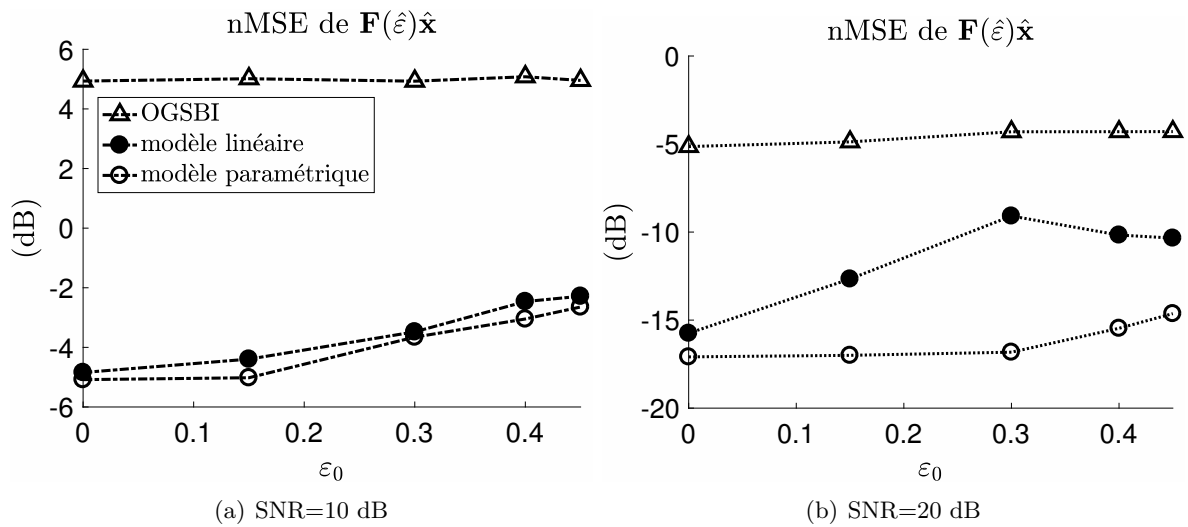


FIGURE 2.10 – nMSE de la scène reconstruite $\mathbf{F}(\hat{\varepsilon})\hat{\mathbf{x}}$ estimée par l’algorithme OGSBI (triangle), par l’algorithme proposé utilisant un modèle linéaire (ronds pleins) et par l’algorithme proposé utilisant un modèle paramétrique (cercles). Le scénario consiste en une cible avec un SNR égal à 10 dB (2.10(a)) ou 20 dB (2.10(b))

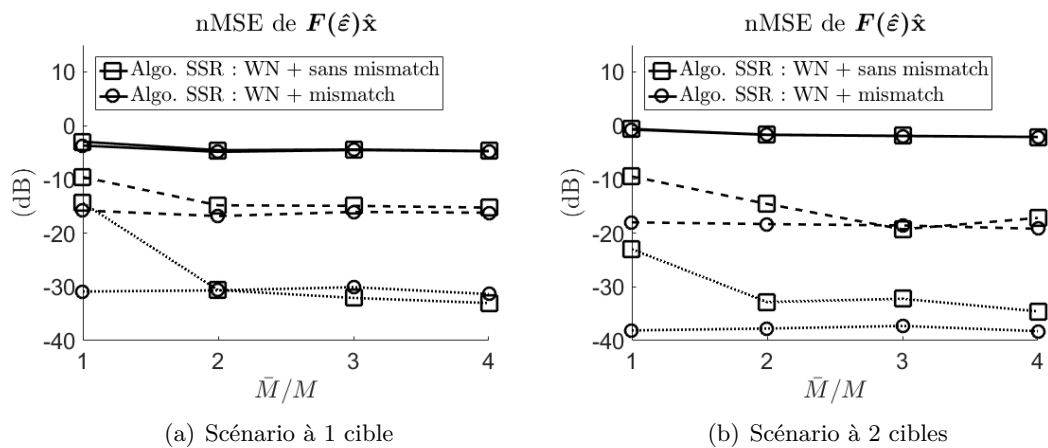


FIGURE 2.11 – Comparaison entre l’algorithme prenant en compte le désalignement (rond) et la version non-robustifiée (carré) de la MSE normalisée de la scène reconstruite $\mathbf{F}(\hat{\varepsilon})\hat{\mathbf{x}}$ 1D en fonction de la finesse de la grille d’analyse. SNR=10 dB (ligne pleine), 20 dB (tirets), 40 dB (pointillés). 2.11(a) Scénario à 1 cible ; 2.11(b) Scénario à 2 cibles ;

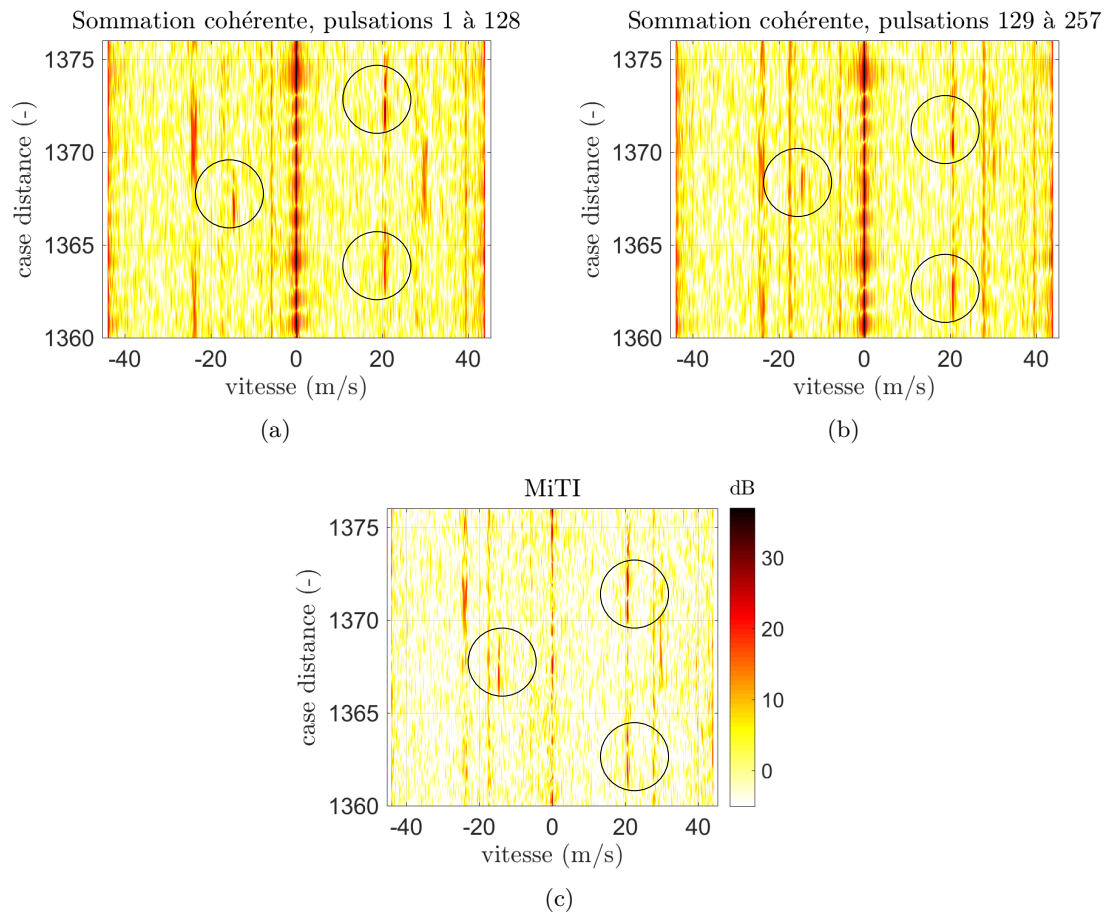


FIGURE 2.12 – Sommatation cohérente du jeu de données expérimentales considérées pour deux trains d’impulsions de longueur $M = 64$ consécutifs (Fig. 2.12(a) et 2.12(b) resp.). 2.12(c) MiTI utilisant ces mêmes données. Par convention, les cibles à vitesse positives s’approchent du radar ; celles à vitesse négative s’en éloignent donc. Les « vraies » cibles sont entourées en noir pour les distinguer des lobes secondaires.

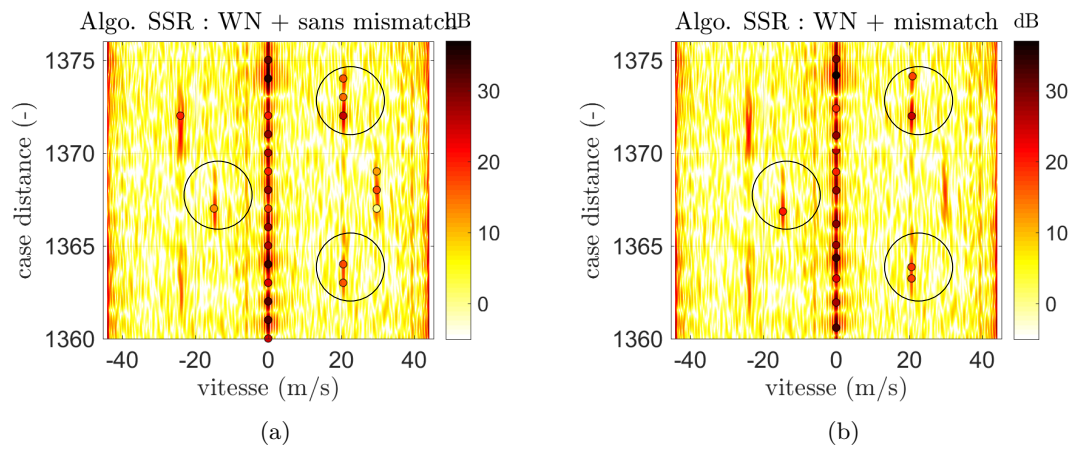


FIGURE 2.13 – Cartes distance-vitesse estimées par les différents algorithmes à partir des données PARSAX 2.13(a) Algorithme Bayésien de SSR non robustifié [24]; 2.13(b) Algorithme Bayésien de SSR prenant en compte le désalignement. Le fond représente la sommation cohérente large bande, les étoiles la vraie scène et les ronds la scène estimée. $K = 16$, $M = 64$, $n_{va} = 2$, $v_a = 45$ m/s, $f_c = 3.315$ GHz, $B = 100$ MHz, $\sigma^2 = 1$.

2.5 Conclusions

Nous avons donc présenté dans ce chapitre un algorithme prenant en compte le potentiel désalignement des cibles avec la grille d'analyse distance-vitesse utilisée dans l'expression du dictionnaire. Cet algorithme ajoute au modèle hiérarchique Bayésien original de [24] deux vecteurs qui modélisent le désalignement en distance et en vitesse. Un échantillonneur de Gibbs modifié est alors mis en place afin d'estimer les paramètres d'intérêt, à savoir le vecteur d'amplitude et les vecteurs de désalignement. L'échantillonnage des vecteurs de désalignement n'est pas une tâche anodine puisqu'il faut tirer des échantillons suivant une loi non classique. L'échantillonneur de Gibbs est alors modifié en conséquence.

Le nouvel algorithme estime correctement une scène radar complexe où les cibles sont désalignées par rapport à la grille d'analyse, avec en outre une cible cachée dans un lobe secondaire d'une autre cible. Il a même été testé avec succès sur des données expérimentales du radar PARSAX de TU Delft. La paramétrisation directe du dictionnaire, par opposition avec un modèle linéaire plus largement répandu, donne de meilleurs résultats malgré une complexité calculatoire plus importante. De plus, les performances de ce nouvel algorithme ne dépendent pas de la valeur du désalignement : la scène radar est correctement estimée même si la cible est en chevauchement total entre deux cases d'analyse (en distance ou en vitesse). Enfin, nous avons vu qu'il n'est pas nécessaire de sur-échantillonner la grille d'analyse pour améliorer les performances de l'algorithme prenant en compte le désalignement, ce qui est intéressant d'un point de vue calculatoire.

On rappelle que cette robustification a été développée sous hypothèse de bruit blanc Gaussien. Il s'agit donc de modifier l'algorithme pour prendre en compte un bruit coloré, et en particulier une composante diffuse de bruit ; c'est l'objet du chapitre suivant.

Algorithme Bayésien de SSR pour des cibles en chevauchement dans du fouillis diffus

Les algorithmes classiques de détection sont souvent mis à mal par la présence de fouillis (*clutter*), et en particulier par ses lobes de repliement qui peuvent masquer des cibles situées aux dites « vitesses aveugles ». Une littérature extensive est donc consacrée à la détection de cibles dans du fouillis, par exemple [91, 113, 94, 111]. En particulier, la méthode classique du *Generalized Likelihood Ratio Test* (GLRT) requiert l'estimation de la matrice de covariance du bruit afin de procéder à un « filtre adapté de blanchiment ». Celle-ci est en général estimée grâce à des données secondaires qui doivent être d'une taille suffisante pour permettre une bonne estimation, tout en conférant au fouillis les mêmes caractéristiques statistiques que dans les données principales. Les données secondaires sont notamment sensées être exemptes de cibles.

Dans le cadre de cette étude, on cherche dans un premier temps à estimer correctement la scène radar, sans mettre en place une procédure de détection à proprement parler. Une solution pour estimer des cibles dans du clutter diffus a été proposée en amont des travaux de thèse dans [21]; elle repose sur une simple modification du schéma hiérarchique Bayésien de l'algorithme non-robustifié de [24] et ne requiert pas de données secondaires. On présente donc dans un premier temps cette méthode, qui sera ensuite utilisée pour modifier l'algorithme robuste au désalignement présenté dans le chapitre précédent et dans [3, 1, 4, 6]. L'algorithme résultant, capable d'estimer des cibles en chevauchement sur la grille d'analyse masquées par du fouillis diffus, est testé sur des données synthétiques et expérimentales.

Sommaire

3.1	Estimation de cibles dans du fouillis diffus	70
3.1.1	Caractéristiques du fouillis de sol	70
3.1.2	Algorithme Bayésien de SSR pour fouillis diffus	72
3.1.3	Exemple de résultats	76
3.2	Modification de l'algorithme robuste au désalignement pour prendre en compte le fouillis diffus	77
3.2.1	Modification du modèle hiérarchique Bayésien	77
3.2.2	Modification du schéma d'estimation	78
3.3	Simulations numériques	82

3.3.1	Exemples de réalisation	82
3.3.2	Performances en fonction du SINR des cibles	83
3.4	Simulations sur données expérimentales	86
3.5	Conclusions	88

3.1 Estimation de cibles dans du fouillis diffus

Dans un premier temps, on rappelle certains résultats sur la caractérisation spectrale du fouillis de sol, en particulier pour introduire la notion de composante diffuse du fouillis. On présente ensuite une modification de l'algorithme de référence de [24] (présenté dans le chapitre 1) qui a pour but de robustifier ce dernier à la présence de fouillis, en particulier lorsque celui-ci possède une composante diffuse. Le modèle hiérarchique Bayésien de [21] présente un intérêt certain car il est facilement intégrable dans celui présenté dans le chapitre précédent, robustifié au désalignement par rapport à la grille d'analyse.

3.1.1 Caractéristiques du fouillis de sol

3.1.1.1 Modèle composite de fouillis

Dans un rapport de 1989 [36], H. C. Chan propose un modèle composite de fouillis, représenté schématiquement sur la Fig. 3.1¹. Ce modèle est déduit de nombreuses mesures de fouillis de sol, en bande S et bande X, sur des zones agricoles et forestières. D'après ce modèle, le spectre du fouillis de sol possède 3 composantes :

- une composante cohérente : cette composante occupe un domaine spectral réduit autour d'un Doppler nul ; sa largeur dépend de la longueur d'onde utilisée, de la vitesse de vent, de l'utilisation du sol, etc. Elle correspond à des objets de fabrication humaine (bâtiments) ou bien à des objets naturels de grande taille (sol, montagnes, troncs d'arbres).
- une composante diffuse lente : cette composante occupe un domaine spectral plus étendu, mais est d'amplitude réduite par rapport à la composante cohérente. Elle n'est localisée ni en temps ni en distance, c'est-à-dire qu'elle apparaît à tout moment de l'observation. Elle correspond à de la végétation, en particulier aux branches d'arbres.
- une composante diffuse rapide : cette dernière composante occupe un domaine spectral encore plus étendu, et possède une amplitude encore plus faible. Cette composante apparaît comme un bruit à bande limitée dans le spectre moyenné du fouillis mais il correspond à des composantes bande étroite qui apparaissent à des fréquences différentes pour des temps d'observation courts. Il peut traduire des feuilles d'arbres agitées par des rafales de vent, des champs cultivés dans une légère brise ou bien des oiseaux.

1. On parle aussi de modèle ICM ou modèle de Billingsley [25], (*Billingsley's Intrinsic Clutter Motion model*).

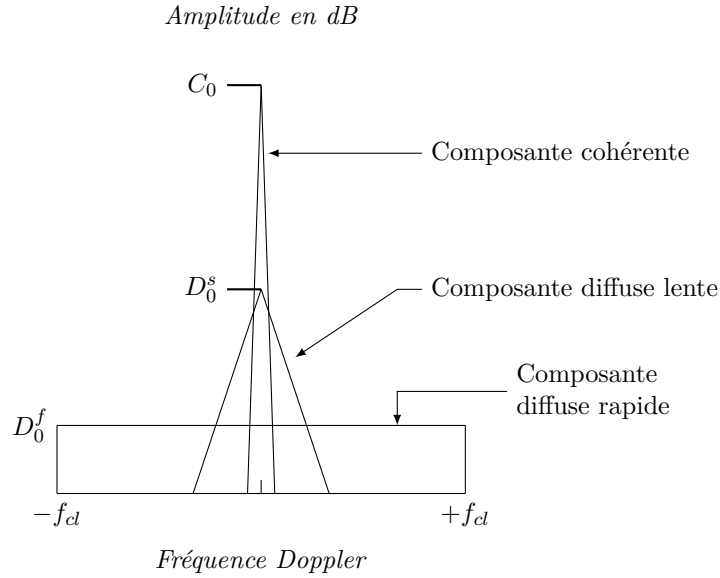


FIGURE 3.1 – Modèle de fouillis composite présenté par Chan [36] qui comporte 3 composantes : une composante cohérente, d'amplitude C_0 , une composante diffuse lente d'amplitude D_0^s , et une composante diffuse rapide d'amplitude D_0^f d'étalement fréquentiel $2f_{cl}$.

Ainsi, étant donné qu'elle est très localisée en Doppler (c'est-à-dire en vitesse), la composante cohérente du fouillis pourrait être interprétée comme une ou plusieurs cibles à vitesse nulle par un algorithme utilisant une hypothèse bruit blanc. En revanche, les composantes diffuses occupent des domaines spectraux plus larges, et ne sont pas prises en compte par un modèle bruit blanc comme celui utilisé dans le chapitre précédent pour les algorithmes [3, 1, 4, 6]. Il est donc nécessaire de modifier le modèle de signal utilisé dans [3, 1, 4, 6] pour prendre en compte de telles caractéristiques.

3.1.1.2 Distribution et densité spectrale du fouillis de sol

Comme dans les chapitres 1 et 2, on considère que le bruit \mathbf{n} suit une distribution complexe Gaussienne centrée :

$$\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}), \quad (3.1)$$

mais $\mathbf{R} \neq \sigma^2 \mathbf{I}_{KM}$ (on n'est plus sous hypothèse de bruit blanc). On dit également que le bruit suit une densité de Rayleigh, qui correspond à la distribution de l'amplitude $[\|\mathbf{n}\|_2]$.

Puisqu'on suppose le bruit \mathbf{n} stationnaire au sens large, on peut étudier sa densité spectrale de puissance $S(f)$ (*Power Spectral Density, PSD*) qui est définie comme la transformée de Fourier de l'autocorrélation. Avec l'approche proposée, $S(f)$ ne contient que la composante continue du spectre ; c'est le vecteur \mathbf{x} qui absorbe la composante discrète. Un modèle Gaussien pour la PSD du fouillis de sol a d'abord été proposé dans [15] puis souvent adopté,

en partie grâce à sa facilité d'implémentation. La PSD Gaussienne s'écrit :

$$S_{Gau}(f) = \frac{1}{\sqrt{2\pi\sigma_f^2}} \exp \left\{ -|f|^2 / (2\sigma_f^2) \right\} \quad (3.2)$$

où σ_f est l'étalement en Doppler du spectre, relié à l'étalement en vitesse σ_v via $\sigma_f = 2f_c\sigma_v/c$. En revanche, des mesures plus récentes de fouillis de sol dû au vent font preuve d'un étalement spectral plus large que le spectre Gaussien, en particulier à cause des branches et des feuilles des arbres agitées par le vent. Un modèle en « loi de puissance » a alors été proposé dans [53] :

$$S_{Pow}(f) = \frac{n \sin(\pi/n)}{2\pi f_c} \frac{1}{1 + (f/f_c)^n}. \quad (3.3)$$

où n est le paramètre de forme de la distribution. En fait, il a été montré dans [25] que le spectre du fouillis de sol dû au vent décroît bien plus rapidement qu'une loi de puissance ; cette décroissance est exponentielle :

$$S_{Exp}(f) = \frac{c\beta_\varepsilon}{4f_c} \exp \left\{ \frac{c\beta_\varepsilon}{2f_c} |f| \right\} \quad (3.4)$$

où β_ε est le paramètre de forme de la fonction exponentielle. On représente sur la Fig. 3.2 la densité spectrale de puissance et le SINR loss associé pour du fouillis avec une PSD Gaussienne, exponentielle, ou de type AR décrite dans la section 3.1.2.1. On fixe l'étalement spectral en vitesse $\sigma_v = 1$ m/s et le rapport entre le fouillis et le bruit thermique, CNR=20 dB. Ici, le SINR loss représente les pertes de SINR (défini dans (3.18)) en sortie de filtre adapté pour \mathbf{R} connu comparées au cas bruit blanc :

$$SINR_{loss}(f) \propto \frac{\mathbf{a}^H(f)\mathbf{R}^{-1}\mathbf{a}(f)}{\sigma^{-2}\mathbf{a}^H(f)\mathbf{a}(f)}. \quad (3.5)$$

On voit que le spectre exponentiel possède effectivement un étalement fréquentiel plus important, en particulier quand $v \rightarrow v_a/2$. Par conséquent, les pertes en SINR autour de $v_a/2$ sont plus importantes avec un spectre exponentiel que Gaussien.

On peut donc élaborer un schéma d'estimation à partir d'un modèle de signal utilisant par exemple un spectre Gaussien. Néanmoins, un tel modèle, plus « physique », peut être difficile à intégrer dans un modèle hiérarchique Bayésien comme celui de [24] présenté dans le chapitre précédent. L'idée est alors d'utiliser un modèle qui permette une intégration simple mais que l'on puisse relier aux modèles physiques, par exemple un modèle autoregressif de bruit.

3.1.2 Algorithme Bayésien de SSR pour fouillis diffus

L'algorithme de référence de [24], développé sous hypothèse de bruit blanc, a été robustifié au cas d'un fouillis ayant une composante diffuse [21]. Pour cela, un modèle autorégressif (AR) de bruit est adopté, associé à différentes hypothèses de stationnarité qui seront détaillées par

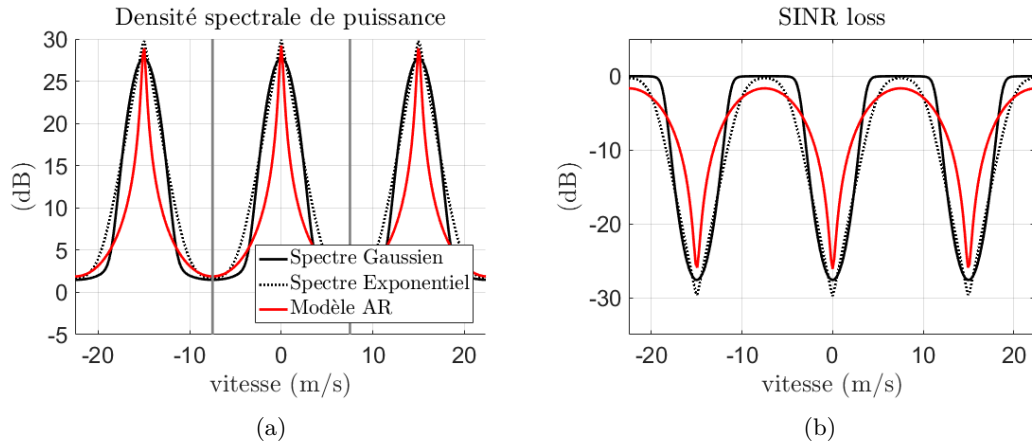


FIGURE 3.2 – Densité spectrale de puissance et SINR loss pour un spectre Gaussien (trait plein noir), exponentiel (trait pointillé noir) ou de type AR d’ordre 1 (rouge). $\sigma_v = 1$ m/s, CNR = 20dB. $f_c = 10$ GHz, $M = 128$, $v_a = 15$ m/s, $n_{va} = 3$; la migration n’est pas prise en compte dans cette représentation.

la suite.

3.1.2.1 Utilisation d’un modèle AR

Le modèle auto-régressif de bruit permet de formaliser une dépendance linéaire entre les différentes valeurs d’une variable au cours du temps (analyse de séries temporelles, par exemple en économie [81]). Ici, on applique le modèle AR au vecteur de bruit \mathbf{n} (sous certaines hypothèses) défini dans le contexte radar, afin d’estimer l’amplitude des cibles de la scène via le vecteur \mathbf{x} . Le modèle AR a également été utilisé à plusieurs reprises avec succès pour estimer directement des matrices de covariance [9, 115, 20].

Plusieurs hypothèses sont formulées dans [21] afin de prendre en compte la composante diffuse du vecteur de bruit \mathbf{n} :

1. le vecteur de bruit \mathbf{n} modélise à la fois le bruit thermique et le fouillis, il est supposé centré Gaussien : $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$;
2. il est décorrélé d’une sous-bande à l’autre ;
3. il est corrélé dans le temps long suivant un processus stationnaire AR d’ordre fini P ;
4. le fouillis est localement homogène, c’est-à-dire que les coefficients AR ne dépendent pas de la sous-bande considérée.

Ces hypothèses plutôt restrictives permettent de modifier simplement le modèle hiérarchique Bayésien « de base » et donnent malgré tout de bonnes performances d’estimation [21]. On peut donc écrire la relation suivante qui formalise le modèle AR d’ordre $P < M$ pour une sous-bande donnée k :

$$[\mathbf{n}_k]_m = \sum_{p=1}^P \phi_p [\mathbf{n}_k]_{m-p} + \eta_m \quad (3.6)$$

où \mathbf{n}_k désigne le vecteur bruit pour la sous-bande k , de taille M . Le paramètre η_m modélise l'innovation à l'instant m , il est défini tel que $\eta_m \sim \mathcal{CN}(0, \sigma_{AR}^2)$ et $\mathcal{E}\{\eta_m^* \eta_n\} = \sigma_{AR}^2 \delta_{m,n}$. Pour une sous-bande k donnée, on peut donc écrire

$$\mathbf{n}_k = \begin{bmatrix} 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \phi_1 & \ddots & & & & & \vdots \\ \vdots & \ddots & \ddots & & & & \vdots \\ \phi_P & & \ddots & \ddots & & & \vdots \\ 0 & \ddots & & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \phi_P & \dots & \phi_1 & 0 \end{bmatrix} \mathbf{n}_k + \boldsymbol{\eta} \triangleq \boldsymbol{\Phi} \mathbf{n}_k + \boldsymbol{\eta} \quad (3.7)$$

ce qui nous permet d'obtenir une expression de la matrice de covariance du vecteur \mathbf{n}_k

$$\boldsymbol{\Gamma} \triangleq \mathcal{E}\{\mathbf{n}_k \mathbf{n}_k^H\} = \sigma_{AR}^2 (\mathbf{I}_M - \boldsymbol{\Phi})^{-1} (\mathbf{I}_M - \boldsymbol{\Phi})^{-H}. \quad (3.8)$$

Étant donné que le fouillis est considéré localement homogène, on peut alors écrire

$$\mathbf{R} = \mathbf{I}_K \otimes \boldsymbol{\Gamma}. \quad (3.9)$$

On s'intéresse plus précisément à l'expression de \mathbf{R}^{-1} qui sera utilisée, entre autres, dans l'expression de la fonction de vraisemblance. On donne tout d'abord l'expression de $\boldsymbol{\Gamma}^{-1}$

$$\boldsymbol{\Gamma}^{-1} = \sigma_{AR}^{-2} (\mathbf{I}_M - \boldsymbol{\Phi})^H (\mathbf{I}_M - \boldsymbol{\Phi}) \quad (3.10)$$

qui est une matrice de taille $M \times M$ avec une bande de largeur P , puis on obtient

$$\begin{aligned} \mathbf{R}^{-1} &= \sigma_{AR}^{-2} \mathbf{I}_K \otimes [(\mathbf{I}_M - \boldsymbol{\Phi})^H (\mathbf{I}_M - \boldsymbol{\Phi})] \\ &= \sigma_{AR}^{-2} [\mathbf{I}_K \otimes (\mathbf{I}_M - \boldsymbol{\Phi})^H] [\mathbf{I}_K \otimes (\mathbf{I}_M - \boldsymbol{\Phi})]. \end{aligned}$$

Grâce à ce nouveau modèle de signal, on écrit la fonction de vraisemblance

$$\begin{aligned} f(\mathbf{y}|\mathbf{x}, \mathbf{R}) &= \frac{1}{\pi^{KM} |\mathbf{R}^{-1}|} \exp \left\{ -(\mathbf{y} - \mathbf{H}(\mathbf{0}, \mathbf{0})\mathbf{x})^H \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}(\mathbf{0}, \mathbf{0})\mathbf{x}) \right\} \\ &= \frac{1}{\pi^{KM} \sigma_{AR}^2} \exp \left\{ -\sigma_{AR}^{-2} \| [\mathbf{I}_K \otimes (\mathbf{I}_M - \boldsymbol{\Phi})] (\mathbf{y} - \mathbf{H}(\mathbf{0}, \mathbf{0})\mathbf{x}) \|^2 \right\} \end{aligned} \quad (3.11)$$

puisque $|\mathbf{R}^{-1}| = |\mathbf{I}_k|^M \times |\boldsymbol{\Gamma}|^K = (\sigma_{AR}^{-2})^{KM}$. On rappelle que l'algorithme de [21] considère que les cibles sont localisées sur la grille d'analyse distance-vitesse, soit $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) = (\mathbf{0}, \mathbf{0})$.

