Comparaison des outils optique et radar en polarimétrie bistatique

Nicolas Trouvé

▶ **To cite this version:**

Nicolas Trouvé. Comparaison des outils optique et radar en polarimétrie bistatique. Electromagnétisme. Ecole Polytechnique X, 2011. Français. NNT : 2011EPXX0078 . tel-00676316v2

HAL Id: tel-00676316

https://theses.hal.science/tel-00676316v2

Submitted on 3 Apr 2012

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

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

Laboratoire de physique
des interfaces et couches minces
ÉCOLE POLYTECHNIQUE

UMR 7647

École Doctorale de Polytechnique

Comparaison des outils optique et radar en polarimétrie bistatique

THÈSE

présentée et soutenue publiquement le 22 Novembre 2011

pour l'obtention du

Doctorat de l'Ecole Polytechnique

(Spécialité Physique)

par

Nicolas Trouvé

Composition du jury

Président :	Mr Vincent Devlaminck, *Université de Lilles*
Rapporteurs :	Mr Eric Pottier, *Université de Rennes 1* Mr François Goudail, *Institut d'optique Graduate School*
Examinateurs :	Mr Antonello de Martino (*Directeur de thèse, Ecole Polytechnique, LPICM*) Mr Shane Cloude, *AELc* Mr Xavier Orlik, *ONERA* Mr Jihad Zallat, *LSIIT* Mrs Elise Colin Koeniguer (*Encadrante de thèse, ONERA*)

Office National d'Etudes et de Recherches Aérospatiales
ONERA
THE FRENCH AEROSPACE LAB

Remerciements

Vous qui lisez ceci et qui vous demandez si j'ai pensé à vous pour mes remerciements : la réponse est oui !
Merci !

"Son, if you really want something in this life, you have to work for it. Now quiet ! They're about to announce the lottery numbers."
Homer J. Simpson

Table des matières

Introduction à la polarimétrie **9**

1 Paramètres de polarisation d'une onde plane 12

1.1 Vecteur de Jones 13

1.2 Vecteur de Stokes 13

2 Paramètres polarimétriques d'une cible, en optique et en radar 14

2.1 Matrice de Sinclair, ou matrice de Jones 15

2.2 Vecteur de diffusion 15

2.3 Matrice de cohérence et de covariance 16

2.4 Matrice de Kennaugh / Mueller 16

3 Vocabulaire radar et optique 17

3.1 Les propriétés d'une cible. 19

3.2 Les matrices de mesure d'une cible 19

3.3 Les conditions géométriques de mesure 20

3.4 Tableau récapitulatif 20

I Influence de la géométrie bistatique sur la mesure **23**

1 Le choix du plan d'analyse d'une cible **25**

1.1 Introduction sur la géométrie de mesure polarimétrique 25

1.1.1 Pourquoi travailler en configuration bistatique ? 26

1.1.2 Conventions 27

1.2 Ensemble des paramètres géométriques imposés et choisis. 29

1.2.1 Monostatisme et géométrie 31

1.2.2 Monostatisme et réciprocité 33

1.2.3 Bistatisme et problèmes principaux 33

1.2.4 Définition des bases polarimétriques bistatiques : méthode classique 36
1.2.5 Définition des bases polarimétriques bistatiques : méthode du plan de diffusion 37
1.2.6 Relations de passage entre les deux conventions 39
1.2.7 Pourquoi choisir la convention du plan de diffusion, et quelles conséquences ? 40

2 Les cibles élémentaires déterministes 43
2.1 Analyse théorique d'une sphère à basse fréquence 43
2.1.1 Définition de quelques paramètres polarimétriques 43
2.1.2 Une sphère diélectrique à basse fréquence dans le plan de diffusion 45
2.1.3 Une sphère diélectrique étudiée à basse fréquence dans la convention classique de mesure 47
2.1.4 La sphère métallique à basse fréquence 47
2.2 Etude théorique du cylindre vertical à basse fréquence 48
2.2.1 Modélisation simplifiée, inspirée de la sphère diélectrique à basse fréquence 48
2.2.2 Modélisation du cylindre infini par résolution des Equations de Maxwell 52
2.3 Analyse et mesure d'une à quatre sphères à basse fréquence 52
2.3.1 La chambre anéchoïque BaBi 52
2.3.2 Théorie 54
2.3.3 Résultats 55
2.4 Conclusion 58

3 Vers les cibles dépolarisantes 63
3.1 Mesure d'un nuage de sphères dans la zone de résonance 63
3.1.1 Objectif et nouveaux paramètres 63
3.1.2 La cible 64
3.1.3 Résultats et analyse 65
3.2 Analyse et mesure d'un nuage de diffuseurs cylindriques proches de la verticale 70
3.2.1 Choix et réalisation de la cible 70
3.2.2 Mesures effectuées en chambre anéchoïque 72

3.2.3 Analyse des résultats obtenus 73

II Estimation statistique en radar 77

1 Modèles statistiques et distances pour une mesure polarimétrique 81

1.1 Modélisation du signal polarimétrique : généralités 81
1.2 Modèle Gaussien 83
1.2.1 Vecteur de diffusion 84
1.2.2 Matrice de covariance ou cohérence 84
1.3 Modèle Non Gaussien 84
1.3.1 Modèle Multiplicatif 85
1.3.2 Modèle à lois connues 86
1.3.3 Modèle à lois inconnues : les SIRV 88
1.3.4 Caractérisation des distributions non gaussiennes : Méthode des Log cumulants 89
1.3.5 Exemple sur des images SAR à haute résolution 93
1.4 Application aux tests statistiques en polarimétrie 98
1.4.1 Rapport de Vraisemblance 98
1.4.2 Effets liés aux nombres d'observations 111
1.4.3 Distances issues du monde de la théorie de l'information 114
1.4.4 Comparaison des performances sur données simulées 119

2 La segmentation 129

2.1 Application des outils statistiques pour la segmentation d'images polarimétriques 129
2.1.1 La segmentation d'image 129
2.1.2 Principe 130
2.1.3 Processus détaillé 130
2.1.4 Utilisation des distances statistiques polarimétriques 134
2.2 Critères images 135
2.2.1 Principe 136
2.2.2 Définition des sous-ensembles flous, et fonction d'appartenance . 139
2.2.3 Règles d'inférence 140

2.2.4 Lignes et bords . . . 144
2.2.5 Fusion des différentes règles et décision. . . . 147
2.2.6 Gestion des grands volumes de données . . . 148
2.3 Résultats et synthèse . . . 149
2.4 Perspectives . . . 151

3 La détection 153
3.1 Détecteurs Gaussiens et amplitude inconnue . . . 154
3.2 Les détecteurs hybrides . . . 162
3.3 Performances . . . 167
3.3.1 Performances théoriques. . . . 167
3.3.2 Performances sur une image SAR synthétique . . . 170
3.3.3 Utilisation de la segmentation d'image pour l'amélioration des processus de détection. . . . 174
3.3.4 Conclusion . . . 177
3.3.5 Perspectives . . . 178

III Les décompositions polarimétriques 181

1 État de l'art sur les décompositions polarimétriques 183
1.1 Valeurs propres et vecteurs propres de cibles dépolarisantes . . . 185
1.1.1 Cloude Pottier . . . 185
1.1.2 Alternative . . . 186
1.1.3 Touzi . . . 187
1.2 Décompositions appuyées sur des modèles. . . . 187
1.2.1 Freeman Durden . . . 188
1.2.2 Yamaguchi . . . 188
1.2.3 Limitations . . . 189
1.3 Lu Chipman et Décomposition Symétrique d'Ossikovski . . . 189
1.3.1 La décomposition polaire des milieux non dépolarisants . . . 189
1.3.2 Lu Chipman . . . 189
1.3.3 Décomposition "reverse" . . . 191
1.3.4 Décomposition symétrique d'Ossikovski . . . 191
1.4 Visualisation et compositions colorée . . . 192

2 Etude des milieux dépolarisants **195**
2.1 Forêts et Agriculture 195
2.1.1 Objectifs 196
2.1.2 Exemple : Banlieue de Brétigny, image monostatique en bande X196
2.1.3 Exemple de forêt en monostatique : Nezer, image monostatique de forêt en bande P 199
2.2 Tissus cancéreux 203
2.3 Apport de la configuration bistatique et perspectives 211
2.3.1 Une forêt de clous 212
2.3.2 Une forêt de nanotubes 213

3 Exploitation du contraste entre milieux déterministes et milieux naturels **219**
3.1 Milieux urbains 219
3.2 PolinSAR 220
3.2.1 Combinaison d'une segmentation polarimétrique avec une méthode d'optimisation de cohérence 220
3.2.2 De la compression des données en segmentation vers la restitution 3D 224

Conclusion Générale **229**

Bibliographie **231**

Introduction à la polarimétrie

La polarimétrie, étude de la polarisation des ondes électromagnétiques, permet d'enrichir la connaissance d'un objet, matériau ou surface, en observant son interaction avec une onde électromagnétique polarisée.

Suivant les longueurs d'onde utilisées, les théories liées à la polarimétrie se sont développées depuis la fin des années 40 en suivant des voies différentes. Pour la polarimétrie radar, il s'agit du formalisme des matrices de Sinclair 2x2 adaptée au cas de la rétrodiffusion (cas dit " monostatique "). Pour la polarimétrie optique dite de "Mueller", on mesure les éléments d'une matrice 4x4 de densité de puissance moyennée, pour toute configuration géométrique des antennes d'émission et de réception.

Polarimétrie optique et radar diffèrent sous les aspects suivants : - Les moyens de mesure (cohérente ou incohérente) et donc les grandeurs mesurées et estimées - Les théories polarimétriques utilisées (décomposition matricielles) - Les configurations géométriques de la mesure (monostatiques à l'heure actuelle en radar, bistatiques en optique) - Les échelles des longueurs d'onde et donc les cibles envisagées.

Cette thèse porte sur le croisement des techniques et théories employées dans les applications de la polarimétrie radar et celles de la polarimétrie optique de Mueller.

Le bénéfice attendu de ces croisements est multiple :

– A partir de mesures réalisées dans le domaine optique (mesure de la matrice de Mueller d'un élément), nous sommes en mesure de déduire tous les paramètres habituellement utilisés en radar dans le cadre des cibles dites " aléatoires " (milieux dépolarisants tels que la forêt) et vice-versa. Nous pouvons donc espérer mieux comprendre la nature physique des phénomènes de dépolarisation, en comparant les différents moyens de mesure de celle-ci.

– Aujourd'hui, l'un des enjeux de la polarimétrie radar concerne la généralisation des outils théoriques existants à la configuration " bistatique ", notamment pour l'étude des milieux aléatoires. Cette configuration se rencontre par exemple lorsque les antennes d'émission et de réception ne se trouvent pas au même endroit, ou bien dans le cas des milieux dits " non réciproques " (rotation de Faraday par exemple). Si cette configuration est "nouvelle" pour les radaristes en polarimétrie, elle est totalement classique en radar. Aussi cette thèse est donc une des rares à notre connaissance qui tente d'apporter des réponses au problème de la polarimétrie radar bistatique en s'inspirant des interprétations traditionnellement réservées au domaine de l'optique, et en comparant les phénomènes physiques aux deux échelles.

En optique comme en radar, nous cherchons à exhiber des scénarios où l'information polarimétrique est suffisamment pertinente par rapport à l'information énergétique ou fréquentielle pour justifier son utilisation. L'emploi de la polarimétrie entraîne en effet un coût matériel et dans le cas du radar, une perte de résolution pour l'analyse des propriétés de dépolarisation. La problématique générale est donc l'extraction de l'information utile à partir d'une mesure polarimétrique. Cette mesure peut, comme toute mesure, être affectée d'une erreur d'étalonnage, d'effets induits par la géométrie, de bruit (bruit thermique, bruit de speckle, etc.).

L'information polarimétrique est fondamentalement semblable pour les micro-ondes (radar), l'infrarouge et le visible (optique). Cependant les moyens de mesure varient très fortement et le contexte fait donc que l'ensemble des outils développés alors (étalonnage, configuration géométrique, filtrage du bruit) divergent entre les différentes communautés.

Ces divergences ont fait que, jusqu'à aujourd'hui, relativement peu de liens ont été tissés entre les deux mondes. Elles s'illustrent notamment par les points suivants :

- La géométrie d'acquisition est classiquement bistatique en optique et encore majoritairement (voire exclusivement) monostatique en radar. De ce fait, l'optique possède des outils très nettement en avance pour traiter les problèmes liés aux géométries particulières. Cette particularité s'illustre sur certains protocoles expérimentaux comme ceux d'étalonnage, ainsi que sur les traitements polarimétriques qui sont déjà adaptés au bistatisme.
- La mesure radar est une mesure cohérente (intensité et phase) alors que la mesure optique est une mesure d'énergie, fréquemment obtenue à l'aide d'un éclairage incohérent. Le niveau et la nature du bruit sont donc complètement différents, ainsi que les grandeurs physiques mesurées : quand l'optique mesure directement des grandeurs quadratiques du champ électrique, le radar doit, lui, effectuer des estimations statistiques de ces mêmes grandeurs. Il s'agit d'une problématique exclusive au radar qui a des conséquences très importantes.
- Les traitements faits a posteriori sur les données suivent aussi des approches sensiblement différentes. En radar ce sont les décompositions polarimétriques en sommes qui sont les plus populaires (Cloude Pottier, Freeman Durden et équivalents, Touzi). Par contre en optique les traitements sont le plus souvent multiplicatifs (Lu et Chipman, Ossikovski symétrique).

De nos jours l'importance des travaux existant dans chaque discipline est telle qu'il est certain que l'utilisation transverse des outils appartenant aux mondes de la polarimétrie optique et de la polarimétrie radar puisse alimenter de façon croisée leurs homologues pour une amélioration commune. C'est ce croisement des outils qui est le sujet de cette thèse.

Dans un premier temps nous nous sommes intéressés à l'impact des géométries particulières sur les mesures polarimétriques. Pour cela nous avons étudié les modèles de diffusion de Mie, ainsi qu'un outil de modélisation du cylindre infini tiré d'un simulateur existant (COSMO). Ces modèles ont tous été confrontés à des mesures en chambre anéchoïque. Nous avons également proposé un modèle géométrique simple pour retrouver les comportements polarimétriques d'un cylindre fin [1]. Ce travail a notamment donné lieu à une publication dans la revue IEEE TGRS [2]. Il fait l'objet de la deuxième partie de cette thèse.

Dans un second temps nous avons travaillé sur l'estimation de la matrice de covariance polarimétrique dans le cadre du radar. L'objectif est ici de proposer des méthodes d'estimations exploitant l'information spatiale tout en étant adaptées aux environnements difficiles tel que l'environnement urbain. Pour cela nous avons proposé une méthode de croissance de régions s'appuyant sur un modèle non gaussien (SIRV) ainsi que sur des critères de forme à l'aide de la logique floue. Ce sujet a notamment donné lieu à soumission dans la revue IEEE TGRS. Il constitue la troisième partie de la thèse.

Enfin, une fois les problèmes liés à la géométrie de mesure, d'estimation statistiques et de prétraitements abordés, nous pouvons considérer les "traitements polarimétriques" proprement dits, et les diverses applications auxquelles ces traitements peuvent servir. C'est ce qui fera l'objet de la dernière partie de cette thèse.

1 Paramètres de polarisation d'une onde plane

La polarisation d'une onde plane est la courbe décrite par le champ électrique **E** dans le plan d'onde (perpendiculaire à la direction de propagation) au cours du temps. Si cette courbe est parfaitement définie nous parlons alors d'onde complètement polarisée (CP) et la description de cette courbe détermine parfaitement la polarisation de l'onde. Si par contre une composante aléatoire intervient dans la courbe en fonction du temps et de l'espace, nous parlons alors d'onde partiellement polarisée (PP). Pour pouvoir travailler avec ce type d'onde nous aurons recours à des techniques de signal aléatoire et nous ferons donc intervenir des moyennes statistiques.

La position initiale du vecteur champ électrique **E**, donnée par la phase absolue, ne contient pas d'information pertinente pour décrire la polarisation d'une onde. Ainsi dans un repère $(\mathbf{H}, \mathbf{V})$ (H pour horizontal, V pour vertical) du plan d'onde le champ **E** peut s'écrire :

$$\mathbf{E} = \begin{pmatrix} E_h \\ E_v \end{pmatrix} = \begin{pmatrix} a_h \cos(\omega t) \\ a_v \cos(\omega t + \delta) \end{pmatrix} \tag{1}$$

Nous pouvons montrer que E_h et E_v vérifient l'équation de l'ellipse définie par :

$$\left(\frac{E_h}{a_h}\right)^2 + \left(\frac{E_v}{a_v}\right)^2 - \frac{2E_v E_h \cos\delta}{a_v a_h} = (\sin\delta)^2 \tag{2}$$

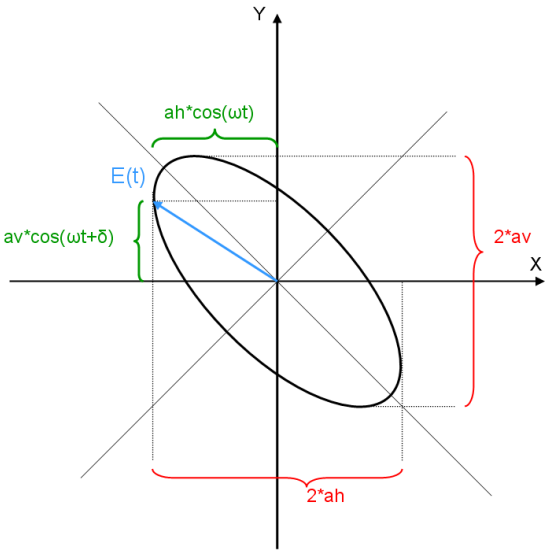


FIGURE 1 – Ellipse de polarisation du champ E dans le cas d'une onde CP

Ainsi la courbe et donc la polarisation de l'onde est une ellipse parfaitement définie par les variables a_h a_v et δ, variables représentées sur la figure 1. Le sens de balayage de l'ellipse définit le sens de polarisation droite ou gauche. .

1.1 Vecteur de Jones

Nous définissons le vecteur de Jones **P** qui contient tous les paramètres de l'état de polarisation de l'onde comme un vecteur complexe de dimension 2 de la manière suivante :

$$\mathbf{P} = \begin{pmatrix} a_h \\ a_v e^{i\delta} \end{pmatrix} \tag{3}$$

ou encore en posant $a = \sqrt{a_h^2 + a_v^2}$:

$$\mathbf{P} = \begin{pmatrix} a\cos\alpha \\ a\sin\alpha e^{i\delta} \end{pmatrix} = \begin{pmatrix} \lambda \\ \nu \end{pmatrix} \tag{4}$$

Le rapport ou coefficient de polarisation est défini par :

$$\rho = \nu/\lambda = \frac{a_v}{a_h} e^{i\delta} = \tan(\alpha) e^{i\delta} \tag{5}$$

et l'état de polarisation normé s'exprime en fonction de ρ :

$$p = \frac{1}{\sqrt{1+\rho\rho^*}} \begin{pmatrix} 1 \\ \rho \end{pmatrix} \tag{6}$$

Le vecteur de Jones suffit parfaitement à décrire une onde CP, mais pas une onde PP. Une onde PP peut être définie comme une combinaison d'une infinité d'ondes CP de polarisations différentes. Par contre, sa définition ne peut être réduite à la moyenne de ces différentes ondes CP. En effet la moyenne de plusieurs ellipses différentes n'est plus nécessairement une ellipse, comme le montre la figure 2. Une onde PP ne peut être décrite par une ellipse et nécessite plus de trois variables pour être parfaitement paramétrée.

1.2 Vecteur de Stokes

Afin de décrire au mieux une onde PP, il est possible d'introduire la matrice de cohérence **C** de son vecteur de Jones, (ici $< . >$ décrit une moyenne statistique), définie par :

$$\mathbf{C} = < \mathbf{P}.\mathbf{P^t} > = \begin{pmatrix} < E_h E_h* > & < E_h E_v* > \\ < E_v E_h* > & < E_v E_v* > \end{pmatrix} \tag{7}$$

Il est également possible de représenter cette matrice de cohérence de manière vectorielle ; le vecteur ainsi obtenu est appelé vecteur de Stokes. Il conviendra aussi bien à la description des ondes CP que des ondes PP. Pour l'obtenir, il faut projeter la matrice de cohérence sur une base de matrices de dimension 2×2.

Parmi ces bases de matrices, plusieurs choix sont possibles, notamment la base lexicographique ou la base des matrices de Pauli dont nous rappelons la définition ci dessous. Le choix d'utiliser la base des matrices de Pauli peut sembler moins naturel ; cependant de nombreuses propriétés que nous ne pouvons étudier en détail ici justifient ce choix pour l'étude de la polarimétrie.

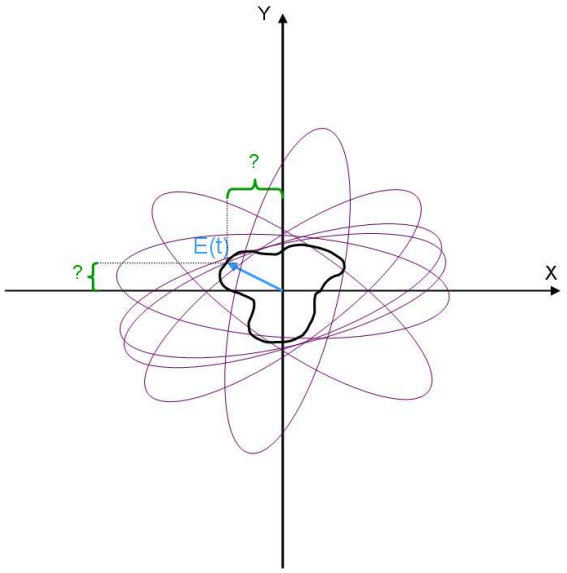


FIGURE 2 – Parcours du champ **E** dans le plan perpendiculaire à l'axe de propagation dans le cas d'une onde PP : la courbe "moyenne" décrite au cours du temps n'est pas une ellipse

Les matrices de Pauli sont les suivantes :

$$\sigma_{\mathbf{0}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \sigma_{\mathbf{1}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \sigma_{\mathbf{2}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \sigma_{\mathbf{3}} = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix} \tag{8}$$

En projetant la matrice dans cette base nous obtenons le vecteur appelé vecteur de Stokes :

$$\mathbf{g_p}(\mathbf{E}) = \begin{pmatrix} < E_h E_h* > + < E_v E_v* > \\ < E_h E_h* > - < E_v E_v* > \\ < E_h E_v* > + < E_v E_h* > \\ -j(< E_h E_v* > - < E_v E_h* >) \end{pmatrix} \tag{9}$$

Par ailleurs, le degré de cohérence d'une onde PP se définit comme :

$$\mu = \frac{< E_h E_v* >}{< E_h E_h* >< E_v E_v* >} \tag{10}$$

Ce scalaire traduit l'importance relative des phénomènes non déterministes par rapport aux phénomènes déterministes.