Densité spectrale de puissance pour un modèle AR La densité spectrale de puissance pour un modèle AR s'écrit :

$$S_{AR}(f, P) = \frac{\sigma_{AR}^2}{|1 - \sum_{p=1}^P \phi_p e^{-2i\pi fp}|^2}, \quad |\phi_p| < 1. \quad (3.12)$$

On considère dans les simulations numériques de la section 3.3 un modèle AR d'ordre $P = 1$. On peut donc chercher à calculer les valeurs de σ_{AR}^2 et ϕ_1 pour obtenir une certaine valeur maximale S_{max} à 0 m/s et minimale S_{min} à $v_a/2$. On a alors, en fonction de $\Delta_S = S_{max}/S_{min}$:

$$\phi_1 = \frac{-2(1 + \Delta_S) + 4\sqrt{\Delta_S}}{2(1 - \Delta_S)}$$

$$\sigma_{AR}^2 = S_{max} (1 - \phi_1)^2.$$

Cette méthodologie est utilisée dans la Fig. 3.2 pour obtenir un spectre AR avec les mêmes extrema que les spectres Gaussien et Exponentiel. On voit que le spectre AR possède alors un étalement spectral très restreint par rapport aux deux autres spectres. En outre, le SINR loss est non-nul à $v = v_a/2$ mais plutôt sélectif autour de 0 m/s et $\pm v_a$. Cette méthode sera utilisée dans les simulations numériques pour générer un fouillis de type AR avec des valeurs spectrales extrémales prédéfinies.

3.1.2.2 Modification de l'algorithme de référence

Modèle hiérarchique Bayésien Les paramètres de perturbation σ_{AR}^2 et $\phi = [\phi_1 \dots \phi_P]^T$ sont intégrés au modèle hiérarchique Bayésien de [24], pour donner le modèle résumé dans la Fig.3.3. Par souci de commodité calculatoire, des distributions a priori conjuguées sont choisies, à savoir une distribution inverse-gamma pour σ_{AR}^2 et une distribution complexe-Gaussienne pour ϕ

$$\sigma_{AR}^2 | \gamma_0, \gamma_1 \sim \mathcal{IG}(\gamma_0, \gamma_1)$$

$$\phi | \mu_\phi, \mathbf{R}_\phi \sim \mathcal{CN}_P(\mu_\phi, \mathbf{R}_\phi)$$

où (γ_0, γ_1) sont les paramètres de forme et d'échelle de $\pi(\sigma_{AR}^2 | \gamma_0, \gamma_1)$ et $(\mu_\phi, \mathbf{R}_\phi)$ sont le vecteur moyenne et la matrice de covariance de $\pi(\phi | \mu_\phi, \mathbf{R}_\phi)$. On se place dans le cas où l'utilisateur ne dispose pas d'information à propos de ces paramètres de perturbation, donc les hyperparamètres (γ_0, γ_1) et $(\mu_\phi, \mathbf{R}_\phi)$ sont ajustés pour obtenir des distributions a priori non-informatives

$$\pi(\sigma_{AR}^2) \propto \frac{1}{\sigma_{AR}^2} \mathbb{I}_{[0, +\infty)}(\sigma_{AR}^2)$$

$$\pi(\phi) \propto 1.$$

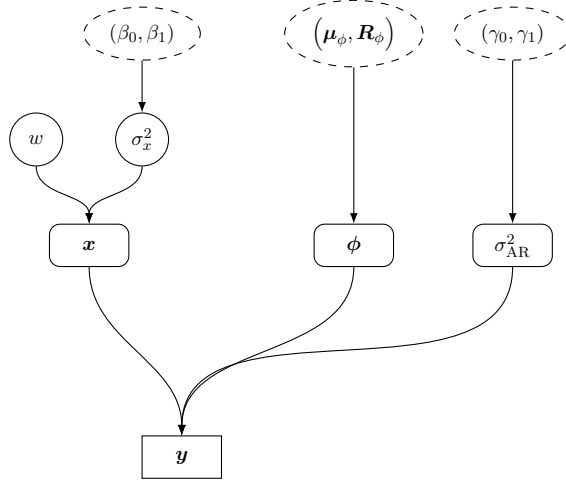


FIGURE 3.3 – Modèle hiérarchique Bayésien de [21] qui propose une robustification à la composante diffuse du fouillis via les paramètres supplémentaires σ_{AR}^2 et ϕ .

Schéma d'estimation L'échantillonneur de Gibbs rappelé dans le chapitre 1 est modifié pour prendre en compte les paramètres de perturbation σ_{AR}^2 et ϕ [21]. Plus précisément, l'échantillonnage des paramètres w et σ_x^2 est inchangé, mais celui du vecteur \mathbf{x} dépend maintenant des paramètres de perturbation à cause de la modification de la vraisemblance. Les paramètres de perturbation sont tirés suivant des lois usuelles : une inverse-Gamma pour σ_{AR}^2 et une Gaussienne complexe multivariée pour ϕ . Le calcul des distributions conditionnelles a posteriori sera détaillé dans la section 3.2.2 lorsque le désalignement est également pris en compte.

3.1.3 Exemple de résultats

On compare l'algorithme robustifié au fouillis diffus de [21] à la version non-robustifiée de [24] sur le scénario à 4 cibles utilisé précédemment pour les Fig. 2.6 et 1.10 ; les paramètres des cibles sont ceux résumés dans le tableau 2.1 mais leur désalignement en distance et en vitesse est nul. On génère pour cela un fouillis suivant un modèle AR simple d'ordre 1 : le fouillis est le même quelle que soit la case distance considérée ; les paramètres σ_{AR}^2 et ϕ_1 sont calculés pour obtenir une valeur maximale du spectre de 40 dB à $v = 0$ m/s et une valeur minimale de -3 dB à $v = v_a/2$, soit $\sigma_{AR}^2 = 1.98$, $\phi_1 = 0.99$.

On représente sur la Fig. 3.4 les cartes distance-vitesse estimées par les deux algorithmes. La Fig. 3.4(a) montre que l'algorithme non-robustifié de [24] ne parvient pas à assimiler la composante diffuse du fouillis, ce qui provoque de nombreuses fausses alarmes aux vitesses aveugles. En outre, les cibles 1 et 4 ne sont pas vues par l'algorithme. En revanche, l'algorithme prenant en compte la composante diffuse du fouillis estime correctement la scène, en termes de localisation et d'amplitude des cibles. Ces résultats convaincants ont motivé le choix du modèle (3.9), (3.10) pour modifier l'algorithme déjà robustifié au désalignement des cibles de [6] présenté dans le chapitre 2.

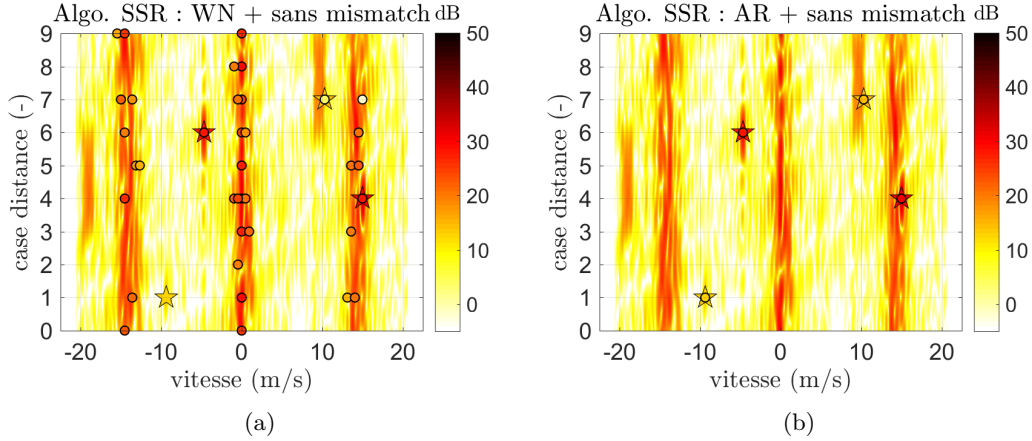


FIGURE 3.4 – Cartes distance-vitesse estimées par les différents algorithmes 3.4(a) Algorithme Bayésien de SSR non robustifié [24] ; 3.4(b) Algorithme Bayésien de SSR prenant en compte la composante diffuse du fouillis [21]. Le fond représente la sommation cohérente large bande, les étoiles la vraie scène et les ronds la scène estimée. $K = 10$, $M = 32$, $n_{va} = 3$, $v_a = 15$ m/s, $B/f_c=10\%$, $\sigma_{AR}^2 = 1.9767$, $\phi_1 = 0.9859$.

3.2 Modification de l’algorithme robuste au désalignement pour prendre en compte le fouillis diffus

Etant donnés les bons résultats obtenus avec l’algorithme de [21], ainsi que la facilité d’intégrabilité du schéma hiérarchique Bayésien associé, on décide de fusionner deux branches des modèles hiérarchiques de [6] (modélisation du désalignement) et [21] (modélisation du fouillis). On obtient alors un algorithme que l’on peut utiliser sur des données comprenant des cibles en chevauchement sur la grille d’analyse dans du fouillis possédant une composante diffuse [7]. Il est alors possible de traiter directement des données expérimentales en ne réglant que 2 hyperparamètres, ceux de σ_x^2 , car on choisit des distributions a priori non-informatives pour les paramètres de perturbation σ_{AR}^2 et ϕ .

3.2.1 Modification du modèle hiérarchique Bayésien

On représente dans la Fig. 3.5 le modèle hiérarchique Bayésien qui résulte de la fusion entre les modèles des références [6] (modélisation du désalignement, branche bleue) et [21] (modélisation du fouillis diffus, branche orange). La fonction de vraisemblance prend maintenant en compte à la fois la modélisation du fouillis (3.9), (3.10) et la paramétrisation du dictionnaire \mathbf{H} par les paramètres de désalignement $(\varepsilon^v, \varepsilon^r)$:

$$f(\mathbf{y}|\mathbf{x}, \varepsilon^v, \varepsilon^r, \sigma_{AR}^2, \phi) = \frac{1}{\pi^{KM} \sigma_{AR}^{2KM}} \exp \left\{ -\sigma_{AR}^2 \| [\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)] (\mathbf{y} - \mathbf{H}(\varepsilon^v, \varepsilon^r) \mathbf{x}) \|^2 \right\}. \quad (3.13)$$

Les distributions a priori assignées à chaque paramètre sont résumées dans le Tableau 3.1.

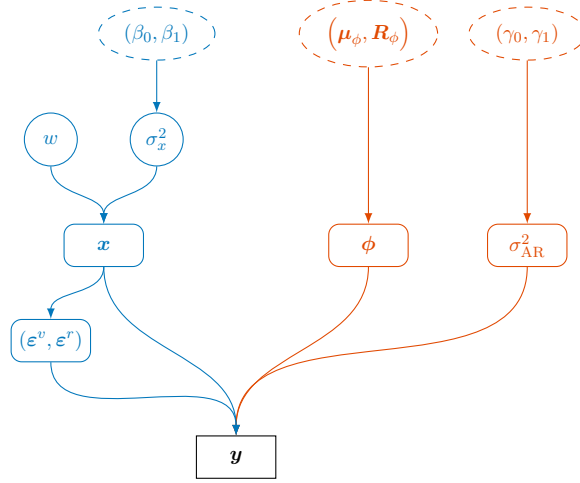


FIGURE 3.5 – Modèle hiérarchique Bayésien utilisé pour la robustification au désalignement (branche bleue) et au fouillis diffus (branche orange).

TABLE 3.1 – Distribution a priori assignée à chaque paramètre du modèle hiérarchique Bayésien utilisé pour la robustification au désalignement et au fouillis diffus.

Paramètre	Distribution a priori
$x_i w, \sigma_x^2$	$\text{BerCN}(w, 0, \sigma_x^2)$
$\varepsilon_i^v, \varepsilon_i^r x_i = 0$	$\delta(\varepsilon_i^v)\delta(\varepsilon_i^r)$
$\varepsilon_i^v, \varepsilon_i^r x_i \neq 0$	$\mathcal{U}_{[-0.5, 0.5] \times [-0.5, 0.5]}(\varepsilon_i^v, \varepsilon_i^r)$
$\sigma_x^2 \beta_0, \beta_1$	$\text{IG}(\beta_0, \beta_1)$
w	$\mathcal{U}_{[0, 1]}$
$\sigma_{AR}^2 \gamma_0, \gamma_1$	$\text{IG}(\gamma_0, \gamma_1)$
$\phi \mu_\phi, \mathbf{R}_\phi$	$\text{CN}_P(\mu_\phi, \mathbf{R}_\phi)$

On rappelle qu'on choisit en fait des distributions a priori non-informatives pour les paramètres de perturbation σ_{AR}^2 et ϕ :

$$\pi(\sigma_{AR}^2) \propto \frac{1}{\sigma_{AR}^2} \mathbb{I}_{[0, +\infty)}(\sigma_{AR}^2) \quad (3.14)$$

$$\pi(\phi) \propto 1. \quad (3.15)$$

3.2.2 Modification du schéma d'estimation

Comme dans le chapitre 2, on cherche à calculer les estimateurs MMSE des paramètres d'intérêt \mathbf{x} et $(\varepsilon^v, \varepsilon^r)$ grâce à un schéma de type MCMC. Ce schéma est cependant modifié par l'expression de la fonction de vraisemblance (3.13) qui implique de prendre en compte des

versions « blanchies » des paramètres :

$$\begin{aligned}\tilde{\mathbf{y}} &= [\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)] \mathbf{y} \\ \tilde{\mathbf{H}} &= [\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)] \mathbf{H} \\ \tilde{\mathbf{e}}_i &= \tilde{\mathbf{y}} - \sum_{\bar{l} \neq i} \tilde{\mathbf{h}}_{\bar{l}}(\varepsilon_{\bar{l}}^v, \varepsilon_{\bar{l}}^r) x_{\bar{l}}.\end{aligned}$$

On définit également les paramètres suivants :

$$\begin{aligned}\check{\mathbf{e}}_i &= [\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)]^H [\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)] \mathbf{e}_i \\ \mathbf{e}_i &= \mathbf{y} - \sum_{\bar{l} \neq i} \mathbf{h}_{\bar{l}}(\varepsilon_{\bar{l}}^v, \varepsilon_{\bar{l}}^r) x_{\bar{l}}.\end{aligned}$$

3.2.2.1 Échantillonnage des paramètres

Vecteur d'amplitude des cibles \mathbf{x} Le vecteur \mathbf{x} est échantillonné terme-à-terme suivant la distribution conditionnelle a posteriori $f(x_i | \mathbf{y}, \mathbf{x}_{-i}, \sigma_{AR}^2, \phi, \varepsilon^v, \varepsilon^r)$, que l'on calcule comme

$$\begin{aligned}f(x_i | \mathbf{y}, \mathbf{x}_{-i}, \sigma_{AR}^2, \phi, \varepsilon^v, \varepsilon^r) &\propto \exp \left\{ -\sigma_{AR}^{-2} \left\| [\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)] (\mathbf{y} - \mathbf{H}(\varepsilon^v, \varepsilon^r) \mathbf{x}) \right\|^2 \right\} \\ &\quad \times \left[(1-w) \delta(|x_i|) + w \frac{1}{\pi \sigma_x^2} \exp \left\{ -\sigma_x^{-2} |x_i|^2 \right\} \right] \\ &\propto \exp \left\{ -\sigma_{AR}^{-2} \left\| \tilde{\mathbf{y}} - \tilde{\mathbf{H}}(\varepsilon^v, \varepsilon^r) \mathbf{x} \right\|^2 \right\} \\ &\quad \times \left[(1-w) \delta(|x_i|) + w \frac{1}{\pi \sigma_x^2} \exp \left\{ -\sigma_x^{-2} |x_i|^2 \right\} \right] \\ &\propto (1-w) \delta(|x_i|) + w \frac{1}{\pi \sigma_x^2} \\ &\quad \exp \left\{ - \left(\sigma_x^{-2} + \sigma_{AR}^{-2} \left\| \tilde{\mathbf{h}}_i(\varepsilon_i^v, \varepsilon_i^r) \right\|^2 \right) |x_i|^2 + 2\sigma_{AR}^{-2} \operatorname{Re} \left\{ x_i^* \tilde{\mathbf{h}}_i(\varepsilon_i^v, \varepsilon_i^r)^H \tilde{\mathbf{e}}_i \right\} \right\}.\end{aligned}$$

Les éléments du vecteur d'amplitude $x_{\bar{i}}$ sont donc tirés suivant la loi Bernoulli-Gaussienne de paramètres $(w_{\bar{i}}, \mu_{\bar{i}}, \eta_{\bar{i}}^2)$:

$$\begin{aligned}x_{\bar{i}} | \mathbf{y}, \mathbf{x}_{-i}, \sigma_{AR}^2, \phi, \varepsilon^v, \varepsilon^r, w, \sigma_x^2 &\sim \operatorname{BerCN} \left(w_{\bar{i}}, \mu_{\bar{i}}, \eta_{\bar{i}}^2 \right) \\ \eta_{\bar{i}}^2 &= \left(\sigma_x^{-2} + \sigma_{AR}^{-2} \left\| \tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r) \right\|^2 \right)^{-1} \\ \mu_{\bar{i}} &= \frac{\sigma_{AR}^{-2} \eta_{\bar{i}}^2 \tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \tilde{\mathbf{e}}_{\bar{i}}}{w} \\ w_{\bar{i}} &= \frac{w}{(1-w) \sigma_x^2 \eta_{\bar{i}}^{-2} \exp \left\{ -\eta_{\bar{i}}^{-2} |\mu_{\bar{i}}|^2 \right\} + w}.\end{aligned}$$

Vecteurs de désalignement $(\varepsilon^v, \varepsilon^r)$ Les vecteurs de désalignement sont également tirés terme-à-terme, suivant la loi jointe conditionnelle a posteriori $f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \varepsilon_{-i}^v, \varepsilon_{-i}^r, \mathbf{x}, \sigma_{AR}^2, \phi; x_i)$.

On omet l'indicatrice $\mathbb{I}_{[-0.5,0.5] \times [-0.5,0.5]}(\varepsilon_i^v, \varepsilon_i^r)$ afin d'alléger les expressions :

$$\begin{aligned}
f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \varepsilon_{-\bar{i}}^v, \varepsilon_{-\bar{i}}^r, \mathbf{x}, \sigma_{AR}^2, \boldsymbol{\phi}; x_{\bar{i}} \neq 0) &\propto \exp \left\{ -\sigma_{AR}^{-2} \|\tilde{\mathbf{y}} - \tilde{\mathbf{H}}(\varepsilon^v, \varepsilon^r) \mathbf{x}\|^2 \right\} \mathbb{I}_{[-0.5,0.5] \times [-0.5,0.5]}(\varepsilon_i^v, \varepsilon_i^r) \\
&\propto \exp \left\{ -\sigma_{AR}^{-2} \|\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_i^v, \varepsilon_i^r)\|^2 |x_{\bar{i}}|^2 + 2\sigma_{AR}^{-2} \operatorname{Re} \left\{ x_{\bar{i}}^* \tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_i^v, \varepsilon_i^r)^H \check{\mathbf{e}}_{\bar{i}} \right\} \right\} \\
&\propto \exp \left\{ -\sigma_{AR}^{-2} \|\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_i^v, \varepsilon_i^r)\|^2 |x_{\bar{i}}|^2 \right\} \\
&\quad \times \exp \left\{ 2\sigma_{AR}^{-2} \operatorname{Re} \left\{ x_{\bar{i}}^* \mathbf{h}_{\bar{i}}(\varepsilon_i^v, \varepsilon_i^r)^H [\mathbf{I}_K \otimes (\mathbf{I}_M - \boldsymbol{\Phi})]^H \check{\mathbf{e}}_{\bar{i}} \right\} \right\} \\
&\propto \exp \left\{ -\sigma_{AR}^{-2} \|\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_i^v, \varepsilon_i^r)\|^2 |x_{\bar{i}}|^2 \right\} \\
&\quad \times \exp \left\{ 2\sigma_{AR}^{-2} \operatorname{Re} \left\{ x_{\bar{i}}^* \mathbf{h}_{\bar{i}}(\varepsilon_i^v, \varepsilon_i^r)^H \check{\mathbf{e}}_{\bar{i}} \right\} \right\}
\end{aligned}$$

On remplace $\|\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_i^v, \varepsilon_i^r)\|^2$ par son expression (calculée dans l'annexe B) car elle dépend de ε_i^v :

$$\begin{aligned}
f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \varepsilon_{-\bar{i}}^v, \varepsilon_{-\bar{i}}^r, \mathbf{x}, \sigma_{AR}^2, \boldsymbol{\phi}; x_{\bar{i}} \neq 0) &\propto \exp \left\{ -2\sigma_{AR}^{-2} |x_{\bar{i}}|^2 \operatorname{Re} \left\{ \sum_{k=0}^{K-1} \sum_{p=1}^P \gamma_{k,p} \exp \left\{ j2\pi p(1 + \mu k) n_{va} \frac{\varepsilon_i^v}{M} \right\} \right\} \right\} \\
&\quad \times \exp \left\{ 2\sigma_{AR}^{-2} \operatorname{Re} \left\{ x_{\bar{i}}^* \sum_{i=0}^{KM-1} [\mathbf{h}_{\bar{i}}(\varepsilon_i^v, \varepsilon_i^r)]_i^* [\check{\mathbf{e}}_{\bar{i}}]_i \right\} \right\} \\
&\propto \exp \left\{ \operatorname{Re} \left\{ \underbrace{\sum_{k=0}^{K-1} \sum_{p=1}^P \left(-2\sigma_{AR}^{-2} |x_{\bar{i}}|^2 \right) \gamma_{k,p} \exp \left\{ j2\pi p(1 + \mu k) n_{va} \frac{\varepsilon_i^v}{M} \right\}}_{T_1} \right\} \right. \\
&\quad \left. + \underbrace{\sum_{k=0}^{K-1} \sum_{m=0}^{M-1} [\mathbf{z}_{\bar{i}}]_i \exp \left\{ j2\pi k \frac{\varepsilon_i^r}{K} \right\} \exp \left\{ -j2\pi m(1 + \mu k) n_{va} \frac{\varepsilon_i^v}{M} \right\}}_{T_0} \right\}
\end{aligned}$$

avec $\mathbf{z}_{\bar{i}} \triangleq 2\sigma_{AR}^{-2} x_{\bar{i}}^* \mathbf{h}_{\bar{i}}(\mathbf{0}, \mathbf{0})^* \odot \check{\mathbf{e}}_{\bar{i}}$. Cette distribution n'appartient pas à une classe de distributions connue, donc l'échantillonneur de Gibbs doit être modifié. Contrairement au cas bruit blanc, il semble difficile de trouver une loi de proposition adéquate pour mettre en place une étape de MH. On tire donc directement les échantillons $\varepsilon_i^v, \varepsilon_i^r$ suivant la loi conditionnelle a posteriori $f(\varepsilon_i^v, \varepsilon_i^r | \mathbf{y}, \varepsilon_{-\bar{i}}^v, \varepsilon_{-\bar{i}}^r, \mathbf{x}, \sigma_{AR}^2, \boldsymbol{\phi}; x_{\bar{i}} \neq 0)$ calculée sur le domaine $[-0.5, 0.5] \times [-0.5, 0.5]$ arbitrairement discrétisé. On remarque que le terme T_1 , qui ne dépend que de ε_i^v peut être directement calculé ; on reconnaît dans T_0 la sommation cohérente large bande de $\mathbf{z}_{\bar{i}}$ calculée aux fréquences $(f_{dCI}, f_{rCI}) = (n_{va} \varepsilon_i^v / \bar{M}, \varepsilon_i^r / \bar{K})$ (rappelé dans l'équation (1.15)).²

2. Cette sommation large bande est restreinte au domaine $(\varepsilon_i^v, \varepsilon_i^r) \in [-.5, .5]^2$, c'est-à-dire $-1/(2\bar{K}) \leq f_{rCI} < 1/(2\bar{K})$ et $-n_{va}/(2\bar{M}) \leq f_{dCI} < n_{va}/(2\bar{M})$.

Paramètres de perturbation σ_{AR}^2 et ϕ On calcule la distribution conditionnelle a posteriori de σ_{AR}^2 dans le cas général où $\sigma_{AR}^2 | \gamma_0, \gamma_1 \sim \mathcal{IG}(\gamma_0, \gamma_1)$ avec (γ_0, γ_1) quelconques :

$$\begin{aligned} f(\sigma_{AR}^2 | \mathbf{y}, \mathbf{x}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) &\propto \sigma_{AR}^{-2KM} \exp \left\{ -\sigma_{AR}^{-2} \|\tilde{\mathbf{y}} - \tilde{\mathbf{H}}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x}\|^2 \right\} \\ &\quad \times \sigma_{AR}^{-2(\gamma_0+1)} \exp \left\{ -\gamma_1 \sigma_{AR}^{-2} \right\} \mathbb{I}_{[0,+\infty)}(\sigma_{AR}^2) \\ &\propto \sigma_{AR}^{-2(KM+\gamma_0+1)} \exp \left\{ -\sigma_{AR}^{-2} \left(\|\tilde{\mathbf{y}} - \tilde{\mathbf{H}}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x}\|^2 + \gamma_1 \right) \right\} \mathbb{I}_{[0,+\infty)}(\sigma_{AR}^2). \end{aligned}$$

Ainsi, σ_{AR}^2 est tiré suivant une loi inverse-Gamma

$$\sigma_{AR}^2 | \mathbf{y}, \mathbf{x}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r \sim \mathcal{IG} \left(KM + \gamma_0, \|\tilde{\mathbf{y}} - \tilde{\mathbf{H}}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x}\|^2 + \gamma_1 \right). \quad (3.16)$$

Ce résultat est semblable au cas bruit blanc, sauf que ce sont des versions blanchies des données \mathbf{y} et du dictionnaire \mathbf{H} qui sont utilisées.

On calcule ensuite la distribution conditionnelle a posteriori du paramètre de perturbation ϕ en se plaçant directement dans le cas d'une distribution a priori non-informative :

$$f(\phi | \mathbf{y}, \mathbf{x}, \sigma_{AR}^2, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \propto \exp \left\{ -\sigma_{AR}^{-2} \|\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)\| (\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x})\|^2 \right\} \times 1.$$

On se focalise d'abord sur le calcul de $\|\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)\| (\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \mathbf{x})\|^2 \triangleq \|\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)\| \mathbf{s}\|^2$, et en particulier sa dépendance en ϕ

$$\begin{aligned} \|\mathbf{I}_K \otimes (\mathbf{I}_M - \Phi)\| \mathbf{s}\|^2 &= \sum_{k=0}^{K-1} \|(\mathbf{I}_M - \Phi) \mathbf{s}_k\|^2 \\ &= \|\mathbf{s}\|^2 + \sum_{k=0}^{K-1} \left(-2 \operatorname{Re} \left\{ \mathbf{s}_k^H \Phi^H \mathbf{s}_k \right\} + \mathbf{s}_k^H \Phi^H \Phi \mathbf{s}_k \right) \\ &= \|\mathbf{s}\|^2 + \sum_{k=0}^{K-1} \left(-2 \operatorname{Re} \left\{ \phi^H \mathcal{V}_k \right\} + \phi^H \mathbf{M}_k \phi \right) \\ &= \|\mathbf{s}\|^2 - 2 \operatorname{Re} \left\{ \phi^H \underbrace{\sum_{k=0}^{K-1} \mathcal{V}_k}_{\mathcal{V}} \right\} + \phi^H \underbrace{\sum_{k=0}^{K-1} \mathbf{M}_k}_{\mathcal{M}} \phi \\ &\triangleq \|\mathbf{s}\|^2 - 2 \operatorname{Re} \left\{ \phi^H \mathcal{V} \right\} + \phi^H \mathcal{M} \phi \end{aligned}$$

où $\mathbf{s}_k \triangleq \left[[\mathbf{s}]_{kM} \quad \dots \quad [\mathbf{s}]_{kM+M-1} \right]^T$ est un sous-vecteur de \mathbf{s} qui contient M éléments ; $\mathcal{V}_k = \left[\mathbf{s}_k^H \mathbf{L}_1^H \mathbf{s}_k \quad \dots \quad \mathbf{s}_k^H \mathbf{L}_P^H \mathbf{s}_k \right]^T$ et $[\mathbf{M}_k]_{p,q} = \mathbf{s}_k^H \mathbf{L}_p^H \mathbf{L}_q \mathbf{s}_k$. \mathbf{L}_p est une matrice dont la p -ième sous-diagonale contient des 1 (les autres éléments étant nuls). On obtient alors la distribution

conditionnelle a posteriori (*pourvu que \mathcal{M} soit inversible*)

$$\begin{aligned} f(\phi|\mathbf{y}, \mathbf{x}, \sigma_{AR}^2, \varepsilon^v, \varepsilon^r) &\propto \exp \left\{ -\sigma_{AR}^{-2} \left(-2 \operatorname{Re} \left\{ \phi^H \mathcal{V} \right\} + \phi^H \mathcal{M} \phi \right) \right\} \\ &\propto \exp \left\{ - \left[\phi - \mathcal{M}^{-1} \mathcal{V} \right]^H \sigma_{AR}^{-2} \mathcal{M} \left[\phi - \mathcal{M}^{-1} \mathcal{V} \right] \right\} \end{aligned}$$

et on tire donc ϕ suivant une distribution complexe-Gaussienne P -variée de moyenne $\mathcal{M}^{-1} \mathcal{V}$ et de matrice de covariance $\sigma_{AR}^2 \mathcal{M}^{-1}$:

$$\phi|\mathbf{y}, \mathbf{x}, \sigma_{AR}^2, \varepsilon^v, \varepsilon^r \sim \mathcal{CN}_P \left(\mathcal{M}^{-1} \mathcal{V}, \sigma_{AR}^2 \mathcal{M}^{-1} \right). \quad (3.17)$$

3.2.2.2 Échantillonnage des hyperparamètres

Connaissant la fonction de vraisemblance (3.13) et les distributions a priori des hyperparamètres w et σ_x^2 , on calcule leur distribution conditionnelle a posteriori respectives :

$$\begin{aligned} \sigma_x^2|\mathbf{y}, \mathbf{x}, \sigma_{AR}^2, \phi, \varepsilon^v, \varepsilon^r &\sim \mathcal{IG} \left(n_1 + \beta_0, \|\mathbf{x}\|^2 + \beta_1 \right) \\ w|\mathbf{y}, \mathbf{x}, \sigma_{AR}^2, \phi, \varepsilon^v, \varepsilon^r &\sim \mathcal{Be} (1 + n_1, 1 + n_0) \end{aligned}$$

qui sont donc inchangées par rapport au cas bruit blanc détaillé dans le chapitre 2.