2 Paramètres polarimétriques d'une cible, en optique et en radar

La polarimétrie s'intéresse à la réponse de cibles vis à vis de la polarisation d'une onde incidente. L'onde émise sera toujours CP, par contre l'onde réfléchie peut être CP

(nous parlons de cible cohérente ou déterministe) ou bien PP (nous parlons alors de cible incohérente ou non-déterministe). Nous avons vu qu'il existait plusieurs façons de décrire une onde polarisée ; il existera donc plusieurs façons de décrire la réponse d'une cible. De manière générale le vecteur servant à la description de l'onde incidente et celui décrivant l'onde réfléchie sont reliés par une matrice dont les dimensions varient en fonction des vecteurs utilisés pour décrire les ondes (Jones, Stokes ...).

2.1 Matrice de Sinclair, ou matrice de Jones

Pour relier les vecteurs de Jones incident et diffusé, nous utilisons la matrice de Sinclair en radar (ou Jones en optique où la convention utilisée est différente). La mesure directe d'une telle matrice n'est possible que dans le cas des ondes radar, car la mesure radar peut s'effectuer de manière "cohérente", c'est à dire en complexe (module et phase) tandis que les mesures optiques sont incohérentes.

$$\mathbf{E_r} = \mathbf{S}.\mathbf{E_i} = \begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix}.\mathbf{E_i} \tag{11}$$

Les termes S_{hh} et S_{vv} désignent les termes copolaires, et les termes S_{hv} et S_{vh} désignent les termes contrapolaires. Entre la matrice de Sinclair et la matrice de Jones il n'existe qu'une différence de convention (sens de l'onde et des vecteurs), nous donnerons le détail de ces conventions par la suite. Pour passer de la matrice de Jones $\mathbf{J}$ à la matrice de Sinclair $\mathbf{S}$, il suffit d'utiliser les expressions :

$$\mathbf{J} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{S}^* \tag{12}$$

$$\mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{J}^* \tag{13}$$

2.2 Vecteur de diffusion

De la même manière que nous avons pu transformer la matrice de cohérence du vecteur de Jones en un vecteur (Stokes) en la projetant sur une base de matrices 2×2, nous pouvons ici aussi considérer le vecteur projection de la matrice de Sinclair, appelé vecteur de diffusion $\mathbf{k}$. Là encore, nous pouvons considérer plusieurs bases de projections différentes (Lexicographique, Pauli).

Pour la base lexicographique nous obtenons :

$$\mathbf{k_l} = \begin{pmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{pmatrix} \tag{14}$$

Pour la base de Pauli nous avons :

$$\mathbf{k_p} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}}} \begin{pmatrix} S_{hh} + S_{vv} \\ S_{hh} - S_{vv} \\ S_{hv} + S_{vh} \\ -j(S_{hv} - S_{vh}) \end{pmatrix} \tag{15}$$

Il faut noter que, pour les vecteurs de diffusion, aucune moyenne n'a encore été introduite car contrairement aux vecteurs de Stokes, ces éléments sont des complexes et ne sont donc pas homogènes à des puissances.

2.3 Matrice de cohérence et de covariance

Le vecteur de diffusion $\mathbf{k}$ peut donc être exprimé classiquement dans deux bases différentes, la base lexicographique et la base des matrices de Pauli, et est alors noté respectivement $\mathbf{k_L}$ ou $\mathbf{k_P}$.

Pour pouvoir interpréter un tel vecteur lorsqu'il est aléatoire, nous nous intéressons à ses matrices de cohérence et covariance. Nous appellerons covariance la matrice de covariance du vecteur de diffusion dans la base lexicographique et la matrice de cohérence cette même matrice si on exprime le vecteur de diffusion dans la base de Pauli.

La matrice de covariance est souvent notée $\mathbf{J}$ en radar, cependant pour ne pas la confondre avec la matrice de Jones nous la noterons $\mathbf{C}$:

$$\mathbf{C} = < \mathbf{k_L} \mathbf{k_L}^\dagger > \tag{16}$$

La matrice de cohérence est notée $\mathbf{T}$:

$$\mathbf{T} = < \mathbf{k_P} \mathbf{k_P}^\dagger > \tag{17}$$

En imagerie radar, le moyennage statistique noté $< . >$ sera souvent réalisé par un moyennage spatial sur des pixels de l'image ; en effet un véritable moyennage temporel (sur plusieurs réalisations) est difficilement envisageable pour des appareils de mesures aéroportés. La matrice de Sinclair étant directement mesurée par les radaristes, c'est donc naturellement autour de la matrice de cohérence qu'une majeure partie de la théorie a été construite. En polarimétrie la matrice $\mathbf{T}$ est plus couramment utilisée, d'une part car la base de Pauli est représentative de la réponse de cibles canoniques, d'autre part parce qu'elle facilite un certain nombre de développements mathématiques. Il faut remarquer au passage que les matrices $\mathbf{T}$ et $\mathbf{C}$ sont hermitiennes (ou auto-adjointes), $\mathbf{T} = \mathbf{T}^\dagger$ propriété mathématique intéressante qui sera très souvent exploitée par la suite.

2.4 Matrice de Kennaugh / Mueller

Si nous voulons relier le vecteur de Stokes (exprimé dans la base de Pauli) de l'onde incidente et de l'onde diffusée, la matrice qui intervient est appelée matrice de Kennaugh (radar) ou de Mueller (optique). Les vecteurs de Stokes ayant déjà été obtenus après une moyenne statistique, les matrices de Kennaugh et de Mueller peuvent donc s'obtenir sans faire de moyenne supplémentaire. La matrice de Mueller est la matrice la plus utilisée

en optique car c'est la seule mesure possible (les vecteurs de Jones étant complexes, leur mesure nécessiterait la connaissance de la phase) et c'est donc essentiellement autour de la matrice de Mueller que la théorie optique a été construite.

La matrice de Kennaugh peut être obtenue sans passer par la mesure des vecteurs de Stokes, comme cela sera souvent le cas en radar ou seules les matrices de Sinclair sont mesurées (La mesure des vecteurs de Stokes n'est pas impossible mais plus coûteuse en nombre).

Le passage de la base lexicographique à la base des matrices de Pauli peut se faire à l'aide de la matrice $\mathbf{A}$:

$$\mathbf{A} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -j \\ 0 & 0 & 1 & j \\ 1 & -1 & 0 & 0 \end{pmatrix} \tag{18}$$

Cette matrice permet également de relier $\mathbf{S}$ à $\mathbf{K}$ par :

$$\mathbf{K} = \mathbf{A^t} < \mathbf{S} \otimes \mathbf{S}^* > \mathbf{A}, \tag{19}$$

ou $\otimes$ corresponds au produit de Kronecker.

Pour passer de la matrice de Mueller à la matrice de Kennaugh, il faut transformer E_h en $-E_h$ dans l'expression du vecteur de Stokes diffusé (ce qui modifie le signe de ses deux dernières composantes), puis prendre le conjugué de ce vecteur (ce qui remodifie le signe de la dernière composante), ce qui revient à :

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \mathbf{K} \tag{20}$$

soit simplement :

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{K} \tag{21}$$

Les principales relations permettant de passer d'un type de mesure à l'autre sont synthétisées sur la figure 3.

3 Vocabulaire radar et optique

Les deux communautés, radar et optique, utilisent leur vocabulaire propre, pour des notions pourtant souvent similaires. Après avoir présenté les différentes grandeurs les plus fréquemment rencontrées, il nous paraît utile de faire un point sur les différentes terminologies consacrées aux études polarimétriques.

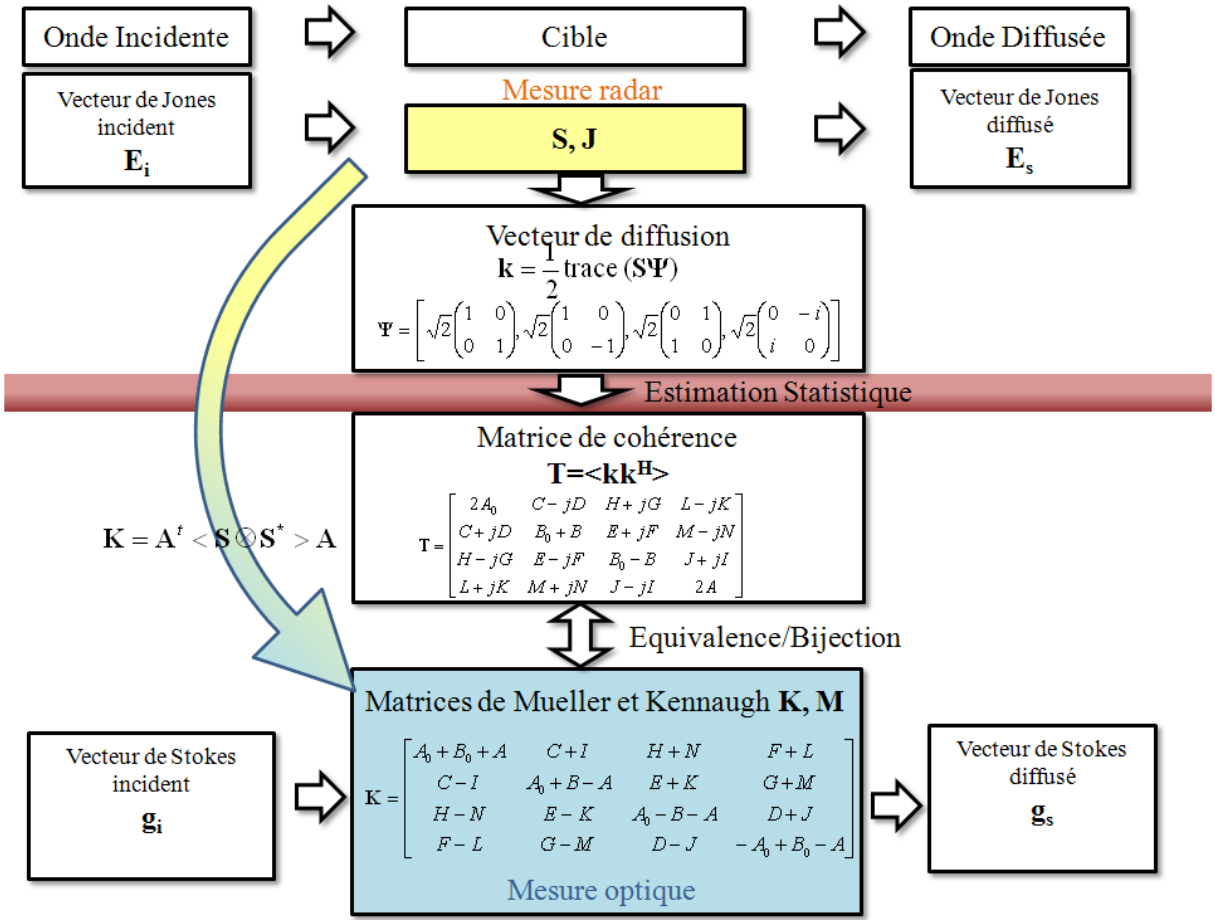


FIGURE 3 – Le lien entre les différentes grandeurs physiques polarimétriques mesurées. Les paramètres $A, A_0, B, B_0, C, D, E, F, G, H, I, J, K, L, M, N$ sont les paramètres de Huynen

3.1 Les propriétés d'une cible.

Lorsqu'une onde électromagnétique interagit avec une cible, son état de polarisation peut changer.

Dans une mesure polarimétrique où l'état initial de l'onde est totalement polarisé, on distingue alors les cibles pour lesquelles l'onde diffusée restera totalement polarisée -autrement dit l'objet est non dépolarisant- des cas où l'état de l'onde diffusée n'est plus totalement polarisé. Dans ce dernier cas, on parle de cible dépolarisante.

Toutefois, on parle parfois en radar de cibles "non déterministes", ou bien cibles naturelles ; tandis que les cibles non dépolarisantes seront qualifiées de cibles déterministes, et parfois par abus de langage, de cibles humaines ou manufacturées. Aux échelles de longueur d'onde radar, ce sont en effet en général ces types de cibles (véhicules, bâtiments) qui ne dépolarisent que peu ou pas les ondes.

3.2 Les matrices de mesure d'une cible

Matrice de diffusion

La matrice de diffusion est une matrice complexe 2x2 qui décrit la transformation de la polarisation d'une onde avant et après diffusion. La matrice de Jones et la matrice de Sinclair sont deux formes de la matrice de diffusion. Elles sont utilisées dans les conventions d'alignement de diffusion avant (FSA) et d'alignement de rétrodiffusion (BSA) respectivement, donc typiquement en optique et en radar. Les conventions d'alignement seront decrites plus en détail dans la première partie du document.

Il faut toutefois prendre garde au fait que la transformation du vecteur de Jones ne peut être utilisée que si les états initial et final sont tous deux totalement polarisés, autrement dit si l'objet est non dépolarisant. En revanche, la transformation du vecteur de Stokes est utilisable dans tous les cas.

Matrice de Stokes

La matrice de Stokes est une matrice réelle 4x4. Elle transforme le vecteur de Stokes de l'onde incidente en vecteur de Stokes de l'onde diffusée, pour un diffuseur particulier illuminé par un radar polarimétrique.

La définition du vecteur de Stokes de l'onde diffusée dépend là encore de la convention utilisée relativement au système de coordonnées. Compte tenu de la confusion possible, deux formes spécialisées de la matrice de Stokes sont en usage, à savoir la matrice de Mueller utilisée avec la convention FSA et la matrice de Kennaugh utilisée avec la convention BSA.

On parle parfois de matrice de Stokes modifiée, ou de matrice de Mueller modifiée. Dans ce cas, la matrice relie une forme modifiée de l'opérateur de diffusion de Stokes : celui-ci utilise les intensités individuelles des composantes horizontale et verticale du champ électrique, plutôt que leur somme et leur différence.

3.3 Les conditions géométriques de mesure

Dans la plupart des cas, l'émetteur et le récepteur du radar partagent une électronique et une antenne commune. On parle alors de radar monostatique. Rien n'empêche cependant de considérer un système radar où l'émetteur et le récepteur sont séparés ; on parle alors de radar bistatique, ou même de configuration multistatique, si l'on a un émetteur et plusieurs récepteurs distincts. Lorsque l'on parle de radar bistatique, on suppose implicitement que l'émetteur et le récepteur sont réellement séparés (soit du point de vue de la distance, soit d'un point de vue angulaire).

Dans le cas d'un radar monostatique, les cibles sont donc implicitement étudiées dans le cadre de la rétrodiffusion.

3.4 Tableau récapitulatif

Les termes consacrés sont récapitulés dans le tableau 1.

Terme générique	radar	optique
Convention	BSA	FSA
Emetteur et récepteur confondus	mesure monostatique	mesure en réctrodiffusion
matrice de diffusion	matrice de Sinclair	matrice de Jones
matrice de Stokes	matrice de Kennaugh	matrice de Mueller
vecteur de Stokes	vecteur de Kennaugh	vecteur de Stokes
Cible dépolarisante	cible non déterministe	cible dépolarisante
Cible non dépolarisante	cible déterministe	cible non dépolarisante
-	cible humaine ou manufacturée	cible de Mueller-Jones

TABLE 1 – Vocabulaire polarimétrique consacré aux domaines radar et optique

Il est important de noter que les outils polarimétriques peuvent se référer à la fois à la description de l'onde électromagnétique (onde incidente et onde diffusée) et à la description de la cible (comment celle ci se comporte par rapport à la polarisation de l'onde incidence). Les différentes grandeurs rencontrées sont ainsi synthétisées sur la figure 4.

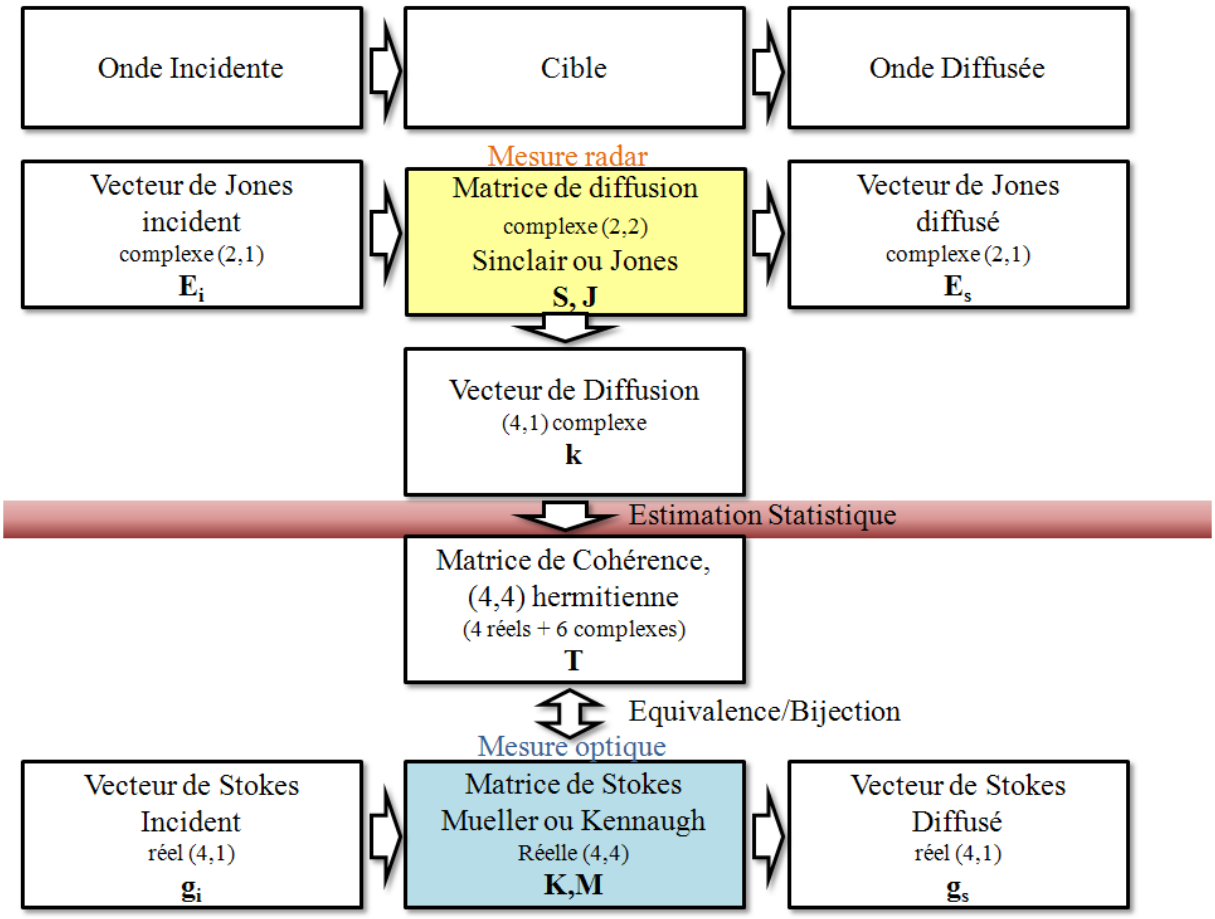


FIGURE 4 – Le lien entre le type de mesures polarimétriques radar et le type de mesure polarimétrique optique

Première partie

Influence de la géométrie bistatique sur la mesure

L'objectif de cette partie est d'arriver à distinguer, dans une mesure bistatique, l'information qui est reliée uniquement aux propriétés intrinsèques de la cible, de celle qui découle de la configuration géométrique imposée.

Pour mener cette étude géométrique, nous devons d'abord définir soigneusement nos références géométriques et les conventions associées.

Cela nous mène à utiliser la convention des repères définis par rapport au plan de diffusion. Dans cette convention, le nombre de paramètres géométriques à considérer pour la mesure est grandement réduit : la connaissance de l'angle de bistatisme β est suffisante pour décrire la situation géométrique des antennes.

L'orientation et les rotations de la cible restent, dans le cas général où il n'existe aucun axe de symétrie, des paramètres qui influencent aussi la mesure polarimétrique. Pour nous dégager de l'influence de l'orientation de la cible, nous avons concentré notre étude sur des cibles ayant justement des axes de symétrie permettant d'étudier uniquement ce phénomène. Par la suite nous avons étudié des cibles simples, sphères et cylindres, en augmentant progressivement le nombre de celles-ci. Cette étude s'est appuyée sur plusieurs modèles existants ou developpés au cours de la thèse : modèles simples géométriques, modèle de Mie, ou modèles issus de l'optique physique. Ces études théoriques ont été complétées par des mesures expérimentales réalisées en chambre anéchoïque. Ces mesures visent à la fois à valider les modèles étudiés mais aussi à démontrer dans certains cas la nécessité de modèles plus complets.

Chapitre 1

Le choix du plan d'analyse d'une cible

1.1 Introduction sur la géométrie de mesure polarimétrique

La métrologie optique, comme la métrologie radar consiste à étudier l'interaction d'un objet ou d'une surface avec une onde électromagnétique. Cette onde peut être mesurée de façon active (laser, antennes) ou passive (soleil, signaux radar d'opportunités GPS/DVB/DAB). Dans tous les cas la mesure de l'onde diffusée ou réfléchie par l'objet est porteuse d'une information utile. De manière générale, la mesure de cette onde est très dépendante de la position de la source, de l'objet, ainsi que de l'instrument de mesure en réception. L'information utile est donc naturellement influencée par la géométrie de la mesure (angle d'incidence, angle de l'onde diffusée, orientation de l'objet dans l'espace, angle de bistatisme, etc.).

La polarimétrie étudie la trajectoire de l'extrémité du champ électrique dans le plan de propagation d'une onde électromagnétique. L'étude de cette trajectoire nécessite la définition d'un repère spatial. Lorsqu'on étudie le comportement polarimétrique d'un objet, il s'agit de comparer les trajectoires du champ avant et après diffusion de l'onde par l'objet. Ainsi dans le cas le plus général la définition de deux repères de polarisation est nécessaire.

La mesure polarimétrique est donc à l'instar de toute mesure électromagnétique, mais aussi à cause de la définition des bases de polarisation, dépendante de la géométrie. L'extraction d'une information utile, fiable, et ne variant pas ou peu avec la géométrie de l'acquisition est une des problématiques majeures et communes à la polarimétrie optique et à la polarimétrie radar.

L'objet de ce chapitre est d'abord d'identifier l'ensemble des paramètres géométriques pouvant influer sur la mesure polarimétrique. Les différences classiques de convention entre la polarimétrie optique et radar seront rappelées et nous ne travaillerons par la suite que dans la convention classiquement utilisée en polarimétrie radar. Ensuite nous rappellerons les hypothèses et les simplifications associées à la configuration monostatique, i.e. lorsque la source et la réception se situent en un même point. Enfin le coeur du travail consiste

en une étude de l'influence du bistatisme, i.e. de la séparation spatiale de la source et de la réception, sur la mesure polarimétrique. Cette étude s'appuie sur trois modèles : un modèle géométrique simple que nous proposons, un modèle de « cylindre infini », et enfin le modèle de diffusion de Mie. Ainsi les résultats de simulations théoriques ainsi que les résultats de mesures en chambre anéchoïque sont comparés et discutés.

L'objectif est de proposer des méthodes et des solutions permettant de minimiser l'influence de la géométrie sur la mesure polarimétrique. En outre cette partie permettra de répondre au choix de la polarisation optimale à utiliser dans une configuration géométrique donnée, pour un système qui ne fonctionne qu'avec une seule polarisation.

1.1.1 Pourquoi travailler en configuration bistatique ?

Les configurations monostatiques impliquent de nombreuses simplifications géométriques au niveau des traitements, et sont beaucoup plus simples à mettre en oeuvre au point de vue expérimental.

Pour la polarimétrie optique, travailler en configuration bistatique n'est pas nécessairement un choix. En effet, l'encombrement des appareils, des sources de lumière polarisée, des analyseurs, est tel qu'il est difficile de concevoir des expérimentations en configuration monostatique. Cependant il est assez courant de rencontrer une configuration appelée "diffusion avant" où l'objet d'étude, la source et la réception sont alignés. Nous retrouvons dans ce cas de figure des simplifications géométriques. Toutefois cette configuration n'est pas majoritaire car elle nécessite alors d'observer les objets d'étude par transparence. De manière générale la polarimétrie optique rencontre donc fréquemment des configurations bistatiques quelconques et possède ainsi une certaine avance sur le monde du radar dans la façon de traiter et d'interpréter l'influence de la géométrie d'acquisition.

En radar nous pouvons vouloir travailler en configuration bistatique pour de multiples raisons :

- D'abord d'un point de vue opérationnel "défense", une source émettant un signal radar est naturellement très peu discrète. Séparer l'émission et la réception permet d'envisager des configurations où la source se situe dans une zone sécurisée et peu sensible, et d'envoyer un récepteur, plus proche de la zone d'étude. On peut ainsi envisager de nombreuses configurations multi-aéroportées ou combinant satellites et réception au sol ou aéroportée. D'ailleurs il est aussi possible d'envisager des configurations où l'émission n'a pas été directement conçue pour cette utilisation. On parle alors de configuration passive, où le signal émis est utilisé pour former une image alors qu'il était destiné à un autre usage tel que les télécommunications ou à la télé-localisation (GPS, Galileo, signaux de télévision : DAB/DVB ...). La furtivité est alors totale et les coûts énergétiques associés à l'émission sont aussi évités.
- une configuration bistatique peut aussi permettre de percevoir des mécanismes physiques différents de ceux observés en monostatique. Par exemple la configuration monostatique est très sensible aux effets dits de "double rebonds" : ce sont les effets qui se forment sur un objet vertical formant un angle droit avec la surface du sol. Il a été montré que si on s'éloigne de cette configuration, on atténue très fortement cet effet. De manière générale, la configuration géométrique octroie un degré de liberté supplémentaire à la mesure que l'on espère plus « riche » pour l'analyse.

Cela dit de manière générale la configuration bistatique n'est que peu voire pas du tout utilisée en pratique. La quasi majorité des mesures bistatiques ont encore lieu en chambre anéchoïque. Jusqu'à aujourd'hui quelques expérimentations ont eu lieu, notamment lors de collaboration entre l'ONERA et le DLR, et entre l'ONERA et le FOI en utilisant deux radars aéroportés, mais les mesures n'ont pas été réalisées en polarimétrie complète. Une expérience à aussi été réalisée en combinant l'émission de TerraSAR-X et un récepteur au sol par REF, mais encore une fois les mesures ont été faites strictement en polarimétrie simple. Le nouveau porteur TanDem-X permettrait de réaliser des mesures en polarimétrie bistatique complète, mais la distance entre les deux satellites est très faible, et les hypothèses de quasi monostatisme resteront probablement valables. Lors du lancement du second satellite, une phase de rattrapage de la première trajectoire a été réalisée, et a permis des angles de bistatisme maximum de quatre degrés en azimut. Malheureusement, cette configuration n'a opérée que dans les polarisations HH et VV. De ce fait les études sur la polarimétrie bistatiques sont rares, et pour beaucoup d'adeptes de la polarimétrie les conséquences pratiques du bistatisme restent mystérieuses ou se limitent à l'hypothèse de réciprocité $HV \neq VH$.

Comme nous l'avons évoqué dans l'introduction de ce chapitre, la polarimétrie est particulièrement sensible à la configuration géométrique de la mesure. C'est pourquoi la polarimétrie bistatique est un challenge supplémentaire. Nous allons d'abord expliquer le cadre général d'une mesure polarimétrique quelconque et l'ensemble des paramètres pouvant évoluer. Puis nous allons rappeler les deux conventions permettant de définir des bases polarimétriques. Nous en choisirons une seule pour toute la suite de ce document, nous éviterons alors d'avoir à exprimer l'ensemble des calculs dans les deux conventions.

1.1.2 Conventions

Dans une configuration quelconque, donc a priori bistatique, il convient de définir deux repères pour décrire les polarisations de l'onde incidente et l'onde diffusée. La trajectoire de l'onde est décrite par le vecteur d'onde traditionnellement appelé $\hat{\mathbf{k}}$. Notons au passage que la notation avec chapeau permet de différencier le vecteur d'onde du vecteur de diffusion polarimétrique. Ce vecteur constitue la première composante du repère ; les deux autres composantes appartiennent au plan d'onde, plan orthogonal à la direction de propagation, et doivent former un repère orthonormé direct. Il suffit donc de choisir un vecteur appartenant au plan d'onde, la dernière composante se déduit alors par un produit vectoriel avec le vecteur $\hat{\mathbf{k}}$. La différence au niveau du choix de convention provient du choix du signe/sens pour les vecteurs d'ondes incident $\hat{\mathbf{k}}_\mathbf{i}$ et diffusé (s pour "scattered") $\hat{\mathbf{k}}_\mathbf{s}$ (ou transmit $\hat{\mathbf{k}}_\mathbf{t}$/ reçu $\hat{\mathbf{k}}_\mathbf{r}$). Ces différentes appellations existent car traditionnellement les noms des vecteurs d'ondes sont associées aux repères "antennes" pour la convention BSA, et aux repères "cible" pour la convention FSA. Nous utiliserons ici et pour le reste de ce document l'appellation $\hat{\mathbf{k}}_\mathbf{i}$ et $\hat{\mathbf{k}}_\mathbf{s}$. Le choix du sens pour ces deux vecteurs définit le choix de convention selon si ils sont alignés avec le sens de propagation de l'onde (FSA), ou vers la cible (BSA).

– Dans la convention FSA $\hat{\mathbf{k}}_\mathbf{i}$ et $\hat{\mathbf{k}}_\mathbf{s}$ pointent dans la direction de propagation de l'onde. Dans la convention BSA ils pointent vers la cible. Ces deux conventions sont

représentées sur la figure 1.1 pour une situation bistatique quelconque.

– Dans les cas particuliers de la rétrodiffusion et de la diffusion avant, ces conventions entraînent de grandes simplifications. En convention FSA et dans le cas de la diffusion avant les vecteurs $\hat{\mathbf{k}}_\mathbf{i}$ et $\hat{\mathbf{k}}_\mathbf{s}$ sont identiques comme représenté sur la figure 1.2 (a). De même avec la convention BSA et le cas de la rétrodiffusion, comme représenté sur la figure 1.2 (b).

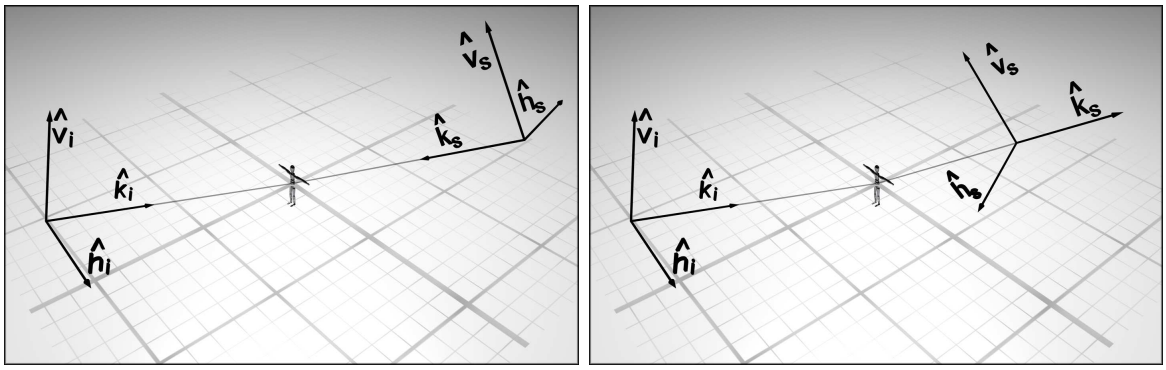


FIGURE 1.1 – Repères polarimétriques dans les conventions BSA (Back Scattering Alignment, à gauche) et FSA (Forward Scattering Alignement, à droite)

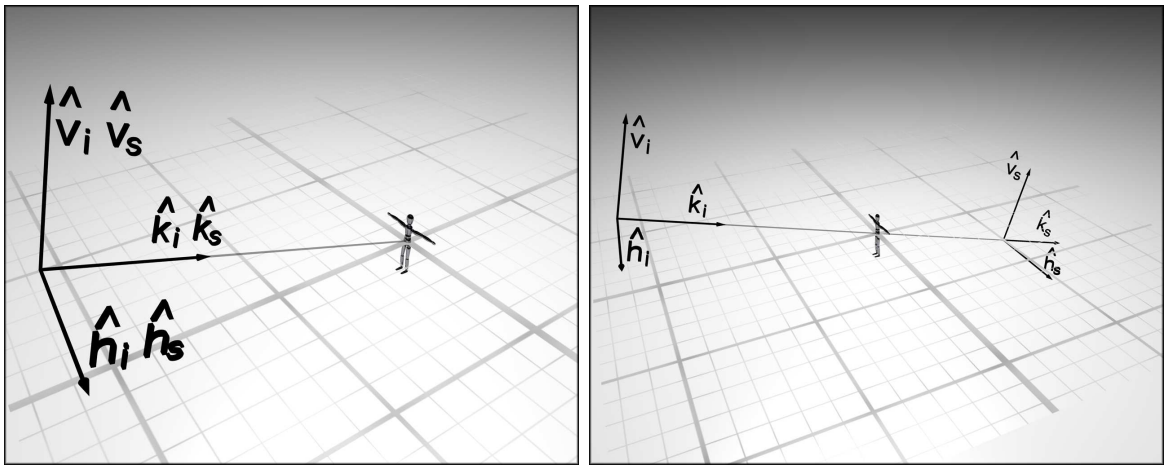


FIGURE 1.2 – Convention BSA (Back Scattering Alignment) et Convention FSA (Forward Scattering Alignement) dans leurs situations classiques de référence (Rétrodiffusion et Diffusion avant). L'intérêt des conventions est donc, dans les deux cas, de conserver le même repère/base polarimétrique à l'émission et à la réception.

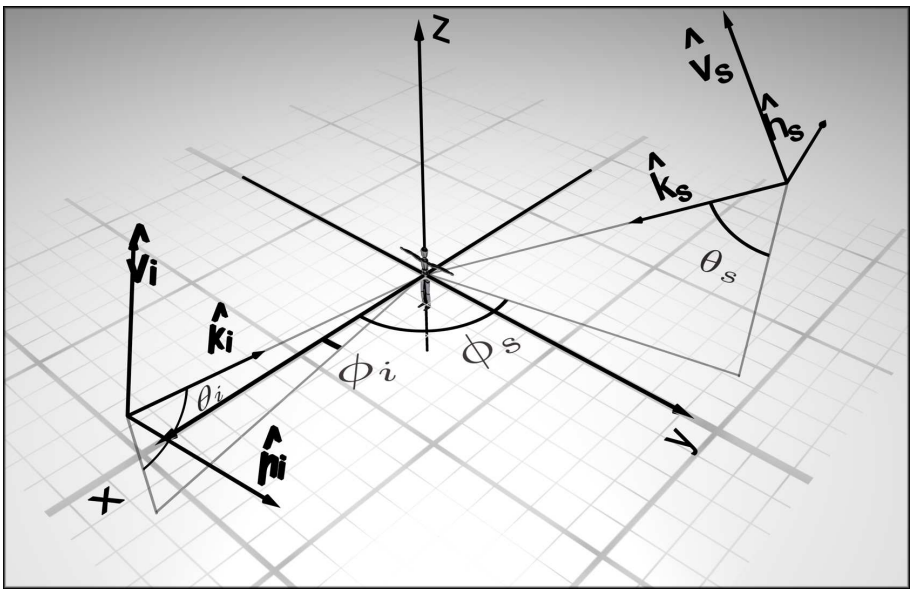


FIGURE 1.3 – Définition des angles θ_s, θ_i, ϕ_s et ϕ_i, qui déterminent la position de l'émission et de la réception.

1.2 Ensemble des paramètres géométriques imposés et choisis.

Dans une mesure polarimétrique bistatique quelconque d'une cible supposée "ponctuelle", un très grand nombre de paramètres sont nécessaires pour décrire sans ambiguïté la configuration géométrique. Il existe des paramètres imposés par la configuration de mesure, comme la position de l'émission et de la réception ou l'orientation de la cible. A cela il faut ajouter des paramètres choisis comme nous venons de le voir c'est notamment le cas d'un des vecteurs de la base de polarisation ($\hat{\mathbf{v}}$ ou $\hat{\mathbf{h}}$). Enfin certains paramètres se déduisent du choix fait pour d'autres paramètres, ainsi par exemple le choix de $\hat{\mathbf{v}}_\mathbf{i}$ induit la valeur de $\hat{\mathbf{h}}_\mathbf{i}$ et inversement. La convention est aussi un choix, mais comme nous l'avons dit plus haut, nous ne considérerons que la convention BSA pour l'ensemble de ce document.

L'ensemble des paramètres est décrit ci-dessous, nous nous plaçons dans un repère ayant pour origine la position de la cible.

1. L'orientation de la cible : elle peut être définie par trois angles d'Euler α, β et γ. Ils sont représentés sur les figures 1.4,1.5 et 1.9 . L'angle α correspond à une rotation autour de l'axe Z, β de l'axe Y, et γ de l'axe Z (les rotations ne sont pas commutatives).

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (1.1)$$

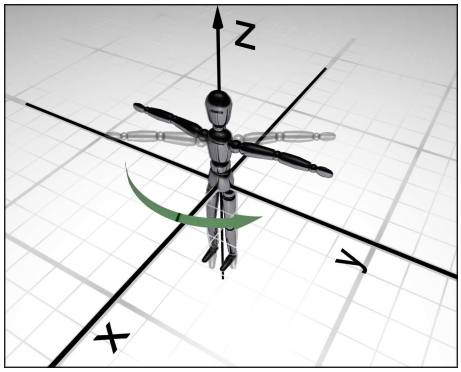


FIGURE 1.4 – Rotation d'angle α autour de l'axe Z

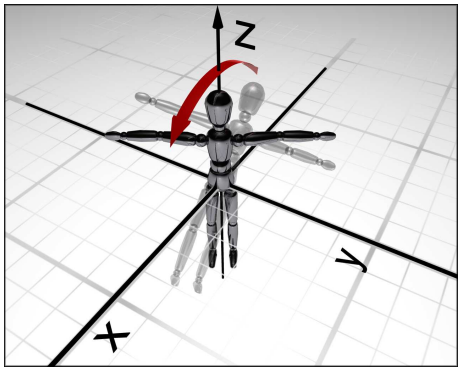


FIGURE 1.5 – Rotation d'angle β autour de l'axe Y

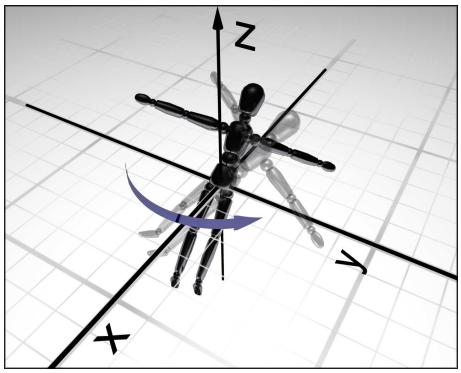


FIGURE 1.6 – Rotation d'angle γ autour de l'axe Z, après application des rotations d'angle α et β

L'orientation de la cible est imposée par la configuration de la mesure. En fonction d'éventuels axes de symétrie, et dans certaines configurations particulières, nous pourrons obtenir des mesures partiellement ou totalement indépendantes de cette orientation.

2. Les positions des antennes sources et capteurs : elles sont repérées chacune par deux angles θ et ϕ, ainsi qu'éventuellement une distance ρ. Cependant la distance n'influe en rien sur la mesure si l'hypothèse champ lointain est respectée. C'est pourquoi il est généralement ignoré. Ces angles et leurs définitions sont décrits sur la figure 1.3. Dans la convention BSA la position des deux antennes définit parfaitement les deux vecteurs d'ondes $\hat{\mathbf{k}}_\mathbf{i}$ et $\hat{\mathbf{k}}_\mathbf{s}$.

$$\hat{\mathbf{k}}_\mathbf{i} = \begin{pmatrix} \sin\theta_i \cos\phi_i \\ \sin\theta_i \sin\phi_i \\ -\cos\theta_i \end{pmatrix}, \quad \hat{\mathbf{k}}_\mathbf{s} = \begin{pmatrix} \sin\theta_s \cos\phi_s \\ \sin\theta_s \sin\phi_s \\ \cos\theta_s \end{pmatrix}$$

3. L'orientation des repères ou bases de polarisation pour l'émission et la réception. La base contient nécessairement le vecteur d'onde $\hat{\mathbf{k}}$, il convient alors de choisir arbitrairement un vecteur $\hat{\mathbf{h}}$ ou $\hat{\mathbf{v}}$, l'autre se déduisant automatiquement par un produit vectoriel, qui garantit l'orthogonalité directe de la base ($\hat{\mathbf{k}}$,$\hat{\mathbf{h}}$,$\hat{\mathbf{v}}$).

$$\hat{\mathbf{v}} = \hat{\mathbf{k}} \times \hat{\mathbf{h}}, \quad \hat{\mathbf{h}} = \hat{\mathbf{v}} \times \hat{\mathbf{k}}.$$

Le choix de $\hat{\mathbf{h}}_\mathbf{i}$ et $\hat{\mathbf{h}}_\mathbf{s}$ est donc nécessaire pour toute mesure polarimétrique, et toute mesure polarimétrique dépend donc de ce choix. Naturellement il convient de différencier les vecteurs "réels" qui correspondent à la position et l'orientation des antennes horizontale et verticale lors de la mesure de celle obtenue après transformation mathématique appliquée aux données brutes. Dans l'ensemble le choix des bases de polarisation est la seule liberté dans une mesure polarimétrique. Cette liberté peut toutefois s'avérer problématique car elle entraîne a priori une non-universalité des résultats, rendant toute comparaison difficile.

1.2.1 Monostatisme et géométrie

Nous allons maintenenant revenir sur les trois points que nous venons de décrire dans le cas particulier d'une configuration monostatique.

1. L'orientation de la cible, autour de l'axe défini par le vecteur $\hat{\mathbf{k}}$ est équivalente à une rotation de la base de polarisation autour de ce même axe. Deux autres axes de rotations sont également possibles mais cela induit alors des effets sur la mesure qu'on ne peut contrôler ni prévoir à moins de disposer des hypothèses supplémentaires sur la cible (invariance, axes de symétrie etc.).