3.3 Simulations numériques

On présente dans cette section les résultats obtenus sur données synthétiques par l'algorithme robustifié à la fois au désalignement par rapport à la grille d'analyse et à la composante diffuse du fouillis. Les données synthétiques sont générées en ajoutant des cibles à un bruit synthétique de type AR d'ordre 1 qui est le même quelle que soit la case distance considérée (sous-section 3.3.1.1) ou bien qui change d'une case distance à l'autre (sous-section 3.3.1.2). Les paramètres de perturbation σ_{AR}^2 et ϕ_1 sont calculés pour obtenir une valeur maximale du spectre arbitraire, notée S_{max} , à $v = 0$ m/s et une valeur minimale S_{min} à $v = v_a/2$.

3.3.1 Exemples de réalisation

3.3.1.1 Fouillis identique quelle que soit la case distance

On considère dans cette partie le scénario à 4 cibles utilisé précédemment ; les caractéristiques des cibles sont détaillées dans le Tableau 2.1. Cette fois-ci la cible 2, située à $+v_a$, est effectivement masquée par les lobes secondaires du fouillis. On génère du fouillis suivant un modèle AR d'ordre 1 qui est le même quelle que soit la case distance considérée ; les paramètres $\sigma_{AR}^2 = 1.98$ et $\phi_1 = 0.99$ correspondent à $S_{max} = 40$ dB à $v = 0$ m/s et

$S_{min} = -3$ dB à $v = v_a/2$. On représente sur la Fig. 3.6 les cartes distance-vitesse estimées par l'algorithme non-robustifié de [24] (dénommé « Algo. SSR : WN+sans mismatch »), par la version ne prenant en compte que le désalignement des cibles de [6] (dénommé « Algo. SSR : WN+mismatch »), par la version ne prenant en compte que la composante diffuse du fouillis de [21] (dénommé « Algo. SSR : AR+sans mismatch »), et enfin par l'algorithme entièrement robustifié qui prend en compte à la fois le désalignement des cibles et la composante diffuse du fouillis [7] (dénommé « Algo. SSR : AR+mismatch »).

Comme dans la Fig. 3.4, l'algorithme ne prenant en compte que la composante diffuse du fouillis (Fig. 3.6(c)) n'estime pas correctement la scène puisque les cibles sont partagées sur les cases distance-vitesse adjacentes (cibles 3 et 4). L'algorithme ne prenant en compte que le désalignement des cibles par rapport à la grille d'analyse (Fig. 3.6(b)) provoque lui de nombreuses fausses-alarmes dans les lobes secondaires du fouillis, et de la cible 3. Comme attendu, l'algorithme de référence non-robustifié (Fig. 3.6(a)) n'est pas capable d'estimer les cibles et provoque de nombreuses fausses-alarmes à la fois dans les lobes secondaires du fouillis et dans ceux des cibles (cibles 3 et 4). Enfin, l'algorithme prenant en compte à la fois le désalignement des cibles et la composante diffuse du fouillis (Fig. 3.6(d)) estime correctement la scène en termes de position et d'amplitude des cibles : les ambiguïtés vitesse sont correctement levées. De plus, si on reconstruit le spectre du fouillis à partir des valeurs estimées de σ_{AR}^2 et ϕ , on obtient une indication de la manière dont l'algorithme interprète le fouillis. On voit sur les Fig. 3.6(e) et 3.6(f) que la composante diffuse du fouillis est bien prise en compte par les algorithmes utilisant le modèle AR.

Pour finir, on représente sur la Fig. 3.7 les histogrammes des échantillons tirés par l'algorithme robuste « AR + mismatch ». On voit que malgré la complexité de l'algorithme, qui nécessite l'échantillonnage de 7 variables aléatoires, les distributions a posteriori empiriques sont piquées autour de la valeur estimée.

3.3.1.2 Fouillis dépendant de la case distance

On génère maintenant un fouillis de type AR d'ordre 1 qui dépend de la case distance considérée. La valeur maximale du spectre est comprise entre 35 et 45 dB, et sa valeur minimale entre -5 et 5 dB. On voit sur la Fig. 3.8 que l'algorithme robustifié décrit dans ce chapitre estime correctement la position et l'amplitude des cibles, en levant les ambiguïtés vitesse. La composante diffuse du fouillis est correctement assimilée, malgré les variations de case distance à case distance.

3.3.2 Performances en fonction du SINR des cibles

On procède enfin à des simulations de Monte-Carlo pour affiner la comparaison entre l'algorithme ne prenant en compte que le désalignement [6] et celui robustifié à la composante diffuse du bruit et au désalignement [7]. On utilise pour cela 3 grandeurs : la MSE normalisée de la scène estimée $\mathbf{H}(\hat{\boldsymbol{\varepsilon}}^v, \hat{\boldsymbol{\varepsilon}}^r)\hat{\mathbf{x}}$, celle de chaque élément $\hat{x}_{\bar{i}}$ et celle sur la zone autour de $+v_a$

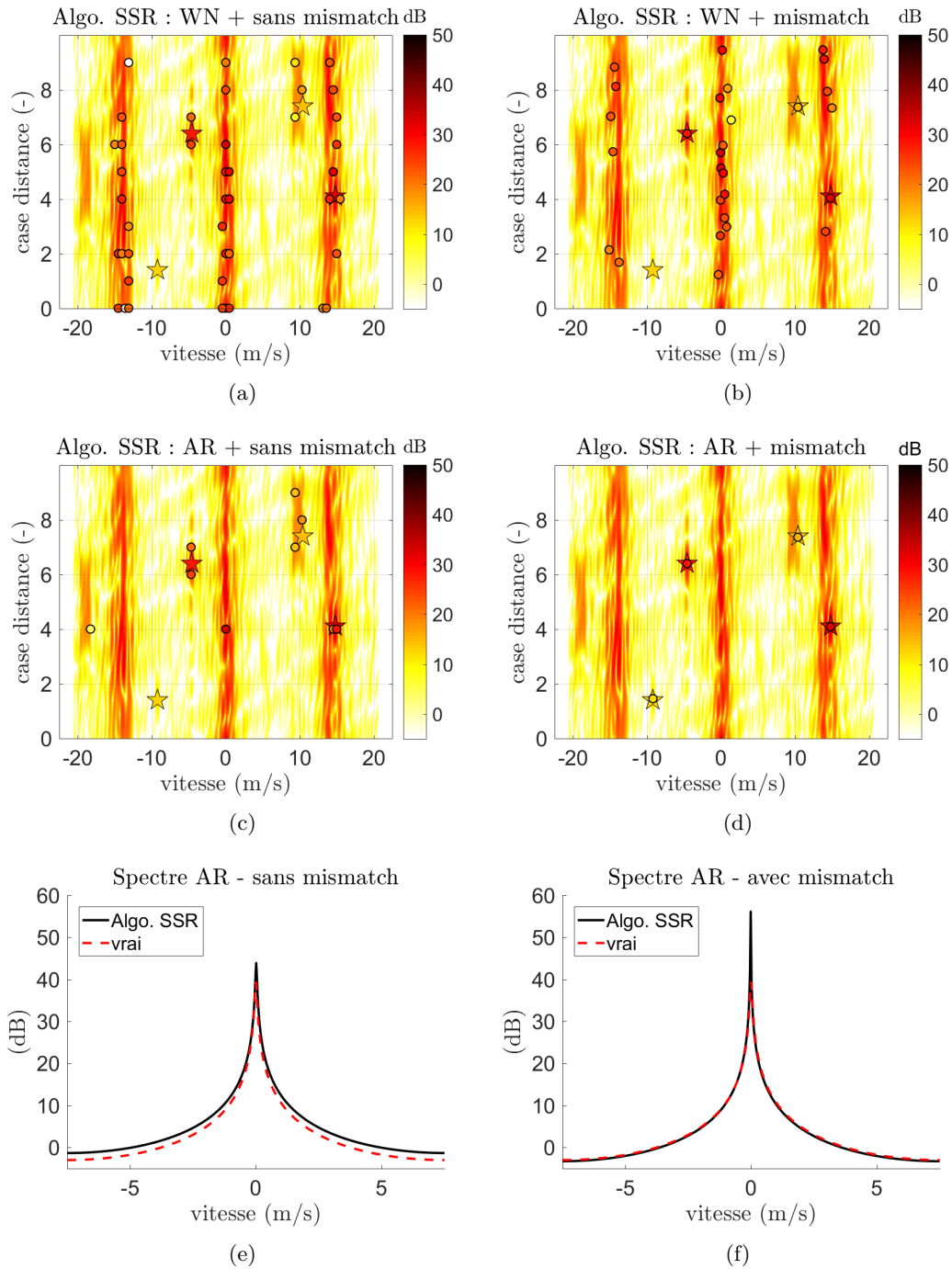


FIGURE 3.6 – Cartes distance-vitesse estimées par les différents algorithmes 3.6(a) Algorithme Bayésien de SSR non robustifié [24] ; 3.6(b) Algorithme Bayésien de SSR prenant en compte le désalignement des cibles par rapport à la grille d’analyse. 3.6(c) Algorithme Bayésien de SSR prenant en compte la composante diffuse du fouillis [24]. 3.6(c) Algorithme Bayésien de SSR prenant en compte le désalignement des cibles et la composante diffuse du fouillis. Le fond représente la sommation cohérente large bande, les étoiles la vraie scène et les ronds la scène estimée. $K = 10$, $M = 32$, $n_{va} = 3$, $v_a = 15$ m/s, $B/f_c=10\%$, $\sigma_{AR}^2 = 1.98$, $\phi_1 = 0.99$, $P = 1$.

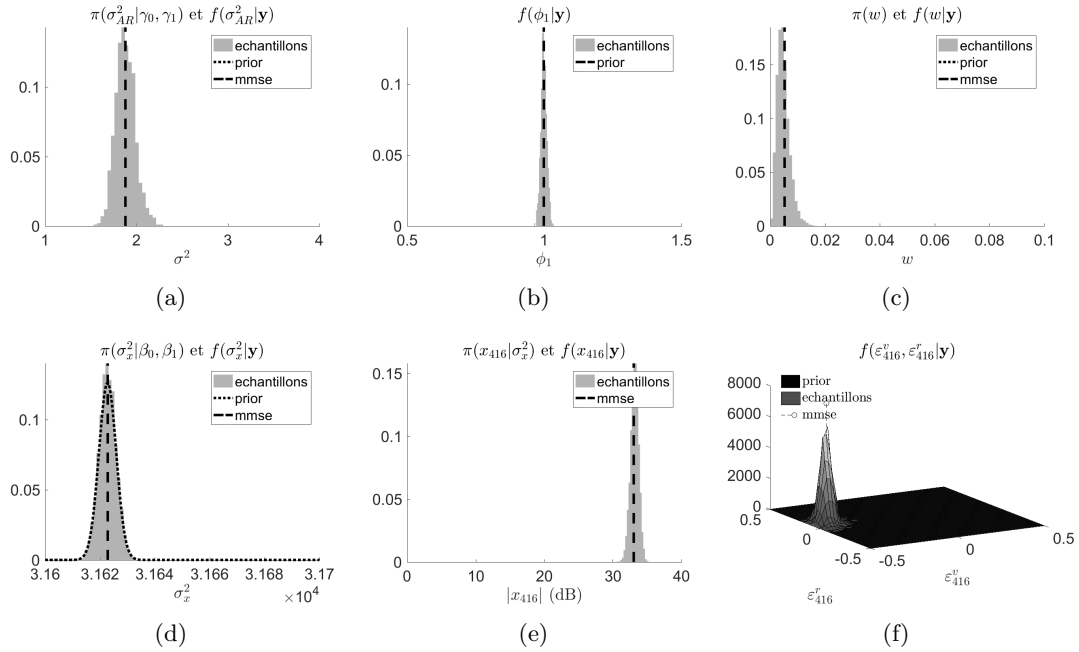


FIGURE 3.7 – Distributions a posteriori empiriques construites à partir des échantillons tirés des paramètres σ_{AR}^2 , ϕ_1 , w et σ_x^2 , ainsi que ceux de l'élément des vecteurs \mathbf{x} et (ϵ^v, ϵ^r) correspondant à la cible 2. La ligne verticale pointillée représente l'estimateur MMSE de chaque paramètre; la courbe pointillée la distribution a priori.

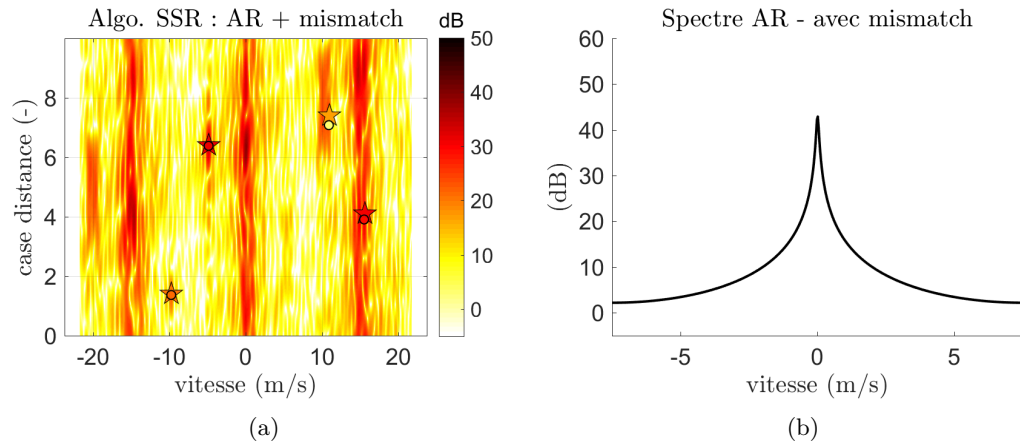


FIGURE 3.8 – Carte distance-vitesse estimée par l'algorithme Bayésien de SSR prenant en compte le désalignement des cibles et la composante diffuse du fouillis. Le fond représente la sommation cohérente large bande, les étoiles les vraies cibles et les ronds la scène estimée. Le fouillis généré de manière synthétique dépend de la case distance considérée (S_{max} compris entre 35 et 45 dB, S_{min} compris entre -5 et 5 dB, $P = 1$). $K = 10$, $M = 32$, $n_{va} = 3$, $v_a = 15$ m/s, $B/f_c = 10\%$, $P = 1$.

(là où se trouve la cible 2) $\left(\mathbf{h}_{\bar{i}}(\hat{\varepsilon}_{\bar{i}}^v, \hat{\varepsilon}_{\bar{i}}^r)x_{\bar{i}}\right)_{v_a^+}$. Cette dernière métrique se focalise sur la zone des vitesses aveugles positives qui est sujette à l'apparition de fausses alarmes. On utilise le scénario à 4 cibles décrit précédemment, en faisant varier à tour de rôle le SINR des cibles 1 (exoclutter), 2 (endoclutter) et 4 (dans le lobe secondaire de la cible 3). N'étant plus sous hypothèse bruit blanc, on définit l'amplitude d'une cible α_t en termes de SINR, *Signal to Interference and Noise Ratio* :

$$\text{SINR} = \mathcal{E} \left\{ |\alpha_t|^2 \right\} \mathbf{a}_t^H \mathbf{R}^{-1} \mathbf{a}_t \quad (3.18)$$

où \mathbf{a}_t désigne le vecteur directionnel associé à la cible.

Les résultats sont résumés dans la Fig. 3.9. On remarque tout d'abord que quelle que soit la métrique considérée, les performances de l'algorithme entièrement robustifié (« AR + mismatch ») sont bien meilleures que celles de l'algorithme ne prenant en compte que le désalignement (« WN + mismatch »), ce à quoi on pouvait s'attendre vu les résultats des figures 3.6 et 3.8. Plus précisément, l'estimation de la scène reconstruite $\mathbf{H}(\hat{\varepsilon}^v, \hat{\varepsilon}^r)\hat{\mathbf{x}}$ ne semble pas dépendre de la valeur du SINR des cibles 1 et 4, pour les deux algorithmes. En revanche, si on augmente le SINR de la cible 2, située à $+v_a$, alors l'estimation de la scène reconstruite $\mathbf{H}(\hat{\varepsilon}^v, \hat{\varepsilon}^r)\hat{\mathbf{x}}$ s'améliore (Fig. 3.9(d)), en particulier dans la zone de vitesses aveugles positives (Fig. 3.9(e)). L'estimation de la cible elle-même s'améliore aussi considérablement (surtout pour l'algorithme « AR + mismatch », Fig. 3.9(f)). Si on se focalise sur l'augmentation du SINR de la cible 1, on voit que l'estimation de la scène reconstruite est bien meilleure avec l'algorithme « AR + mismatch » (Fig. 3.9(a)), alors qu'il n'y a pas de différence marquée entre les deux algorithmes en termes d'estimation de la cible elle-même (Fig. 3.9(c)), puisqu'elle se situe en-dehors des lobes secondaires du fouillis. En ce qui concerne la cible 4, on remarque les mêmes tendances mais l'écart entre les deux algorithmes en termes d'estimation de la cible elle-même se creuse quand on augmente son SINR. En fait, avec un modèle bruit blanc de nombreuses fausses alarmes apparaissent (en particulier au niveau des vitesses aveugles) ce qui fait augmenter le plancher thermique. La cible 4 apparaît alors encore plus faible qu'elle ne l'est vraiment, et n'est pas vue par l'algorithme « WN+mismatch ».

En conclusion, l'algorithme entièrement robustifié permet d'obtenir des performances d'estimation très satisfaisantes. En outre, c'est l'estimation des cibles situées aux vitesses aveugles ou masquées par une cible plus forte qui bénéficie le plus de la prise en compte d'un modèle de fouillis diffus par rapport à une modèle bruit blanc.

3.4 Simulations sur données expérimentales

Pour finir, on teste l'algorithme proposé prenant en compte le désalignement des cibles et la composante diffuse du fouillis sur des données expérimentales du PARSAX [70]. Ces données sont obtenues en ajoutant les 4 cibles décrites dans le Tableau 2.1 à du fouillis expérimental correspondant à une zone boisée au sud de Delft. Ces données semi-expérimentales ont été construites car le fouillis présent dans les jeux de données avec cibles (correspondant à une

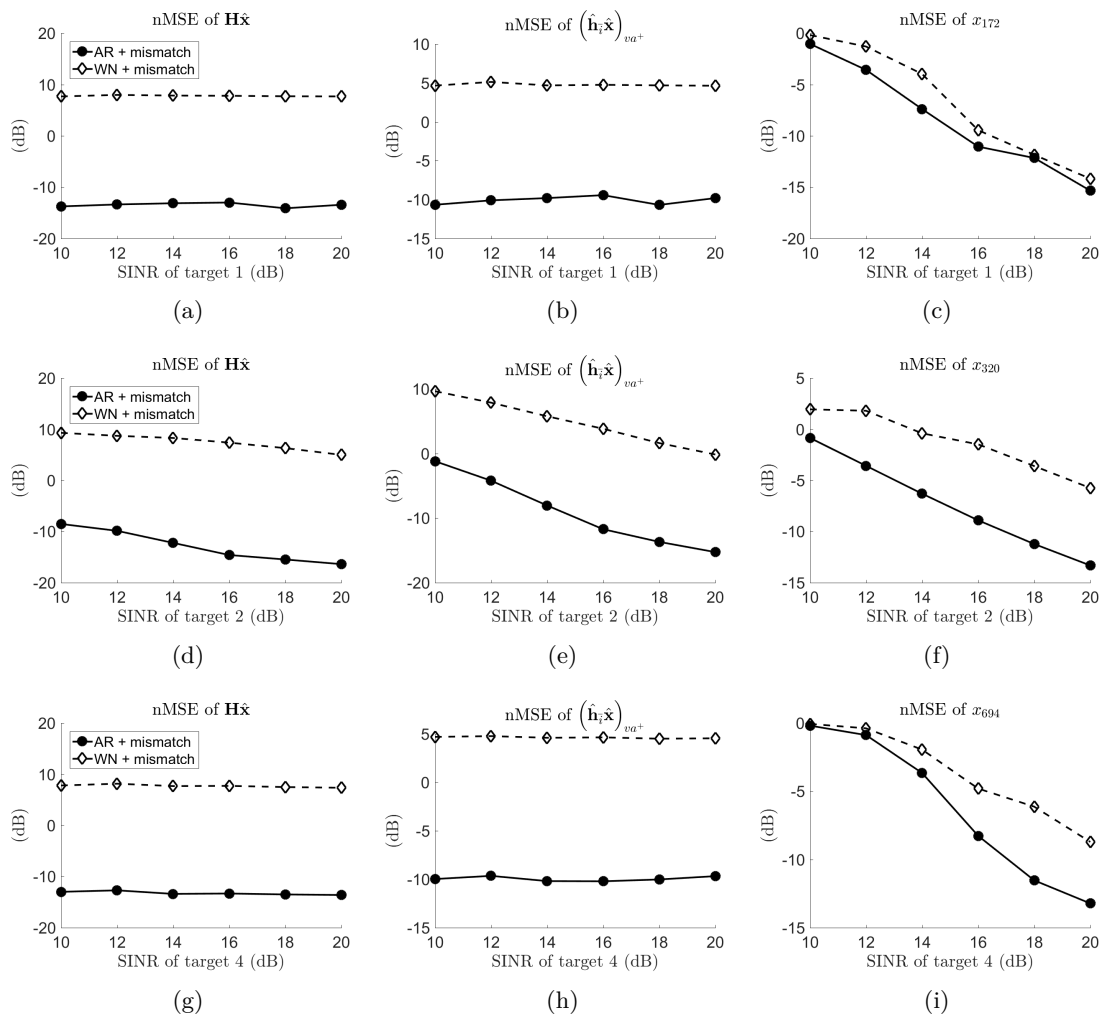


FIGURE 3.9 – Comparaison entre les performances de l’algorithme ne prenant en compte que le désalignement (« WN + mismatch », tirets) et celui prenant en compte le désalignement et le fouillis diffus (« AR + mismatch », ligne pleine). Pour la première ligne on fait varier le SINR de la cible 1 (Fig. 3.9(a), 3.9(b), 3.9(c)), celui de la cible 2 pour la deuxième ligne (Fig. 3.9(d), 3.9(e), 3.9(f)), et celui de la cible 4 pour la troisième ligne (Fig. 3.9(g), 3.9(h), 3.9(i)).

autoroute) n'est pas très fort et ne comporte pas de composante diffuse. L'idée était alors de viser une zone de fouillis potentiellement diffus (une zone boisée) et d'ajouter des cibles coopératives a posteriori.

On voit sur la Fig. 3.10 que seul l'algorithme prenant en compte à la fois le désalignement des cibles et la composante diffuse du fouillis est capable d'estimer correctement la scène. Les algorithmes ne prenant pas en compte le désalignement sont sujets à de nombreuses fausses alarmes dans les lobes secondaires des cibles et du fouillis (Fig. 3.10(a), 3.10(a)). L'algorithme ne prenant en compte que le désalignement des cibles souffre lui de quelques fausses alarmes dans les lobes secondaires du fouillis (Fig. 3.10(b)). Il est également intéressant de voir que les algorithmes utilisant le modèle AR interprètent ce fouillis expérimental comme la somme entre une composante diffuse (visible sur les spectres AR estimés, Fig. 3.6(e), 3.6(f)) et des éléments discrets à vitesse nulle visibles sur les cartes distance-vitesse (Fig. 3.10(c), 3.10(d)), ce qui correspond bien au modèle de Chan de la Fig. 3.1).

3.5 Conclusions

Dans ce chapitre, nous avons donc présenté une robustification de l'algorithme de référence de [24] à la fois au désalignement des cibles par rapport à la grille d'analyse, et à la composante diffuse du fouillis de sol. Pour cela, le schéma de robustification au désalignement des cibles présenté dans le chapitre précédent et dans [3, 1, 4, 6] est fusionné avec le schéma de robustification à la composante diffuse du fouillis présenté dans [21]. Ce schéma utilise un modèle autoregressif de bruit d'ordre fini appliqué à chaque sous-bande ; il implique la prise en compte de deux nouveaux paramètres de perturbation dans le schéma hiérarchique Bayésien et la procédure d'estimation MCMC. L'échantillonneur de Gibbs utilisé pour estimer les paramètres d'intérêt est alors modifié, en particulier l'estimation des paramètres de désalignement et du vecteur d'amplitude. Le nouvel algorithme entièrement robustifié estime correctement une scène radar comportant des cibles désalignées par rapport à la grille d'analyse dans du clutter diffus. En particulier, l'une des cibles masquée par un lobe secondaire du fouillis est correctement estimée. Nous avons vu que l'utilisation du modèle AR permet d'améliorer considérablement les performances de l'algorithme (par rapport à l'algorithme bruit blanc du chapitre précédent) et que l'estimation de la scène globale ne dépend que du SINR de la cible cachée dans un lobe secondaire du fouillis (pas du SINR des autres cibles). Enfin, l'algorithme a été testé avec succès sur des données semi-expérimentales obtenues à partir de fouillis mesuré par le radar PARSAX. En particulier, les ambiguïtés vitesse sont correctement levées et l'algorithme est capable d'estimer des cibles situées aux vitesses aveugles.

Nous avons rappelé que l'utilisateur de l'algorithme robustifié doit régler deux hyperparamètres, qui ajustent la distribution a priori de la puissance des cibles. En fait, ce réglage peut considérablement influencer les performances de l'algorithme sur des scénarios à forte dynamique de puissance où des cibles fortes cohabitent avec des cibles faibles. L'objet du chapitre suivant est de proposer une nouvelle distribution a priori pour le vecteur d'amplitude afin d'estimer de telles scènes.

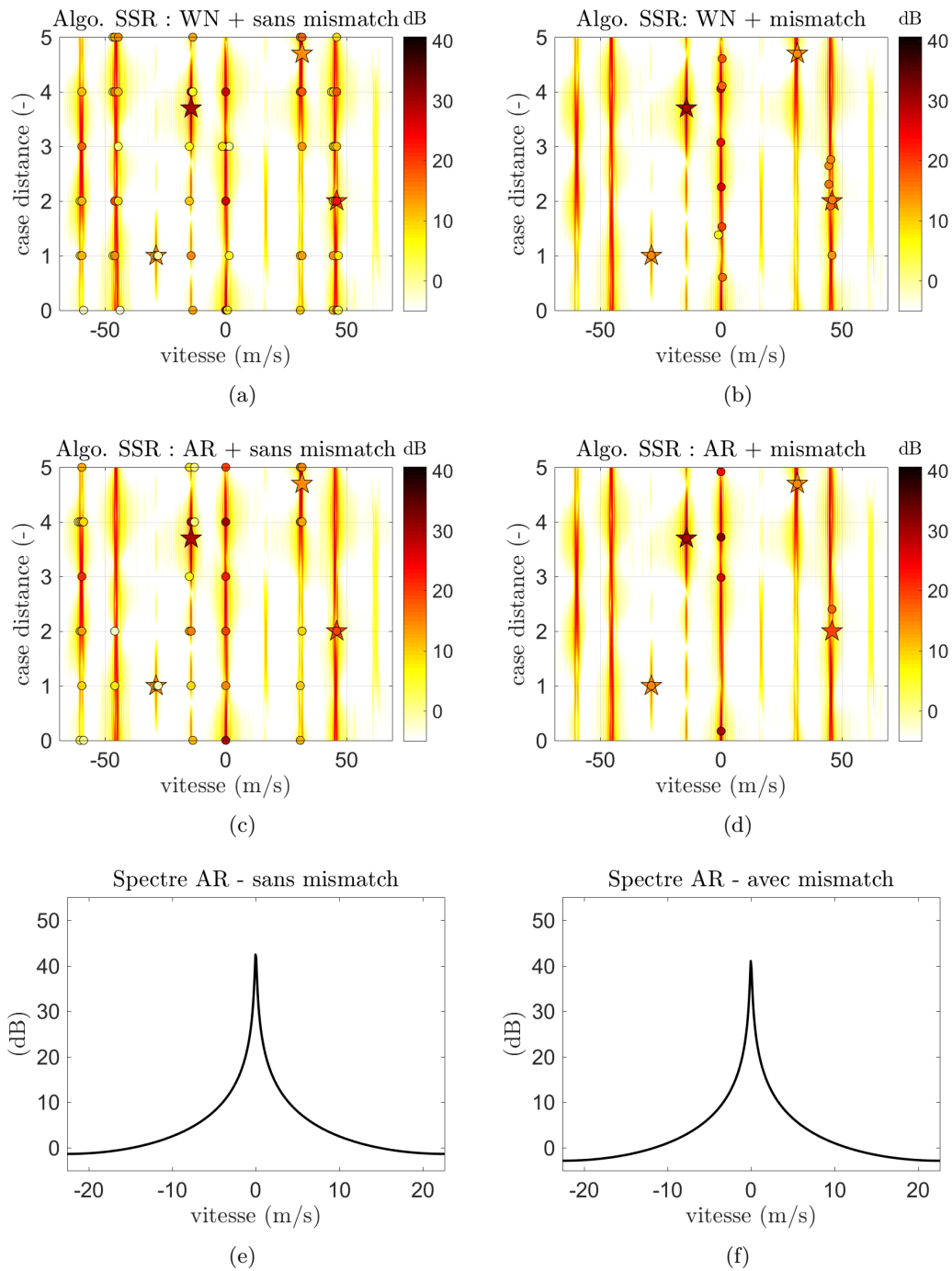


FIGURE 3.10 – Cartes distance-vitesse estimées par les différents algorithmes 3.6(a) Algorithme Bayésien de SSR non robustifié [24]; 3.6(b) Algorithme Bayésien de SSR prenant en compte le désalignement des cibles par rapport à la grille d’analyse. 3.6(c) Algorithme Bayésien de SSR prenant en compte la composante diffuse du fouillis [24]. 3.6(c) Algorithme Bayésien de SSR prenant en compte le désalignement des cibles et la composante diffuse du fouillis. Le fond représente la sommation cohérente large bande, les étoiles les vraies cibles et les ronds la scène estimée. $K = 6$, $M = 64$, $F_c = 3.315$ GHz, $B = 100$ MHz, $n_{va} = 3$, $v_a = 45$ m/s.