2. Le choix d'un des vecteurs de la base de polarisation $\hat{\mathbf{h}}$ ou $\hat{\mathbf{v}}$, est le seul degré de liberté dans la définition des bases polarimétriques. En effet le même repère est utilisé pour l'émission et la réception. Le choix traditionnel pour la définition des

bases de polarisation définit le vecteur horizontal $\hat{\mathbf{h}}$ comme étant orthogonal à l'axe z, orienté suivant la normale de la surface du sol. Nous avons ainsi,

$$\hat{\mathbf{h}} = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}}{|\hat{\mathbf{z}} \times \hat{\mathbf{k}}|}, \quad \hat{\mathbf{v}} = \hat{\mathbf{k}} \times \hat{\mathbf{h}} \tag{1.2}$$

$$\hat{\mathbf{h}} = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix}, \quad \hat{\mathbf{v}} = \begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ \sin\theta \end{pmatrix}, \quad \hat{\mathbf{k}} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ -\cos\theta \end{pmatrix}. \tag{1.3}$$

L'existence de ce choix soulève donc de nombreuses questions dans l'utilisation du signal polarimétrique. En effet nous souhaiterions idéalement avoir des outils polarimétriques, dont l'utilisation et l'interprétation ne dépendent pas du choix qui a été fait au niveau du repère polarimétrique. C'est un désir très naturel car il garantit que l'information extraite n'est pas biaisée par une erreur, ou une mauvaise connaissance du dispositif de mesure. A priori deux solutions peuvent alors être envisagées :

1. Développer des outils produisant des paramètres indépendant du choix de la base de polarisation. C'est ce qu'on appelle traditionnellement la "Roll Invariance". Ainsi un paramètre "Roll Invariant" sera indépendant du choix de base polarimétrique, mais aussi et il est important de le noter, de l'orientation de la cible dans le plan d'onde. Ainsi ces paramètres ne sont pas porteurs de l'intégralité de l'information donnée par la mesure polarimétrique. L'orientation de la cible peut en effet être un paramètre pertinent pour l'utilisateur.
2. Estimer l'angle de rotation du repère polarimétrique et appliquer une rotation visant à "désorienter" la mesure polarimétrique. L'information résultante est alors traitée, et tous les résultats issus de ces matrices "désorientées" sont indépendants du choix qui a été fait pour la base de polarisation. Le problème de cette approche, est que l'estimation de cet angle est une estimation relative. Elle dépend aussi de l'axe principal d'orientation de la cible. Lorsque cette méthode est employée, toute information sur l'orientation de la cible est confondue avec l'orientation de la base polarimétrique. Dans une configuration monostatique, l'orientation de la base polarimétrique est constante avant et après diffusion. Aussi l'angle estimé correspond à l'orientation de la cible, à une constante près. Par contre, dans une configuration bistatique, les deux repères liés à l'onde incidente et à l'onde diffusée peuvent ne pas avoir les mêmes orientations.

De manière générale en monostatique la nuance dans la "Roll Invariance" est souvent mal comprise. Elle apparaît comme une condition quasiment essentielle à tout traitement. A notre avis il est important de distinguer l'orientation de la cible de l'orientation des antennes en choisissant des repères polarimétriques adéquats. Ainsi l'estimation de l'angle telle qu'elle est faite lorsque les matrices sont "désorientées" sera un paramètre qui dépendra alors uniquement de la cible, et non de la configuration de mesures. Encore une fois, en monostatique la nuance est difficile à percevoir car l'angle estimé ne dépend, à une constante près, que de l'orientation de la cible ce qui convient pour la totalité des configurations rencontrées. Nous allons voir qu'en bistatique cette nuance est importante.

1.2.2 Monostatisme et réciprocité

La propriété majeure des configurations monostatiques, et plus particulièrement de la rétrodiffusion radar est la réciprocité. Dans une configuration monostatique les trajets "aller" et "retour" sont identiques, ainsi dans une position monostatique l'hypothèse de réciprocité est valide et implique une simplification importante de l'information polarimétrique. Dans cette configuration, et quand l'hypothèse de réciprocité est vérifiée, les mesures des termes croisées, S_{hV} et S_{vH} de la matrice de Sinclair sont égaux. Les conséquences pour l'ensemble des matrices d'ordre deux, et principalement des matrices de cohérences, est une réduction des degrés de liberté, ainsi que pour certaines matrices de la dimension de celles ci.

En pratique les mesures de S_{hV} et de S_{vH} diffèrent, parce qu'elles sont affectées de deux réalisations de bruit différentes. Traditionnellement il est courant de remplacer les deux valeurs par leurs moyennes. C'est effectivement commode pour une utilisation "physique" des mesures polarimétriques, lorsqu'il s'agit d'en extraire des paramètres purement polarimétriques, dont le sens et l'interprétation seraient délicats si l'hypothèse de réciprocité attendue n'était pas respectée à cause d'un bruit trop important.

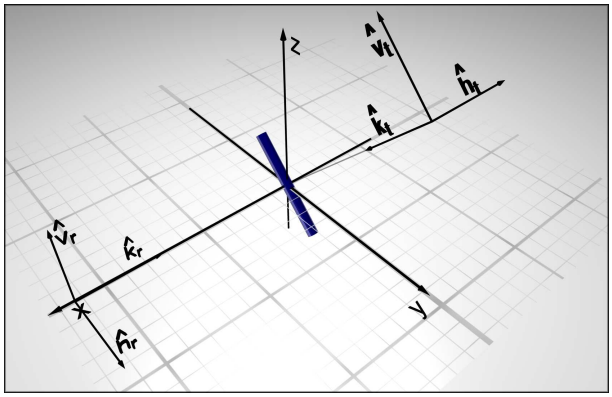
Par contre quand le signal polarimétrique est utilisé à des fins statistiques, pour des classifications, segmentations ou détections, il n'est pas évident que de remplacer brutalement les termes croisés par une moyenne soit optimal au niveau des traitements statistiques. Nous aborderons cette question dans la deuxième partie de ce manuscrit.

1.2.3 Bistatisme et problèmes principaux

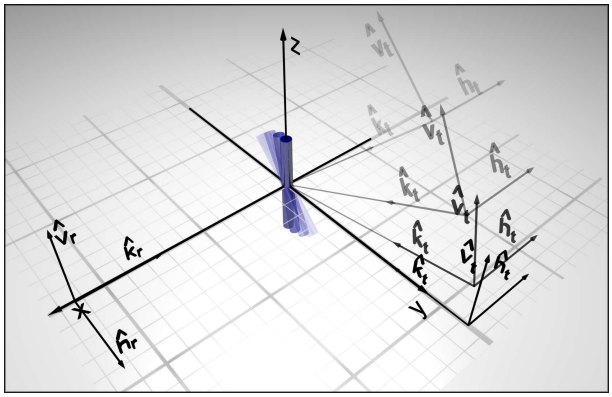
Dans une configuration bistatique, le choix d'un des axes des deux bases de polarisation, à l'émission et à la réception est nécessaire. Ici il faut remarquer qu'une rotation de la cible, quelque soit l'axe de cette rotation, n'est pas équivalente à une rotation des bases de polarisation. Pour illustrer ce phénomène nous montrons un exemple sur une cible simple illustré sur la Figure 1.7.

(a) La cible est un cylindre très fin, assimilé à un polariseur, placé dans le plan d'incidence Oyz, qui présente un angle d'inclinaison par rapport à la verticale de 45°. Cet angle est appelé couramment angle de "tilt". L'effet d'un tel cylindre sur la polarisation est ce qu'on appelle un polariseur linéaire : une des polarisations, perpendiculaire à l'axe du cylindre est totalement atténuée, alors que la polarisation parallèle au cylindre n'est pas affectée. Cela ce traduit, au niveau de la matrice de Sinclair, par des valeurs singulières 0 et 1. Ici nous sommes dans la convention classique radar liée aux plans d'incidences.

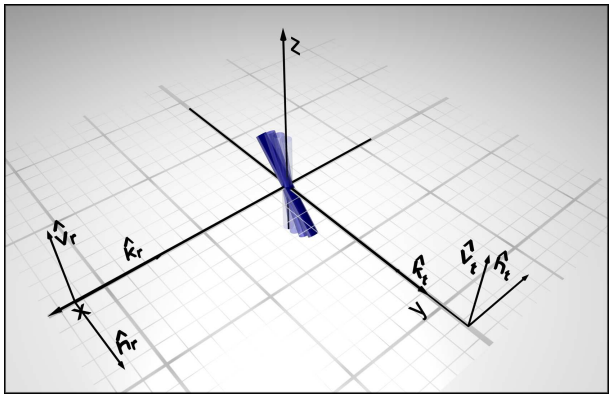
(b) Nous considérons la même cible dans la même configuration Géométrique que précédemment. Mais cette fois, nous choisissons le plan de diffusion comme plan de référence pour l'analyse polarimétrique. Le plan de diffusion est le plan contenant l'émission, la cible et la réception. Nous voulons que la cible conserve la même orientation par rapport à ce plan de diffusion que dans la situation précédente. Cela revient à effectuer une rotation autour de l'axe-x afin de placer ce plan en position horizontale. Cette rotation est équivalente à une rotation de l'antenne d'émission, donc ne concerne que le choix de base de polarisation pour l'émission. Cet effet ce traduit par une rotation



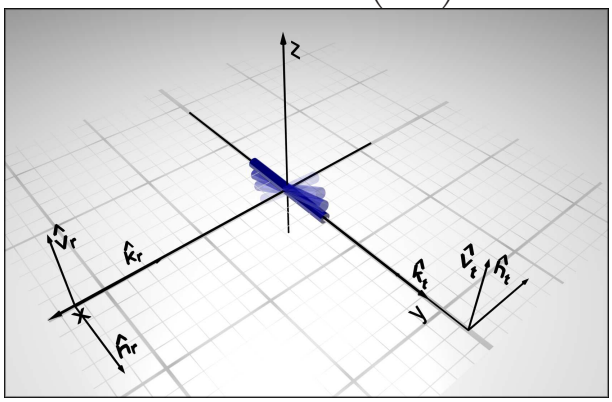
(a) $\mathbf{S} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}$



(b) Changement de plan de convention : passage au plan de diffusion $\mathbf{S^P} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$



(c) Modification de l'orientation de la cible $\mathbf{S^P} = \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}$



(d) Modification de l'orientation de la cible $\mathbf{S^P} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

FIGURE 1.7 – Exemples de transformations géométriques pour une situation bistatique à $\beta = 90°$

mathématique de la matrice de Sinclair, qui n'a pas d'effet sur les valeurs singulières de celle-ci. La matrice de Sinclair est bien modifiée, mais les valeurs singulières restent 0 et 1. (Notons par contre que les valeurs propres sont, elles, modifiées : elles valent 0 et $-\frac{1}{\sqrt{(2)}}$ dans la situation précédente, et 0 et 1 dans cette nouvelle situation).

(c) Ici nous gardons le plan de diffusion comme plan de référence. Nous faisons tourner la cible avec une rotation d'angle de 45° autour de l'axe de l'onde incidente. C'est une rotation qui concerne donc l'orientation de la cible. Celle ci a deux effets, sur l'angle de "tilt" de l'antenne d'émission, mais aussi sur les valeurs singulières de la matrice mesurée. Cette rotation de la cible ne correspond pas à une rotation mathématique de la matrice de Sinclair mesurée. En effet elle influe aussi sur la projection de la cible sur le plan d'onde diffusée. Les valeurs singulières sont alors 0 et $\frac{1}{\sqrt{2}}$.

(d) Enfin nous terminons de tourner la cible de 45° supplémentaires, autour du même axe que précédemment. Nous obtenons alors un cas particulier où les deux valeurs singulières sont nulles. Dans cet exemple il est évident qu'une rotation de la cible ne peut avoir d'équivalent mathématique indépendant de la cible pour expliquer l'évolution des valeurs singulières.

Revenons donc aux approches traditionnellement utilisées en monostatique et leurs applications en bistatique,

1. L'utilisation de paramètres "Roll Invariant" garde son sens, mais il convient de faire attention à ce qu'on entend par "Roll". Nous pouvons effectivement déduire des paramètres indépendants du choix des deux bases de polarisation. Il s'agit de trouver des paramètres qui sont mathématiquement invariants par rotation de la base de l'émission et de la réception. Certains paramètres "Roll Invariant" en monostatique ne le sont pas en bistatique. Mais comme en monostatique il s'agit d'une extraction partielle de l'information, qui ne renseigne pas sur l'orientation de la cible.
2. La désorientation des cibles est toujours possible mais dangereuse ; d'abord parce que le terme peut induire en erreur. En bistatique, il est impossible d'obtenir une mesure indépendante de l'orientation de la cible (sauf bien sur pour des cibles ayant des axes de symétries).

Une cible "désorientée" en bistatique est une matrice dont on a retiré à la fois l'influence du choix des bases de polarisation d'émission et de réception, et aussi partiellement l'influence de l'orientation de la cible. J'insiste sur le terme "partiellement" car comme nous venons de le voir dans notre exemple, l'orientation de la cible a des effets à la fois de rotations au sens mathématique sur les matrices de Sinclair mesurées, mais aussi sur les valeurs singulières de celles-ci. Ainsi en bistatique, lorsque le plan d'onde incident et le plan d'onde diffusé diffèrent, une cible ne peut être que partiellement désorientée. Les angles d'inclinaison estimés comprennent alors, comme en monostatique, le mélange d'un angle d'orientation de la cible et d'un angle provenant du choix de base de polarisation. Le problème est alors que contrairement au monostatique, le choix de la base de polarisation dépend de tous les angles qui définissent la position relative de la cible par rapport à l'émission et la réception θ_i, ϕ_i, θ_s et ϕ_s. Ces angles, pour des positions d'antennes fixes, dépendent de la position de la cible. Ainsi sur une image bistatique entière, les bases de

polarisation évoluent sensiblement d'un pixel à l'autre. Ainsi si on cherche à exploiter l'information concernant l'orientation de la cible, après avoir partiellement désorienté la cible mathématiquement, il est impossible de discerner la contribution de la cible de celle de la configuration géométrique dans l'angle d'inclinaison que l'on estime. C'était déjà le cas en monostatique, mais la contribution de la configuration géométrique était constante, donc souvent ignorée. C'est malheureusement impossible dans une configuration bistatique quelconque. Dans ce cas, que l'on désire travailler sur des matrices désorientées, des paramètres roll invariant ou non, il est nécessaire de séparer l'influence du choix des bases de polarisation de l'orientation propre de la cible. Ainsi l'estimation des angles qui pourra être faite par la suite, ne dépendra alors que de la cible et plus de la configuration de mesure, propre à chaque expérimentation et à chaque image.

1.2.4 Définition des bases polarimétriques bistatiques : méthode classique

Le système de coordonnée communément utilisé en radar est le plus souvent imposé par les conditions de mesures. Comme en monostatique le plan "horizontal" est défini par le plan tangent à la surface de la terre. La normale à ce plan est noté $\hat{\mathbf{z}}$. La définition des bases de polarisation se fait alors en imposant que :

- $(\hat{\mathbf{k}}_\mathbf{i}, \hat{\mathbf{h}}_\mathbf{i}, \hat{\mathbf{v}}_\mathbf{i})$ et $(\hat{\mathbf{k}}_\mathbf{s}, \hat{\mathbf{h}}_\mathbf{s}, \hat{\mathbf{v}}_\mathbf{s})$ forment une base orthonormale directe.
- Les deux vecteurs $\hat{\mathbf{v}}$ sont dans le même sens que $\hat{\mathbf{z}}$.
- Les deux vecteurs $\hat{\mathbf{h}}$ sont parallèles au plan horizontal.

Ces notations correspondent à celles décrites dans [3] à l'exception des indices des vecteurs que nous avons voulu uniformiser pour l'ensemble du manuscrit. Nous obtenons donc,

$$\begin{aligned}
\hat{\mathbf{h}}_\mathbf{i} &= \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_\mathbf{i}}{|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_\mathbf{i}|}, \quad \hat{\mathbf{v}}_\mathbf{i} = \hat{\mathbf{k}}_\mathbf{i} \times \hat{\mathbf{h}}_\mathbf{i}, \\
\hat{\mathbf{h}}_\mathbf{s} &= \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_\mathbf{s}}{|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_\mathbf{s}|}, \quad \hat{\mathbf{v}}_\mathbf{s} = \hat{\mathbf{k}}_\mathbf{s} \times \hat{\mathbf{h}}_\mathbf{s}
\end{aligned} \tag{1.4}$$

$$\begin{aligned}
\hat{\mathbf{k}}_\mathbf{i} &= \begin{pmatrix} \sin\theta_i \cos\phi_i \\ \sin\theta_i \sin\phi_i \\ -\cos\theta_i \end{pmatrix}, \quad \hat{\mathbf{h}}_\mathbf{i} = \begin{pmatrix} -\sin\phi_i \\ \cos\phi_i \\ 0 \end{pmatrix}, \quad \hat{\mathbf{v}}_\mathbf{i} = \begin{pmatrix} \cos\theta_i \cos\phi_i \\ \cos\theta_i \sin\phi_i \\ \sin\theta_i \end{pmatrix}, \\
\hat{\mathbf{k}}_\mathbf{s} &= \begin{pmatrix} \sin\theta_s \cos\phi_s \\ \sin\theta_s \sin\phi_s \\ \cos\theta_s \end{pmatrix}, \quad \hat{\mathbf{h}}_\mathbf{s} = \begin{pmatrix} \sin\phi_s \\ -\cos\phi_s \\ 0 \end{pmatrix}, \quad \hat{\mathbf{v}}_\mathbf{s} = \begin{pmatrix} -\cos\theta_s \cos\phi_s \\ -\cos\theta_s \sin\phi_s \\ \sin\theta_s \end{pmatrix}.
\end{aligned} \tag{1.5}$$

On définit l'angle de bistatisme β par,

$$\cos\beta = \hat{\mathbf{k}}_\mathbf{i} \cdot \hat{\mathbf{k}}_\mathbf{s}, \quad \sin\beta = |\hat{\mathbf{k}}_\mathbf{i} \times \hat{\mathbf{k}}_\mathbf{s}|. \tag{1.6}$$

L'ensemble des vecteurs est représenté sur la Figure 1.8.

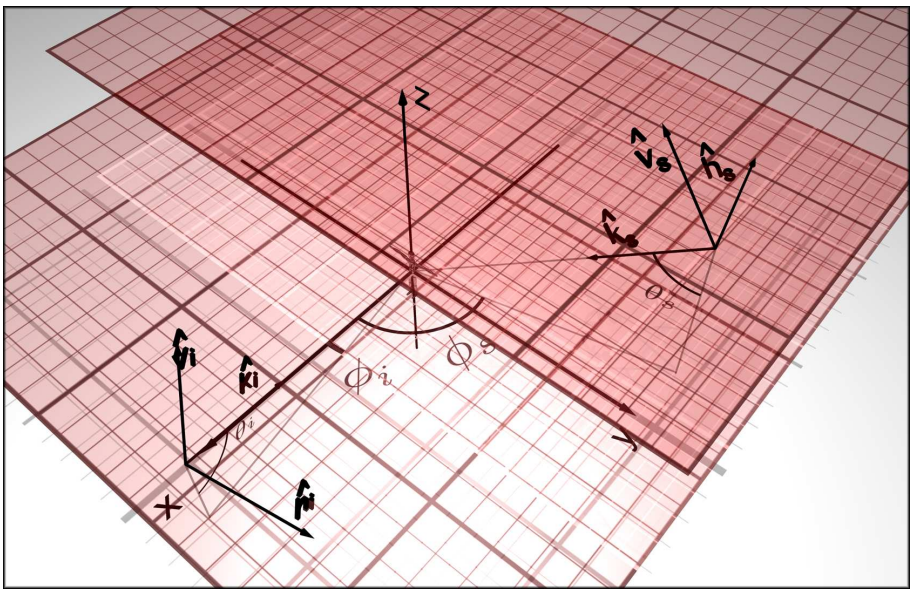


FIGURE 1.8 – Définition des bases polarimétriques en utilisant la convention imposée par le plan horizontal de la surface de la terre

Ce choix correspond à celui qui est fait en monostatique, le plan d'incidence correspondant alors au plan de l'onde diffusé. Cela dit ce choix traditionnel n'est pas aussi adapté aux configurations bistatiques et peut amener à des difficultés d'interprétation.

1.2.5 Définition des bases polarimétriques bistatiques : méthode du plan de diffusion

Nous proposons d'utiliser des bases polarimétriques plus adaptées aux situations bistatiques. C'est ce choix qui est traditionnellement fait pour une mesure optique. Cette convention a notamment été utilisée dans [4]. Nous appelons plan de diffusion le plan contenant $\hat{\mathbf{k}}_\mathbf{i}, \hat{\mathbf{k}}_\mathbf{s}$, les vecteurs de l'onde incidente et diffusée. Le plan de diffusion est représenté sur la Figure 1.10.

Si on appelle $\hat{\mathbf{n}}$ le vecteur unitaire, normal à ce plan de diffusion on a alors :

$$\hat{\mathbf{n}} = \frac{\hat{\mathbf{k}}_\mathbf{i} \times \hat{\mathbf{k}}_\mathbf{s}}{|\hat{\mathbf{k}}_\mathbf{i} \times \hat{\mathbf{k}}_\mathbf{s}|}. \tag{1.7}$$

Dans cette convention, la normale au plan de diffusion $\hat{\mathbf{n}}$ remplace le vecteur $\hat{\mathbf{z}}$ qui était utilisé dans la convention précédente. Nous noterons pour les différencier les vecteurs avec l'exposant p . Ainsi de la même manière que dans la section précédente, les bases de polarisation $(\hat{\mathbf{k}}_\mathbf{i}, \hat{\mathbf{h}}_\mathbf{i}^\mathbf{p}, \hat{\mathbf{v}}_\mathbf{i}^\mathbf{p})$ et $(\hat{\mathbf{k}}_\mathbf{s}, \hat{\mathbf{h}}_\mathbf{s}^\mathbf{p}, \hat{\mathbf{v}}_\mathbf{s}^\mathbf{p})$ forment aussi une base orthonormale directe, et $\hat{\mathbf{v}}_\mathbf{i}^\mathbf{p} = \hat{\mathbf{v}}_\mathbf{s}^\mathbf{p} = \hat{\mathbf{n}}$. On peut ainsi écrire les vecteurs horizontaux :

$$\hat{\mathbf{h}}_\mathbf{i}^\mathbf{p} = \frac{\hat{\mathbf{n}} \times \hat{\mathbf{k}}_\mathbf{i}}{|\hat{\mathbf{n}} \times \hat{\mathbf{k}}_\mathbf{i}|}, \tag{1.8}$$

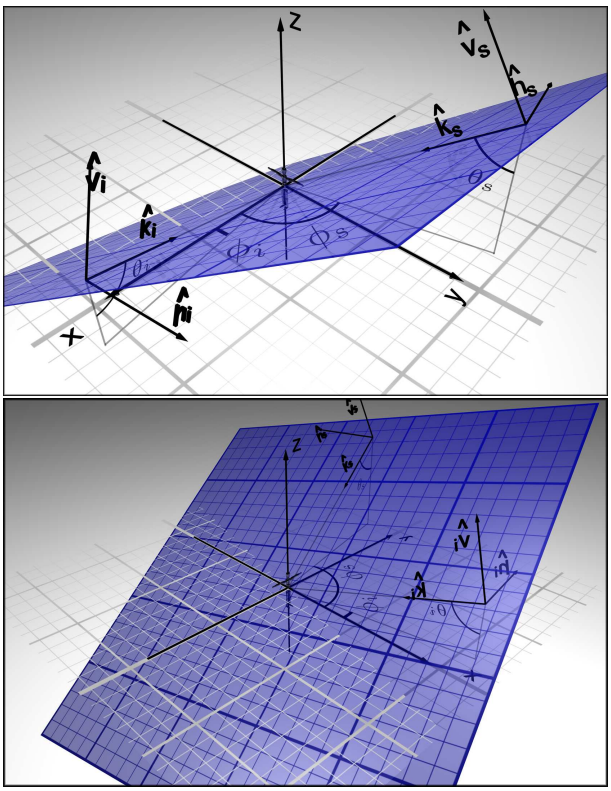


FIGURE 1.9 – Représentations de plans de diffusion

$$\hat{\mathbf{h}}_{\mathbf{s}}^{\mathbf{p}} = \frac{\hat{\mathbf{n}} \times \hat{\mathbf{k}}_{\mathbf{s}}}{|\hat{\mathbf{n}} \times \hat{\mathbf{k}}_{\mathbf{s}}|}. \tag{1.9}$$

Ainsi $\hat{\mathbf{h}}_{\mathbf{i}}^{\mathbf{p}}$ et $\hat{\mathbf{h}}_{\mathbf{s}}^{\mathbf{p}}$ jouent le rôle de la composante parallèle, et $\hat{\mathbf{v}}_{\mathbf{i}}^{\mathbf{p}}$ et $\hat{\mathbf{v}}_{\mathbf{s}}^{\mathbf{p}}$ représentent les composantes verticales. Ces vecteurs sont représentés sur la Figure 1.10.