Estimation de scènes radar à forte dynamique de puissance

Nous avons vu dans les chapitres 2 et 3 que l'utilisation de l'algorithme prenant en compte le désalignement des cibles de [3, 1, 4, 6], ou bien de l'algorithme entièrement robustifié de [7] nécessite le réglage de deux hyperparamètres seulement, (β_0, β_1) , qui ajustent la distribution a priori de la puissance des cibles σ_x^2 . A priori, l'utilisateur ne dispose pas d'information sur ce paramètre, mais on peut envisager de procéder au préalable à une sommation cohérente large bande rapide [22] afin d'estimer le niveau moyen de puissance des cibles et ajuster les hyperparamètres (β_0, β_1) en conséquence. Malgré ce réglage, les scènes à forte dynamique de puissance ne sont pas correctement estimées par l'algorithme de [6], comme en témoigne la Fig. 1.10(c).

On propose dans ce chapitre une nouvelle distribution a priori pour le vecteur \mathbf{x} (qui contient l'information d'amplitude et de phase des cibles) qui repose sur l'utilisation de classes de puissance. Cette nouvelle distribution suppose un nouveau modèle hiérarchique Bayésien, associé à un nouveau schéma d'estimation, qui ont été présentés dans [5, 2]. Ce nouvel algorithme est testé sur des données synthétiques et expérimentales. Plusieurs simulations de Monte-Carlo sur données synthétiques permettent en outre d'évaluer ses performances sur des scénarios complexes à forte dynamique.

Sommaire

4.1	Mise en évidence du problème	92
4.2	Nouvelle distribution a priori et modèle hiérarchique Bayésien associé	93
4.2.1	Paramétrisation du vecteur d'amplitude des cibles \mathbf{x}	94
4.2.2	Distribution a priori des autres paramètres	95
4.2.3	Distribution a priori des hyperparamètres	95
4.3	Nouveau schéma d'estimation	96
4.3.1	Échantillonnage du vecteur amplitude $\boldsymbol{\rho}$	97
4.3.2	Échantillonnage du vecteur phase $\boldsymbol{\phi}$	99
4.3.3	Échantillonnage des autres paramètres et hyperparamètres	99
4.3.4	Implémentation de l'algorithme	100
4.4	Exemples de réalisations de l'algorithme avec classes	101
4.4.1	Données synthétiques	102
4.4.2	Données semi-expérimentales	103
4.5	Simulations de Monte-Carlo sur données synthétiques	104

4.5.1	Intérêt des classes de puissance	106
4.5.2	Choix des classes de puissance	108
4.5.3	Robustesse du réglage des classes par rapport au niveau de bruit	109
4.6	Conclusions	109

4.1 Mise en évidence du problème

Avec le développement des techniques de furtivité, les algorithmes de détection doivent être capables de détecter des cibles faibles, dépassant de quelques décibels le niveau de bruit. En outre, ces cibles faibles peuvent être masquées par des cibles plus fortes, particulièrement si elles sont situées dans leurs lobes secondaires. On peut par exemple imaginer qu'un drone puisse être masqué par un avion de ligne, ou bien une petite embarcation par un gros navire. Un tel scénario est donc doublement contraignant puisqu'il s'agit de concevoir un algorithme capable à la fois d'estimer une cible qui ne sort pas notablement du bruit, mais également capable d'encaisser une forte dynamique de puissance. L'objectif de ce chapitre est de pousser les performances des algorithmes proposés pour estimer des scènes avec plus de 40 dB de dynamique (typiquement une cible forte avec un SNR de 60 dB et une faible avec un SNR inférieur à 20 dB).

Nous avons vu que les algorithmes de représentation parcimonieuse présentent un intérêt particulier lorsqu'il s'agit d'estimer de telles scènes. En particulier, l'algorithme CLEAN décrit dans le chapitre 2 permet en théorie d'encaisser une forte dynamique de puissance et de détecter une cible faible. Néanmoins, nous avons vu qu'il est susceptible de provoquer des fausses alarmes [2], en particulier lorsque le nombre de cibles est inconnu. Une approche similaire au CLEAN a été proposée dans [104] : il s'agit de détecter des cibles faibles masquées par des cibles fortes, tout en prenant en compte la migration de ces cibles ; l'algorithme utilise pour y parvenir la transformée de Keystone associée à un détecteur CFAR. L'écart maximal entre les cibles est néanmoins d'environ 20 dB, ce qui correspond finalement aux performances de l'algorithme de référence de [24]. Mentionnons également l'utilisation de la transformée de Radon-Fourier [116] et de la transformée de Hough [35] pour la détection de cibles faibles (et fluctuantes), sans que ne soit pourtant mentionnée la compétition avec des cibles fortes.

Dans le cadre de ces travaux de thèse, on utilise les principes de l'estimation Bayésienne. Le but est donc de proposer un nouveau schéma hiérarchique Bayésien, ou bien de modifier en partie celui utilisé dans les chapitres précédents. En particulier, l'utilisation d'une autre distribution a priori pour forcer la parcimonie du vecteur \mathbf{x} pourrait améliorer les performances de l'algorithme sur des scénarios à forte dynamique de puissance. Nous avons vu dans la section 1.3 que différents types de distributions a priori peuvent forcer la parcimonie. En particulier, les priors hiérarchiques continus tels que le prior de Laplace [57], de Student-t [106] ou double-Pareto généralisé [14] sont généralement utilisés dans le cadre des représentations parcimonieuses. L'inconvénient de ces distributions est qu'elles utilisent un seul paramètre pour régler à la fois le niveau de parcimonie de la scène à estimer et la puissance des cibles. Ce

réglage limité est d'autant plus contraignant que la dynamique de puissance est importante : si on souhaite favoriser l'estimation de cibles fortes, via une loi de mélange de moyenne élevée, alors on augmente la variance de la distribution (Laplace ou Student-t par exemple). Celle-ci devient alors moins piquée, ce qui revient à favoriser les cibles faibles (mais pas forcément nulles) et peut contrecarrer la parcimonie du vecteur \mathbf{x} .

Nous avons vu que les distributions mixtes, comme la distribution Bernoulli-Gaussienne (1.56) utilisée dans [24] et les algorithmes robustifiés correspondants des chapitres précédents, permet de découpler le réglage de la parcimonie et de la puissance des cibles. Néanmoins, la structure mixte de cette distribution a priori ne suffit pas pour estimer des scènes à forte dynamique de puissance, comme en témoigne la Fig. 1.10(c). Pour faire écho aux développements présentés par la suite, on s'intéresse à la distribution de $\rho_i \triangleq |x_i|$. Si on considère que w est uniformément réparti sur $[0, 1]$, on obtient

$$\begin{aligned} \pi(\rho_i | \beta_0, \beta_1) &= \int \int \pi(\rho_i | \sigma_x^2) \pi(\sigma_x^2) \pi(w) d\sigma_x^2 dw \\ &= \frac{1}{2} \delta(\rho_i) + \rho_i \frac{\beta_0}{\beta_1} \left[1 + \frac{\rho_i^2}{\beta_1} \right]^{-(\beta_0+1)}. \end{aligned}$$

On représente sur la Fig. 4.1 la composante continue de cette distribution pour différents réglages de la loi de mélange $\pi(\sigma_x^2 | \beta_0, \beta_1)$: une loi informative centrée sur les faibles puissances (« lpower ») ou les fortes puissances (« hpower »), ou bien une loi non-informative (« flat »). Notons que la loi non-informative semblerait la plus appropriée lorsque l'utilisateur ne dispose pas d'information sur la puissance des cibles présentes sur la scène. On voit qu'aucun des trois réglages proposé ne permet de balayer de manière satisfaisante l'intégralité de l'intervalle de puissance. On comprend alors pourquoi la distribution a priori Bernoulli-Gaussienne de x_i , bien que découplant le réglage de la parcimonie et de la puissance des cibles, ne permet pas d'encaisser une forte dynamique de puissance. L'utilisation d'un prior mixte semble donc intéressante dans le cadre des SSR, mais la composante continue du prior doit être capable d'encaisser une très forte dynamique afin d'être utilisée sur les scénarios complexes considérés dans ce chapitre.

4.2 Nouvelle distribution a priori et modèle hiérarchique Bayésien associé

L'approche utilisée ici pour estimer des scènes à forte dynamique est d'abord de paramétrer le vecteur \mathbf{x} en coordonnées polaires : $\mathbf{x} \triangleq \boldsymbol{\rho} \odot e^{i\phi}$. L'idée est ensuite d'utiliser une nouvelle distribution a priori pour le vecteur $\boldsymbol{\rho}$ en utilisant des classes de puissance afin d'estimer des cibles avec des amplitudes très différentes¹.

Le modèle hiérarchique proposé dans [5, 2] et décrit dans ce chapitre peut être utilisé pour des applications variées. Ici, on se place dans le cadre de l'estimation de cibles migrantes

1. La notation ϕ est réutilisée, mais désigne $\angle \mathbf{x}$ dans ce chapitre.

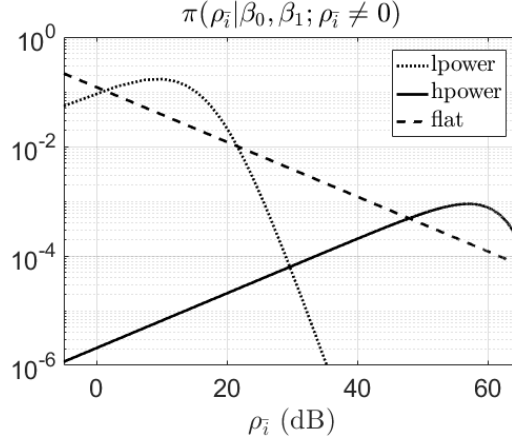


FIGURE 4.1 – Composante continue de la distribution a priori de $\rho_{\bar{i}} = |x_{\bar{i}}|$ lorsque le prior de $x_{\bar{i}}$ est « Bernoulli-Student-t hiérarchique », pour différents réglages de la loi de mélange : « flat » loi non-informative, « lpower » loi informative centrée sur les faibles puissances : $(m_{\sigma_x^2}, \sqrt{\text{var}_{\sigma_x^2}}) = (15, 15)$ dB, « hpower » loi informative centrée sur les fortes puissances : $(m_{\sigma_x^2}, \sqrt{\text{var}_{\sigma_x^2}}) = (60, 15)$ dB.

potentiellement désalignées avec la grille d'analyse 2D. On utilise donc le modèle de signal (2.9) qui s'écrit

$$\mathbf{y} = \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}_r)\mathbf{x} + \mathbf{n}. \quad (4.1)$$

On considère que le bruit est blanc Gaussien, soit $\mathbf{n}|\sigma^2 \sim \mathcal{CN}_M(\mathbf{0}, \sigma^2 \mathbf{I}_{KM})$, ce qui donne l'expression de la fonction de vraisemblance

$$f(\mathbf{y}|\mathbf{x}, \sigma^2) = \frac{1}{\pi^M \sigma^{2M}} \exp \left\{ -\frac{\|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}_r)\mathbf{x}\|_2^2}{\sigma^2} \right\}. \quad (4.2)$$

4.2.1 Paramétrisation du vecteur d'amplitude des cibles \mathbf{x}

Comme expliqué précédemment, il s'agit de définir une nouvelle distribution a priori pour le paramètre d'amplitude $\boldsymbol{\rho}$. Pour cela, on utilise d'abord le fait que l'amplitude des cibles est techniquement limitée par la plage de fonctionnement du récepteur [77, chap.11]. L'intervalle de puissance est ensuite divisé en plusieurs classes, au sein de cette plage de fonctionnement.

On rappelle que l'on définit $\boldsymbol{\rho}, \boldsymbol{\phi}$ comme $\mathbf{x} = \boldsymbol{\rho} \odot e^{i\boldsymbol{\phi}}$, soit $\forall \bar{i} \in \{0, \dots, \bar{K}\bar{M}-1\}, x_{\bar{i}} = \rho_{\bar{i}} e^{i\phi_{\bar{i}}}$. Les éléments de $\boldsymbol{\rho}$ et $\boldsymbol{\phi}$ sont a priori i.i.d. On définit alors la nouvelle distribution a priori du vecteur d'amplitude $\boldsymbol{\rho}$ qui consiste en un atome à zero et un mélange de distributions uniformes

$$\pi(\rho_{\bar{i}}|\mathbf{w}) = w_0 \delta(\rho_{\bar{i}}) + \sum_{c=1}^C w_c k_c \mathbb{I}_{[\rho_c^-, \rho_c^+]}(\rho_{\bar{i}}) \quad (4.3)$$

où $\mathbf{w} = [w_c]_{c=0, \dots, C}$ et le vecteur de probabilité d'appartenance à une classe, et $w_0 = 1 - \sum_{c=1}^C w_c$. k_c est la constante de mise à l'échelle que l'on calcule facilement pour la distribution

uniforme comme $k_c = 1/(\rho_c^+ - \rho_c^-)$. La nouvelle distribution (4.3) est interprétée de la manière suivante : l'élément $\rho_{\bar{i}}$ est nul avec une probabilité w_0 ; sinon, il appartient à la classe de puissance c avec une probabilité w_c et est uniformément distribué au sein de cette classe. Les classes de puissance sont alors définies sachant que l'amplitude des cibles est limitée par l'amplitude maximale que peut acquérir le récepteur, représentée par la limite haute ρ_C^+ . Des classes de longueur $\rho_c^+ - \rho_c^-$ sont alors créées, et associées à une probabilité (un poids) w_c afin de balayer l'intervalle de puissance comme désiré.

D'autre part, on conditionne le vecteur de phase ϕ au vecteur d'amplitude ρ : si l'amplitude est nulle, alors la phase l'est également. Sinon, on assigne une distribution a priori non-informative au vecteur de phase ϕ :

$$\begin{cases} \pi(\phi_{\bar{i}} | \rho_{\bar{i}} = 0) &= \delta(\phi_{\bar{i}}) \\ \pi(\phi_{\bar{i}} | \rho_{\bar{i}} \neq 0) &= \frac{1}{2\pi} \mathbb{I}_{[0, 2\pi]}(\phi_{\bar{i}}). \end{cases} \quad (4.4)$$

4.2.2 Distribution a priori des autres paramètres

La fonction de vraisemblance (4.2) dépend du vecteur \mathbf{x} , mais également des paramètres σ^2 , $(\varepsilon^v, \varepsilon^r)$. On choisit donc de leur assigner des distributions a priori, qui sont celles utilisées dans le modèle prenant en compte le désalignement des cibles du Chapitre 2. On a donc une distribution inverse-Gamma de paramètres (γ_0, γ_1) pour la puissance du bruit σ^2 :

$$f(\sigma^2 | \gamma_0, \gamma_1) \propto \frac{e^{-\gamma_1/\sigma^2}}{(\sigma^2)^{\gamma_0+1}} \mathbb{I}_{[0, +\infty)}(\sigma^2). \quad (4.5)$$

Les paramètres de désalignement sont maintenant conditionnés par la valeur d'amplitude à la case d'analyse considérée, $\rho_{\bar{i}}$:

$$\begin{cases} \pi(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \rho_{\bar{i}} = 0) &= \delta(\varepsilon_{\bar{i}}^v) \delta(\varepsilon_{\bar{i}}^r) \\ \pi(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \rho_{\bar{i}} \neq 0) &= \mathbb{I}_{[-0.5, 0.5]}(\varepsilon_{\bar{i}}^v) \mathbb{I}_{[-0.5, 0.5]}(\varepsilon_{\bar{i}}^r). \end{cases} \quad (4.6)$$

4.2.3 Distribution a priori des hyperparamètres

La nouvelle distribution a priori du vecteur amplitude ρ dépend d'un hyperparamètre, le vecteur de probabilité d'appartenance \mathbf{w} . On lui assigne donc une distribution a priori, pour constituer le dernier étage du modèle hiérarchique Bayésien. On utilise pour cela une distribution a priori classiquement utilisée pour les vecteurs de probabilité : une loi de Dirichlet multivariée de paramètres de concentration $\theta_0, \dots, \theta_C > 0$ [69], notée $\mathbf{w} \sim \mathcal{Dir}(\theta_0, \dots, \theta_C)$

$$\pi(\mathbf{w} | \theta_0, \dots, \theta_C) \propto w_0^{\theta_0-1} \mathbb{I}_{[0,1]}(w_0) \times \prod_{c=1}^C w_c^{\theta_c-1} \mathbb{I}_{[0,1]}(w_c). \quad (4.7)$$

Les paramètres de concentration peuvent être choisis pour favoriser certaines classes de puissance devant les autres, ou bien être tous égaux à 1 pour donner une loi de Dirichlet symétrique, qui est non-informative.

On résume le nouveau modèle hiérarchique Bayésien dans la Fig. 4.2. Ce nouveau modèle nécessite le réglage de plusieurs hyperparamètres, (le nombre de classes C , les limites des classes $\{\rho_c^-, \rho_c^+\}_{c=1, \dots, C}$, les paramètres de concentration $\theta_0, \dots, \theta_C$), plus nombreux que pour le modèle sans classes de puissance qui ne nécessite que le réglage des hyperparamètres de la puissance des cibles $(\beta_0, \beta_1)^2$. En fait, les paramètres des classes de puissance se règlent facilement, et un réglage a même été proposé dans [2] qui permet d'obtenir de bonnes performances d'estimation sur des scénarios variés à forte dynamique de puissance. Les paramètres de concentration $\{\theta_c\}_{c=0 \dots C}$ sont ajustés pour obtenir une loi non-informative.

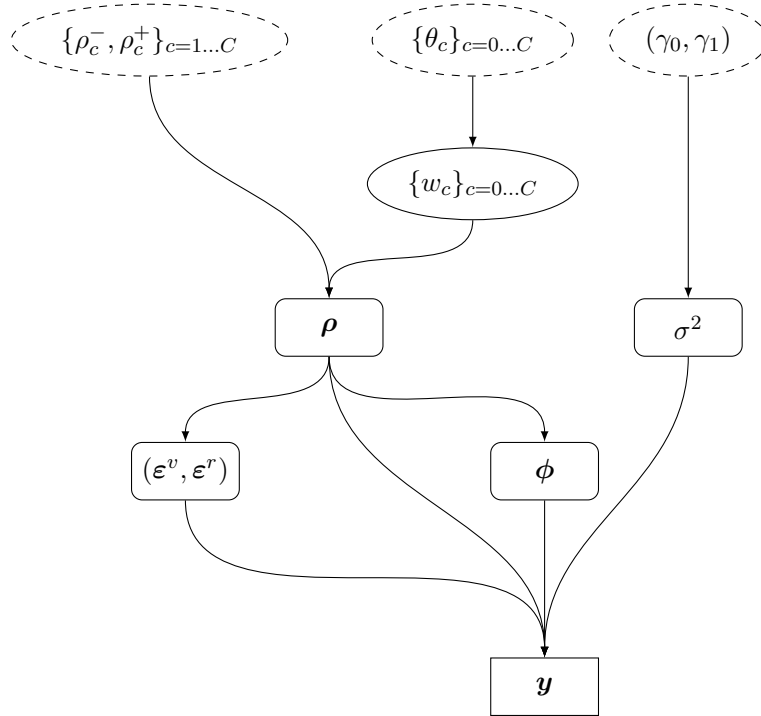


FIGURE 4.2 – Nouveau schéma hiérarchique Bayésien mis en place pour estimer des scènes à forte dynamique de puissance. Le vecteur complexe \mathbf{x} est désormais paramétré en termes d'amplitude ρ et de phase ϕ ; le support de ρ est divisé en C classes de puissance.

4.3 Nouveau schéma d'estimation

On cherche à mettre en place un schéma d'estimation des paramètres d'intérêt $\mathbf{x} = \rho e^{i\phi}$ et (ϵ^v, ϵ^r) . Dans les algorithmes présentés dans les chapitres précédents, c'est l'estimateur

2. On ne mentionne pas les hyperparamètres de la puissance du bruit, (γ_0, γ_1) , qui apparaissent dans les deux modèles.

MMSE du paramètre \mathbf{x} qui est calculé. Étant donnée la nouvelle paramétrisation, on choisit de calculer l'estimateur suivant :

$$\hat{\mathbf{x}}_{\text{class}} = \mathcal{E} \left\{ \boldsymbol{\rho} \odot e^{i\phi} | \mathbf{y} \right\} \quad (4.8)$$

$$= \int_{\boldsymbol{\rho}, \phi} \{ \boldsymbol{\rho} \odot e^{i\phi} \} f(\boldsymbol{\rho}, \phi | \mathbf{y}) d\boldsymbol{\rho} d\phi. \quad (4.9)$$

qui est la moyenne a posteriori du vecteur \mathbf{x} . Comme on ne peut pas calculer cette intégrale de manière analytique, on la marginalise par rapport aux autres paramètres du problème :

$$\hat{\mathbf{x}}_{\text{class}} = \int_{\sigma^2, \boldsymbol{\rho}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \mathbf{w}} \{ \boldsymbol{\rho} \odot e^{i\phi} \} f(\sigma^2, \boldsymbol{\rho}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \mathbf{w} | \mathbf{y}) d\sigma^2 d\boldsymbol{\rho} d\phi d\boldsymbol{\varepsilon}^v d\boldsymbol{\varepsilon}^r d\mathbf{w}. \quad (4.10)$$

On calcule d'autre part l'estimateur MMSE des paramètres de désalignement d'après (2.16). Un échantillonneur de Gibbs est implémenté pour calculer ces estimateurs comme

$$\begin{aligned} \hat{\mathbf{x}}_{\text{class}} &= \frac{1}{N_r} \sum_{t=1}^{N_r} \boldsymbol{\rho}^{(t+N_{bi})} \odot e^{i\phi^{(t+N_{bi})}} \\ \hat{\boldsymbol{\varepsilon}}_{\text{MMSE}}^v &= \frac{1}{N_r} \sum_{t=1}^{N_r} \boldsymbol{\varepsilon}^{v(t+N_{bi})} \\ \hat{\boldsymbol{\varepsilon}}_{\text{MMSE}}^r &= \frac{1}{N_r} \sum_{t=1}^{N_r} \boldsymbol{\varepsilon}^{r(t+N_{bi})} \end{aligned}$$

qui sont les moyennes empiriques des échantillons $\mathbf{x}^{(t)} = \boldsymbol{\rho}^{(t)} \odot e^{i\phi^{(t)}}$, $\boldsymbol{\varepsilon}^{v(t)}$ et $\boldsymbol{\varepsilon}^{r(t)}$. On tire donc des échantillons $\boldsymbol{\zeta}^{(t)} = [\sigma^{2(t)}, \boldsymbol{\rho}^{(t)}, \phi^{(t)}, \boldsymbol{\varepsilon}^{v(t)}, \boldsymbol{\varepsilon}^{r(t)}, \mathbf{w}^{(t)}]$ suivant la distribution jointe a posteriori

$$f(\sigma^2, \boldsymbol{\rho}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \mathbf{w} | \mathbf{y}) \propto f(\mathbf{y} | \sigma^2, \boldsymbol{\rho}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) \times \pi(\boldsymbol{\rho} | \mathbf{w}) \pi(\phi | \boldsymbol{\rho}) \pi(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r | \boldsymbol{\rho}) \pi(\mathbf{w}) \pi(\sigma^2).$$

4.3.1 Échantillonnage du vecteur amplitude $\boldsymbol{\rho}$

Le paramètre $\boldsymbol{\rho}$ est échantillonné terme-à-terme, donc il s'agit de calculer la distribution conditionnelle a posteriori de $\rho_{\bar{i}}$

$$\begin{aligned} f(\rho_{\bar{i}} | \mathbf{y}, \sigma^2, \boldsymbol{\rho}_{-\bar{i}}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \mathbf{w}) &\propto f(\mathbf{y} | \boldsymbol{\rho}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \sigma^2) \pi(\rho_{\bar{i}} | \mathbf{w}) \\ &\propto \exp \left\{ -\sigma^{-2} \left[\rho_{\bar{i}}^2 \|\mathbf{h}_{\bar{i}}(\boldsymbol{\varepsilon}_{\bar{i}}^v, \boldsymbol{\varepsilon}_{\bar{i}}^r)\|^2 - 2\rho_{\bar{i}} \operatorname{Re} \left\{ \mathbf{h}_{\bar{i}}(\boldsymbol{\varepsilon}_{\bar{i}}^v, \boldsymbol{\varepsilon}_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}} e^{-i\phi_{\bar{i}}} \right\} \right] \right\} \\ &\quad \times \left[w_0 \delta(\rho_{\bar{i}}) + \sum_{c=1}^C w_c k_c \mathbb{I}_{[\rho_c^-, \rho_c^+]}(\rho_{\bar{i}}) \right] \end{aligned}$$

ce qui donne finalement la distribution mixte

$$f(\rho_{\bar{i}} | \mathbf{y}, \sigma^2, \boldsymbol{\rho}_{-\bar{i}}, \phi, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r, \mathbf{w}) \propto w_{\bar{i},0} \delta(\rho_{\bar{i}}) + \sum_{c=1}^C \frac{w_{\bar{i},c} k_{\bar{i},c}}{\sqrt{2\pi\eta_{\bar{i}}^2}} \exp \left\{ -\frac{\eta_{\bar{i}}^{-2}}{2} (\rho_{\bar{i}} - \mu_{\bar{i}})^2 \right\} \mathbb{I}_{[\rho_c^-, \rho_c^+]}(\rho_{\bar{i}})$$

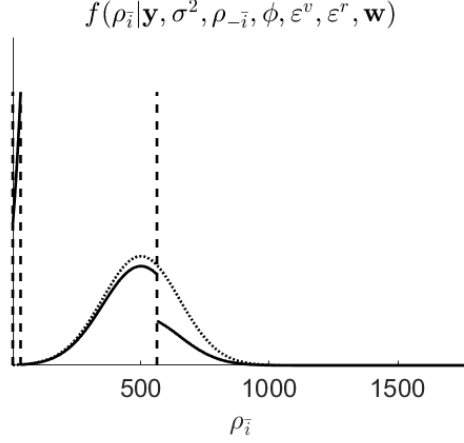


FIGURE 4.3 – Composante continue de la distribution conditionnelle a posteriori de $\rho_{\bar{i}}$ (4.3.1); $\mu_{\bar{i}} = 500$, $\eta_{\bar{i}} = 150$ and $\mathbf{w}_{\bar{i}} = [10, 15, 60, 15]\%$. Les tirets verticaux représentent les limites des 3 classes de puissance non-nulles. La courbe pointillée est la distribution Gaussienne $\mathcal{N}(\mu_{\bar{i}}, \eta_{\bar{i}}^2)$.

où l'on a $\forall c = 1, \dots, C$

$$\eta_{\bar{i}}^2 = \frac{1}{2} \sigma^2 \|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^{-2} \quad (4.11)$$

$$\mu_{\bar{i}} = \|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^{-2} \operatorname{Re} \left\{ \mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}} e^{-i\phi_{\bar{i}}} \right\} \quad (4.12)$$

$$w_{\bar{i},c} = \frac{w_c \times \frac{k_c}{k_{\bar{i},c}} (2\pi\eta_{\bar{i}}^2)^{1/2} \exp \left\{ \frac{1}{2} \eta_{\bar{i}}^{-2} \mu_{\bar{i}}^2 \right\}}{w_0 + \sum_{c=1}^C w_c \frac{k_c}{k_{\bar{i},c}} (2\pi\eta_{\bar{i}}^2)^{1/2} \exp \left\{ \frac{1}{2} \eta_{\bar{i}}^{-2} \mu_{\bar{i}}^2 \right\}} \quad (4.13)$$

et $w_{\bar{i},0} = 1 - \sum_{c=1}^C w_{\bar{i},c}$. Le facteur $k_{\bar{i},c}$ est un facteur d'échelle qui résulte de la troncature de la Gaussienne de moyenne $\mu_{\bar{i}}$ et de variance $\eta_{\bar{i}}^2$ sur l'intervalle $[\rho_c^-, \rho_c^+]$; on utilise la notation $\mathcal{N}_{[\rho_c^-, \rho_c^+]}(\mu_{\bar{i}}, \eta_{\bar{i}}^2)$ pour désigner cette Gaussienne tronquée.