1.2.6 Relations de passage entre les deux conventions

De la même manière nous noterons $\mathbf{S}$ les matrices de Sinclair dans la convention classique et $\mathbf{S}^{\mathbf{p}}$ dans la convention utilisant le plan de diffusion comme plan de référence.

Les deux matrices sont liées par la relation :

$$\mathbf{S} = \mathbf{R}_{\alpha_{\mathbf{s}}}^{\mathbf{t}} \mathbf{S}^{\mathbf{p}} \mathbf{R}_{\alpha_{\mathbf{i}}}. \tag{1.10}$$

$\mathbf{R}_{\alpha_{\mathbf{i}}}$ et $\mathbf{R}_{\alpha_{\mathbf{s}}}$ sont les rotations à appliquer dans le plan de l'onde incidente et diffusée :

$$\mathbf{R}_{\alpha_{\mathbf{i,s}}} = \begin{pmatrix} \cos\alpha_{i,s} & \sin\alpha_{i,s} \\ -\sin\alpha_{i,s} & \cos\alpha_{i,s} \end{pmatrix}. \tag{1.11}$$

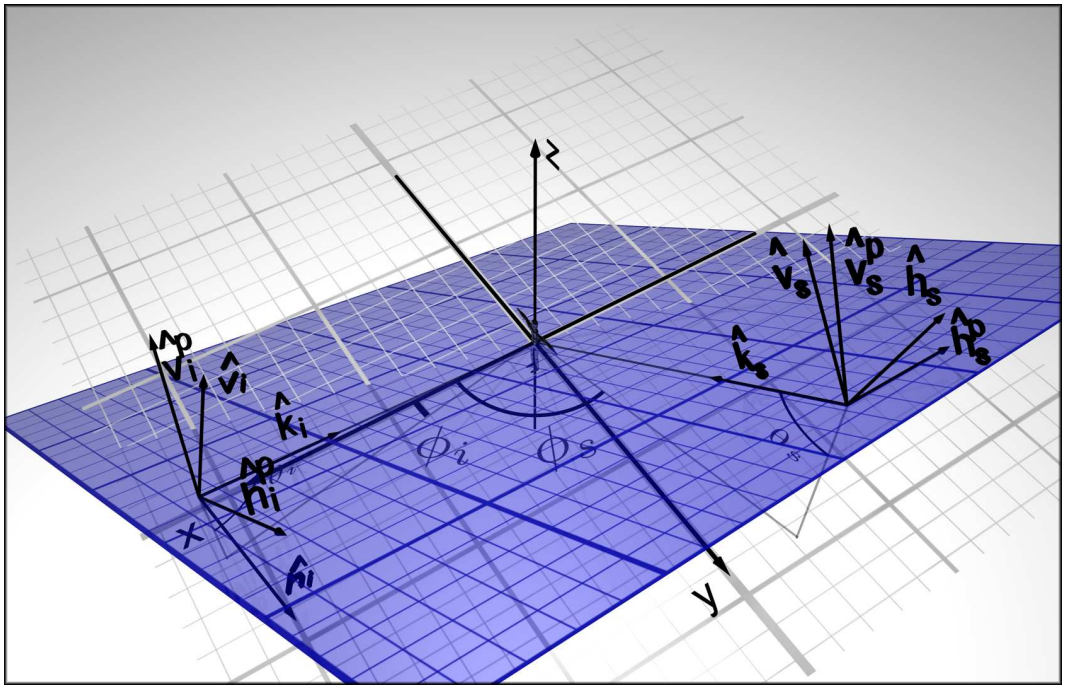


FIGURE 1.10 – Définition des bases de polarisation dans le plan de diffusion

La rotation R_{α_i} transforme $\hat{\mathbf{h}}_{\mathbf{i}}^{\mathbf{p}}$ en $\hat{\mathbf{h}}_{\mathbf{i}}$ et $\hat{\mathbf{v}}_{\mathbf{i}}^{\mathbf{p}}$ en $\hat{\mathbf{v}}_{\mathbf{i}}$. De la même manière R_{α_s} transforme $\hat{\mathbf{h}}_{\mathbf{s}}^{\mathbf{p}}$ en $\hat{\mathbf{h}}_{\mathbf{s}}$ et $\hat{\mathbf{v}}_{\mathbf{s}}^{\mathbf{p}}$ en $\hat{\mathbf{v}}_{\mathbf{s}}$ (voir Figure 1.10).

Les expressions de α_s et α_i sont alors :

$$\cos\alpha_i = \mathbf{h}_{\mathbf{i}}^{\mathbf{p}} \cdot \mathbf{h}_{\mathbf{i}}, \quad \sin\alpha_{\mathbf{i}} = \mathbf{h}_{\mathbf{i}}^{\mathbf{p}} \cdot \mathbf{v}_{\mathbf{i}}, \tag{1.12}$$

$$\cos\alpha_s = \mathbf{h_s^p} \cdot \mathbf{v_s}, \quad \sin\alpha_s = \mathbf{h_s^p} \cdot \mathbf{v_s}. \tag{1.13}$$

Il est aussi possible d'exprimer α_s and α_i en fonction de β, $\theta_{i,s}$ and $\phi_{i,s}$. Il suffit de remplacer $\hat{\mathbf{n}}$ dans les équations 1.8 et 1.9 par son expression dans l'équation 1.7. Ce qui donne :

$$\hat{\mathbf{h}}_\mathbf{i}^\mathbf{p} = \frac{\mathbf{1}}{\sin\beta}(\hat{\mathbf{k}}_\mathbf{s} - \cos\beta\hat{\mathbf{k}}_\mathbf{i}), \tag{1.14}$$

$$\hat{\mathbf{h}}_\mathbf{s}^\mathbf{p} = \frac{\mathbf{1}}{\sin\beta}(\cos\beta\hat{\mathbf{k}}_\mathbf{s} - \hat{\mathbf{k}}_\mathbf{i}), \tag{1.15}$$

Les cosinus s'écrivent alors :

$$\cos\alpha_i = \frac{1}{\sin\beta}\sin\theta_s\sin(\phi_i - \phi_s), \tag{1.16}$$

$$\cos\alpha_s = \frac{1}{\sin\beta}\sin\theta_i\sin(\phi_i - \phi_s). \tag{1.17}$$

De même les sinus :

$$\sin\alpha_i = \frac{1}{\sin\beta}(\sin\theta_s\cos\theta_i\cos(\phi_s - \phi_i) + \sin\theta_i\cos\theta_s), \tag{1.18}$$

$$\sin\alpha_s = -\frac{1}{\sin\beta}(\sin\theta_i\cos\theta_s\cos(\phi_s - \phi_i) + \cos\theta_i\sin\theta_s). \tag{1.19}$$

1.2.7 Pourquoi choisir la convention du plan de diffusion, et quelles conséquences ?

Le choix d'un plan de référence adéquat est important car il conditionne les analyses polarimétriques faites a posteriori. Nous montrons par exemple une image sur la Figure 1.11 obtenue pour 4 sphères, pour les deux conventions différentes. L'angle bistatique moyen pour cette configuration est proche de $\pi/2$.

Nous distinguons dans la suite trois domaines de fréquence pour une sphère (et plus généralement pour une cible quelconque) :

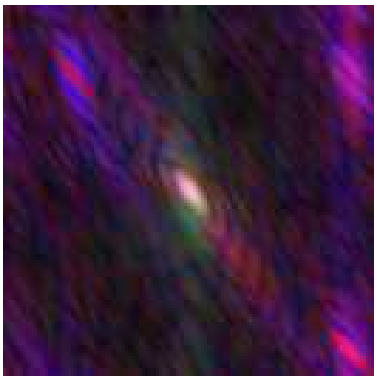
Le domaine optique ou champ éloigné s'applique lorsque $2\pi r/\lambda > 10$ où r est le rayon de la sphère. $2\pi r/\lambda$ est appelé paramètre de taille.

La région de Mie, ou zone de résonance, concerne la région où la longueur d'onde utilisée est du même ordre de grandeur que les dimensions de la cible.

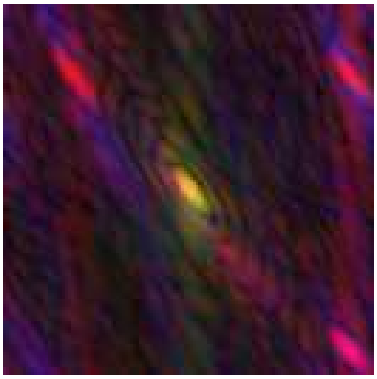
La zone de Rayleigh, ou domaine basse fréquence, est la zone pour laquelle la longueur d'onde est beaucoup plus grande que les dimensions de la cible.

Dans l'exemple suivant, les sphères modélisées de l'image se situent dans la région de résonance. Par ailleurs, notons que la synthèse d'image utilisée est une synthèse d'image radar bistatique, ce qui explique que la tache de résolution (ou fonction d'étalement d'un objet ponctuel) n'est pas un cercle mais peut s'approximer par une ellipse.

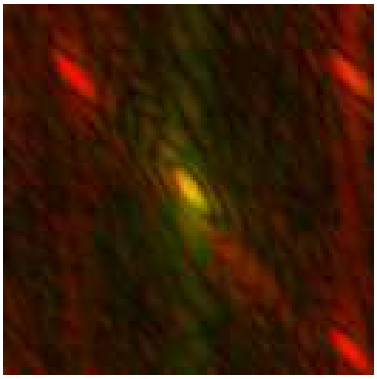
La figure 1.11 nous permet de constater que :



(a) image synthétisée dans la convention de mesure radar classique



(b) image synthétisée dans la convention du plan de diffusion pour le centre de la scène



(c) image synthétisée dans la convention du plan de diffusion recalculé pour chacun des pixels de l'image

FIGURE 1.11 – Image polarimétrique de 4 sphères de Mie avec différentes conventions de repères polarimétriques. - Rouge : $|hH+vV|^2$, Vert : $|hH-vV|^2$, Bleu : $|hV-vH|^2$

- Dans le cas de la synthèse d'image avec la convention classique radar liée au plan d'incidence et de diffusion, l'image polarimétrique obtenue montre que les niveaux de polarisations croisées sont bien visibles (ils apparaissent en bleu sur l'image) et donc ne sont pas négligeables.
- Lorsque l'on passe à une image polarimétrique calculée avec comme plan de référence le plan de diffusion lié au centre de la scène, ces niveaux sont grandement diminués.
- Enfin, le plan de diffusion "Émetteur - cible - Récepteur" varie d'un pixel à l'autre de l'image si les dimensions de celles-ci ne sont pas négligeables devant la distance à l'émetteur ou la distance au récepteur. En redéfinissant un repère adéquat de polarisation pour chaque pixel de l'image, alors on constate que les signaux croisés diminuent encore.

Par cet exemple nous voulons illustrer deux grands points que nous synthétisons ici :

- Le choix du plan de diffusion nous semble plus approprié pour analyser les comportements polarimétriques tels que les symétries présentées par la cible, et l'évaluation de ce que nous appelons "angles d'orientation" de la cible. Ces angles d'inclinaison seront par contre cette fois porteurs d'une information qui dépend uniquement de la cible et non plus aussi de la configuration géométrique.
- Dans tous les cas il est possible que les bases de polarisations varient pour chaque point de trajectoires des antennes. Il est alors plus commode d'appliquer une correction géométrique, compensant les variations des bases de polarisation en chacun des pixels, et en recalculant chaque plan de diffusion localement, avant toute étude polarimétrique bistatique.

Chapitre 2

Les cibles élémentaires déterministes

2.1 Analyse théorique d'une sphère à basse fréquence

Afin d'illustrer l'impact du choix de convention nous commençons par étudier une sphère à basse fréquence dans une configuration bistatique quelconque. Puisqu'une sphère est un objet intrinsèquement symétrique, nous nous affranchissons ainsi de l'impact de l'orientation de la cible. Seuls les choix de repères de polarisation ainsi que la position relative des deux antennes affectent la mesure polarimétrique. Nous comparons ici la convention dite "classique" radar, où les vecteurs horizontaux sont perpendiculaires à un plan d'incidence vertical, avec la convention du plan de diffusion. Nous commençons par l'étude faite avec la convention du plan de diffusion car les modèles électromagnétiques sont justement décrits dans ce plan. A partir des expressions dans cette convention, nous pouvons obtenir les expressions dans le cas de la convention classique, imposée par la mesure.

2.1.1 Définition de quelques paramètres polarimétriques

Cette section est dédiée à la définition de quelques paramètres polarimétriques de la littérature que nous allons utiliser comme moyen d'étude dans la suite de ce chapitre. La description de ces paramètres est volontairement succincte dans la mesure où un chapitre de la dernière partie est intégralement dédié aux décompositions et aux paramètres polarimétriques optique et radar. Ici le choix se réduit donc à deux décompositions matricielles. La première, celle de Cloude et Pottier souvent appelée "H-alpha" [5] est très couramment utilisée en radar. Il s'agit d'une décomposition additive de la matrice de cohérence, basée sur une décomposition en valeurs propres / vecteurs propres. La seconde, est une décomposition multiplicative, de Lu et Chipman [6], fondée sur la décomposition de la matrice de Mueller en un produit de matrices de Mueller canoniques. Ces deux décompositions ont été choisies pour cette étude parce qu'elles sont les plus répandues dans leurs domaines respectifs (radar et optique). Dans le chapitre dédié aux décompositions nous aurons l'occasion de revenir plus en détail sur les différents choix possibles de paramètres polarimétriques, et nous présenterons d'ailleurs les résultats d'un plus grand nombre d'outils sur des données bistatiques.

Dans la première partie de notre étude, seules des cibles canoniques (et donc déterministes ou non dépolarisantes) sont étudiées. Nous nous restreindrons aux paramètres pertinents pour ce type de cible, et nous ne nous intéresserons pas à la dépolarisation/entropie.

Nous utilisons les notations suivantes : $\mathbf{S}$ est la matrice de Sinclair de la cible, les éléments de cette matrice de diffusion sont notés S_{hH}, S_{hV}, S_{vH} et S_{vV}. Le premier indice, en minuscule, correspond à la polarisation de l'onde diffusée, et le second, en majuscule à la polarisation de l'onde incidente.

$$\mathbf{S} = \begin{pmatrix} S_{\mathrm{hH}} & S_{\mathrm{hV}} \\ S_{\mathrm{vH}} & S_{\mathrm{vV}} \end{pmatrix}. \tag{2.1}$$

La matrice de Jones $\mathbf{J}$ est définie dans la convention FSA, elle est reliée à la matrice de Sinclair $\mathbf{S}$ par $\mathbf{S} = \mathrm{diag}([\mathbf{1}\ -\mathbf{1}])\mathbf{J}$. Le vecteur de diffusion polarimétrique $\hat{\mathbf{k}}$ exprimé dans la base de Pauli s'écrit : $\hat{\mathbf{k}} = \mathbf{1}/\sqrt{(\mathbf{2})}[\mathbf{S}_{\mathrm{hH}} + \mathbf{S}_{\mathrm{vV}}, \mathbf{S}_{\mathrm{hH}} - \mathbf{S}_{\mathrm{vV}}, \mathbf{S}_{\mathrm{hV}} + \mathbf{S}_{\mathrm{vH}}, \mathbf{j}(\mathbf{S}_{\mathrm{hV}} - \mathbf{S}_{\mathrm{vH}})]^{\mathbf{t}}$.

Les paramètres que nous allons étudier sont :

- L'angle α, donné par l'arcosinus de la première composante du vecteur de diffusion $\hat{\mathbf{k}}$.
- La diatténuation, définie comme un contraste entre les intensités de deux polarisations propres (polarisation associée aux vecteurs propres de la matrice de diffusion).
- La retardance R, définie comme le déphasage entre deux polarisations propres, traditionnellement exprimé en convention FSA. Dans cette convention on a donc par exemple $R = 180°$ pour un miroir dans le cas de la rétrodiffusion, et $R = 0°$ dans le vide dans le cas de la diffusion avant.

Nous rappelons ici l'origine de ces paramètres dans le cas d'une cible déterministe. Comme décrit dans [6], les éléments ayant un effet sur la polarisation, de même que leurs matrices de Jones, se divisent en deux groupes : les éléments homogènes, ayant des polarisations propres orthonormales, et les non-homogènes, ayant des polarisations propres non orthogonales. Mathématiquement les éléments homogènes possèdent des matrices de Jones dites normales, vérifiant donc $\mathbf{J}.\mathbf{J}^{\dagger} = \mathbf{J}^{\dagger}.\mathbf{J}$.

Les propriétés de polarisation que sont la diatténuation et la retardance, dans le cas d'un élément homogène, sont caractérisées par les valeurs propres et les vecteurs de polarisation propre. Ainsi, D, la diatténuation ou dichroïsme de la cible s'écrit $D = \frac{\sigma^2_{\max} - \sigma^2_{\min}}{\sigma^2_{\max} + \sigma^2_{\min}}$ où σ_{max} et σ_{min} sont les valeurs singulières de $\mathbf{J}$.

$\mathbf{J}$ et $\mathbf{S}$ possèdent les même valeurs singulières, parce qu'elles sont égales aux valeurs propres de $\mathbf{S}^{\dagger}\mathbf{S} = \mathbf{J}^{\dagger}\mathbf{J}$. Ces valeurs singulières σ^2_{max} et σ^2_{min} correspondent à la transmission des intensités maximum et minimum.

- Dans le cas particulier où la matrice de Jones est normale, alors $\sigma^2_{\max} = |\lambda_{\max}|^2$ et $\sigma^2_{\min} = |\lambda_{\min}|^2$ où $\lambda_{\max}$ et $\lambda_{\min}$ sont les valeurs propres de $\mathbf{J}$.
  La diatténuation s'écrit alors :

$$D = \frac{|\lambda_{\max}|^2 - |\lambda_{\min}|^2}{|\lambda_{\max}|^2 + |\lambda_{\min}|^2} \tag{2.2}$$

  et la retardance :

$$R = |\arg\lambda_{\max} - \arg\lambda_{\min}|. \tag{2.3}$$

Il est possible d'utiliser les valeurs propres de la matrice de Sinclair **S** aussi bien que celles de **J** pour calculer la diatténuation. Cela n'est pas vrai par contre pour la retardance.

– La diatténuation et la retardance d'un élément inhomogène sont plus difficiles à calculer, puisque leurs polarisations propres ne sont pas orthonormales. Les valeurs propres ne définissent donc pas une valeur correcte de diatténuation et de retard. Pour les éléments inhomogènes une étape préalable de décomposition polaire de la matrice de Jones est nécessaire. Cette étape est décrite en détail dans [6].

L'angle α est plus facile à calculer. Il est donné par l'arcosinus de la première composante du vecteur de diffusion $\hat{\mathbf{k}}$, ce qui correspond à l'expression, en fonction des termes de la matrice de diffusion :

$$\alpha = \arccos \sqrt{\frac{|S_{\mathrm{hH}} + S_{\mathrm{vV}}|^2}{2(|S_{\mathrm{hH}}|^2 + |S_{\mathrm{vV}}|^2 + |S_{\mathrm{vH}}|^2 + |S_{\mathrm{hV}}|^2)}}. \tag{2.4}$$

L'origine de la popularité de l'angle α est sa propriété dite de "roll-invariance" dans les configurations monostatiques. Cette propriété signifie en pratique que l'angle alpha demeure inchangé après rotation de la cible dans son plan d'onde incident. Mathématiquement, il est donc indépendant d'un changement de base de polarisation. Ce n'est plus le cas en bistatique, cependant si une correction des rotations des bases polarimétriques est effectuée son interprétation est toujours intéressante.

2.1.2 Une sphère diélectrique à basse fréquence dans le plan de diffusion

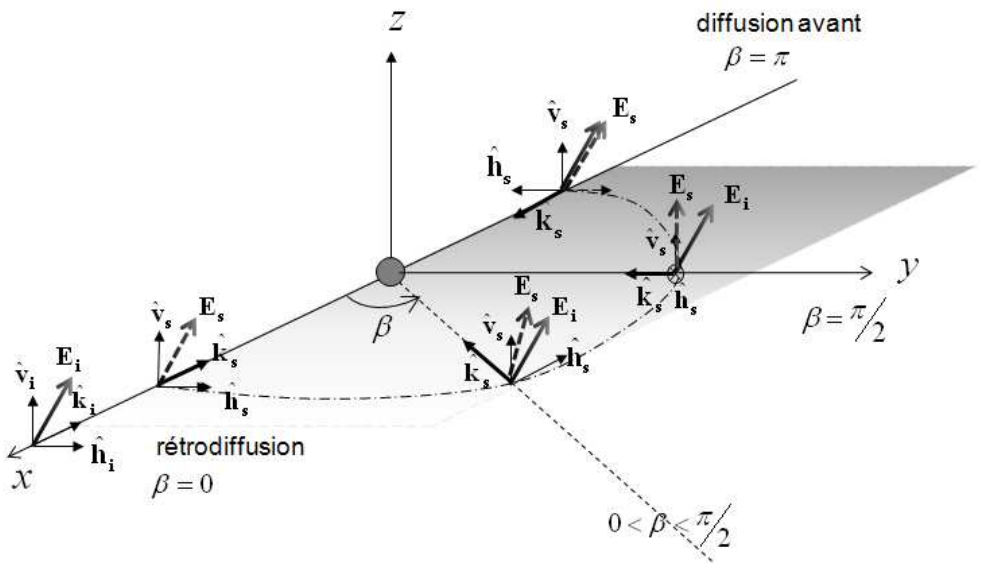


FIGURE 2.1 – Evolution des vecteurs polarisations dans le plan de diffusion en fonction de l'angle bistatique

Le modèle de diffusion de Rayleigh s'applique quand le rayon r d'une sphère diffusante est négligeable devant la longueur d'onde λ de l'onde électromagnétique incidente.

Une particule diélectrique, isotrope, vérifiant cette hypothèse a, dans le plan de diffusion, une matrice de Sinclair [7] :

$$\mathbf{S^P} = \begin{pmatrix} \cos\beta & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.5}$$

L'expression de cette matrice peut être vue comme une transformation géométrique qui affecte les axes de polarisation entre le plan d'incidence et le plan de diffusion. L'effet sur la polarisation est donc géométrique. Ainsi, le champ électrique incident $\mathbf{E_i}$ est projeté sur le plan de l'onde diffusé, laissant E_v inchangé et transformant E_h en $\cos\beta E_h$. Ces champs sont représentés sur la figure 2.1.

Le vecteur de diffusion $\mathbf{k}$ s'écrit après normalisation (ce qui revient à retirer l'influence du *span*, ou somme des intensités de tous les canaux polarimétriques) :

$$\mathbf{k} = \frac{1}{\sqrt{2(1+\cos^2\beta)}} \begin{pmatrix} \cos\beta + 1 \\ \cos\beta - 1 \\ 0 \\ 0 \end{pmatrix}. \tag{2.6}$$

De cette expression, nous pouvons déduire les paramètres polarimétriques décrits plus haut :

– La diatténuation :

$$D = \frac{1-\cos^2\beta}{1+\cos^2\beta}, \tag{2.7}$$

– L'angle α :

$$\alpha = \arccos\frac{1+\cos\beta}{\sqrt{2(1+\cos^2\beta)}}, \tag{2.8}$$

– La retardance de la sphère diélectrique est alors égale à 180° pour $\beta \in [0, 90°]$ et égale à 0 pour $\beta \in [90°, 180°]$.

Dans l'expression de ces paramètres polarimétriques on peut distinguer les propriétés intrinsèques de la cible, des effets géométriques dus à la configuration bistatique. Ainsi la diatténuation illustre parfaitement ce phénomène. Elle représente l'anisotropie de la cible : une forte diatténuation implique un fort changement d'amplitude entre deux polarisations. Pour la sphère à basse fréquence, cet effet est inexistant dans une situation monostatique de rétrodiffusion ou de diffusion avant, ce qui est un résultat conforme à l'intuition car cet objet est totalement isotrope. Cet effet de diatténuation est maximum quand on s'approche de l'angle $\beta = \frac{\pi}{2}$. La matrice de diffusion devient alors un polariseur vertical parfait, ce qui ne correspond pas à une propriété de la cible, mais est uniquement du à la géométrie particulière d'acquisition.

2.1.3 Une sphère diélectrique étudiée à basse fréquence dans la convention classique de mesure

Dans la convention classique radar, liée aux conditions de mesure, la matrice de Sinclair peut être exprimée à l'aide de l'équation 1.10, et le vecteur de diffusion **k** associé s'écrit :

$$\mathbf{k} = \frac{1}{\sqrt{2\left(1+\cos^2\beta\right)}} \begin{pmatrix} (\cos\beta+1)\cos(\alpha_i+\alpha_s) \\ (\cos\beta-1)\cos(\alpha_i-\alpha_s) \\ (\cos\beta-1)\sin(\alpha_i-\alpha_s) \\ j(\cos\beta+1)\sin(\alpha_i+\alpha_s) \end{pmatrix}, \tag{2.9}$$

Ce qui donne pour l'angle α :

$$\alpha = \arccos \frac{|(1+\cos\beta)\cos(\alpha_i+\alpha_s)|}{\sqrt{2(1+\cos^2\beta)}}. \tag{2.10}$$

On peut alors remarquer, que dans cette convention, n'importe quelle valeur peut être obtenue pour l'angle α pour un angle de bistatisme β donné. Il suffit de "jouer" sur les angles θ_i, ϕ_i, θ_s etϕ_s dans ce but. Cela permet d'exhiber par exemple des cas très particuliers, où la géométrie uniquement permet d'exhiber des réponses non symétriques pour une sphère.

Ainsi par exemple $\alpha_i = 45°, \alpha_s = 0°$ peuvent être obtenus avec $\theta_i = 90°, \theta_s = 45°, \phi_s = 0°, \phi_i = 90°$ ce qui résulte en une matrice de Sinclair :

$$\mathbf{S} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{2.11}$$

2.1.4 La sphère métallique à basse fréquence

La différence entre une sphère métallique et une sphère diélectrique est en fait assez importante sur les calculs théoriques qui s'y rapportent.

En effet lorsque l'on considère une sphère métallique parfaitement conductrice, on est amené à faire des hypothèses concernant les champs dans le volume qui ne sont plus vrais en diélectrique. Nous pourrions penser qu'en menant les calculs sur une sphère diélectrique, puis en faisant tendre la partie imaginaire de la permittivité diélectrique vers l'infini, on retrouve les résultats obtenus pour la sphère métallique. Malheureusement ce n'est pas le cas. L'expression relative à une sphère métallique diffère de celle d'une sphère diélectrique [7]. En effet, l'hypothèse basse fréquence et le modèle de diffusion de Rayleigh qui en découle ne permet pas une continuité entre le modèle parfaitement conducteur et le modèle diélectrique. Il s'agit d'un problème de "double limite" avec comme deux hypothèses limites "permittivité à partie imaginaire infinie" et "rapport rayon de la cible sur longueur d'onde infiniment petit", et ces deux limites ne commutent pas.

Nous étudions donc à la fois le cas métallique et le cas diélectrique. En effet :

– en pratique, il est plus facile de réaliser des cibles métalliques à échelle réduite en chambre anéchoïque, et les mesures ainsi réalisées bénéficient d'un meilleur rapport signal à bruit.

– il est très difficile de maîtriser parfaitement les permittivités complexes des matériaux diélectriques utilisés en mesures.
– les cibles naturelles d'intérêt en configuration réelle sont la plupart du temps diélectriques (végétaux, glace).

Ainsi une sphère métallique dans l'approximation de Rayleigh a, dans le plan de diffusion, la matrice Sinclair suivante :

$$\mathbf{S^P} = \begin{pmatrix} \frac{1}{2} + \cos\beta & 0 \\ 0 & \frac{1}{2}\cos\beta + 1 \end{pmatrix}. \tag{2.12}$$

De la même manière que précédemment on peut en déduire l'expression des paramètres polarimétriques suivants :

– la diattenuation est
$$D = \frac{\frac{3}{4}(1 - \cos^2\beta)}{\frac{5}{4}(\cos^2\beta + 1) + 2\cos\beta}. \tag{2.13}$$
Elle est nulle pour $\beta = 0°$ et $\beta = 180°$, et maximum pour $\beta = 120°$.
– α est égal à :
$$\alpha = \arccos\frac{3|1 + \cos\beta|}{\sqrt{10\cos^2\beta + 10 + 16\cos\beta}}. \tag{2.14}$$
L'angle α est nul pour $\beta = 0°$, et est égal à $90°$ pour $\beta = 180°$, et on a $\alpha = 45°$ pour $\beta = 120°$.
– la retardance de la sphère métallique sera alors égale à $180°$ pour $\beta \in [0°, 120°]$ et égale à $0°$ pour $\beta \in [120°, 180°]$.

Ces expressions de α, de la retardance ainsi que de la diattenuation sont présentées, pour la sphère théorique à basse fréquence sur la figure 2.2. Ces expressions sont données dans la convention du plan de diffusion. Nous pourrions, comme pour la sphère diélectrique, présenter les résultats dans le repère lié à la mesure, mais l'impact d'un mauvais choix de convention reste similaire, en outre les expressions littérales sont sensiblement plus compliquées et présentent peu d'intérêt. Notons par contre cette différence théorique de l'angle de bistatisme de transition entre le cas de la sphère métallique et celui de la sphère diélectrique : dans le second cas, cette zone de transition de situe en $\beta = 120°$, alors que le cas de la sphère diélectrique est symétrique autour de $\beta = 90°$.

2.2 Etude théorique du cylindre vertical à basse fréquence

2.2.1 Modélisation simplifiée, inspirée de la sphère diélectrique à basse fréquence

Le cylindre est l'une des cibles élémentaires les plus courantes après la sphère. On s'intéresse ici au cylindre "long", c'est à dire dont la hauteur est grande devant le rayon. Il présente des propriétés polarisantes intéressantes. L'origine du phénomène s'explique assez simplement ; dans un des axes, parallèle à l'axe du cylindre, la circulation du courant induit n'est pas perturbée puisque le cylindre se comporte comme un fil conducteur. Inversement

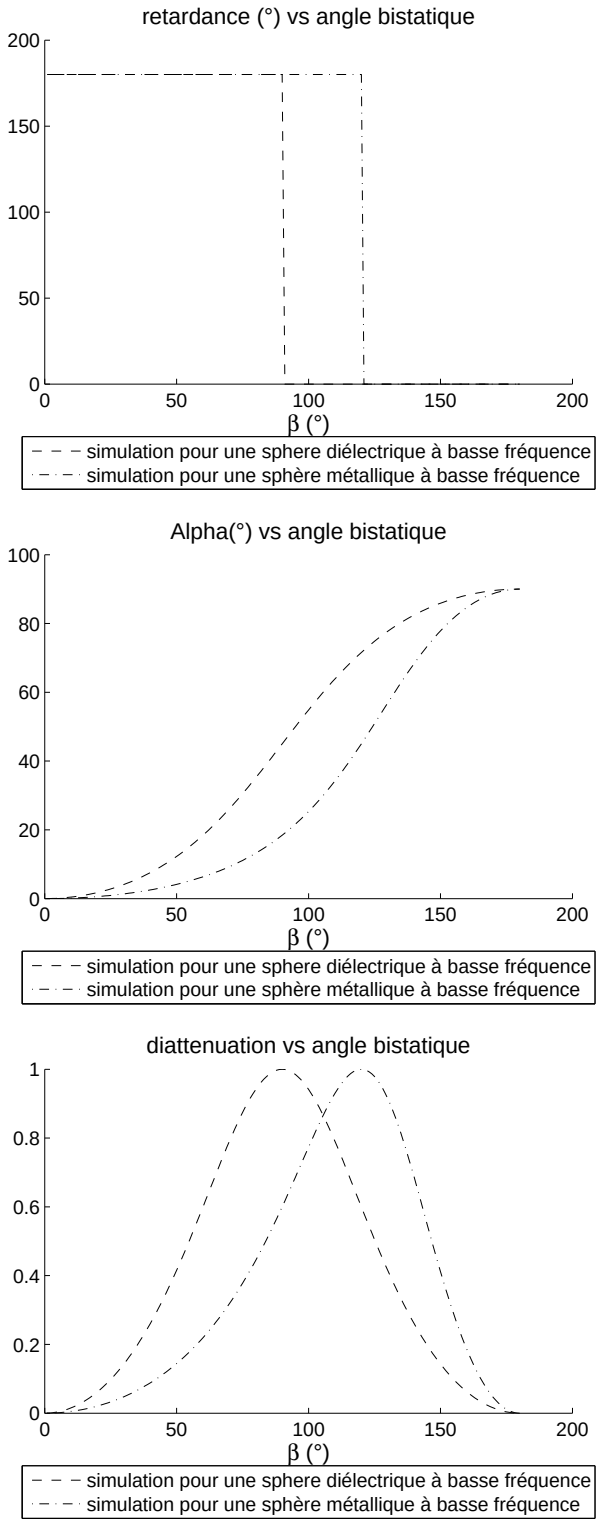


FIGURE 2.2 – Paramètres polarimétriques pour la sphère à basse fréquence : Retardance, α et diatténuation

l'oscillation d'électrons dans l'axe orthogonal à l'axe du cylindre est perturbée, puisque contrainte dans la section transversale du cylindre. Dans le cas particulier du cylindre infiniment fin, cette oscillation est même impossible. Ainsi quand une onde polarisée vient exciter ce cylindre, une polarisation parallèle au cylindre sera non affectée, alors qu'une polarisation perpendiculaire à celui ci sera atténuée d'un facteur proportionnel à la largeur de celui ci. Dans le cas d'un cylindre infiniment fin, il s'agit d'un polariseur parfait, car seule la polarisation parallèle est diffusée. C'est d'ailleurs sur ce principe que sont conçu la plupart des polariseurs/analyseurs linéaires.

Par rapport à une sphère, le cylindre a des propriétés polarimétriques propres, qui se combinent aux effets induits par la géométrie. Nous proposons ici de le modéliser simplement en utilisant deux paramètres a et b dont la différence dépend de sa finesse ou anisotropie. Ainsi si $\frac{a}{b}$ tend vers l'infini alors le cylindre est infiniment fin, et si $a = b$ on obtient une cible similaire à la sphère.

En monostatique la matrice de Sinclair du cylindre vertical est simplement :

$$\mathbf{S} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \tag{2.15}$$

En bistatique on suppose maintenant qu'un cylindre diélectrique vertical s'exprimera dans le plan de diffusion de manière similaire à la sphère, par la matrice suivante :

$$\mathbf{S^P} = \begin{pmatrix} a\cos\beta & 0 \\ 0 & b \end{pmatrix} \tag{2.16}$$

La condition pour que ce modèle approché convienne est que le rayon du cylindre soit petit devant la longueur d'onde utilisée.

De la même manière on peut déduire les paramètres polarimétriques :

– diatténuation :

$$D = \frac{1 - (\frac{a}{b})^2 \cos^2\beta}{1 + (\frac{a}{b})^2 \cos^2\beta}, \tag{2.17}$$

– angle α :

$$\alpha = \arccos \frac{1 + \frac{a}{b}\cos\beta}{\sqrt{2(1 + (\frac{a}{b})^2 \cos^2\beta)}}, \tag{2.18}$$

– La retardance du cylindre diélectrique sera similaire à la sphère et égale à 180° pour $\beta \in [0, 90°]$ et égale à 0 pour $\beta \in [90°, 180°]$.

Les expressions de l'angle α, de la retardance ainsi que de la diatténuation sont données pour le cylindre théorique à basse fréquence, et les courbes sont présentées sur la figure 2.3. La convention choisie est celle du plan de diffusion.

On observe ici que les effets sur la diatténuation sont combinés avec la diatténuation propre liée à la cible. L'important est de constater que l'estimation du rapport a/b est perturbé par la configuration bistatique.

Ces expressions simplifiées du cylindre à basse fréquence restent valables pour des cylindres légèrement inclinés, tant que les dimensions de sa section dans le plan de diffusion restent faibles devant la longueur d'onde. L'angle d'inclinaison du cylindre par rapport à la verticale doit donc rester faible.

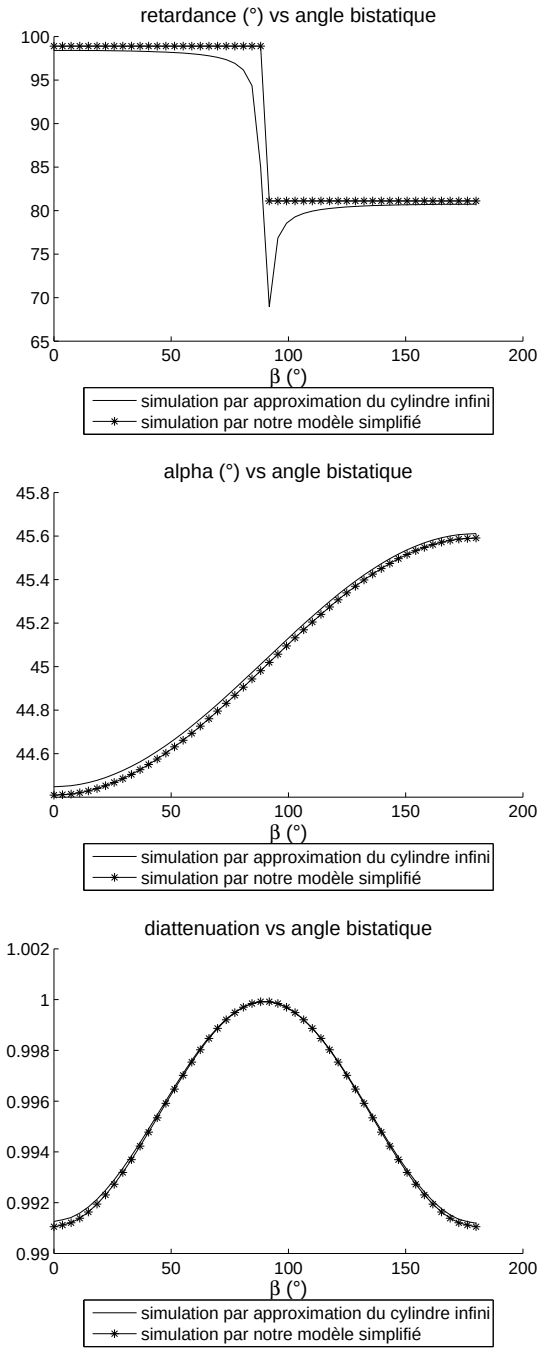


FIGURE 2.3 – Paramètres polarimétriques pour le cylindre infini à basse fréquence : Retardance, α et diatténuation

2.2.2 Modélisation du cylindre infini par résolution des Equations de Maxwell

Nous cherchons à valider nos approximations en les confrontant aux résultats de modèles électromagnétismes moins approximatifs. Pour des diffuseurs cylindriques, il n'existe pas de calcul exact permettant de déterminer l'amplitude de diffusion. Cependant, il est possible d'obtenir une bonne approximation pour le cylindre infini.

Puisque nous disposons déjà de tels modèles, par le biais du simulateur COSMO, nous avons pu comparer les courbes tirées de notre modélisation simplifiées, aux courbes données par ces modèles.

Nous avons superposé ces courbes à nos précédentes sur la figure 2.3. Leur correspondance est très satisfaisante.

2.3 Analyse et mesure d'une à quatre sphères à basse fréquence

Nous allons maintenant présenter et analyser la mesure polarimétrique réalisée sur une, deux, puis quatre sphères métalliques à basse fréquence. Ces mesures seront comparées à des simulations effectuées à l'aide du modèle de Mie pour des sphères métalliques.

2.3.1 La chambre anéchoïque BaBi

La chambre anéchoïque BaBi est un des outils de mesures présents au sein du Département ElectroMagnétique et Radar. Elle est située dans le centre de Palaiseau et permet d'effectuer un grand nombre de mesures à échelle réduite afin de préparer des campagnes à grandes échelles, valider des modèles électromagnétiques ou des traitements. Les antennes d'émission et de réception sont placées sur des rails, les moteurs servant à déplacer les antennes permettent d'atteindre des précisions de position de l'ordre de 10^{-2} degrés. Deux rails peuvent être utilisés, le premier est un rail demi-circulaire de 5.5m de rayon placé à 2.5m de hauteur. Le second rail est linéaire et est fixé sur le plafond de la chambre. Le dispositif expérimental est illustré sur la figure 2.6. Dans cette section les objets d'études sont des sphères métalliques de 38mm de rayon. Une, deux puis quatre sphères ont été utilisées, et ont été placées au centre d'une plate-forme en polystyrène (transparent aux fréquences considérées). La hauteur de cette plate-forme peut être ajustée pour faire varier les angles d'élévation de 70° à 90°.

Dans l'ensemble des mesures sur les sphères à basse fréquence, c'est une configuration bistatique "complexe" qui a été choisie. En effet dans la chambre BaBi il est possible de se placer sur le rail circulaire et ainsi de se placer directement dans la configuration dite du plan de diffusion. Ici nous chercherons à mettre en avant l'avantage de ce plan et vérifier notamment l'intérêt des relations de passage ; nous choisissons donc volontairement une configuration plus difficile. Pour obtenir à la fois un angle d'élévation non nul et un angle de bistatisme en azimut, l'émetteur est placé sur le rail linéaire en hauteur dans la chambre. Il aura une position fixe définie par $\theta_i = 55°$ et $\phi_i = -60°$. L'antenne de réception est placée sur un rail circulaire horizontal (figure 3.6). La trajectoire de l'antenne de réception

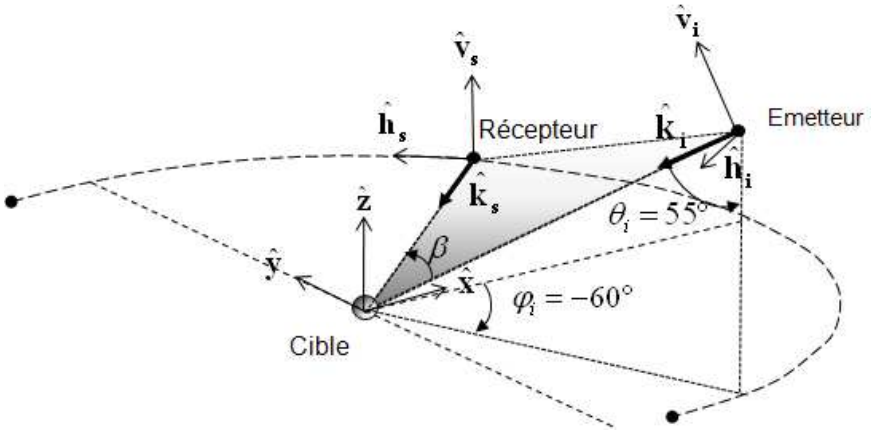


FIGURE 2.4 – Position des antennes et de la cible pour la mesure de une à quatre sphères métalliques

peut être décrite par $\theta_s = 90°$ et $\phi_s \in [-85°, 100°]$. La gamme de fréquence qui peut être utilisée dans la chambre anéchoïque est très importante. Ici nous voulons travailler dans la zone de résonance ; les fréquences choisies vont de 1.47 GHz à 1.68 GHz avec 801 points d'échantillonnage.