Ainsi, la distribution conditionnelle a posteriori de $\rho_{\bar{i}}$ est une loi mixte : elle possède un atome à 0 et une composante non-nulle qui est la somme de lois Gaussiennes tronquées. On remarque que les paramètres $(\mu_{\bar{i}}, \eta_{\bar{i}}^2)$ de chaque Gaussienne tronquée $\mathcal{N}_{[\rho_c^-, \rho_c^+]}(\mu_{\bar{i}}, \eta_{\bar{i}}^2)$ ne dépendent pas de la classe c ; par contre, la probabilité $w_{\bar{i},c}$ associée à chaque Gaussienne tronquée dépend effectivement de la classe considérée. Autrement dit, le support de la Gaussienne $\mathcal{N}(\mu_{\bar{i}}, \eta_{\bar{i}}^2)$ est découpé en plusieurs classes et les tronçons de Gaussienne associés à chaque classe sont pondérés différemment. Cette caractéristique est illustrée sur la Fig. 4.3 où on représente la composante continue de $f(\rho_{\bar{i}} | \mathbf{y}, \sigma^2, \rho_{-\bar{i}}, \phi, \varepsilon^v, \varepsilon^r, \mathbf{w})$.

Remarque. Dans tous les calculs, on peut utiliser $\|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2 = 1$ puisque les colonnes du dictionnaire $\mathbf{H}(\varepsilon^v, \varepsilon^r)$ sont unitaires. On laisse néanmoins le terme $\|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2$ au cas où un autre dictionnaire serait utilisé.

4.3.2 Échantillonnage du vecteur phase ϕ

Le paramètre ϕ est également échantillonné terme-à-terme, donc on calcule la distribution conditionnelle a posteriori de $\phi_{\bar{i}}$, conditionnellement à la valeur correspondante $\rho_{\bar{i}}$:

$$\begin{aligned} f(\phi_{\bar{i}}|\mathbf{y}, \sigma^2, \phi_{-\bar{i}}, \boldsymbol{\rho}, \mathbf{w}; \rho_{\bar{i}} = 0) &\propto f(\mathbf{y}|\boldsymbol{\rho}, \phi, \sigma^2)\pi(\phi_{\bar{i}}|\rho_{\bar{i}} = 0) \\ &= \delta(\phi_{\bar{i}}). \end{aligned}$$

D'autre part, on a

$$\begin{aligned} f(\phi_{\bar{i}}|\mathbf{y}, \sigma^2, \phi_{-\bar{i}}, \boldsymbol{\rho}, \boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r; \rho_{\bar{i}} \neq 0) &\propto f(\mathbf{y}|\boldsymbol{\rho}, \phi, \sigma^2)\pi(\phi_{\bar{i}}|\rho_{\bar{i}} \neq 0) \\ &\propto \exp\left\{2\sigma^{-2}\rho_{\bar{i}}\operatorname{Re}\left\{\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}} e^{-i\phi_{\bar{i}}}\right\}\right\} \times \mathbb{I}_{[0,2\pi)} \\ &\propto \exp\left\{\kappa_{\bar{i}}\cos(\phi_{\bar{i}} - \psi_{\bar{i}})\right\} \mathbb{I}_{[0,2\pi)} \end{aligned}$$

avec $\kappa_{\bar{i}} = 2\sigma^{-2}\rho_{\bar{i}}|\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}}|$ et $\psi_{\bar{i}} = \angle\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)^H \mathbf{e}_{\bar{i}}$. On reconnaît alors une distribution de von Mises-Fisher de paramètre de concentration $\kappa_{\bar{i}}$ et de direction moyenne $\psi_{\bar{i}}$, que l'on note $\mathcal{VM}(\kappa_{\bar{i}}, \psi_{\bar{i}})$.

4.3.3 Échantillonnage des autres paramètres et hyperparamètres

L'échantillonnage des autres paramètres et hyperparamètres est inchangé par rapport au cas sans classes de puissance, en exprimant $\mathbf{x} = \boldsymbol{\rho} \odot e^{i\phi}$ ou $x_{\bar{i}} = \rho_{\bar{i}}e^{i\phi_{\bar{i}}}$ suivant les cas.

4.3.3.1 Paramètres de désalignement $(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)$

Comme dans l'algorithme prenant en compte le désalignement de [6] présenté dans le chapitre 2, les paramètres de désalignement sont tirés conjointement terme-à-terme. Le calcul de la distribution jointe conditionnelle a posteriori $f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r|\mathbf{y}, \sigma^2, \phi, \boldsymbol{\rho}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r, \mathbf{w})$ donne

$$\begin{aligned} f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r|\mathbf{y}, \sigma^2, \phi, \boldsymbol{\rho}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r, \mathbf{w}; \rho_{\bar{i}} \neq 0) &\propto \exp\left\{\operatorname{Re}\left\{\sum_{i=0}^{KM-1} 2\sigma^{-2}\rho_{\bar{i}}e^{-i\phi_{\bar{i}}}[\mathbf{h}_{\bar{i}}^H(0,0)]_i[\mathbf{e}_{\bar{i}}]_i\right.\right. \\ &\quad \left.\left.\exp\left\{j2\pi\left(\frac{\varepsilon_{\bar{i}}^r}{K}k - \frac{\varepsilon_{\bar{i}}^v}{M}n_{va}(1 + \mu k)m\right)\right\}\right\}\right\} \\ &\quad \times \mathbb{I}_{[-.5,.5]}(\varepsilon_{\bar{i}}^v)\mathbb{I}_{[-.5,.5]}(\varepsilon_{\bar{i}}^r) \\ f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r|\mathbf{y}, \sigma^2, \phi, \boldsymbol{\rho}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r, \mathbf{w}; \rho_{\bar{i}} = 0) &\propto \delta(\varepsilon_{\bar{i}}^v)\delta(\varepsilon_{\bar{i}}^r) \end{aligned}$$

qui est la même distribution conditionnelle a posteriori que dans le chapitre 2 avec $x_{\bar{i}} = \rho_{\bar{i}}e^{i\phi_{\bar{i}}}$. On utilise donc la procédure décrite dans le chapitre 2 pour tirer conjointement $(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)$.

4.3.3.2 Puissance du bruit σ^2

La distribution conditionnelle a posteriori de σ^2 est simplement calculée comme

$$\sigma^2 | \mathbf{y}, \boldsymbol{\rho}, \boldsymbol{\phi} \sim \mathcal{IG} \left(\gamma_0 + KM, \gamma_1 + \|\mathbf{y} - \mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r)(\boldsymbol{\rho} \odot e^{i\boldsymbol{\phi}})\|_2^2 \right).$$

4.3.3.3 Vecteur de probabilités \mathbf{w}

Le calcul de la distribution conditionnelle a posteriori de \mathbf{w} donne

$$\begin{aligned} f(\mathbf{w} | \mathbf{y}, \boldsymbol{\rho}) &\propto \pi(\boldsymbol{\rho} | \mathbf{w}) \pi(\mathbf{w}) \\ &\propto \prod_{\bar{i}=0}^{\bar{K}\bar{M}-1} \left\{ w_0 \delta(\rho_{\bar{i}}) + \sum_{c=1}^C w_c k_c \mathbb{I}_{[\rho_c^-, \rho_c^+]}(\rho_{\bar{i}}) \right\} \times \prod_{c=0}^C w_c^{\theta_c - 1} \\ &\propto w_0^{n_0 + \theta_0 - 1} \times \prod_{c=1}^C w_c^{n_c + \theta_c - 1} \end{aligned}$$

où $n_c = \#\{\bar{i} | \rho_{\bar{i}} \in [\rho_c^-, \rho_c^+]\}$, c'est-à-dire le nombre d'éléments du vecteur $\boldsymbol{\rho}$ qui appartiennent à la classe c , et $n_0 = \bar{K}\bar{M} - \sum_{c=1}^C n_c$. On reconnaît donc une distribution de Dirichlet multivariée, de paramètres de concentration $(n_0 + \theta_0, n_1 + \theta_1, \dots, n_C + \theta_C)$. On tire \mathbf{w} via la méthode de [46, chap.11] qui utilise une loi Gamma de paramètres $(n_c + \theta_c, 1)$.

4.3.4 Implémentation de l'algorithme

Afin de faciliter l'implémentation de ce nouvel algorithme, des « étiquettes » $\alpha_{\bar{i}} = [\boldsymbol{\alpha}]_{\bar{i}}$ sont utilisées. Ces étiquettes indiquent dans quelle classe se situe l'élément d'amplitude correspondant $\rho_{\bar{i}} : \alpha_{\bar{i}} \in \{0, 1, \dots, C\}$. Les étiquettes $\alpha_{\bar{i}}$ sont alors utilisées pour tirer les échantillons $\rho_{\bar{i}}$, $\phi_{\bar{i}}$ et $(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)$ en utilisant le formalisme suivant

$$\Pr[\alpha_{\bar{i}} = c] = w_{\bar{i},c}, \quad c = 0, \dots, C$$

$$\begin{cases} f(\rho_{\bar{i}} | \alpha_{\bar{i}} = 0) &= \delta(\rho_{\bar{i}}) \\ f(\rho_{\bar{i}} | \alpha_{\bar{i}} = c) &= \mathcal{N}_{[\rho_c^-, \rho_c^+]}(\mu_{\bar{i}}, \eta_{\bar{i}}^2), \quad c = 1, \dots, C. \end{cases}$$

$$\begin{cases} f(\phi_{\bar{i}} | \alpha_{\bar{i}} = 0) &= \delta(\phi_{\bar{i}}) \\ f(\phi_{\bar{i}} | \alpha_{\bar{i}} = c) &= \mathcal{VM}(\kappa_{\bar{i}}, \psi_{\bar{i}}), \quad c = 1, \dots, C. \end{cases}$$

$$\begin{cases} f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \alpha_{\bar{i}} = 0) &= \delta(\varepsilon_{\bar{i}}^v) \delta(\varepsilon_{\bar{i}}^r) \\ f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \alpha_{\bar{i}} = c) &\propto \exp \left\{ \sum_{i=0}^{KM-1} \kappa_i \cos \left(2\pi \left[\frac{\varepsilon_{\bar{i}}^v}{M} n_{va} (1 + \mu k) m - k \frac{\varepsilon_{\bar{i}}^r}{K} \right] - \phi_i \right) \right\} \mathbb{I}_{[-.5, .5]}(\varepsilon_{\bar{i}}^v) \mathbb{I}_{[-.5, .5]}(\varepsilon_{\bar{i}}^r). \end{cases}$$

En pratique, on tire donc l'étiquette $\alpha_{\bar{i}}$ sur $\{0, \dots, C\}$ avec les probabilités associées $\{w_{\bar{i},0}, \dots, w_{\bar{i},C}\}$. Ensuite, suivant la valeur de $\alpha_{\bar{i}}$, on tire les échantillons $\rho_{\bar{i}}$ suivant une loi gaussienne tronquée en utilisant par exemple la méthode de [41]; les échantillons $\phi_{\bar{i}}$ sont eux tirés suivant une loi

de von-Mises-Fisher and utilisant la méthode de [18]. Enfin, les échantillons $(\varepsilon_i^v, \varepsilon_i^r)$ sont tirés en utilisant la méthode décrite dans le chapitre 2.

On résume le nouvel échantillonneur de Gibbs sur la Fig. 4.4. On rappelle que cet échantillonneur permet non seulement d'obtenir un estimé du vecteur d'intérêt \mathbf{x} , mais également les estimateurs MMSE des autres paramètres, en particulier des paramètres de désalignement $(\varepsilon^v, \varepsilon^r)$ qui sont utilisés pour reconstruire la scène estimée.

Require: \mathbf{y} , (γ_0, γ_1) , $(\rho_c^-, \rho_c^+)_{c=1, \dots, C}$, $[\theta_c]_{c=0, \dots, C}$
Ensure: $\hat{\mathbf{x}}_{\text{class}}$, $(\hat{\varepsilon}^v, \hat{\varepsilon}^r)_{\text{MMSE}}$
 {Initialisation}
 $\boldsymbol{\rho}^{(0)}$, $\boldsymbol{\phi}^{(0)}$, $\boldsymbol{\alpha}^{(0)}$, $(\varepsilon^{v(0)}, \varepsilon^{r(0)})$
 {Itérations}
for $n = 1$ à $N_{bi} + N_r$ **do**
 $\sigma^{2(n)} \sim \mathcal{IG}(KM + \gamma_0, \|\mathbf{y} - \mathbf{H}(\varepsilon^{v(n-1)}, \varepsilon^{r(n-1)})\mathbf{x}^{(n-1)}\|_2^2 + \gamma_1)$
 $\mathbf{w}^{(n)} | \boldsymbol{\rho}^{(n-1)} \sim \text{Dir}(\theta_0 + n_0, \dots, \theta_C + n_C)$
for $\bar{i} = 0$ to $\bar{K}\bar{M} - 1$ **do**
 $\alpha_{\bar{i}}^{(n)} \sim \text{discrete}(\{0, \dots, C\}, \mathbf{w}_{\bar{i}})$
if $\alpha_{\bar{i}}^{(n)} = c$ with $c \neq 0$ **then**
 $\rho_{\bar{i}}^{(n)} | \mathbf{y}, \sigma^{2(n)}, \boldsymbol{\rho}_{-\bar{i}}^{(n)}, \boldsymbol{\phi}^{(n)}, \varepsilon^{v(n)}, \varepsilon^{r(n)} \sim \mathcal{N}_{[\rho_c^-, \rho_c^+]}(\mu_{\bar{i}}, \eta_{\bar{i}}^2)$
 $\phi_{\bar{i}}^{(n)} | \mathbf{y}, \sigma^{2(n)}, \boldsymbol{\phi}_{-\bar{i}}^{(n)}, \boldsymbol{\rho}^{(n)}, \varepsilon^{v(n)}, \varepsilon^{r(n)} \sim \mathcal{VM}(\kappa_{\bar{i}}, \psi_{\bar{i}})$
 $(\varepsilon_{\bar{i}}^{v(n)}, \varepsilon_{\bar{i}}^{r(n)})$ via une étape de MH ou tirage direct.
else
 $\rho_{\bar{i}}^{(n)} = 0$
 $\phi_{\bar{i}}^{(n)} = 0$
end if
 $x_{\bar{i}}^{(n)} = \rho_{\bar{i}}^{(n)} e^{i\phi_{\bar{i}}^{(n)}}$
end for
end for
 {Estimateurs}
 $\hat{\mathbf{x}}_{\text{class}} = \frac{1}{N_r} \sum_{n=1}^{N_r} \boldsymbol{\rho}^{(n+N_{bi})} \odot e^{i\boldsymbol{\phi}^{(n+N_{bi})}}$
 $\hat{\zeta}_{\text{MMSE}} = \frac{1}{N_r} \sum_{n=1}^{N_r} \boldsymbol{\theta}^{(n+N_{bi})}$

FIGURE 4.4 – Échantillonneur de Gibbs modifié associé au nouveau schéma hiérarchique Bayésien utilisant des classes de puissance. $\alpha_{\bar{i}}$ représente « l'étiquette » associée au \bar{i} -ième élément de $\boldsymbol{\rho}$; on la tire sur $\{0, \dots, C\}$ avec les probabilités associées $\{w_{\bar{i},0}, \dots, w_{\bar{i},C}\}$.

4.4 Exemples de réalisations de l'algorithme avec classes

L'algorithme avec classes de puissance est ensuite testé sur des données synthétiques et des données semi-expérimentales construites à partir de mesures du radar PARSAX. On le compare à l'algorithme sans classes de puissance de [24]. Dans cette section, on ne montre que des exemples de réalisations (1D et 2D); la section suivante sera consacrée aux résultats de simulations de Monte-Carlo.

4.4.1 Données synthétiques

4.4.1.1 Scénario unidimensionnel

Dans un premier temps, on teste l'algorithme avec classes sur un scénario unidimensionnel ($K = 1$) qui sera utilisé pour les simulations de Monte-Carlo de la section 4.5. Ce scénario consiste en une cible forte (SNR=60 dB) entourée de deux cibles faibles (SNR=15 dB) situées sur les cases d'analyse adjacentes : $f_{d_{i_1}} = 0$, $f_{d_s} = f_{d_{i_1}} + 1/\bar{M} = f_{d_{i_2}} - 1/\bar{M}$. Les paramètres de l'algorithme avec classes de puissance sont résumés dans le Tableau 4.1 ; le choix de ces différents paramètres est justifié dans la Section 4.5 (prior de σ^2 et classes de puissance). Les résultats obtenus sont présentés dans la Fig. 4.5 : on compare l'estimation de l'algorithme avec classes de puissance à celle de l'algorithme sans classes pour différentes formes de la loi de mélange $\pi(\sigma_x^2|\beta_0, \beta_1)$. On voit que seul l'algorithme avec classes est capable de distinguer les cibles faibles de la cible forte. Ce n'est d'ailleurs pas le cas si on utilise l'algorithme de référence APES [72] pourtant réputé pour sa bonne résolution. On voit également que l'estimation de l'algorithme sans classes dépend du réglage de la loi de mélange ; sur cette réalisation aucun des trois réglages proposé ne permet pourtant d'estimer les cibles faibles.

TABLE 4.1 – Paramètres de traitement

Grille d'analyse	$\bar{M} = M$
Nombre d'itérations	$N_{bi} = 200, N_r = 1000$
Initialisation	vraies valeurs
Prior de σ^2	$(\gamma_0, \gamma_1) = (0, 0)$
Classes de puissance	$(-\infty, \sigma_{dB}^2)$ dB, $[\sigma_{dB}^2, 30 + \sigma_{dB}^2)$ dB, $[30 + \sigma_{dB}^2, 65 + \sigma_{dB}^2)$ dB
Prior de \mathbf{w}	$\theta_c = 1, c = 0, \dots, C$

4.4.1.2 Scénario bidimensionnel avec cibles en chevauchement

L'algorithme avec classes a également été développé pour l'estimation de cibles migrantes 2D potentiellement désalignées par rapport à la grille d'analyse distance-vitesse. On le compare cette fois à l'algorithme de [6] présenté dans le chapitre 2. On utilise le même scénario que dans les chapitres précédents, à savoir 4 cibles dont les paramètres sont résumés dans le Tableau 2.1, mais on le modifie pour avoir une forte dynamique de puissance. La cible 3 possède donc un SNR de 60 dB et la cible 4 un SNR de 14 dB. La Fig. 4.5 montre que l'algorithme avec classes de puissance estime correctement la scène, alors que l'algorithme sans classes ne voit ni la cible 1 ni la cible 4 (masquée dans le lobe secondaire de la cible 3). Notons que l'algorithme sans classes a été ajusté sur une loi de mélange $\pi(\sigma_x^2|\beta_0, \beta_1)$ centrée sur des cibles fortes, puisque c'est ce réglage qui donne les meilleures performances d'après

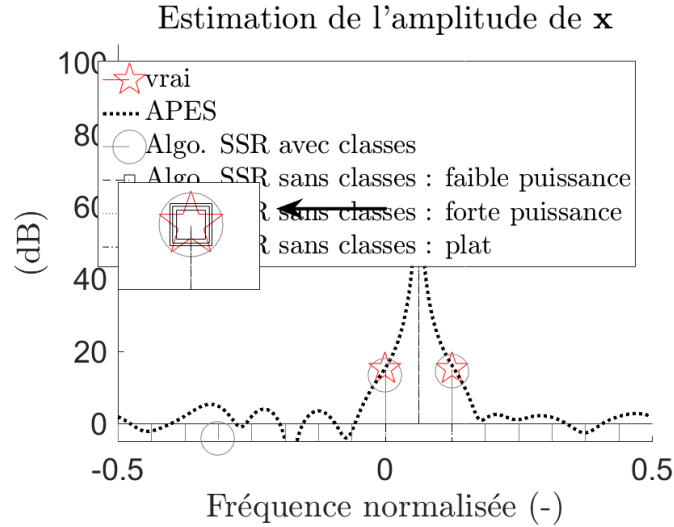


FIGURE 4.5 – Exemple de réalisation 1D de l’algorithme utilisant des classes de puissance (cercle, ligne pleine) comparées à celles de l’algorithme sans classes de puissance de [24] (carrés) pour différents réglages de la loi de mélange $\pi(\sigma_x^2|\beta_0, \beta_1)$: loi centrée sur des cibles faibles, loi centrée sur des cibles fortes ou loi non-informative. Le scénario consiste en une cible forte (SNR=60 dB) entourée de deux cibles faibles (SNR=15 dB) situées sur les cases d’analyse adjacentes. $M = 16$, $f_c = 10$ GHz, $\sigma^2 = 1$.

les simulations de Monte-Carlo de la section 4.5.

Il semble donc que l’algorithme avec classes de puissances permette de distinguer les cibles faibles masquées par une cible forte, que ce soit dans le cas unidimensionnel ou dans le cas bidimensionnel avec migration et désalignement. Ces tendances sont confirmées par les simulations de Monte-Carlo de la section 4.5.

4.4.2 Données semi-expérimentales

On compare ensuite les algorithmes avec et sans classes sur des données semi-expérimentales. Celles-ci sont construites en ajoutant des cibles synthétiques à du bruit thermique mesuré expérimentalement par le PARSAX de TU Delft [70]. Le but est de confronter l’algorithme avec classes de puissance à des conditions réelles, tout en ayant des cibles coopératives sur une scénario à forte dynamique de puissance. Les cibles sont bidimensionnelles et sujettes à la migration. En revanche, on les place sur la grille d’analyse et on ne prend pas en compte leur potentiel désalignement, $\mathbf{H}(\boldsymbol{\varepsilon}^v, \boldsymbol{\varepsilon}^r) = \mathbf{H}(\mathbf{0}, \mathbf{0})$, ce qui permet de se concentrer sur l’intérêt du nouveau modèle avec classes de puissance. L’algorithme sans classes de [24] est réglé pour favoriser des cibles fortes pour les mêmes raisons que précédemment. Par souci de comparaison, les données sont également traitées par l’algorithme CLEAN [60, 108] adapté au cas des cibles bidimensionnelles migrantes. On utilise pour cela la sommation cohérente adaptée à la forme d’onde large bande, dans sa version rapide [22]. L’algorithme s’arrête dès qu’il n’identifie plus de cibles dépassant un certain SNR (ici, 10 dB post-traitement).

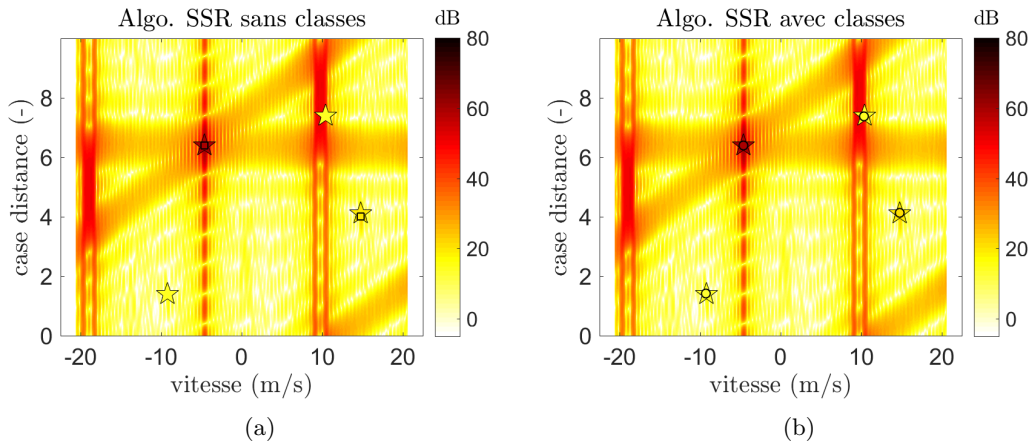


FIGURE 4.6 – Exemple de réalisation 2D de l’algorithme utilisant des classes de puissance (cercle) comparées à celles de l’algorithme sans classes de puissance de [24] (carrés) pour une loi de mélange centrée sur des cibles fortes. Le scénario est celui des chapitres précédents (cf Tableau 2.1) mais la cible 3 possède un SNR de 60 dB et la cible 4 un SNR de 14 dB. $K = 10$, $M = 32$, $n_{va} = 3$, $\bar{M} = n_{va}M$, $\bar{K} = K$, $v_a = 15$ m/s, $N_{bi} = 500$, $N_r = 1000$

Les cartes distance-vitesse estimées par les différents algorithmes sont représentées sur la Fig. 4.7. On voit une fois de plus que l’algorithme sans classes de puissance n’est pas capable d’estimer les cibles faibles masquées par les cibles fortes. Au contraire, l’algorithme avec classes estime correctement la localisation et l’amplitude des cibles, même les plus faibles. L’algorithme CLEAN est capable d’estimer toutes les cibles, y compris les plus faibles. En revanche, une fausse-alarme apparaît ($v \approx 14$ m/s, case distance # 6526). On remarque qu’une cible à vitesse nulle d’environ 12 dB est estimée par les trois algorithmes entre les cases distance #6524 et 6525 ; elle correspond probablement à un biais continu du circuit de réception.

4.5 Simulations de Monte-Carlo sur données synthétiques

On se focalise ensuite sur le scénario unidimensionnel de la Fig. 4.5 où deux cibles faibles entourent une cible forte (sur les cases d’analyse adjacentes). On rappelle qu’il s’agit d’un scénario bruit blanc, où les cibles sont situées sur la case d’analyse (pas de désalignement). On procède alors à des simulations de Monte-Carlo en faisant varier le SNR des deux cibles faibles, et on compare les performances obtenues à celles de l’algorithme sans classes de puissance de [24] pour différentes formes de la loi de mélange $\pi(\sigma_x^2 | \beta_0, \beta_1)$: loi centrée sur des cibles faibles, loi centrée sur des cibles fortes ou loi non-informative. On utilise dans ces simulations les paramètres décrits dans le Tableau 4.1 (en particulier les hyperparamètres de l’algorithme avec classes de puissance).

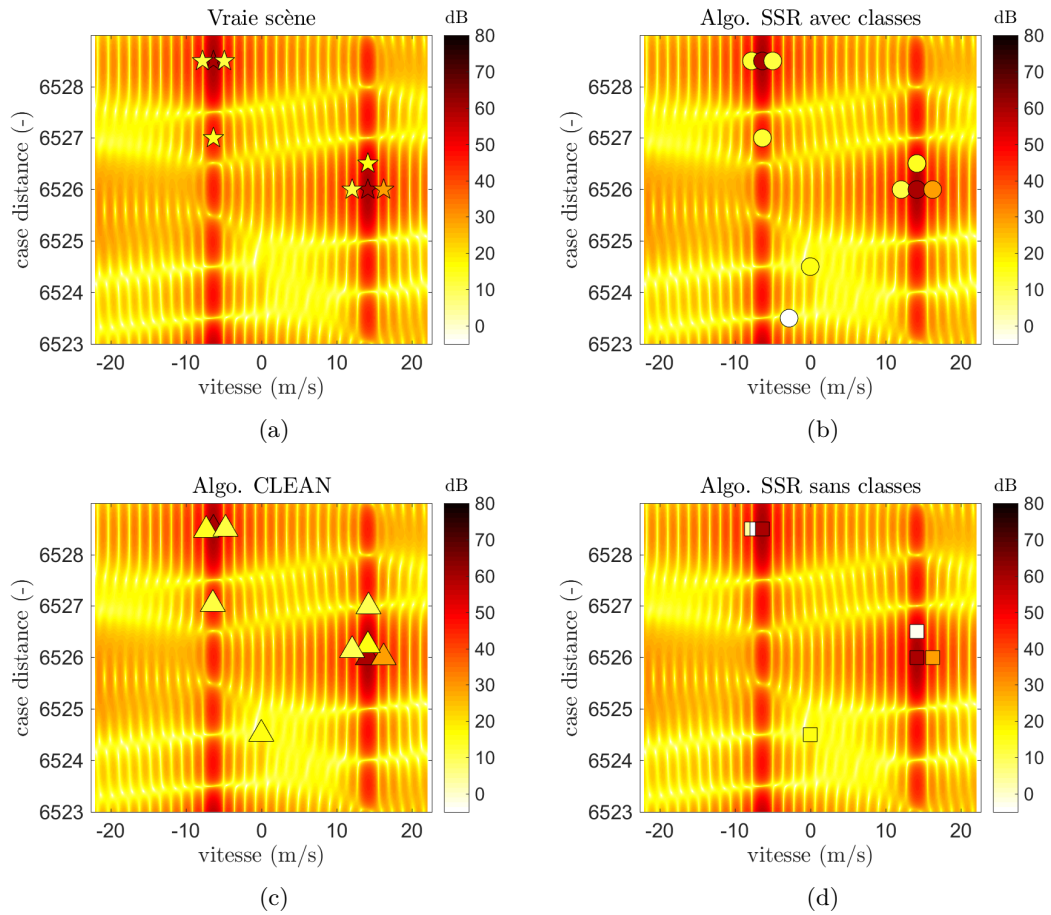


FIGURE 4.7 – Estimation sur données semi-expérimentales construites à partir d’une mesure de bruit enregistrée par le PARSAX 4.7(a) Vraie scène, 4.7(b) Algorithme avec classes de puissance, 4.7(c) Algorithme CLEAN, 4.7(d) Algorithme sans classes de puissance de [24]. Le fond représente la sommation cohérente large bande et les étoiles les vraies cibles. $K = 6$, $M = 32$, $n_{va} = 1$, $v_a = 45$ m/s, $\bar{M} = 2M$, $\bar{K} = 2K$.

4.5.1 Intérêt des classes de puissance

Dans un premier temps, on compare l'algorithme avec classes de puissance présenté dans ce chapitre à celui sans classes de puissance via la MSE normalisée de la scène reconstruite $\mathbf{F}\hat{\mathbf{x}}_{\text{class}}$ (Fig. 4.8(a)), la MSE normalisée de la cible forte (x_s , Fig. 4.8(c)) et la MSE normalisée de la cible faible située sur la case d'analyse 0 (à gauche de la cible forte, x_l) mais les résultats sont les mêmes pour l'autre cible faible (Fig. 4.8(b)). On voit que l'utilisation de classes de puissance améliore considérablement les performances de reconstruction de la scène, et ce quel que soit le réglage de l'algorithme sans classes. Plus précisément, l'estimation de la cible forte est la même mais l'estimation des cibles faibles est bien meilleure pour l'algorithme avec classes que pour l'algorithme sans classes. On remarque également le caractère non-monotonique de la MSE normalisée de $\mathbf{F}\mathbf{x}$ pour l'algorithme sans classes :

- l'erreur augmente d'abord avec le SNR des cibles faibles (SNR=6 à 14 dB). En fait, la valeur de nMSE obtenue correspond à la nMSE théorique quand les deux cibles faibles ne sont pas vues par l'algorithme (ligne verte), soit $10 \log_{10} (2|x_l|^2 / (|x_s|^2 + 2|x_l|^2))$.
- l'erreur diminue ensuite avec le SNR, jusqu'à rattraper les performances de l'algorithme avec classes.

On peut donc en conclure que les cibles faibles ne sont quasiment jamais vues par l'algorithme sans classes si leur SNR est inférieur à 14 dB. Lorsque leur SNR dépasse cette valeur, les cibles faibles sont plus souvent vues par l'algorithme et leur amplitude correctement estimée. L'utilisation de classes de puissance permet donc de diminuer le nombre de non-détection des cibles faibles, sur des scénarios complexes à forte dynamique de puissance.

Afin de soutenir ce propos, on compare les deux algorithmes, avec ou sans classes, sur un autre scénario à forte dynamique de puissance. Le scénario consiste ici en 5 cibles aléatoirement situées sur l'axe fréquentiel. Trois d'entre-elles possèdent une amplitude uniformément distribuée sur l'intervalle [10 14] dB (cibles faibles); l'amplitude des deux autres est uniformément distribuée sur l'intervalle [30 65] dB (cibles fortes). On utilise les mêmes métriques que précédemment pour comparer les deux algorithmes et on résume les résultats dans le Tableau 4.2. Les valeurs de MSE obtenues confirment les tendances observées sur le scénario à trois cibles utilisé précédemment : l'erreur obtenue avec l'algorithme avec classes de puissance sur la reconstruction de la scène est inférieure de 4 dB à celle obtenue par l'algorithme sans classes, quel que soit le réglage de la loi de mélange. L'estimation de la cible forte n'est pas influencée par le choix de l'algorithme, mais les cibles faibles sont mieux estimées par l'algorithme avec classes (écart de 4 dB environ).

TABLE 4.2 – Performances d'estimation de l'algorithme avec classes (« pw classes ») comparées à celles de l'algorithme sans classes (« npwclasses ») pour différents réglages de la loi de mélange. Le scénario consiste en 5 cibles aléatoirement localisées sur l'axe fréquentiel (3 cibles faibles et 2 cibles fortes). Les résultats sont exprimés en décibels.

	pw classes	npwclasses : lpower	npwclasses : hpower	npwclasses : flat
MSE de $\mathbf{F}\hat{\mathbf{x}}_{\text{class}}$	12.14	16.50	16.95	16.75
MSE de x_s	-0.34	-0.33	-0.34	-0.33
MSE de x_l	6.71	10.54	10.93	10.73

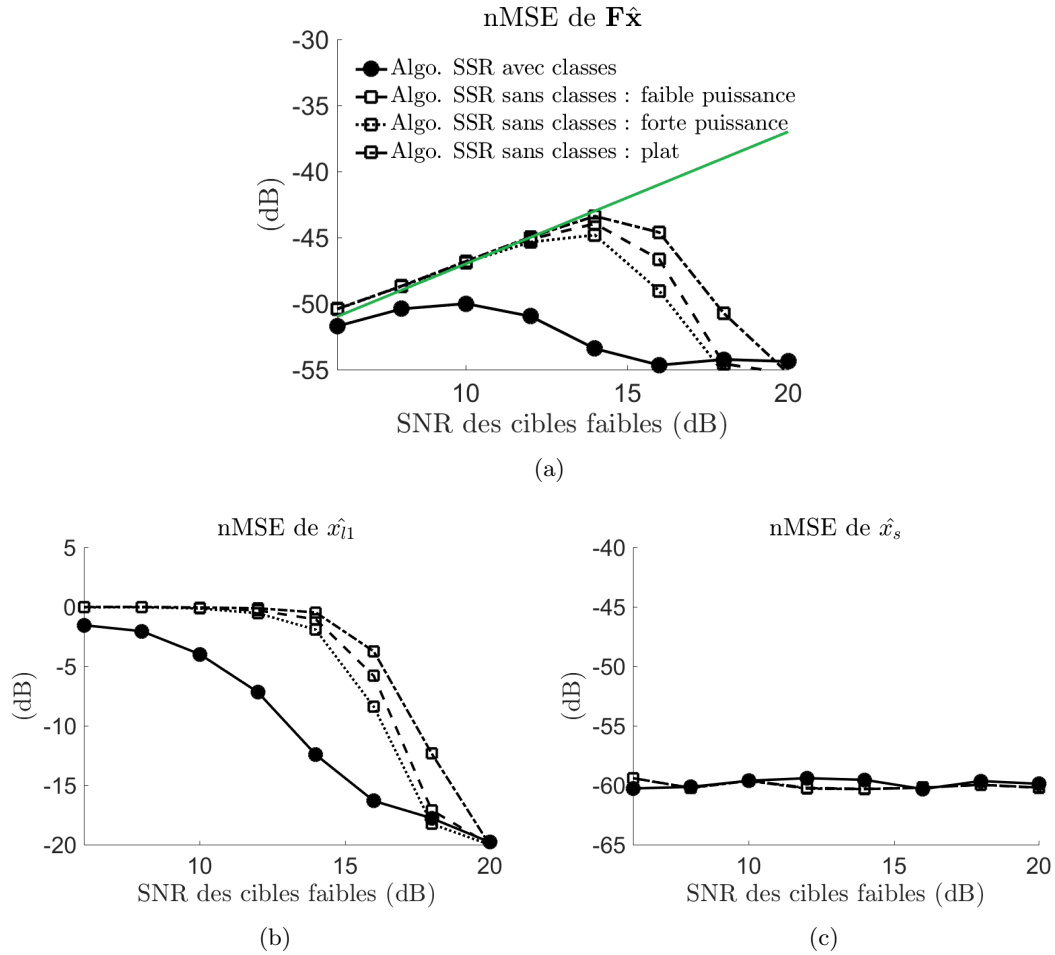


FIGURE 4.8 – Performances de l’algorithme utilisant des classes de puissance (cercle, ligne pleine) comparées à celles de l’algorithme sans classes de puissance de [24] (carrés) pour différents réglages de la loi de mélange $\pi(\sigma_x^2|\beta_0, \beta_1)$: loi centrée sur des cibles faibles (tirets), loi centrée sur des cibles fortes (pointillés) ou loi non-informative (tirets-pointillés). On se limite au cas unidimensionnel ($K = 1$) et on considère le scénario de la Fig. 4.5 qui consiste en une cible forte entourée de deux cibles faibles sur les cases adjacentes dont on fait varier le SNR. 4.8(a) nMSE de la scène reconstruite $\mathbf{F}\hat{\mathbf{x}}$, 4.8(b) nMSE de la cible faible \hat{x}_{l1} , 4.8(c) nMSE de la cible forte \hat{x}_s .

4.5.2 Choix des classes de puissance

Ce nouvel algorithme utilisant des classes de puissance nécessite le choix du nombre de classes C , et celui des limites de ces classes $\{(\rho_c^-, \rho_c^+)\}_{c=1\dots C}$. Des simulations numériques ont été effectuées sur le scénario d'intérêt à trois cibles, pour différents choix de classes de puissance :

- a) $(-\infty \quad 65)$ dB ($C = 1$)
- b) $(-\infty \quad 0), [0 \quad 30), [30 \quad 65)$ dB ($C = 3$)
- c) $(-\infty \quad 0), [0 \quad 30), [30 \quad 55), [55 \quad 65)$ dB ($C = 4$)
- d) $[0 \quad 30), [30 \quad 65)$ dB ($C = 2$)
- e) $(-\infty \quad 30), [30 \quad 65)$ dB ($C = 2$).

La comparaison entre ces différents choix de classes permettra de tirer plusieurs conclusions : doit-on diviser la plage de puissance en plusieurs classes (choix a) ? Est-ce que la première classe non-nulle $(-\infty \quad 0)$ dB est bien nécessaire (choix d / e) ? Est-ce qu'une division fine dans la plage de puissances élevées est essentielle (choix b / c) ?

On compare les performances obtenues avec ces différentes classes via le calcul de la MSE normalisée de la scène reconstruite $\mathbf{F}\hat{\mathbf{x}}_{\text{class}}$, représentée sur la Fig. 4.9. Tout d'abord on voit que le choix d'une seule classe $(-\infty \quad 65)$ dB (choix a) donne les moins bonnes performances (similaires à celles obtenues par l'algorithme sans classes, Fig. 4.8(a)) : les cibles faibles ne sont pas vues. En comparant ensuite les choix b et c on peut conclure qu'une division fine de la partie haute de l'intervalle de puissances n'est pas nécessaire. En fait, les cibles fortes qui appartiennent aux classes $[30 \quad 55), [55 \quad 65)$ dB sont facilement estimées par l'algorithme, quelles que soient les classes utilisées, puisqu'elles sont bien au-dessus du niveau de bruit. Enfin, l'écart entre les performances obtenues avec les choix b et d valide l'utilisation de la première classe de puissance $(0 \quad 1)$, i.e., $(-\infty \quad 0)$ dB puisqu'elle permet d'améliorer les performances. En effet, cette classe offre une alternative à l'algorithme quand celui-ci hésite entre la décision « pas de cible » (classe zéro) et « une cible au-dessus du niveau de bruit » (classe $[0 \quad 30)$ dB). On voit en outre en comparant les choix b et e qu'il ne suffit pas d'utiliser une classe unique $(-\infty \quad 30)$ pour estimer les cibles faibles, mais qu'il faut bien diviser la partie basse de l'intervalle de puissance en différentes classes. Le choix suivant de classes de puissance est donc préconisé :

1. $(-\infty \quad \sigma_{dB}^2)$,
2. $[\sigma_{dB}^2 \quad \sigma_{dB}^2 + 30)$,
3. $[\sigma_{dB}^2 + 30 \quad \mathcal{A}_{\text{max}})$,

où \mathcal{A}_{max} est l'amplitude maximale détectable (qui correspond à la plage de fonctionnement du récepteur). On règle les limites des classes par rapport au niveau de bruit estimé, $\sigma_{dB}^2 = 10 \log_{10}(\sigma^2)$; dans les simulations numériques de la Fig. 4.9 $\sigma_{dB}^2 = 0$ dB. Notons que ce choix a donné de bonnes performances d'estimation sur différents scénarios à forte dynamique de puissance : scénario synthétique 1D à 3 cibles (Fig. 4.5, Fig. 4.8), scénario synthétique 1D aléatoire à 5 cibles (Tableau 4.2), scénario synthétique 2D à 4 cibles (Fig. 4.6), scénario semi-expérimental 2D (Fig. 4.7).

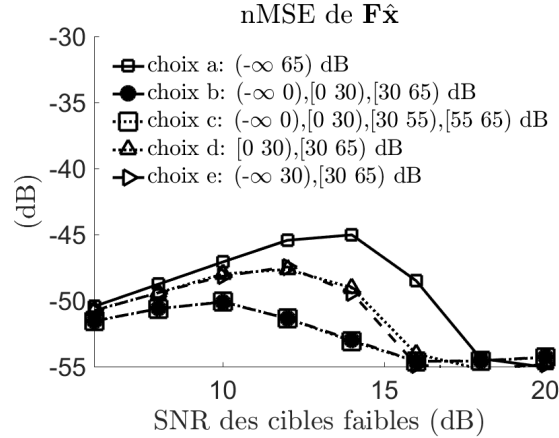


FIGURE 4.9 – Performance de l'algorithme avec classes de puissance pour différents choix de classes. Le scénario unidimensionnel considéré consiste en une cible forte entourée de deux cibles faibles sur les cases adjacentes et dont on fait varier le SNR.

4.5.3 Robustesse du réglage des classes par rapport au niveau de bruit

Étant donné le réglage préconisé, l'utilisateur doit estimer le niveau de bruit σ^2 au préalable, ce qui est plutôt commun dans le cadre de l'application radar. On étudie néanmoins la robustesse de l'algorithme à une erreur de ± 5 dB sur l'estimation σ_{dB}^2 avant le réglage des classes de puissance. On compare donc les classes de puissance suivantes :

- $(-\infty, 5), [5, 35), [35, 65)$ dB, ce qui correspond à une erreur de $\delta_{\sigma^2} = +5$ dB,
- $(-\infty, -5), [-5, 25), [25, 65)$ dB, ce qui correspond à une erreur de $\delta_{\sigma^2} = -5$ dB.

On voit sur la Fig. 4.10 qu'une erreur de ± 5 dB sur l'estimation de σ_{dB}^2 ne modifie pas radicalement les performances de l'algorithme (quelques décibels d'écart). Néanmoins, il semble plus avantageux de sous-estimer σ^2 .

D'autre part, une distribution a priori non-informative pour la puissance de bruit σ^2 a été choisie (voir Tableau 4.1). Les résultats précédents montrent qu'un tel choix donne de bonnes performances, mais on souhaite quantifier les possibles pertes de performance qu'il entraîne. On compare donc les performances de l'algorithme lorsqu'on utilise une distribution a priori non-informative pour échantillonner σ^2 à celles obtenues lorsque ce paramètre est connu, c'est-à-dire à une estimation clairvoyante vis-à-vis de σ^2 . La Fig. 4.11 montre que l'algorithme clairvoyant vis-à-vis de σ^2 améliore les performances de +5 dB au maximum. Le choix d'une distribution non-informative pour σ^2 implique donc une certaine perte de performances, qui n'empêche pas toutefois de supplanter l'algorithme sans classes de puissance.

4.6 Conclusions

Nous avons présenté dans ce chapitre une nouvelle distribution a priori pour l'estimation parcimonieuse d'une scène à forte dynamique de puissance, ainsi que le modèle hiérarchique et

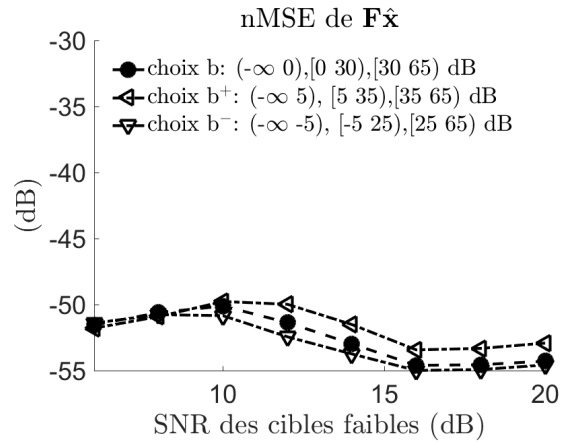


FIGURE 4.10 – Performance de l’algorithme avec classes lorsque l’estimation du niveau de bruit σ^2 est entachée d’une erreur de ± 5 dB. Le scénario unidimensionnel considéré consiste en une cible forte entourée de deux cibles faibles sur les cases adjacentes et dont on fait varier le SNR.

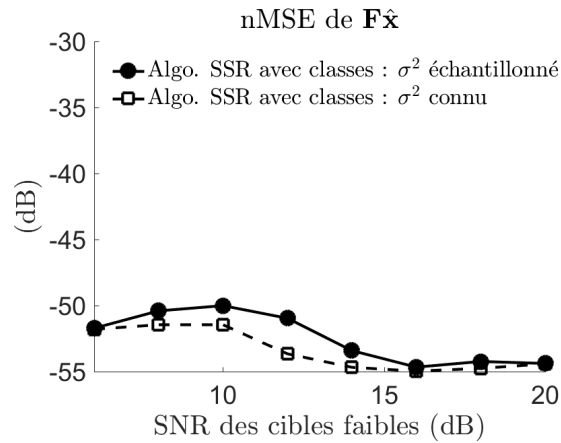


FIGURE 4.11 – Performance de l’algorithme avec classes lorsqu’on échantillonne le paramètre σ^2 comparé à celles obtenues lorsque le niveau de bruit σ^2 est connu (estimation clairvoyante vis-à-vis de σ^2). Le scénario unidimensionnel considéré consiste en une cible forte entourée de deux cibles faibles sur les cases adjacentes et dont on fait varier le SNR.

le schéma d'estimation MCMC associés. Cette nouvelle distribution renforce la parcimonie du vecteur d'intérêt et la décorrèle du niveau de puissance des cibles grâce à sa structure mixte, ce qui n'est pas le cas des priors hiérarchiques « classiques » tels que le prior de Laplace. L'originalité de cette distribution réside dans la division de son support en plusieurs classes qui correspondent à la puissance des cibles. Un réglage optimal de ces classes de puissance a été proposé et testé sur différents scénarios où cohabitent des cibles fortes et faibles.

L'algorithme avec classes de puissance a été testé sur des données synthétiques, et s'est révélé plus performant que l'algorithme sans classes pour estimer des scènes radar à forte dynamique de puissance. Plus précisément, ce n'est pas l'estimation des cibles fortes qui est améliorée, mais bien celles des cibles faibles masquées par les cibles fortes (par exemple dans leurs lobes secondaires). Ce comportement a été confirmé sur différents scénarios, et sur données semi-expérimentales (cibles synthétiques ajoutées à du bruit thermique mesuré expérimentalement).

Conclusion

Les travaux présentés dans ce manuscrit s'inscrivent dans la problématique générale de l'estimation et de la détection de cibles en environnement complexe. Finalement, nous avons bien développé un algorithme capable d'estimer des cibles de manière non-ambigüe, même en présence de fouillis diffus ou de forte dynamique de puissance. Dans un premier temps, les contributions apportées par ces travaux sont résumées, puis des perspectives de travaux futurs sont proposées.

Contributions

Comme expliqué dans le chapitre 1, l'objectif de ces travaux de thèse est de proposer un (ou des) algorithme(s) d'estimation non-ambigüe de cibles grâce à une forme d'onde large bande à PRF fixe. Dans le premier chapitre, nous avons motivé l'utilisation d'une forme d'onde large bande, classiquement utilisée en imagerie SAR, pour estimer/détecter des cibles. Pour cela, nous avons présenté quelques traitements classiques utilisant une bande étroite, et leurs inconvénients. Ensuite, en décrivant des méthodes de représentation spectrale adaptées au large bande nous avons introduit le choix des représentations parcimonieuses pour obtenir une représentation sans lobes secondaires du signal d'intérêt ; nous avons ensuite précisé le cadre algorithmique Bayésien utilisé dans les algorithmes de SSR développés. Cet exposé a mené au point de départ de ces travaux de thèse : un algorithme Bayésien de SSR développé dans [24] qui, sous certaines hypothèses, permet la représentation non-ambigüe de cibles grâce à un signal large bande. En soulignant ses limites, nous avons motivé les travaux menés lors de cette thèse, à savoir la robustification de cet algorithme de référence à des conditions plus « réalistes » : cibles en chevauchement sur la grille d'analyse, fouillis de sol avec une composante diffuse, scènes à très forte dynamique de puissance.

Dans le deuxième chapitre, nous avons décrit la prise en compte du possible désalignement des cibles par rapport à la grille d'analyse distance-vitesse intrinsèque au dictionnaire utilisé dans l'algorithme de SSR de référence. Plus précisément, nous avons vu que le modèle paramétrique proposé (paramétrisation directe du dictionnaire, par opposition à un modèle linéaire utilisant un développement limité) permet une estimation non-ambigüe de la scène radar : les cibles ne sont pas « dédoublées » dans les lobes secondaires (en distance et/ou en vitesse). Nous avons également montré qu'un raffinement de la grille d'analyse n'améliore pas les performances de l'algorithme proposé, bien que cette méthode intuitive soit souvent utilisée pour résoudre le problème de désalignement.

L'algorithme du chapitre 2 prenant en compte le désalignement est ensuite étendu pour prendre en compte du fouillis de sol ayant une composante diffuse. Nous avons d'abord rappelé le modèle hiérarchique Bayésien proposé dans [21] utilisant un modèle auto-regressif de bruit pour prendre en compte le fouillis diffus. Celui-ci est fusionné au modèle prenant en compte

le désalignement du chapitre 2 afin d'obtenir un algorithme final qui estime de manière non-ambigüe des cibles en chevauchement sur la grille d'analyse, qu'elles soient situées en-dehors des lobes secondaires du fouillis ou dans les vitesses aveugles. Cet algorithme a été testé avec succès sur des données synthétiques, avec et sans homogénéité du fouillis entre les cases distance, mais également sur du fouillis expérimental mesuré par le radar PARSAX.

Enfin, nous avons décrit dans le chapitre 4 un algorithme capable d'estimer des scènes à très fortes dynamique de puissance, c'est-à-dire où cohabitent des cibles faibles et des cibles très fortes. Afin de rester dans le même cadre algorithmique que pour l'algorithme de référence, une nouvelle distribution a priori a été introduite, qui repose sur l'utilisation de classes de puissance. Nous avons vu que l'utilisation de classes de puissance améliore considérablement les performances d'estimation de scènes à forte dynamique de puissance, et proposé un réglage de ces classes en fonction de la puissance de bruit thermique de la scène. Une erreur d'estimation du niveau de bruit thermique de ± 5 dB ne met pas en péril les performances de l'algorithme, mais il est préférable de le sous-estimer. L'algorithme a également été validé sur données synthétiques et semi-expérimentales (cibles synthétiques ajoutées à du bruit expérimental).

Perspectives

Comme souligné tout au long du manuscrit, de nombreux travaux futurs pourront être menés à l'issue de ces travaux. On se focalise ici sur certaines perspectives, qui semblent primordiales.

Confirmation des performances sur des données expérimentales supplémentaires

Nous avons vu dans les chapitres 2, 3 et 4 que tous les algorithmes développés ont été testés et validés sur des données expérimentales ou semi-expérimentales provenant du radar PARSAX de l'université de Delft. Cependant, l'inconvénient des jeux de données enregistrés avec ce radar jusqu'à présent est que la bande fractionnelle B/f_c est plutôt faible. Autrement dit, il faut utiliser un long temps d'intégration pour obtenir un phénomène de migration notable. De plus, le fouillis de sol mesuré avec le PARSAX possède une composante diffuse plutôt faible, alors que l'on cherche à valoriser l'algorithme « complet » du chapitre 3 sur du fouillis diffus. Ainsi, il serait intéressant de tester les différents algorithmes sur des données avec un plus fort coefficient de migration afin de tester l'algorithme dans des scénarios encore plus complexes. Il serait aussi nécessaire de tester l'algorithme « complet » du chapitre 3 sur différents types de fouillis : fouillis ayant une composante diffuse plus marquée, mais également fouillis suivant une distribution possédant une plus longue queue (ce qui devrait être le cas en augmentant la bande instantanée). Suite à ces expérimentations, il sera peut-être nécessaire de revoir le modèle auto-régressif de bruit utilisé pour prendre en compte la composante diffuse de bruit. D'autre part, nous avons souligné dans le chapitre 4 la nécessité de tester l'algorithme avec classes de puissance sur des données expérimentales à très forte dynamique de puissance. Pour l'instant, des cibles synthétiques ont été ajoutées à des mesures expérimentales de bruit ther-

mique, mais des données entièrement expérimentales permettraient sans doute de souligner de nouvelles problématiques.

Diminution de la charge calculatoire Comme expliqué dans le chapitre 1, les méthodes MCMC permettent d’obtenir une estimation précise des paramètres d’intérêt pourvu que la convergence des chaînes d’échantillonnage soit respectée. L’algorithme PCGS rencontre aujourd’hui un certain succès [110, 83, 48] puisqu’il permet de conserver les qualités d’estimation des méthodes d’approximation stochastiques, tout en réduisant la charge calculatoire. Suivant la méthode utilisée dans [48], on pourra expliciter la variable binomiale associée à la distribution Bernoulli-Gaussienne de \mathbf{x} et marginaliser la distribution conditionnelle a posteriori de \mathbf{x} par rapport à cette nouvelle variable. Une autre solution consiste à utiliser une méthode d’approximation déterministe comme par exemple une méthode variationnelle Bayésienne. Cette méthode a déjà été appliquée à l’algorithme de référence de [24] dans [23]. Par contre, elle nécessite une reformulation de la loi a posteriori qui est non triviale dès lors que le modèle hiérarchique prend en compte le désalignement des cibles (chapitre 2) ou la composante diffuse de bruit via un modèle auto-régressif de bruit (chapitre 3). Cette étude permettrait alors de quantifier les pertes d’estimation dues à l’approximation variationnelle Bayésienne afin de décider si elles sont rédhibitoires.

Algorithmes de détection Étant donnés les objectifs de la thèse, la prochaine étape serait bien sûr de mettre en place une procédure de détection des cibles présentes dans la scène radar. Cette étape permettrait de caractériser les algorithmes développés dans ces travaux en termes de probabilité de détection et de fausse-alarme, et donc de se comparer à d’autres algorithmes de la littérature en utilisant des métriques autres que celles utilisées dans le manuscrit et généralement utilisées dans les SSR seulement (à savoir l’erreur quadratique moyenne du vecteur parcimonieux ou de la scène reconstruite). En particulier, il serait intéressant de comparer les algorithmes présentés dans ce manuscrit, utilisant le paradigme Bayésien dans le cadre des SSR, à celles développées en parallèle à TU Delft (algorithme W-IAA [87] décrit dans le chapitre 1). Jusqu’à présent, les deux approches n’ont été comparées qu’en termes d’estimation de la scène radar, alors que l’algorithme W-IAA est en principe suivi d’une étape de détection [87]. Cependant, même si la problématique de détection de cible est largement explorée dans la littérature, par exemple [67, 113, 94], ce n’est pas encore le cas dans le cadre spécifique des représentations parcimonieuses. Cette thématique a toutefois connu de récentes avancées, voir [82] et les références de la section I-B dont la plus ancienne remonte à 2004. Dans [82], le test d’hypothèse composite classique

$$\begin{cases} \mathcal{H}_0 : \mathbf{y} &= \mathbf{n} \\ \mathcal{H}_1 : \mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{n} \end{cases}$$

est reformulé pour tirer parti de la parcimonie du vecteur \mathbf{x} comme

$$\begin{cases} \mathcal{H}_0 : \mathbf{x} = \mathbf{0}, \text{ i.e., } x_{\bar{i}} = 0, \forall \bar{i} \in \{0, \dots, \bar{K}\bar{M} - 1\} \\ \mathcal{H}_1 : \mathbf{x} \neq \mathbf{0}, \text{ i.e., } \exists x_{\bar{i}} \neq 0, \text{ et } \mathbf{x} \text{ est parcimonieux.} \end{cases}$$

Plusieurs tests de détection sont alors comparés, qui utilisent les estimateurs ML ou MAP du paramètre d'intérêt \mathbf{x} : test de vraisemblance généralisée (GLR), test de densité a posteriori (PDR), test de vraisemblance utilisant le MAP (LRMAP). Ces approches ne semblent pas directement applicables à nos travaux, puisque nous nous sommes focalisés sur des estimateurs MMSE. En effet, dans les algorithmes implémentés nous avons calculé l'estimateur MMSE comme la moyenne des échantillons tirés ; ainsi le vecteur $\hat{\mathbf{x}}_{\text{MMSE}}$ possède des composantes non-nulles très faibles dues à des erreurs occasionnelles d'estimation et qui font baisser la parcimonie. Sous hypothèse de bruit blanc, il serait possible de seuiller les composantes faibles du vecteur $\hat{\mathbf{x}}_{\text{MMSE}}$ en utilisant par exemple $\hat{\sigma}_{\text{MMSE}}^2$ pour pouvoir utiliser le test d'hypothèse composite basé sur la parcimonie du vecteur \mathbf{x} .

L'algorithme utilisant des classes de puissance présenté dans le chapitre 4 permettrait d'utiliser une philosophie similaire. En effet, les classes de puissance sont réglées par rapport au niveau de bruit σ_{dB}^2 supposé. En utilisant le réglage proposé, on sait que les cibles appartenant aux classes $c = 0$ et $c = 1$ sont en-dessous du niveau de bruit thermique, et ne correspondent donc pas à de « vraies » cibles, à l'inverse des cibles appartenant aux classes $c = 2$ et $c = 3$. Cette idée ouvre la voie à une procédure d'estimation *ad-hoc*, mais de nombreux travaux devront être menés pour parvenir à prédire les performances de l'algorithme en termes de probabilité de détection et de fausse-alarme.

Amélioration de l'échantillonnage des paramètres de désalignement ϵ^v, ϵ^r Nous avons expliqué dans les chapitres 2 et 3 que les paramètres de désalignement ϵ^v, ϵ^r sont échantillonnés via une méthode de « tirage direct ». Cette méthode d'échantillonnage *ad-hoc*, satisfaisante en termes de qualité d'estimation, ne l'est pas en termes de charge calculatoire et de garanties théoriques de convergence. Peut-être faudrait-il revoir le modèle même de prise en compte du désalignement, qui prétend atteindre une résolution infinie (deux cibles sur des cases adjacentes avec des valeurs opposées de désalignement). Le développement récent des méthodes se passant de grille d'analyse et reposant sur l'utilisation de la norme atomique ouvrent à ce sujet de nouvelles perspectives, et limitent d'ailleurs la résolution des algorithmes. A ce jour, seules des approches déterministes ont été développées mais il serait intéressant d'envisager une approche Bayésienne du problème.