La position des sphères est représentée sur la figure 2.5 en vue de dessous. Lorsqu'une seule sphère est mesurée, elle est placée à la position 1 du centre de la chambre. Lorsque deux sphères sont utilisées, elles sont placées aux positions 1 et 2, et elles sont alors placées dans le plan d'incidence. Lorsque les quatre sphères sont utilisées, il n'existe aucune symétrie particulière.

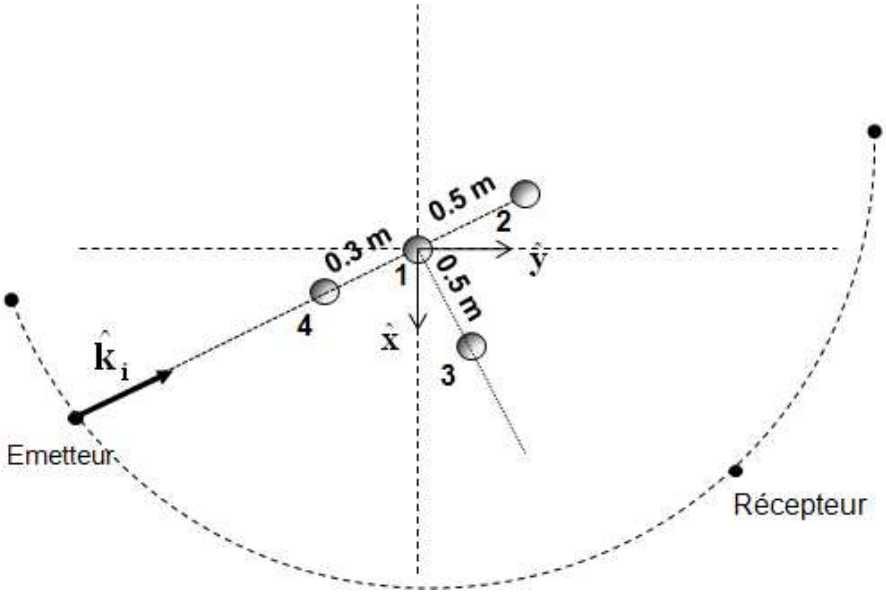


FIGURE 2.5 – Positions des quatre spheres

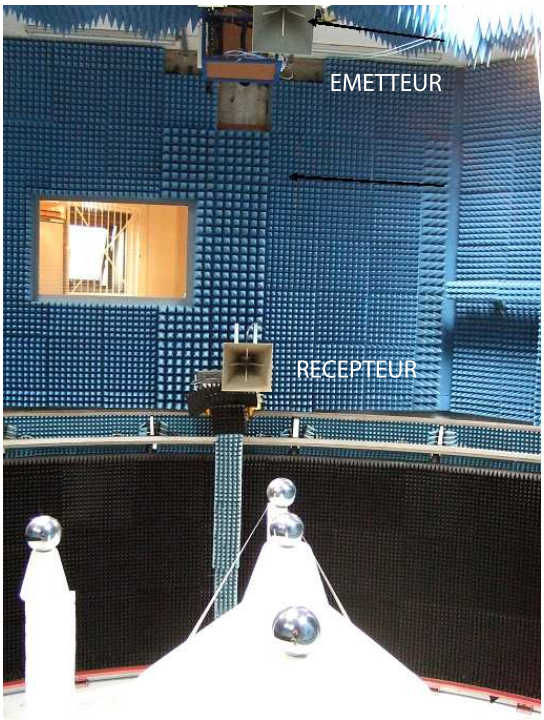


FIGURE 2.6 – Photographie de l'experience

La calibration des antennes est effectuée en position monostatique sur un certain nombre de cibles (l'utilisation d'un plus grand nombre de cibles canoniques que le nombre minimal requis permet de minimiser les erreurs). Plusieurs sphères étalons de différentes tailles, ainsi que des dièdres orientés sont utilisés. Pour chaque cible la réponse théorique est parfaitement connue. Les termes de correction estimés par un étalonnage monostatique sont alors appliqués pour la mesure bistatique. Une dernière mesure dite "contre-étalon" est effectuée sur une cible connue pour vérifier la précision de calibration. Dans notre cas l'erreur maximum de calibration est de l'ordre de 1.2 dB.m^2 dans toute la bande de fréquence.

2.3.2 Théorie

Compte tenu du rayon des sphères et de la bande de fréquence utilisée, le modèle de diffusion de Rayleigh n'est pas suffisant pour décrire l'interaction entre l'onde électromagnétique et la cible. Le modèle de Mie, [8], plus général, permet de décrire l'onde diffusée par une sphère quel que soit le rapport entre le diamètre et la longueur d'onde. L'expression de ces champs se fait en somme infinie de termes et ne présente pas une expression aussi simple que dans le cas de Rayleigh [7]. Plus précisément la solution de Rayleigh correspond au premier terme de la somme donnée par la solution de Mie. Nous allons donc ici confronter nos mesures au modèle le plus précis, donc celui de Mie. Il est important par contre de noter que comme la solution de Mie correspond à un raffinement de la solution de Rayleigh, on s'attend à observer des effets du même type que ceux que

nous avions décrits dans la partie précédente.

2.3.3 Résultats

Représentation polarimétrique colorée

Nous proposons sur les figures suivantes une représentation en couleurs du signal polarimétrique. Les ordonnées correspondent à une variation du paramètre de taille de la cible, défini par kr ou $2\pi r/\lambda$. En pratique ces différentes ordonnées sont obtenues par les différentes mesures fréquentielles effectuées en chambre. En abscisse, c'est l'angle de bistatisme qui varie.

Trois composantes peuvent être utilisées dans une représentation colorée (Rouge, Vert et Bleu). Dans une configuration monostatique, la base de Pauli est couramment utilisée pour ce genre de représentation colorée. Ainsi la couleur rouge correspond à la puissance pour la polarisation |hH+vV|, le vert à la polarisation |hH-vV|, et le bleu à |hV+vH|. En monostatique comme l'hypothèse de réciprocité est valide, on a $S_{\mathrm{hV}} = S_{\mathrm{vH}}$. Ainsi les trois composantes sont suffisantes pour représenter l'information polarimétrique. Pour une configuration bistatique ce n'est plus le cas, et il faudrait a priori une manière alternative de représenter l'information polarimétrique, avec quatre paramètres.

Cela dit pour notre exemple, comme le plan de diffusion est utilisé pour représenter l'information polarimétrique, et que nous étudions une cible symétrique, les polarisations croisées sont nulles [9], et la représentation de Pauli est donc encore suffisante.

Une sphère dans la zone de résonance

Nous présentons ainsi un exemple de ce type de représentation sur la figure 2.7 pour une sphère dans la zone de résonance. Rappelons que cette zone de résonance correspond simplement à un paramètre de taille proche de 1. Les franges présentes sur cette représentation correspondent à des interférences avec un autre point brillant de la scène, et sont donc probablement dues à une interaction parasite avec un élément de la chambre. Nous ne pouvons malheureusement pas filtrer celle ci compte tenu de la taille de la bande de fréquence. Cela dit elle ne perturbe pas de façon importante l'information présentée ici.

Comme pour le comportement théorique de la sphère à basse fréquence dans l'approximation de Rayleigh, nous observons un changement continu du comportement polarimétrique en fonction de β. Le signal évolue d'une situation où S_{hH} et S_{vV} ont le même signe (induisant une retardance, $R = 180°$, et un angle $\alpha = 0°$) à un signal où ces composantes ont des signes opposés. On voit donc sur la figure 2.7 que la transition s'effectue maintenant aux alentours de $\beta = 110°$. Mathématiquement on retrouve qu'au niveau de cette transition on a $\alpha = R = 90°$. Cela correspond à un minimum pour $\mathbf{S}_{\mathrm{hH}}$ comme on peut le voir sur la figure 2.8.

La figure 2.8 nous permet de vérifier qu'en utilisant le plan de diffusion comme plan de référence on retrouve bien des termes croisés égaux et nuls pour la sphère, ce qui était attendu pour une cible symétrique. Cela n'est plus le cas lorsque c'est le plan d'incidence qui est utilisé comme plan de référence. Cela confirme des résultats déjà présentés dans [10]. Dans la suite nous ne considérerons plus que le plan de diffusion comme plan de

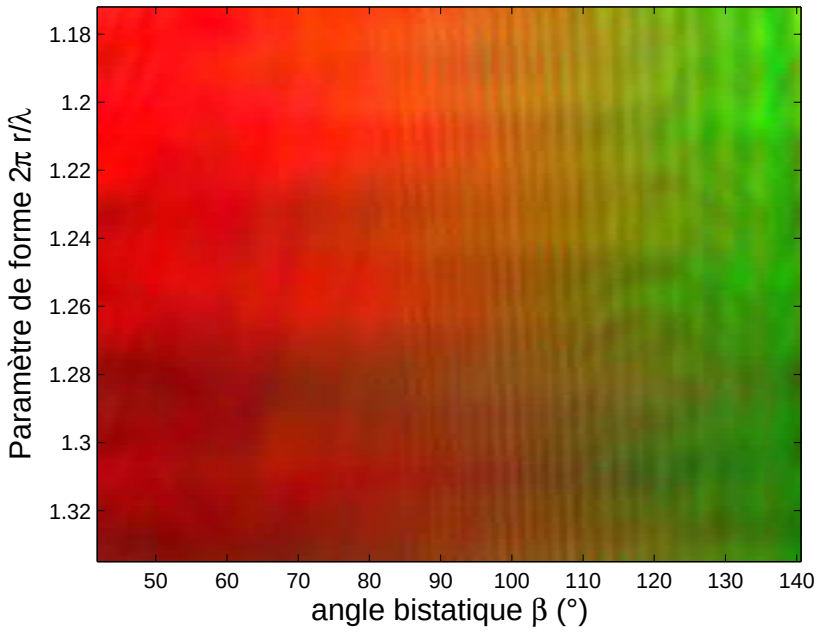


FIGURE 2.7 – Représentation polarimétrique colorée du champ diffusé par une sphère en fonction du paramètre de forme et de l'angle de bistatisme β, - Rouge : $|hH+vV|^2$, Vert : $|hH-vV|^2$, Bleu : $|hV+vH|^2$

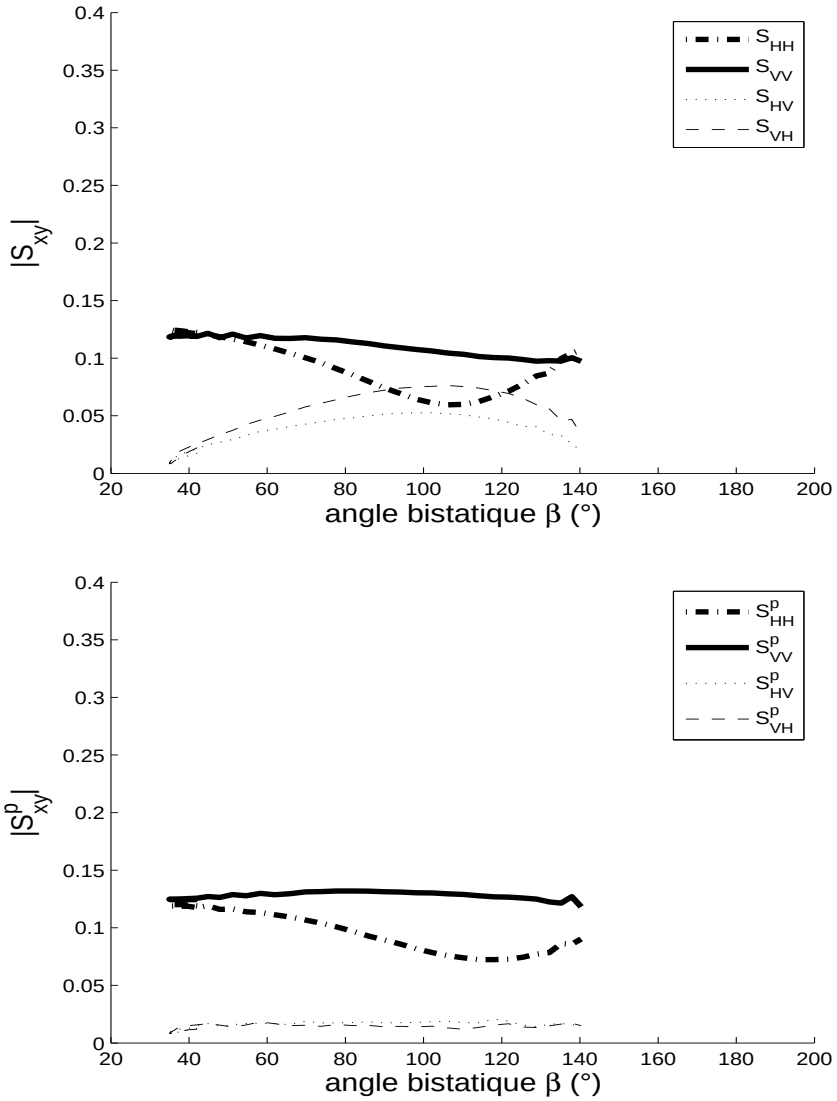


FIGURE 2.8 – Amplitude du signal polarimétrique du champ diffusé par une sphère : comparaison des deux conventions (plan d'incidence et plan de diffusion) pour $r/\lambda = 0.2$

référence, comme dans [9].

Nous considérons maintenant les trois paramètres que nous avons déjà étudiés pour la sphère dans l'approximation de Rayleigh. Nous calculons ainsi les paramètres α, la retardance et la diatténuation pour une sphère dans la zone de résonance, pour des données simulées ainsi que sur les résultats de la mesure en chambre. Les courbes sont représentées sur la figure 2.9. Nous ajoutons par ailleurs les résultats présentés précédemment pour les sphères à basse fréquence dans l'approximation de Rayleigh.

Nous pouvons en déduire :

- Dans la zone de résonance, la discontinuité sur la retardance n'est plus présente.
- La transition située à $R = 90°$ est maintenant obtenue pour $\beta \simeq 110°$.
- La transition située à $\alpha = 45°$ est obtenue pour le même angle de bistatisme, ie. $\beta \simeq 110°$.
- Il existe deux effets sur la diatténuation : la valeur maximum atteinte est de l'ordre de 0.5 et non plus de 1. Cette valeur maximum est aussi atteinte pour $\beta \simeq 110°$.

Deux et quatre sphères dans la zone de résonance

Nous allons maintenant introduire des sphères supplémentaires, qui introduiront des interférences constructives et destructives. Leurs positions ont été décrites plus haut, et il s'agit de sphères de même rayon. Les polarisations croisées restent nulles lorsque le plan de diffusion est utilisé. Les résultats des deux mesures sont présentés sur les figures 2.10, pour les deux sphères, et 2.11 pour les quatre sphères. Nous observons des interférences liées aux différences de marche comme nous pouvions l'attendre. Celles ci n'affectent par contre que l'énergie totale du signal mais pas ou très peu l'information polarimétrique, déterminée par l'énergie relative dans les différent canaux. Nous pouvons encore une fois constater l'intérêt de la convention du plan de diffusion sur les deux figures. Nous pouvons par contre conclure que dans la mesure ou aucun changement sur la polarisation n'a été observé, l'influence des diffusions multiples est a priori négligeable.

2.4 Conclusion

Dans ce chapitre nous avons présenté les résultats de mesures et de simulations concernant des cibles élémentaires simples en configuration bistatique. Nous avons ainsi illustré l'intérêt du choix du plan de diffusion comme plan de référence pour le choix des bases polarimétriques à l'émission et la réception. Concernant l'impact de la configuration géométrique sur l'information polarimétrique, nous observons un accord entre les différents modèles étudiés et la mesure. Nous remarquons les principaux effets suivants :

- La diatténuation, nulle en configuration monostatique pour une sphère, augmente avec l'angle de bistatisme, et prend une valeur maximale entre $[90-100]^\circ$ en fonction du milieu (métallique ou diélectrique).
- Le retard évolue entre les deux valeurs extremes (180° en rétrodiffusion et 0° en diffusion avant), de manière continue dans le cas métallique mais discontinue pour le cas diélectrique.

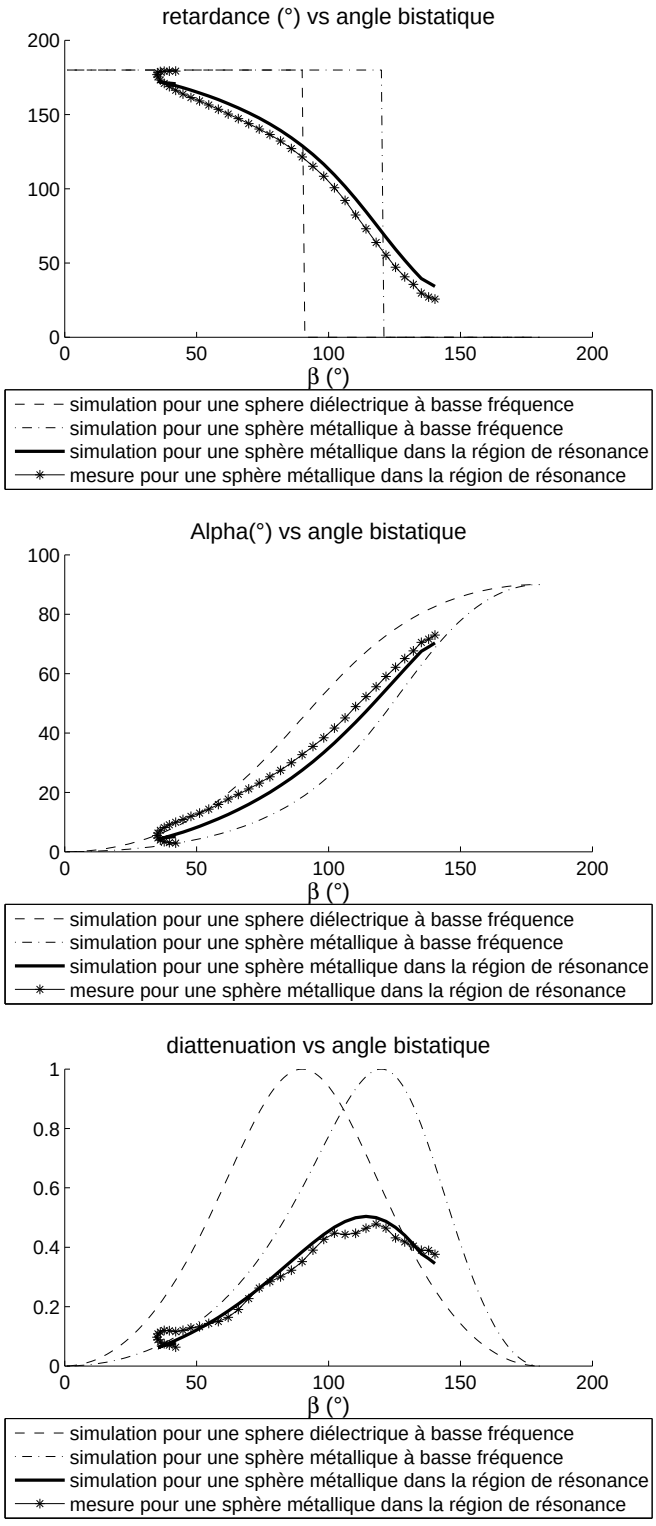


FIGURE 2.9 – Paramètres polarimétriques pour la sphère en région de résonance : Retardance, α et diatténuation

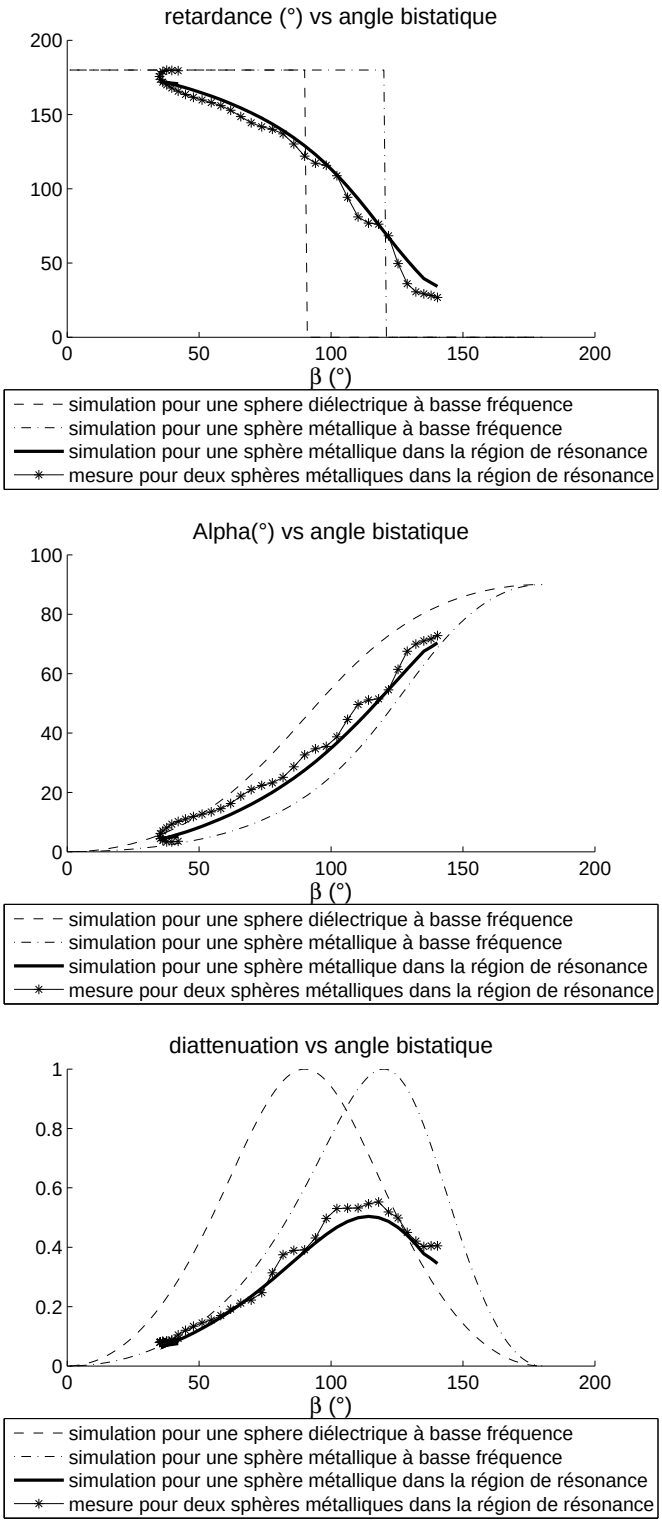


FIGURE 2.10 – Paramètres polarimétriques pour deux sphères en region de résonance : Retardance, α et diatténuation