Étude de la distribution conditionnelle a posteriori du désalignement

Dans cette partie, on se focalise sur la distribution conditionnelle a posteriori des paramètres de désalignement dans le but de choisir une loi de proposition adaptée dans l'algorithme de MH utilisé pour leur échantillonnage. Cette étude est limitée au cas unidimensionnel (une seule case distance considérée, $K = 1$). Par conséquent, on ne considère que le désalignement par rapport à la grille d'analyse en vitesse, et on note $\varepsilon \triangleq \varepsilon^v$. Le dictionnaire \mathbf{H} correspond alors à un simple dictionnaire de Fourier 1D, noté \mathbf{F} , de taille $M \times \bar{M}$. Ainsi, le vecteur parcimonieux \mathbf{x} et le vecteur de désalignement ε sont de taille $\bar{M} \times 1$ et indicés par \bar{m} .

A.1 Reformulation de la distribution conditionnelle a posteriori

Si on se limite à une seule case distance, alors la distribution conditionnelle a posteriori s'écrit ¹

$$f(\varepsilon_{\bar{m}} | \mathbf{y}, \sigma^2, \mathbf{x}, \varepsilon_{-\bar{m}}) \propto \exp \left\{ \sum_{m=1}^{M-1} \kappa_m \cos \left(2\pi \frac{\varepsilon_{\bar{m}}}{M} m - \phi_m \right) \right\} \times \mathbb{I}_{[-.5, .5]}(\varepsilon_{\bar{m}}) \quad (\text{A.1})$$

avec $\kappa_m = |[\mathbf{z}_{\bar{m}}]_m|$, $\phi_m = \angle[\mathbf{z}_{\bar{m}}]_m$, et $\mathbf{z}_{\bar{m}} = 2\sigma^{-2} x_{\bar{m}}^* [\mathbf{f}_{\bar{m}}^*(0) \odot \mathbf{e}_{\bar{m}}]$. Cette distribution ressemble alors à une distribution de Von Mises Généralisée (GvM), définie dans [56, 54, 55, 97] comme :

$$f(\omega | \mu_1, \dots, \mu_k, \kappa_1, \dots, \kappa_k) \propto \exp \left\{ \sum_{j=1}^k \kappa_j \cos j(\omega - \mu_j) \right\} \mathbb{I}_{[0, 2\pi)}(\omega) \quad (\text{A.2})$$

où $\omega \in [0, 2\pi)$, $\kappa_1, \dots, \kappa_k \geq 0$ et $\mu_1 \in [0, 2\pi)$, $\mu_2 \in [0, \pi)$, \dots , $\mu_k \in [0, 2\pi/k)$. On remarque cependant que comme $\varepsilon_{\bar{m}} \in [-0.5, 0.5]$, alors $2\pi \frac{\varepsilon_{\bar{m}}}{M} \in [-\pi/\bar{M}, \pi/\bar{M}]$ donc la distribution a posteriori de $\varepsilon_{\bar{m}}$ est une GvM dilatée et tronquée.

Dans un premier temps, on se focalise sur l'allure des « pseudo-modes » $\kappa_m \cos \left(2\pi \frac{\varepsilon_{\bar{m}}}{M} m - \phi_m \right)$.

1. on considère implicitement que $x_{\bar{m}} \neq 0$.

On remarque qu'ils possèdent au maximum deux extrema sur l'intervalle $\varepsilon_{\bar{m}} \in [-.5, .5]$, mais suivant la nature de ces extrema, l'allure du pseudo-mode change considérablement. Il semble donc difficile d'approximer les pseudo-modes par des fonctions simples (polynômes de degré 2 par exemple) et d'utiliser cette approximation dans l'expression de la distribution conditionnelle a posteriori $f(\varepsilon_{\bar{m}}|\mathbf{y}, \sigma^2, \mathbf{x}, \varepsilon_{-\bar{m}})$.

Suite à ces premiers résultats, on cherche quand même à évaluer l'influence des paramètres κ_m et ϕ_m sur la distribution conditionnelle a posteriori $f(\varepsilon_{\bar{m}}|\mathbf{y}, \sigma^2, \mathbf{x}, \varepsilon_{-\bar{m}})$. On voit sur la Figure A.1 l'influence des paramètres ϕ_m et κ_m sur la posterior conditionnelle. Quand ϕ_m augmente de 0 à π le mode de la posterior conditionnelle se déplace vers 0.5, jusqu'à être « partagé » entre -0.5 et 0.5 lorsque $\phi_m = \pi$. Quand κ_m augmente, la distribution conditionnelle a posteriori devient plus « piquée » et en conséquence son maximum augmente.

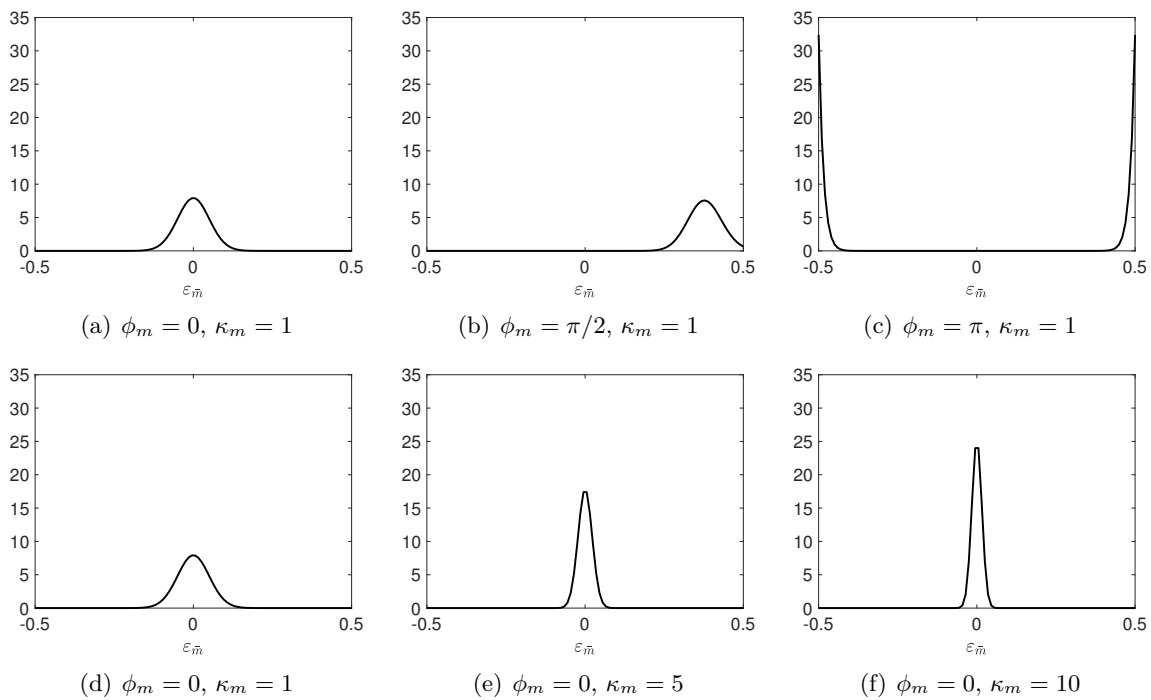


FIGURE A.1 – Distribution conditionnelle a posteriori de $\varepsilon_{\bar{m}}$ pour différentes valeurs de ϕ_m et κ_m .

En conclusion, il semble délicat d'approximer les pseudo-modes par une fonction simple et donc on ne peut pas trouver d'approximation simple de la distribution conditionnelle a posteriori de $\varepsilon_{\bar{m}}$, ce qui nous aurait donné une loi de proposition pour l'algorithme de MH.

On remarque cependant que pour de fortes valeurs de κ_m , la distribution conditionnelle a posteriori semble pouvoir être approximée par une Gaussienne. On recherche dans ce cas la moyenne $\mu_{\varepsilon_{\bar{m}}}$ et la variance $\sigma_{\varepsilon_{\bar{m}}}^2$ pour « coller » au mieux à la posterior conditionnelle; cette Gaussienne pourrait alors servir de loi de proposition dans l'algorithme de MH.

A.2 Approximation de la posterior par une loi normale

On se place dans le cas où κ_m possède de fortes valeurs, c'est-à-dire à fort SNR. En effet, si on reprend l'expression de κ_m :

$$\begin{aligned}\kappa_m &= |2\sigma^{-2}x_{\bar{m}}^* [\mathbf{f}_{\bar{m}}^*(0) \odot \mathbf{e}_{\bar{m}}]_m| \\ &= 2\sigma^{-2}|x_{\bar{m}}| \times |[\mathbf{f}_{\bar{m}}^*(0)]_m| \times |[\mathbf{e}_{\bar{m}}]_m|.\end{aligned}$$

En négligeant le bruit puisqu'on est à fort SNR, on écrit $\mathbf{e}_{\bar{m}} \approx x_{\bar{m}}\mathbf{f}_{\bar{m}}(\varepsilon_{\bar{m}})$, d'où

$$\begin{aligned}\kappa_m &\approx 2\sigma^{-2}|x_{\bar{m}}| \times \frac{1}{\sqrt{M}} \times |x_{\bar{m}} \frac{1}{\sqrt{M}}| \\ &\approx \frac{2}{M}|x_{\bar{m}}|^2\sigma^{-2} \\ &\approx \frac{2}{M}\text{SNR}.\end{aligned}\tag{A.3}$$

En effet, pour une cible d'amplitude α_t se situant à la \bar{i} -ième case d'analyse, $\text{SNR} = KM \times \mathcal{E}\{|\alpha_t|^2\}/\sigma^2 \approx |\alpha_t|^2/\sigma^2 = x_{\bar{m}}/\sigma^2$ puisque $x_{\bar{m}} = \sqrt{KM}\alpha_t$. En conclusion, à fort SNR, κ_m est proportionnel au SNR post-traitement $|x_{\bar{m}}|^2/\sigma^2$.

Pour approximer la distribution conditionnelle a posteriori de $\varepsilon_{\bar{m}}$ par une loi normale, on se place dans le cas d'un fort SNR et on suppose que $\theta_m = \phi_m\bar{M}/m \in [-\pi, \pi]$. On reprend alors son expression dans l'équation (A.1) et on procède à l'approximation $\cos x \approx 1 - \frac{x^2}{2}$. On obtient, avec $\psi_m = \frac{1}{m}\phi_m$:

$$\begin{aligned}f(\varepsilon_{\bar{m}}|\mathbf{y}, \mathbf{x}, \sigma^2, \boldsymbol{\varepsilon}_{-\bar{m}}) &\propto \exp\left\{\sum_{m=1}^{M-1} \kappa_m \cos\left(m\left(2\pi\frac{\varepsilon_{\bar{m}}}{M} - \psi_m\right)\right)\right\} \mathbb{I}_{[-.5,.5]}(\varepsilon_{\bar{m}}) \\ &\propto \exp\left\{-\sum_{m=1}^{M-1} \kappa_m \frac{m^2}{2} \left(\frac{2\pi}{M}\right)^2 \left(\varepsilon_{\bar{m}} - \frac{\bar{M}}{2\pi}\psi_m\right)^2\right\} \mathbb{I}_{[-.5,.5]}(\varepsilon_{\bar{m}})\end{aligned}$$

κ_m étant indépendant de $\varepsilon_{\bar{m}}$ à fort SNR. On factorise alors la dernière équation pour la mettre sous la forme d'une loi normale de moyenne $\mu_{\varepsilon_{\bar{m}}}$ et variance $\sigma_{\varepsilon_{\bar{m}}}^2$:

$$f(\varepsilon_{\bar{m}}|\mathbf{y}, \mathbf{x}, \sigma^2, \boldsymbol{\varepsilon}_{-\bar{m}}) \propto \exp\left\{-\frac{(\varepsilon_{\bar{m}} - \mu_{\varepsilon_{\bar{m}}})^2}{2\sigma_{\varepsilon_{\bar{m}}}^2}\right\} \mathbb{I}_{[-.5,.5]}(\varepsilon_{\bar{m}})$$

avec

$$\begin{aligned}\sigma_{\varepsilon_{\bar{m}}}^2 &= \left(\sum_{m=1}^{M-1} \kappa_m m^2 \left(\frac{2\pi}{M}\right)^2\right)^{-1} \\ \mu_{\varepsilon_{\bar{m}}} &= \sigma_{\varepsilon_{\bar{m}}}^2 \sum_{m=1}^{M-1} \kappa_m m^2 \left(\frac{2\pi}{M}\right) \psi_m\end{aligned}$$

On représente sur la Figure A.2 la loi conditionnelle a posteriori de $\varepsilon_{\bar{m}}$ et son approximation par une loi normale de paramètres $\mu_{\varepsilon_{\bar{m}}}$ et $\sigma_{\varepsilon_{\bar{m}}}^2$ pour différentes valeurs de SNR (post-traitement). On voit que l'approximation semble relativement correcte pour des valeurs de SNR post-traitement supérieures à 10 dB. Pour des valeurs en-dessous de 10 dB la loi normale ne « colle plus » à la posterior.

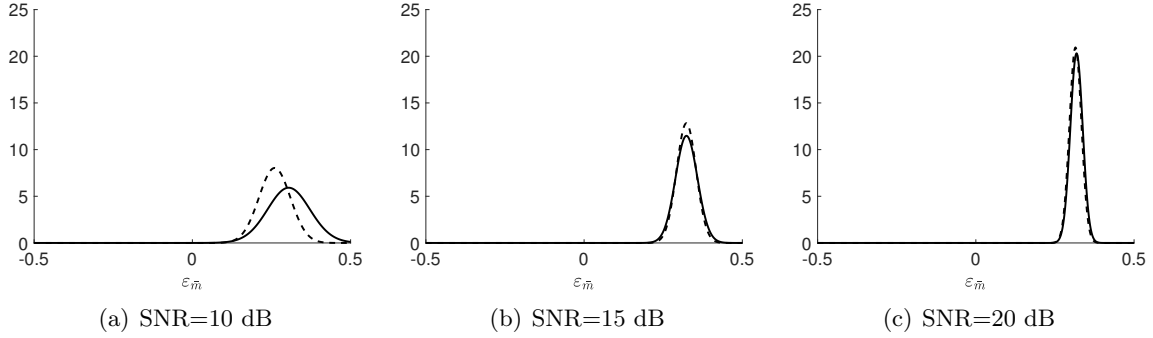


FIGURE A.2 – Comparaison entre la posterior conditionnelle de $\varepsilon_{\bar{m}}$ (trait plein) et son approximation par une loi normale (pointillés) pour différentes valeurs de SNR (post-traitement) : $\varepsilon_{\bar{m}} = 0.3$, $\sigma^2 = 1$

A.2.1 Choix de loi de proposition pour l'échantillonnage de ε

D'après les observations faites précédemment, le choix de la loi de proposition semble dépendre du SNR. Plus précisément, la classe même de la loi de proposition change suivant le SNR estimé : à fort SNR, une loi normale semble adaptée, mais à faible SNR on lui préférerait une loi uniforme sur $[-0.5, 0.5]$. De plus, les moments d'ordre 1 et 2 de la loi de proposition Gaussienne changent à chaque itération et doivent être estimés au préalable.

Pour illustrer le choix de loi de proposition, on choisit d'échantillonner ε en considérant tous les autres paramètres connus (\mathbf{x} , σ^2 , w , σ_x^2). On procède pour cela à 1000 itérations de Metropolis Hastings décrit en 2.2.4.2 (temps de chauffe de 100 itérations) avec une loi de proposition normale de moyenne $\mu_{\varepsilon_{\bar{m}}}$ et variance $\sigma_{\varepsilon_{\bar{m}}}^2$, puis une loi de proposition uniforme sur $[-0.5, 0.5]$. L'algorithme MH est initialisé en tirant une variable aléatoire suivant les lois de propositions. L'échantillonnage est fait pour une seule cible de SNR=10, 15 ou 20 dB.

Les résultats sont représentés dans la Figure A.3 et le Tableau A.1 via les valeurs du taux d'acceptation moyen τ et de l'erreur quadratique moyenne (MSE) sur $\hat{\varepsilon}_{\bar{m}}$. On remarque que le taux d'acceptation moyen pour la loi uniforme décroît lorsque le SNR augmente (cette loi est de moins en moins adaptée), alors que celui pour la loi normale augmente avec le SNR (cette loi est de mieux en mieux adaptée). Il est clair qu'à fort SNR la loi de proposition Gaussienne de paramètres $\mu_{\varepsilon_{\bar{m}}}$ et $\sigma_{\varepsilon_{\bar{m}}}^2$ est mieux adaptée. A faible SNR, c'est la loi de proposition uniforme qui donne la MSE la plus faible ; le taux d'acceptation est plus faible, pourtant la loi de proposition Gaussienne est décalée par rapport à la loi cible.

On décide donc de faire le choix de la loi de proposition en fonction de la valeur du SNR

post-traitement, qui sera estimée à chaque réalisation n de l'échantillonneur de Gibbs :

- $\widehat{\text{SNR}}^{(n)} > 10$ dB : loi de proposition normale de moyenne $\mu_{\varepsilon_{\bar{m}}}$ et variance $\sigma_{\varepsilon_{\bar{m}}}^2$
- $\widehat{\text{SNR}}^{(n)} \leq 10$ dB : loi de proposition uniforme sur l'intervalle $[-0.5, 0.5]$

Pour estimer la valeur du SNR à la n -ième réalisation, on utilise sa définition et les valeurs courantes de $x_{\bar{m}}$ et σ^2 : $\widehat{\text{SNR}}^{(n)} = |x_{\bar{m}}^{(n)}|^2 / \sigma^{2(n)}$.

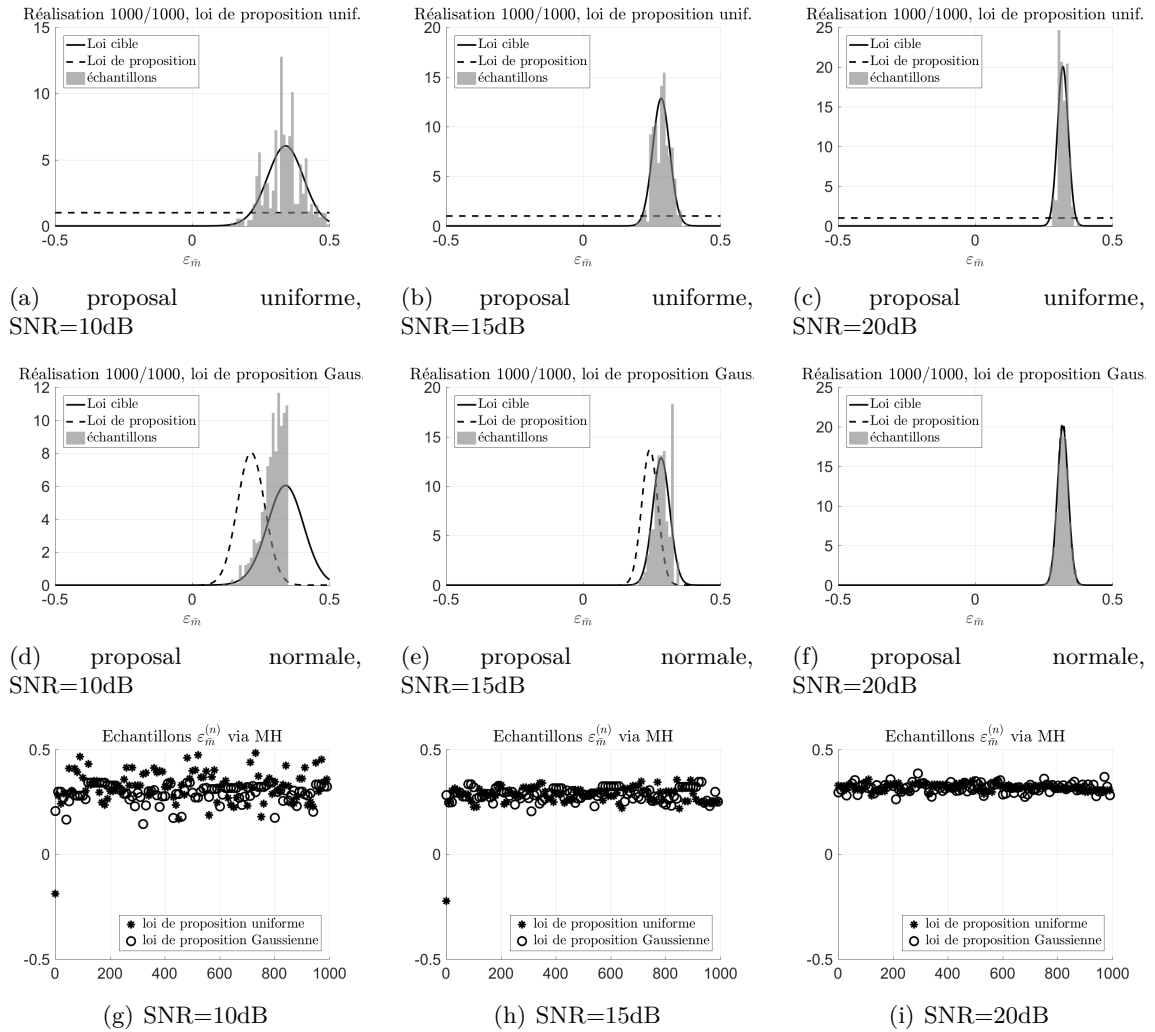


FIGURE A.3 – Échantillonnage de $\varepsilon_{\bar{m}}$ pour une loi de proposition uniforme ou normale pour différentes valeurs de SNR (post-traitement) : les échantillons sont représentés sous forme d'histogramme et comparés à la loi cible $f(\varepsilon_{\bar{m}}|\mathbf{y}, \mathbf{x}, \sigma^2, \varepsilon_{-\bar{m}})$ (trait plein) ; la loi de proposition est en pointillés.

TABLE A.1 – Performances de l'échantillonnage de $\varepsilon_{\bar{m}}$ avec une loi de proposition uniforme ou normale suivant le SNR post-traitement : taux d'acceptation moyen τ et MSE sur $\hat{\varepsilon}_{\bar{m}}$

SNR (dB)	8		10		15		20	
Loi de prop.	unif.	norm.	unif.	norm.	unif.	norm.	unif.	norm.
MSE($\hat{\varepsilon}_{\bar{m}}$)	4.4e-03	7.0e-03	3.5e-03	4.0e-03	1.4e-03	1.4e-03	4.1e-04	3.9e-04
$\tau(\%)$	26	32	20	38	11	64	6	94

Calcul de $\|\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2$

Dans cette section, on calcule la norme $\|\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2$ et en particulier sa dépendance en $(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)$ pour obtenir une expression concise de la distribution conditionnelle a posteriori $f(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r | \mathbf{y}, \boldsymbol{\varepsilon}_{-\bar{i}}^v, \boldsymbol{\varepsilon}_{-\bar{i}}^r, \mathbf{x}, \sigma_{AR}^2, \boldsymbol{\phi}; x_{\bar{i}} \neq 0)$. Dans un premier temps, on se focalise sur le vecteur $\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)$ lui-même :

$$\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r) = [\mathbf{I}_K \otimes (\mathbf{I}_M - \boldsymbol{\Phi})] \mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r) = \begin{bmatrix} (\mathbf{I}_M - \boldsymbol{\Phi})[\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{0:M-1} \\ \vdots \\ (\mathbf{I}_M - \boldsymbol{\Phi})[\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{kM:kM+M-1} \\ \vdots \\ (\mathbf{I}_M - \boldsymbol{\Phi})[\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{(K-1)M:KM-1} \end{bmatrix}$$

avec

$$(\mathbf{I}_M - \boldsymbol{\Phi})[\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{kM:kM+M-1} = \begin{bmatrix} [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{kM} \\ \vdots \\ \sum_{p=1}^P (-\phi_p) [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{m+kM-p} + [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{m+kM} \\ \vdots \\ \sum_{p=1}^P (-\phi_p) [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{M-1+kM-p} + [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{M-1+kM} \end{bmatrix}$$

On peut alors calculer la norme $\|\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2$:

$$\begin{aligned} \|\tilde{\mathbf{h}}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)\|^2 &= \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} \left| [(\mathbf{I}_M - \boldsymbol{\Phi}) [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{kM:kM+M-1}]_m \right|^2 \\ &= \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} \left(\left| [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{m+kM} \right|^2 + \sum_{p \leq m} |\phi_p [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{m+kM-p}|^2 \right. \\ &\quad \left. + \sum_{\substack{p \leq m \\ p \leq P}} 2 \operatorname{Re} \left\{ [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{m+kM}^* (-\phi_p) [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{m+kM-p} \right\} \right. \\ &\quad \left. + \sum_{\substack{q < p \leq m \\ p \leq P}} 2 \operatorname{Re} \left\{ \phi_p^* \phi_q [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{m+kM-p}^* [\mathbf{h}_{\bar{i}}(\varepsilon_{\bar{i}}^v, \varepsilon_{\bar{i}}^r)]_{m+kM-q} \right\} \right) \end{aligned}$$

$$\begin{aligned}
\|\tilde{\mathbf{h}}_i(\varepsilon_i^v, \varepsilon_i^r)\|^2 &= 1 + \frac{1}{M} \sum_{p=1}^P (M-p) |\phi_p|^2 \\
&\quad + \underbrace{\sum_{k=0}^{K-1} \sum_{p=1}^P (M-p) 2 \operatorname{Re} \left\{ (-\phi_p) \exp \left\{ -j2\pi p(1 + \mu k) n_{va} \frac{\varepsilon_i^v + \bar{m}'}{M} \right\} \times \frac{1}{KM} \right\}}_{T_3} \\
&\quad + \underbrace{\sum_{k=0}^{K-1} \sum_{p=1}^P (M-p) \sum_{q=1}^{p-1} 2 \operatorname{Re} \left\{ \phi_p^* \phi_q \exp \left\{ j2\pi(p-q)(1 + \mu k) n_{va} \frac{\varepsilon_i^v + \bar{m}'}{M} \right\} \times \frac{1}{KM} \right\}}_{T_4}
\end{aligned}$$

On se focalise sur les termes T_3 et T_4 , en commençant par l'expression de T_4

$$\begin{aligned}
T_4 &= \sum_{k=0}^{K-1} \sum_{p=1}^P \sum_{q=1}^{p-1} (M-p) 2 \operatorname{Re} \left\{ \phi_p^* \phi_q \exp \left\{ j2\pi(p-q)(1 + \mu k) n_{va} \frac{\varepsilon_i^v + \bar{m}'}{M} \right\} \times \frac{1}{KM} \right\} \\
&= \sum_{k=0}^{K-1} \sum_{d=1}^{P-1} \sum_{p=d+1}^P (M-p) 2 \operatorname{Re} \left\{ \phi_p^* \phi_{p-d} \exp \left\{ j2\pi d(1 + \mu k) n_{va} \frac{\varepsilon_i^v + \bar{m}'}{M} \right\} \times \frac{1}{KM} \right\}
\end{aligned}$$

puisque $\{(p, q) | p = 1 : P, q = 1 : p-1\} = \{(p, d) | p = d+1 : P, d = 1 : P-1\}$. La somme des deux termes $T_3 + T_4$ s'écrit alors

$$\begin{aligned}
T_3 + T_4 &= \sum_{k=0}^{K-1} \left(\sum_{d=1}^{P-1} 2 \operatorname{Re} \left\{ \left[(M-d)(-\phi_d^*) + \sum_{p=d+1}^P \phi_p^* \phi_{p-d} \right] \frac{1}{KM} \exp \left\{ j2\pi d(1 + \mu k) n_{va} \frac{\varepsilon_i^v + \bar{m}'}{M} \right\} \right\} \right. \\
&\quad \left. + 2 \operatorname{Re} \left\{ (M-P)(-\phi_P^*) \frac{1}{KM} \exp \left\{ j2\pi P(1 + \mu k) n_{va} \frac{\varepsilon_i^v + \bar{m}'}{M} \right\} \right\} \right) \\
&= \sum_{k=0}^{K-1} \sum_{d=1}^P 2 \operatorname{Re} \left\{ \beta_d \times \frac{1}{KM} \exp \left\{ j2\pi d(1 + \mu k) n_{va} \frac{\bar{m}'}{M} \right\} \times \exp \left\{ j2\pi d(1 + \mu k) n_{va} \frac{\varepsilon_i^v}{M} \right\} \right\} \\
&= \sum_{k=0}^{K-1} \sum_{d=1}^P 2 \operatorname{Re} \left\{ \gamma_{k,d} \exp \left\{ j2\pi d(1 + \mu k) n_{va} \frac{\varepsilon_i^v}{M} \right\} \right\}
\end{aligned}$$

On obtient finalement un terme qui ne dépend que de ε_i^v (indépendant de ε_i^r) :

$$\|\tilde{\mathbf{h}}_i(\varepsilon_i^v, \varepsilon_i^r)\|^2 = 1 + \frac{1}{M} \sum_{p=1}^P (M-p) |\phi_p|^2 + 2 \operatorname{Re} \left\{ \sum_{k=0}^{K-1} \sum_{d=1}^P \gamma_{k,d} \exp \left\{ j2\pi d(1 + \mu k) n_{va} \frac{\varepsilon_i^v}{M} \right\} \right\}$$

avec $\gamma_{k,d} = \beta_d \times \frac{1}{KM} \exp \left\{ j2\pi d(1 + \mu k) n_{va} \frac{\bar{m}'}{M} \right\}$, $\beta_d = \sum_{p=d}^P (M-p) \phi_p^* \phi_{p-d}$ et $\phi_0 = -1$.

Densités de probabilité utilisées

Loi Beta $\mathcal{B}e(\beta_0 > 0, \beta_1 > 0)$

$$f(x|\beta_0, \beta_1) = \frac{x^{\beta_0-1}(1-x)^{\beta_1-1}}{B(\beta_0, \beta_1)} \mathbb{I}_{[0,1]}(x),$$

Loi Gaussienne $N_P(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$$f(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}}{\sqrt{|\boldsymbol{\Sigma}|(2\pi)^P}}$$

 $CN_P(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$$f(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\exp\left\{-(\mathbf{x} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}}{|\boldsymbol{\Sigma}| \pi^P}$$

Loi Bernoulli-Gaussienne $\mathcal{B}erCN(w, \mu, \sigma^2)$

$$f(x|\mu, \sigma^2) = (1-w)\delta(x) + w \frac{\exp\{-\sigma^{-2}|x - \mu|^2\}}{\pi \sigma^2}$$

Loi Gamma $\mathcal{G}a(\gamma_0 > 0, \gamma_1 > 0)$

$$f(x|\gamma_0, \gamma_1) = \frac{\gamma_1^{\gamma_0}}{\Gamma(\gamma_0)} x^{\gamma_0-1} e^{-\gamma_1 x} \mathbb{I}_{[0,+\infty)}(x)$$

Loi Exponentielle $\mathcal{E}xp(\lambda) = \mathcal{G}a(1, \lambda > 0)$

$$f(x|\lambda) = \lambda e^{-\lambda x} \mathbb{I}_{[0,+\infty)}(x)$$

Loi Inverse Gamma $\mathcal{I}\mathcal{G}(\gamma_0 > 0, \gamma_1 > 0)$

$$f(x|\gamma_0, \gamma_1) = \frac{\gamma_1^{\gamma_0}}{\Gamma(\gamma_0)} \frac{e^{-\gamma_1/x}}{x^{\gamma_0+1}} \mathbb{I}_{[0,+\infty)}(x)$$

Loi de Student's t $\mathcal{T}_P(\nu > 0, \theta \in \mathbb{R}^P, \boldsymbol{\Sigma})$

$$f(x|\nu, \theta, \boldsymbol{\Sigma}) = \frac{1}{(\nu\pi)^{P/2} |\boldsymbol{\Sigma}|^{1/2}} \frac{\Gamma(\frac{\nu+P}{2})}{\Gamma(\frac{\nu}{2})} \left[1 + \frac{(x - \theta)^t \boldsymbol{\Sigma}^{-1} (x - \theta)}{\nu} \right]^{-\frac{\nu+P}{2}}$$

 $\mathcal{T}_P(\nu > 0, \theta \in \mathbb{C}^P, \boldsymbol{\Sigma})$

$$f(x|\nu, \theta, \boldsymbol{\Sigma}) = \frac{1}{\pi |\boldsymbol{\Sigma}| \nu^P} \frac{\Gamma(\nu + P)}{\Gamma(\nu)} \left[1 + \frac{(x - \theta)^H \boldsymbol{\Sigma}^{-1} (x - \theta)}{\nu} \right]^{-(\nu+P)}$$

Loi de Laplace $\mathcal{L}aplace(\mu, b)$

$$f(x|\mu, b) = \frac{1}{2b} \exp\left\{-\frac{|x - \mu|}{b}\right\}$$

Loi de Dirichlet $\mathcal{D}_k(\theta_1, \dots, \theta_k > 0)$

$$f(x|\theta_1, \dots, \theta_k) = \frac{\Gamma(\theta_1 + \dots + \theta_k)}{\Gamma(\theta_1) \dots \Gamma(\theta_k)} x_1^{\theta_1-1} \dots x_k^{\theta_k-1} \mathbb{I}_{\{\sum x_i = 1\}}$$

Loi de Von Mises - Fisher $\mathcal{V}\mathcal{M}(\kappa, \boldsymbol{\mu}), \kappa > 0, \|\boldsymbol{\mu}\| = 1$

$$f(\mathbf{x}|\kappa, \boldsymbol{\mu}) = \frac{\kappa^{P/2-1}}{(2\pi)^{P/2} I_{P/2-1}(\kappa)} \exp\left\{\kappa \boldsymbol{\mu}^T \mathbf{x}\right\} \mathbb{I}_{[0,2\pi)}(\mathbf{x})$$

Liste des publications

- [1] M. Lasserre, S. Bidon, O. Besson, and F. Le Chevalier, “Bayesian sparse Fourier representation of off-grid targets with application to experimental radar data,” *Signal Processing*, vol. 111, pp. 261–273, 2015.
- [2] M. Lasserre, S. Bidon, and F. Le Chevalier, “New sparse-promoting prior for the estimation of a radar scene with weak and strong targets,” *IEEE Transactions on Signal Processing*, vol. 64, no. 17, pp. 4634–4643, Sept 2016.
- [3] M. Lasserre, S. Bidon, O. Besson, and F. Le Chevalier, “Bayesian Sparse Fourier Representation of Off-Grid Targets,” in *2014 International Radar Conference (RADAR)*, Lille, France, October 2014.
- [4] S. Bidon, M. Lasserre, O. Besson, and F. Le Chevalier, “Bayesian sparse estimation of targets with range-Doppler grid mismatch,” in *2015 IEEE Radar Conference (RadarCon)*, Arlington, VA, May 2015.
- [5] M. Lasserre, S. Bidon, O. Besson, and F. Le Chevalier, “Bayesian sparse estimation of a radar scene with weak and strong targets,” in *Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa), 2015 3rd International Workshop on*, June 2015, pp. 51–55.
- [6] M. Lasserre, S. Bidon, and F. Le Chevalier, “Velocity ambiguity mitigation of off-grid range migrating targets via bayesian sparse recovery,” in *2016 IEEE Statistical Signal Processing Workshop (SSP)*, June 2016, pp. 1–5.
- [7] —, “An unambiguous radar mode with a single PRF wideband waveform,” in *2017 IEEE Radar Conference (RadarCon)*, Seattle, WA, May 2017.
- [8] S. Bidon, M. Lasserre, and F. Le Chevalier, “Bayesian sparse estimation of radar targets in the compressed sensing framework,” in *Compressive Sensing of Earth Observations*, C. Chen, Ed. CRC Press, 2017, ch. 7, to be published.

Bibliographie

- [9] Y. I. Abramovich, N. K. Spencer, and M. D. E. Turley. Order estimation and discrimination between stationary and time-varying (TVAR) autoregressive models. *IEEE Trans. Signal Process.*, 55(6) :2861–2876, June 2007.
- [10] V. Anastassopoulos, G.A. Lampropoulos, A. Drosopoulos, and M. Key. High resolution radar clutter statistics. *IEEE Trans. Aerosp. Electron. Syst.*, 35(1) :43–60, 1999.
- [11] D.F. Andrews and C.L. Mallows. Scale mixtures of normal distributions. *Journal of the Royal Statistical Society*, 36(1) :99–102, 1974.
- [12] A. Armagan, D.B. Dunson, and J. Lee. Generalized Double Pareto Shrinkage. *Statistical Sinica*, 23(1) :119–143, 2013.
- [13] S. D. Babacan, S. Nakajima, and M. N. Minh. Bayesian group-sparse modeling and variational inference. *IEEE Trans. Signal Process.*, 62(11) :2906–2921, 2014.
- [14] R. G. Baraniuk, V. Cevher, and M. B. Wakin. Low-dimensional models for dimensionality reduction and signal recovery : A geometric perspective. *Proceedings of the IEEE*, 98(6) :959–971, June 2010.
- [15] E. J. Barlow. Doppler radar. *Proceedings of the IRE*, 37(4) :340–355, April 1949.
- [16] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1) :183–202, 2009.
- [17] S. Becker, J. Bobin, and E. J. Candes. NESTA : A fast and accurate first-order method for sparse recovery. *SIAM Journal on Imaging Sciences*, 4(1) :1–39, 2011.
- [18] D. J. Best and N. I. Fisher. Efficient simulation of the von Mises distribution. *Journal of the Royal Statistical Society*, (2) :152–157.
- [19] B.N. Bhaskar, G. Tang, and B. Recht. Atomic norm denoising with applications to line spectral estimation. *IEEE Transactions on Signal Processing*, 61(23) :5987–5999, 2013.
- [20] P. J. Bickel and E. Levina. Regularized estimation of large covariance matrices. *The Annals of Statistics*, 36(1) :199–227, 2008.
- [21] S. Bidon, O. Besson, J. Y. Tournet, and F. Le Chevalier. Bayesian sparse estimation of migrating targets in autoregressive noise for wideband radar. In *2014 IEEE Radar Conference (RadarCon)*, Cincinnati, OH, May 2014.
- [22] S. Bidon, L. Savy, and F. Deudon. Fast coherent integration for migrating targets with velocity ambiguity. In *2011 IEEE Radar Conference (RadarCon)*, pages 027–032, May 2011.
- [23] S. Bidon, A. Tamalet, and J. Y. Tournet. Variational Bayesian inference for sparse representation of migrating targets in wideband radar. In *2013 IEEE Radar Conference (RadarCon)*, pages 1–5, April 2013.
- [24] S. Bidon, J. Y. Tournet, L. Savy, and F. Le Chevalier. Bayesian sparse estimation of migrating targets for wideband radar. *IEEE Trans. Aerosp. Electron. Syst.*, 50(2) :871–886, April 2014.

- [25] J.B. Billingsley. *Low-angle Radar Land Clutter : Measurements and Empirical Models*. William Andrew Pub., 2002.
- [26] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [27] T. Blumensath and M. E. Davies. Iterative thresholding for sparse approximations. *Journal of Fourier Analysis and Applications*, 14(5) :629–654, 2008.
- [28] T. Blumensath and M. E. Davies. Iterative hard thresholding for compressed sensing. *Applied and Computational Harmonic Analysis*, 27(3) :265–274, November 2009.
- [29] S. Bourguignon and H. Carfantan. Bernoulli-Gaussian spectral analysis of unevenly spaced astrophysical data. *IEEE 13th Workshop on Statistical Signal Processing (SSP 2005)*, pages 811–816, 2005.
- [30] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathématique*, 346(9) :589 – 592, 2008.
- [31] E. J. Candès and M. B. Wakin. An Introduction To Compressive Sampling. *IEEE Signal Processing magazine*, (March 2008) :21–30, 2008.
- [32] E.J. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6) :906–956, 2014.
- [33] E.J. Candès, J.K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8) :1207–1223, 2006.
- [34] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proc. of the IEEE*, 57(8) :1408–1418, August 1969.
- [35] B. D. Carlson, E. D. Evans, and S. L. Wilson. Search radar detection and track with the hough transform. III. Detection performance with binary integration. *IEEE Trans. Aerosp. Electron. Syst.*, 30(1) :116–125, Jan 1994.
- [36] H. C. Chan. Spectral characteristics of low-angle radar ground clutter. Technical Report 1020, Defence Research Establishment Ottawa, Canada, December 1989.
- [37] V. Chandrasekaran, B. Recht, P.A. Parrilo, and A.S. Willsky. The Convex Geometry of Linear Inverse Problems. *Foundations of Computational Mathematics*, 12(6) :805–849, 2012.
- [38] S. Chen and D. Donoho. Basis Pursuit. In *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 41–44 vol.1, Oct 1994.
- [39] Y. Chen and Y. Chi. Robust spectral compressed sensing via structured matrix completion. *IEEE Trans. Inf. Theory*, 60(10) :6576–6601, Oct 2014.
- [40] Y. Chi and Y. Chen. Compressive two-dimensional harmonic retrieval via atomic norm minimization. *IEEE Trans. Signal Process.*, 63(4) :1030–1042, Feb 2015.
- [41] N. Chopin. Fast simulation of truncated Gaussian distributions. *Statistics and Computing*, 21(2) :275–288, 2011.
- [42] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11) :1413–1457, 2004.

- [43] G. Davidson, H.D. Griffiths, and S. Ablett. Analysis of high-resolution land clutter. *IEE Proceedings - Vision, Image and Signal Processing*, 151(1) :86–91, 2004.
- [44] F. Deudon, S. Bidon, O. Besson, and J. Y. Tourneret. Velocity dealiased spectral estimators of range migrating targets using a single low-PRF wideband waveform. *IEEE Trans. Aerosp. Electron. Syst.*, 49(1) :244–265, Jan 2013.
- [45] F. Deudon, S. Bidon, O. Besson, J. Y. Tourneret, M. Montécot, and F. Le Chevalier. Modified Capon and APES for spectral estimation of range migrating targets in wideband radar. In *2010 IEEE Radar Conference (RadarCon)*, pages 1280–1285, May 2010.
- [46] L. Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, 1986.
- [47] N. Dobigeon, A. O. Hero, and J. Y. Tourneret. Hierarchical Bayesian sparse image reconstruction with application to MRFM. *IEEE Trans. Image Process.*, 18(9) :2059–2070, September 2009.
- [48] N. Dobigeon and J.Y. Tourneret. Bayesian orthogonal component analysis for sparse representation. *IEEE Trans. Signal Process.*, 58(5) :2675–2685, 2010.
- [49] D. L. Donoho, Y. Tsaig, I. Drori, and J. L. Starck. Sparse solution of underdetermined systems of linear equations by Stagewise Orthogonal Matching Pursuit. *IEEE Transactions on Information Theory*, 58(2) :1094–1121, Feb 2012.
- [50] M.F. Duarte and R.G. Baraniuk. Spectral compressive sensing. *Appl. Comput. Harmon. Anal.*, 35 :111–129, 2013.
- [51] A. Fannjiang and W. Liao. Coherence pattern-guided compressive sensing with unresolved grids. *SIAM Journal on Imaging Sciences*, 5(1) :179–202, 2012.
- [52] A. Farina, F. Gini, M.V. Greco, and L. Verrazzani. High resolution sea clutter data : statistical analysis of recorded live data. *IEE Proceedings - Radar, Sonar and Navigation*, 144(3) :121–130, 1997.
- [53] W. Fishbein, S.W. Graveline, and O.E. Rittenbach. Clutter attenuation analysis. Technical report, ECOM-2808, March 1967.
- [54] R. Gatto. Some computational aspects of the generalized von Mises distribution. *Statistics and Computing*, 18(3) :321–331, 2008.
- [55] R. Gatto. Information theoretic results for circular distributions. *Statistics*, 43(4) :409–421, 2009.
- [56] R. Gatto and S.R. Jammalamadaka. The generalized von Mises distribution. *Statistical Methodology*, 4(3) :341–353, 2007.
- [57] R. Giri and B. Rao. Type I and Type II Bayesian methods for sparse signal recovery using scale mixtures. *IEEE Trans. Signal Process.*, 64(13) :3418–3428, July 2016.
- [58] A. Gretsistas and M.D. Plumley. An alternating descent algorithm for the off-grid DoA estimation problem with sparsity constraints. In *Proc. of the 20th European Signal Processing Conference (EUSIPCO 2012)*, pages 874–878, 2012.
- [59] M.A. Herman and T. Strohmer. General deviants : An analysis of perturbations in compressed sensing. *IEEE Journal of Selected Topics in Signal Processing*, 4(2) :342–349, April 2010.

- [60] J.A. Hogbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement*, 15 :417–426, 1974.
- [61] L. Hu, Z. Shi, J. Zhou, and Q. Fu. Compressed sensing of complex sinusoids : An approach based on dictionary refinement. *IEEE Trans. Signal Process.*, 60(7) :3809–3822, July 2012.
- [62] L. Hu, J. Zhou, Z. Shi, and Q. Fu. A fast and accurate reconstruction algorithm for compressed sensing of complex sinusoids. *IEEE Trans. Signal Process.*, 61(22) :5744–5754, November 2013.
- [63] R. Jagannath and K.V.S. Hari. Block sparse estimator for grid matching in single snapshot DoA estimation. *IEEE Signal Process. Lett.*, 20(11), November 2013.
- [64] R. Jagannath, G. Leus, and R. Pribić. Grid matching for sparse signal recovery in compressive sensing. In *Proc. of the 9th European Radar Conference*, pages 111–114, 2012.
- [65] S. Ji, Y. Xue, and L. Carin. Bayesian compressive sensing. *IEEE Trans. Signal Process.*, 56(6) :2346–2356, June 2008.
- [66] S.M. Kay. *Fundamentals of Statistical Signal Processing, Volume I : Estimation Theory*. Signal Processing series. Prentice-Hall PTR, 1993.
- [67] S.M. Kay. *Fundamentals of Statistical Signal Processing, Volume II : Detection Theory*. Signal Processing series. Prentice-Hall PTR, 1998.
- [68] J. Kormylo and J. Mendel. Maximum likelihood detection and estimation of Bernoulli-Gaussian processes. *IEEE Transactions on Information Theory*, 28(3) :482–488, May 1982.
- [69] S. Kotz, N. Balakrishnan, and N. L. Johnson. *Continuous Multivariate Distributions*, volume 1. Wiley Series in Probability and Statistics, 2 edition, 2000.
- [70] O.A. Krasnov, G.P. Babur, Z. Wang, L. P. Lighthart, and F. van der Zwan. Basics and first experiments demonstrating isolation improvements in the agile polarimetric FM-CW radar – PARSAX. *International Journal of Microwave and Wireless Technologies*, 2 :419–428, 8 2010.
- [71] F. Le Chevalier. Procédé et dispositif de détection de cibles pour radar Doppler à impulsions non ambiguës à large bande, 1996. French Patent 96 08509.
- [72] J. Li and P. Stoica. An adaptive filtering approach to spectral estimation and SAR imaging. *IEEE Trans. Signal Process.*, 44(6) :1469–1484, June 1996.
- [73] Y. Li and Y. Chi. Off-the-grid line spectrum denoising and estimation with multiple measurement vectors. *IEEE Transactions on Signal Processing*, 64(5) :1257–1269, 2016.
- [74] D. Malioutov, M. Cetin, and A.S. Willsky. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.*, 53(8) :3010–3022, August 2005.
- [75] S.G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12) :3397–3415, 1993.
- [76] P. A. C. Marques and J. M. B. Dias. Velocity estimation of fast moving targets using a single SAR sensor. *IEEE Trans. Aerosp. Electron. Syst.*, 41(1) :75–89, Jan 2005.

- [77] W. Melvin and J. Scheer, editors. *Principles of Modern Radar : Advanced Techniques*. Institution of Engineering and Technology, Edison, NJ, 2012.
- [78] T.J. Mitchell and J.J. Beauchamp. Bayesian Variable Selection in Linear Regression. *Journal of the American Statistical Association*, 83(404) :1023–1032, 1988.
- [79] D. Needell and J.A. Tropp. CoSaMP : Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3) :301 – 321, 2009.
- [80] D. Needell and R. Vershynin. Signal recovery from incomplete and inaccurate measurements via regularized Orthogonal Matching Pursuit. *IEEE Journal of Selected Topics in Signal Processing*, 4(2) :310–316, April 2010.
- [81] S. Ni and D. Sun. Bayesian Estimates for Vector Autoregressive Models. *Journal of Business & Economic Statistics*, 23(1) :105–117, jan 2005.
- [82] S. Paris, D. Mary, and A. Ferrari. Detection tests using sparse models, with application to hyperspectral data. *IEEE Trans. Signal Process.*, 61(6) :1481–1494, March 2013.
- [83] T. Park and D.A. van Dyk. Partially collapsed Gibbs samplers : Illustrations and applications. *Journal of Computational and Graphical Statistics*, 18(2) :283–305, 2009.
- [84] J.T. Parker and L.C. Potter. A Bayesian perspective on sparse regularization for STAP post-processing. In *2010 IEEE Radar Conference (RadarCon)*, pages 1471–1475, 2010.
- [85] Y.C. Pati, R. Rezaifar, and P.S. Krishnaprasad. Orthogonal Matching Pursuit : recursive function approximation with applications to wavelet decomposition. *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, pages 40–44, 1993.
- [86] N. Petrov and F. Le Chevalier. Fast implementation of Iterative Adaptive Approach for wideband unambiguous radar detection. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, pages 1207–1211, Aug 2015.
- [87] N. Petrov and F. Le Chevalier. Iterative Adaptive Approach for unambiguous wideband radar target detection. In *2015 European Radar Conference (EuRAD)*, pages 45–48, Sept 2015.
- [88] J. Qian, X. Lv, M. Xing, L. Li, and Z. Bao. Motion parameter estimation of multiple ground fast-moving targets with a three-channel Synthetic Aperture Radar. *IET Radar, Sonar Navigation*, 5(5) :582–592, June 2011.
- [89] C.P. Robert. *Le choix Bayésien*. Springer-Verlag Paris, 2006.
- [90] C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2004.
- [91] H. Rohling. Radar CFAR thresholding in clutter and multiple target situations. *IEEE Trans. Aerosp. Electron. Syst.*, AES-19(4) :608–621, July 1983.
- [92] M. Rosenbaum and A.B. Tsybakov. Sparse recovery under matrix uncertainty. *The Annals of Statistics*, 38(5) :2620–2651, 2010.
- [93] T. Sakamoto, A. Matsuoka, and H. Yomo. Estimation of Doppler velocities from sub-Nyquist ultra-wideband radar measurements. *IEEE Sensors Journal*, 16(23) :8557–8565, Dec 2016.
- [94] K. J. Sangston, F. Gini, and M. S. Greco. Coherent radar target detection in heavy-tailed compound-gaussian clutter. *IEEE Trans. Aerosp. Electron. Syst.*, 48(1) :64–77, Jan 2012.

- [95] D.C. Schleher. *MTI radar*. Radar Library. Artech House, 1978.
- [96] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1) :10–21, Jan 1949.
- [97] S.Pfyffer and R. Gatto. An efficient simulation algorithm for the generalized von Mises distribution of order two. *Computational Statistics*, 28(1) :255–268, 2013.
- [98] P.J.F. Swart and P. Snoeij. Modeling Radar Scatter from Distributed Targets Using a Coupled Scatterer Approach. *IEEE Trans. Geosci. Remote Sens.*, 32(2) :296–306, 1994.
- [99] Z. Tan, Y. Eldar, and A. Nehorai. Direction of arrival estimation using co-prime arrays : A super resolution viewpoint. *IEEE Trans. Signal Process.*, 62(21) :5565–5576, 2014.
- [100] Z. Tan and A. Nehorai. Sparse direction of arrival estimation using co-prime arrays with off-grid targets. *IEEE Signal Process. Lett.*, 21(1), January 2014.
- [101] Z. Tan, P. Yang, and A. Nehorai. Joint sparse recovery method for compressed sensing with structured dictionary mismatches. *IEEE Trans. Signal Process.*, 62(19) :4997–5008, 2014.
- [102] G. Tang, B.N. Bhaskar, P. Shah, and B. Recht. Compressed sensing off the grid. *IEEE Trans. Inf. Theory*, 59(11) :7465–7490, November 2013.
- [103] O. Teke, A.C. Gurbuz, and O. Arikan. Perturbed Orthogonal Matching Pursuit. *IEEE Trans. Signal Process.*, 61(24), December 2013.
- [104] R. Q. Tian, Z. P. Fan, C. Y. Lin, Q. L. Bao, and Z. P. Chen. A detection algorithm for radar weak targets under the interference of strong targets. In *2016 Progress in Electromagnetic Research Symposium (PIERS)*, pages 560–565, Aug 2016.
- [105] R. Tibshirani. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society*, 58(1) :267–288, 1996.
- [106] M.E. Tipping. Sparse Bayesian Learning and the Relevance Vector Machine. *Journal of Machine Learning Research*, 1 :211 – 244, 2001.
- [107] J.A. Tropp and A.C. Gilbert. Signal recovery from random measurements via Orthogonal Matching Pursuit. *IEEE Trans. Inf. Theory*, 53(12) :4655–4666, 2007.
- [108] J. Tsao and B.D. Steinberg. Reduction of sidelobe and speckle artifacts in microwave imaging : the CLEAN technique. *IEEE Trans. Antennas Propag.*, 36(4) :543–557, 1988.
- [109] L.M.H. Ulander, H. Hellsten, and G. Stenström. Synthetic-Aperture Radar Processing using Fast Factorised Back-Projection. *IEEE Trans. Aerosp. Electron. Syst.*, 39(3) :1–32, 2003.
- [110] D.A. van Dyk and T. Park. Partially collapsed Gibbs samplers : Theory and methods. *Journal of the American Statistical Association*, 103(482) :790–796, 2008.
- [111] P. Wang, H. Li, and B. Himed. Moving target detection using distributed mimo radar in clutter with nonhomogeneous power. *IEEE Trans. Signal Process.*, 59(10) :4809–4820, Oct 2011.
- [112] K.D. Ward, C.J. Baker, and S. Watts. Maritime surveillance radar. I. Radar scattering from the ocean surface. *IEE Proceedings*, 137(2) :51–62, 1990.
- [113] S. Watts. Radar detection prediction in k-distributed sea clutter and thermal noise. *IEEE Trans. Aerosp. Electron. Syst.*, AES-23(1) :40–45, Jan 1987.

- [114] P. Welch. The use of Fast Fourier Transform for the estimation of power spectra : A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.*, 15(2) :70–73, Jun 1967.
- [115] A. Wiesel, O. Bibi, and A. Globerson. Time varying autoregressive moving average models for covariance estimation. *IEEE Trans. Signal Process.*, 61(11) :2791–2801, 2013.
- [116] J. Xu, J. Yu, Y. N. Peng, and X. G. Xia. Radon-Fourier transform for radar target detection, I : Generalized Doppler filter bank. *IEEE Trans. Aerosp. Electron. Syst.*, 47(2) :1186–1202, April 2011.
- [117] Z. Yang and L. Xie. On gridless sparse methods for line spectral estimation from complete and incomplete data. *IEEE Trans. Signal Process.*, 63(12) :3139–3153, 2015.
- [118] Z. Yang and L. Xie. Exact joint sparse frequency recovery via optimization methods. *IEEE Trans. Signal Process.*, 64(19) :5145–5157, 2016.
- [119] Z. Yang, L. Xie, and C. Zhang. Off-grid direction of arrival estimation using sparse Bayesian inference. *IEEE Trans. Signal Process.*, 61(1) :38–43, January 2013.
- [120] T. Yardibi, J. Li, P. Stoica, M. Xue, and A. B. Baggeroer. Source localization and sensing : A nonparametric iterative adaptive approach based on weighted least squares. *IEEE Trans. Aerosp. Electron. Syst.*, 46(1) :425–443, Jan 2010.
- [121] L. Yu, Hong H. Sun, , J.P. Barbot, and G. Zheng. Bayesian Compressive Sensing for clustered sparse signals. In *Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2011)*, number 1, pages 3948–3951, 2011.
- [122] C. Yuejje, L. L. Scharf, A. Pezeshki, and A.R. Calderbank. Sensitivity to basis mismatch in compressed sensing. *IEEE Trans. Signal Process.*, 59(5) :2182–2195, May 2011.
- [123] S. Zhang, Y. Liu, and X. Li. Autofocusing for sparse aperture ISAR imaging based on joint constraint of sparsity and minimum entropy. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 10(3) :998–1011, 2017.
- [124] Y. Zhang, Z. Ye, X. Xu, and N. Hu. Off-grid DoA estimation using array covariance matrix and block-sparse Bayesian learning. *Signal Process.*, 98 :197–201, 2014.
- [125] H. Zhu, G. Leus, and G.B. Giannakis. Sparsity-cognizant total least-squares for perturbed compressive sampling. *IEEE Trans. Signal Process.*, 59(5) :2002–2016, 2011.
- [126] X. Zhuge and A.G. Yarovoy. A sparse aperture MIMO-SAR-based UWB imaging system for concealed weapon detection. *IEEE Trans. Geosci. Remote Sens.*, 49(1) :509–518, 2011.

Résumé — Les travaux menés lors de cette thèse s’inscrivent dans le cadre général de la détection de cibles en utilisant une forme d’onde non-conventionnelle large bande. L’utilisation d’une forme d’onde large bande à faible PRF a été proposée par le passé une alternative aux traitements multi-PRF qui limitent le temps d’illumination de la scène. En effet, l’augmentation de la bande instantanée permet d’obtenir une meilleure résolution distance ; les cibles rapides sont alors susceptibles de migrer lors du temps de traitement, mais ce phénomène de couplage distance-vitesse peut être mis à profit pour lever les ambiguïtés. L’objectif de la thèse est alors de développer, pour une forme d’onde large bande avec faible PRF, des traitements prenant en compte la migration des cibles et capables de lever les ambiguïtés vitesse dans des scénarios réalistes. Les travaux se basent sur un algorithme de représentation parcimonieuse non-ambigüe de cibles migrantes, dans un cadre algorithmique Bayésien. Cet algorithme est en revanche développé sous certaines hypothèses, et des travaux de robustification sont alors entrepris afin de l’utiliser sur des scénarios plus réalistes. Dans un premier temps, l’algorithme est robustifié au désalignement des cibles par rapport à la grille d’analyse, puis modifié pour prendre également en compte une possible composante diffuse de bruit. Il est également remanié pour estimer correctement une scène comportant une forte diversité de puissance, où des cibles fortes masquent potentiellement des cibles faibles. Les différents algorithmes sont validés à la fois sur des données synthétiques et expérimentales.

Mots clés : traitement du signal, radar large bande, migration distance, estimation Bayésienne, représentations parcimonieuses.

Abstract — The work conducted during this PhD falls within the general context of radar target detection using a non-conventional wideband waveform. More precisely, the use of a low-PRF wideband waveform has been proposed in the past as an alternative to the classical staggered-PRF processing used to mitigate velocity ambiguities that limits dwell time. Increasing the instantaneous bandwidth improves range resolution ; fast moving targets are then likely to migrate during the coherent processing interval. This range-velocity coupling can then be used to mitigate velocity ambiguities. This PhD thesis aims at developing an algorithm able to provide unambiguous estimation of migrating targets using a low-PRF wideband waveform. It is based on a sparse representation algorithm able to unambiguously estimate migrating targets, within a Bayesian framework. However, this algorithm is developed under some hypothesis, and then requires robustification to be used on more realistic scenarii. First, the algorithm is robustified to the case of off-grid targets, and then upgraded to take into account a possible diffuse clutter component. On the other hand, the reference algorithm is modified to accurately estimate high dynamic range scenes where weak targets compete with strong targets. All the developed algorithms have been validated on synthetic and experimental data recorded by the PARSAX radar from the Technical University of Delft, The Netherlands.

Keywords : signal processing, wideband radar, range migration, Bayesian inference, sparse representations.