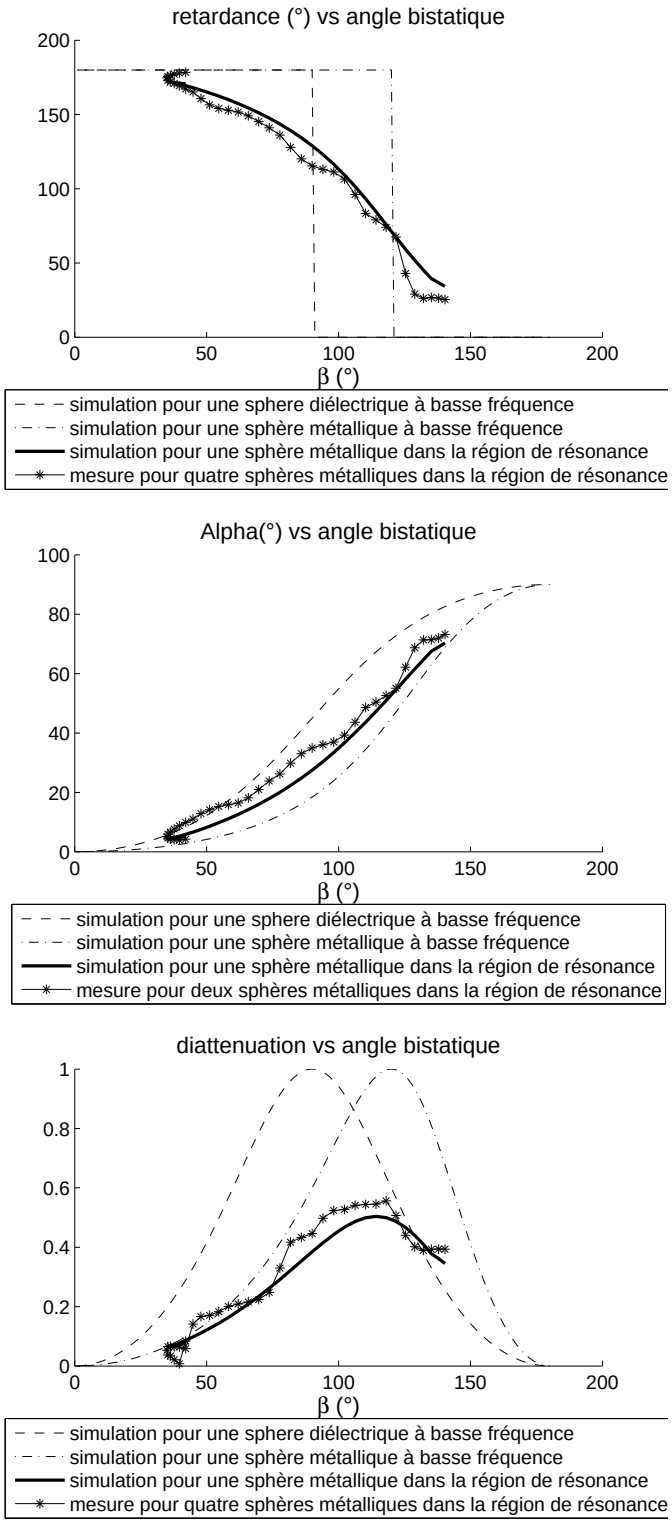


FIGURE 2.11 – Paramètres polarimétriques pour quatre sphères en region de résonance : Retardance, α et diatténuation

– L'angle alpha prend l'ensemble des valeurs entre 0° et 90° de manière continue quel que soit le matériaux composant les sphères.

Ces effets peuvent donc amener à une interprétation erronée des paramètres quand la géométrie de la scène est mal connue : il est possible de confondre une cible anisotrope (cylindre) et une sphère par exemple. Mais lorsque la géométrie de la scène est parfaitement contrôlée, ces effets peuvent permettre d'identifier les caractéristiques électromagnétiques des particules diffusantes.

Chapitre 3

Vers les cibles dépolarisantes

3.1 Mesure d'un nuage de sphères dans la zone de résonance

3.1.1 Objectif et nouveaux paramètres

Nous avons observé qu'avec quatre sphères dans la zone de résonance, l'influence des diffusions multiples était négligeable. Nous cherchons maintenant à mettre en évidence des comportements polarimétriques qui ne peuvent être déduits du simple modèle de Mie, mais qui traduiraient l'existence de trajets multiples et d'effets de type volumique. Nous cherchons à approcher une cible ayant un comportement de type "naturel" donc ayant notamment des propriétés de dépolarisation.

Nous considérons donc maintenant un nuage de l'ordre d'un millier de sphères. Pour des raisons d'ordre pratique, les billes mesurées sont maintenant beaucoup plus petites, et les longueurs d'onde utilisées également.

Nous pouvons travailler sur les mêmes paramètres polarimétriques que nous avons pu utiliser pour la sphère ou le cylindre, cependant il convient d'introduire des paramètres supplémentaire capables de mesurer des propriétés de dépolarisation. Nous introduisons donc les deux paramètres supplémentaires suivants :

- La dépolarisation, obtenue par la décomposition de Lu et Chipman [11] de la matrice de Mueller $\mathbf{M}$. Cette décomposition correspond à une décomposition polaire généralisée pour des matrices de Mueller dépolarisantes. Elle transforme la matrice de Mueller en un produit de trois matrices de Mueller canoniques. Lu et Chipman ont démontré dans [11] que toute matrice de Mueller peut se décomposer comme $\mathbf{M} = \mathbf{M_\Delta M_R M_D}$ où les matrices élémentaires $\mathbf{M_\Delta}$, $\mathbf{M_R}$, $\mathbf{M_D}$ représentent respectivement un dépolariseur pur, un retardeur et un diatténuateur. Ainsi à travers cette décomposition nous pouvons continuer à utiliser les paramètres de diatténuation et de retard pour des cibles non déterministes. La dépolarisation est déduite de la matrice du dépolariseur pur :

$$\mathbf{M_\Delta} = \begin{pmatrix} 1 & \mathbf{0^t} \\ \mathbf{0} & \mathbf{m_\Delta} \end{pmatrix}, \tag{3.1}$$

où :

$$\Delta = 1 - \frac{\text{trace } (\mathbf{m}_\Delta))}{3}. \tag{3.2}$$

– L'entropie, introduite par Shane Cloude et Eric Pottier [5]. Ce paramètre est introduit dans une décomposition en vecteur propre de la matrice de cohérence. Cette matrice, définie, rappelons-le, par $\mathbf{T} = \left\langle \mathbf{k}\mathbf{k}^\dagger \right\rangle$ est hermitienne et peut donc se décomposer en une somme :

$$\mathbf{T} = \lambda_1 \mathbf{k_1} \mathbf{k_1^\dagger} + \lambda_2 \mathbf{k_2} \mathbf{k_2^\dagger} + \lambda_3 \mathbf{k_3} \mathbf{k_3^\dagger}, \tag{3.3}$$

elle est donc interprétée comme une somme incohérente de trois mécanismes déterministes, décrits par les vecteurs propres $\mathbf{k_i}$ dont les poids relatifs est donné par les valeurs propres λ_i. L'entropie H correspond au degré d'incertitude des mécanismes présents dans la zone étudiée. On interprète donc p_i comme la probabilité de chaque événement $\mathbf{k_i}$: $H = -\sum_{i=1}^{3} p_i \log_3 p_i$, où $p_i = \frac{\lambda_i}{\sum_{i=1}^{3} \lambda_i}$.

De la même manière nous pouvons étendre la définition de l'angle $\bar{\alpha}$ au sens de Bernoulli [12], dans le cas d'une cible non déterministe, par une somme affectée de poids relatifs aux valeurs propres :

$$\bar{\alpha} = \sum_{i=1}^{3} p_i \alpha_i \tag{3.4}$$

où chaque α_i correspond au vecteur propre $\mathbf{k_i}$.

Un moyenne statistique est nécessaire pour estimer les matrices de Mueller ou de cohérence dont sont extraits les paramètres que nous venons de présenter. En SAR, cette moyenne statistique est traditionnellement calculée par une moyenne spatiale dans les pixels au voisinage du pixel d'intérêt. Ici nous travaillons sur des mesures brutes, donc nous choisissons plutôt une moyenne fréquentielle sur toute la bande disponible, dans la mesure où la bande choisie est assez restreinte et où nous observons très peu de variations significatives des paramètres déterministes (diatténuation, retard, et α) sur cette plage. Cela dit cela reste une approximation. Une vraie moyenne statistique demanderait de multiplier les mesures un grand nombre de fois pour pouvoir disposer de plusieurs réalisations.

3.1.2 La cible

Pour construire une cible ayant un comportement volumique, et dont nous pourrions modéliser le comportement de chaque particule, nous avons construit un cylindre de polystyrene, contenant un millier de petites sphères. Ce cylindre se décompose en 26 tranches de 3 à 10mm de hauteur et de 19.5 cm de rayon. Le cylindre a ainsi une hauteur totale de 13 cm. Chaque tranche contient 30 ± 2 sphères de 4.5mm de diamètre en métal distribuées de manière aléatoire dans ce disque. En tout, environ 780 sphères ont été reparties aléatoirement dans le cylindre, ainsi la distance moyenne entre elles est de 2.7cm. L'épaisseur des tranches a volontairement été choisie aléatoirement pour éviter tout effet de type réseau qui aurait pu survenir. La cible est présentée sur la figure 3.1. Nous rappelons que, le polystyrène étant transparent aux longueurs d'ondes considérées, la cible correspond à un ensemble de sphères réparties aléatoirement dans du vide. Ceci à été vérifié par la mesure

préalable d'un même cylindre ne contenant aucune sphère, dont la mesure a donné des résultats étant 20dB en dessous de la mesure de la cible contenant les sphères jusqu'a un angle de bistatisme de 140°. Ici nous utilisons directement le rail circulaire de 5.5m de rayon sur lequel sont placées les deux antennes. Nous n'avons maintenant plus "besoin" d'effectuer de transformation géométrique : la mesure se fait directement dans le plan de diffusion. La précision de mesure atteinte est inférieure à 0.3dB.m^2.

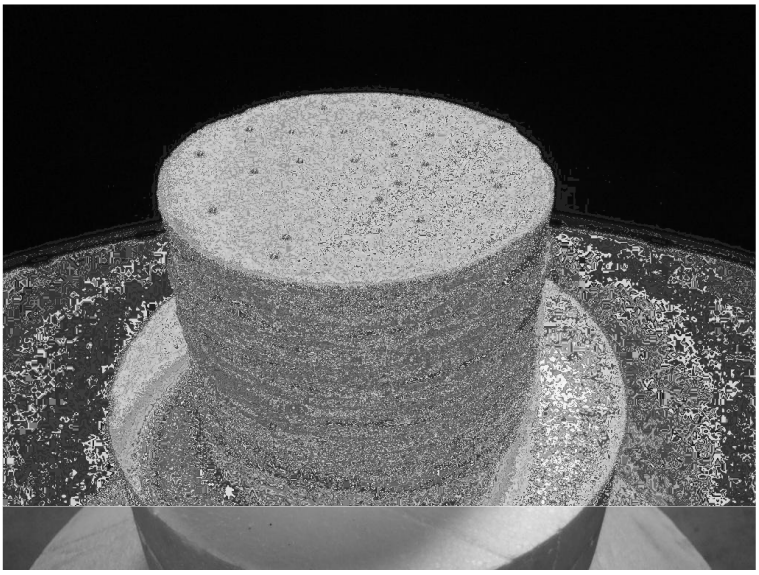


FIGURE 3.1 – Cible constituée d'un nuage de billes métalliques réparties dans un volume cylindrique.

Nous sommes déjà dans le plan de diffusion, nous avons donc $\theta_i = \theta_s = 90°$. Nous fixons l'émission à $\phi_i = 0$, et la réception évolue sur le rail, allant de $\phi_s = 4°$ à $\phi_s = 190°$ avec un pas de 0.5°. Les fréquences utilisées sont dans la bande [29 GHz, 31 GHz] avec un pas de 15 MHz. Dans la configuration du rail demi circulaire nous pouvons a priori atteindre toute les configurations bistatiques envisageables. L'encombrement des antennes d'émission et de réception fait que la configuration de rétrodiffusion n'est pas rigoureusement possible, ainsi l'angle minimum de bistatisme est de 4°. Compte tenu de la taille des cibles et de la largeur du lobe d'antenne il est aussi difficile d'obtenir des configurations dites de diffusion avant. En effet pour des angles de bistatisme supérieurs à 140° les signaux allant directement de l'antenne d'émission à la réception, sans être diffusés par la cible, aveuglent la réception, noyant l'information utile dans le bruit. Nous ne pouvons donc garantir les résultats que pour $\beta < 140°$.

3.1.3 Résultats et analyse

Nous présentons les résultats expérimentaux sur la figure 3.2 et nous pouvons observer que l'amplitude des polarisations croisées est, contrairement à toutes les mesures précédentes, non négligeable comme nous pouvons le voir sur la représentation colorée où du

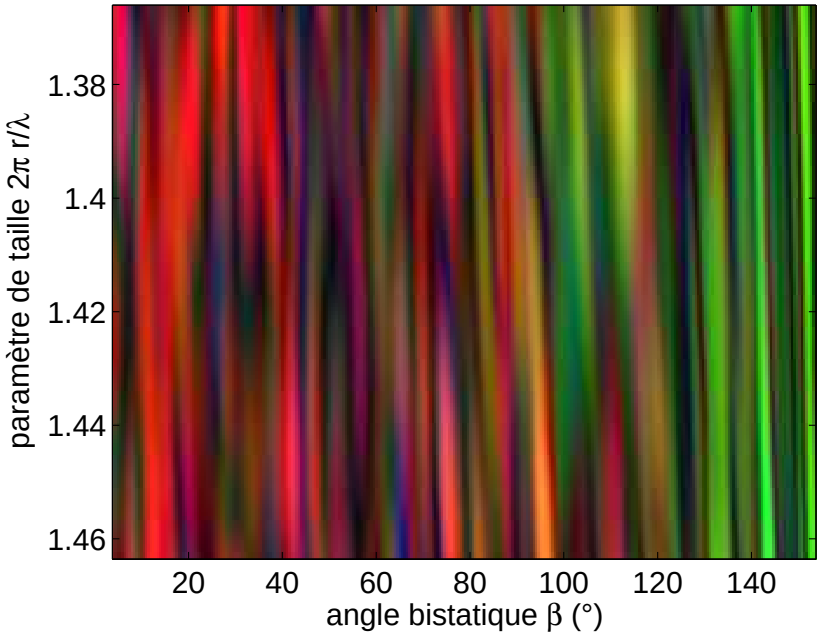


FIGURE 3.2 – Représentation polarimétrique colorée du champ diffusé par un nuage de sphères en fonction du paramètre de taille et de l'angle de bistatisme β, - Rouge : $|hH+vV|^2$, Vert : $|hH\text{-}vV|^2$, Bleu : $|hV+vH|^2$

bleu apparaît. Cette représentation n'est alors plus suffisante pour représenter tous les degrés de liberté. Ceci confirme l'importance des polarisations croisées hV et vH.

Nous voulons démontrer l'existence d'interaction et de trajet multiples. Pour cela nous allons rapidement simuler la réponse de ce nuage en l'absence d'interactions et comparer aux résultats de la mesure. Si on néglige les interactions, la diffusion par l'ensemble du cylindre correspond à une somme cohérente des champs diffusés par chacune des sphères indépendamment. Chaque contribution de chaque sphère est affectée d'une différence de phase induite par les positions relatives des billes dans le cylindre. Si on fait l'hypothèse d'une onde plane, on peut alors exprimer la matrice de Sinclair $\mathbf{S}$:

$$\mathbf{S}(f,\beta) = \mathbf{S}_{\text{une sphère}}(f,\beta) \sum_i \exp(j2\pi \frac{f}{c}(x_i \cos\beta + y_i \sin\beta)), \tag{3.5}$$

ou x_i et y_i sont les coordonnées des sphères, et $\mathbf{S}_{\text{une sphère}}$ est leur matrice de Sinclair.

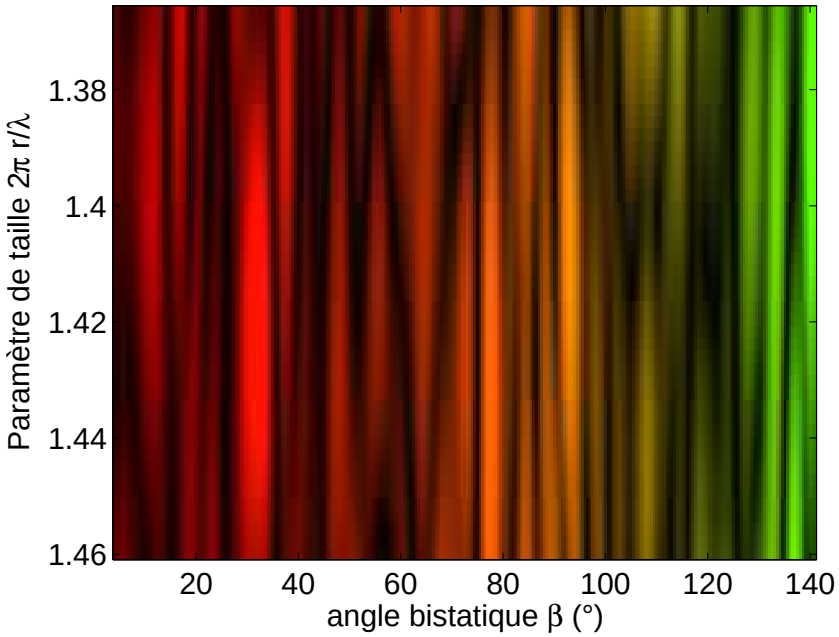


FIGURE 3.3 – Représentation polarimétrique colorée du champ diffusée par un nuage de sphères en fonction de la fréquence et de l'angle de bistatisme β, Résultats de simulation en négligeant les interactions - Rouge : |hH+vV|^2, Vert : |hH-vV|^2, Bleu : |hV+vH|^2

Les champs résultants sont calculés et représentés sur la figure 3.3. Cette figure confirme la présence de franges. Ainsi la différence de marche entre les rayons diffusée par les différentes sphères explique la différence d'intensité pour des sphères situées notamment aux extrémités du cylindre. Ces franges peuvent être considérées comme une figure de speckle. L'angle sous lequel sont vues les sphères varie avec leurs positions ; et la longueur d'onde de l'ordre du rayon du cylindre, fait que la différence de phase entre une sphère située au centre et une situé sur les bords est significative. Nous pouvons du

coup estimer grossièrement la largeur angulaire de ces franges de speckle par la formule λ/D, où λ correspond à la longueur d'onde et D au diamètre du cylindre. Ici nous avons $\lambda/D = 2.9°$ ce qui correspond visuellement à ce que nous pouvons observer sur la largeur angulaire des franges en fonction de l'angle de bistatisme β.

Pour ce qui est des mesures radiométrique et polarimétrique du premier ordre, cette simple simulation semble satisfaisante. Pour ces paramètres, l'influence des multi trajets semble toujours négligeable. Cependant si on observe la polarisation croisée, d'après cette simulation, nous nous attendrions à observer des polarisations croisées Hv et Vh nulles, or nous avons vu que ce n'était pas le cas notamment sur la figure 3.3.

Si nous regardons maintenant les effets du second ordre, c'est à dire le niveau des paramètres dédiés uniquement aux cibles non déterministes que sont l'entropie et la dépolarisation, nous observons des différences entre la simulation et la mesure. Nous rappelons ici que le moyennage statistique nécessaire pour estimer ces paramètres est fait sur la bande de fréquence disponible. La figure 3.4 montre par exemple l'entropie H pour la mesure et la simulation. Le modèle qui ne prend pas en compte les interactions entre les sphères prédit une entropie nulle, alors que l'entropie mesurée évolue autour de 0.5. Nous ne pouvons par contre pas conclure en ce qui concerne l'évolution de cette entropie compte tenu du fort bruit lié au speckle. Nous pouvons constater la même chose si on observe la dépolarisation sur cette figure. Comme pour l'entropie, la dépolarisation est nulle dans la simulation et elle est loin d'être négligeable sur les résultats de mesure. Encore une fois l'évolution de cette dépolarisation reste trop bruitée pour en déduire autre chose que sa non nullité. Les évolutions de l'entropie et de la dépolarisation semblent toutefois très corrélées.

Nous pouvons aussi observer l'évolution des paramètres $\bar{\alpha}$, la retardance ainsi que la diatténuation sur la figure 3.4. Comme nous l'avions attendu, ces paramètres du premier ordre font assez bien correspondre simulation et mesure. Ce qui montre que, bien que le modèle ne prenant pas compte des interactions soit insuffisant pour expliquer les paramètres non déterministes, il convient pour l'ensemble des paramètres communs avec les cibles déterministes.

De cette mesure sur le nuage de sphères nous avons pu déduire les éléments suivants :

- Une modélisation simple, ne tenant pas compte des interactions, permet de prévoir les effets du bistatisme sur les paramètres d'ordre un, tel que l'angle α, la diatténuation ou le retard.
- Les interactions entre les sphères créent de la dépolarisation, cette dépolarisation semble évoluer avec l'angle de bistatisme mais nos mesures ne permettent pas de conclure quant à cette évolution.
- Des modèles plus complexes, prenant en compte les interactions sont nécessaires pour anticiper cette évolution.
- Comme pour les cibles déterministes, l'évolution des paramètres d'ordre 1 avec l'angle de bistatisme est importante.

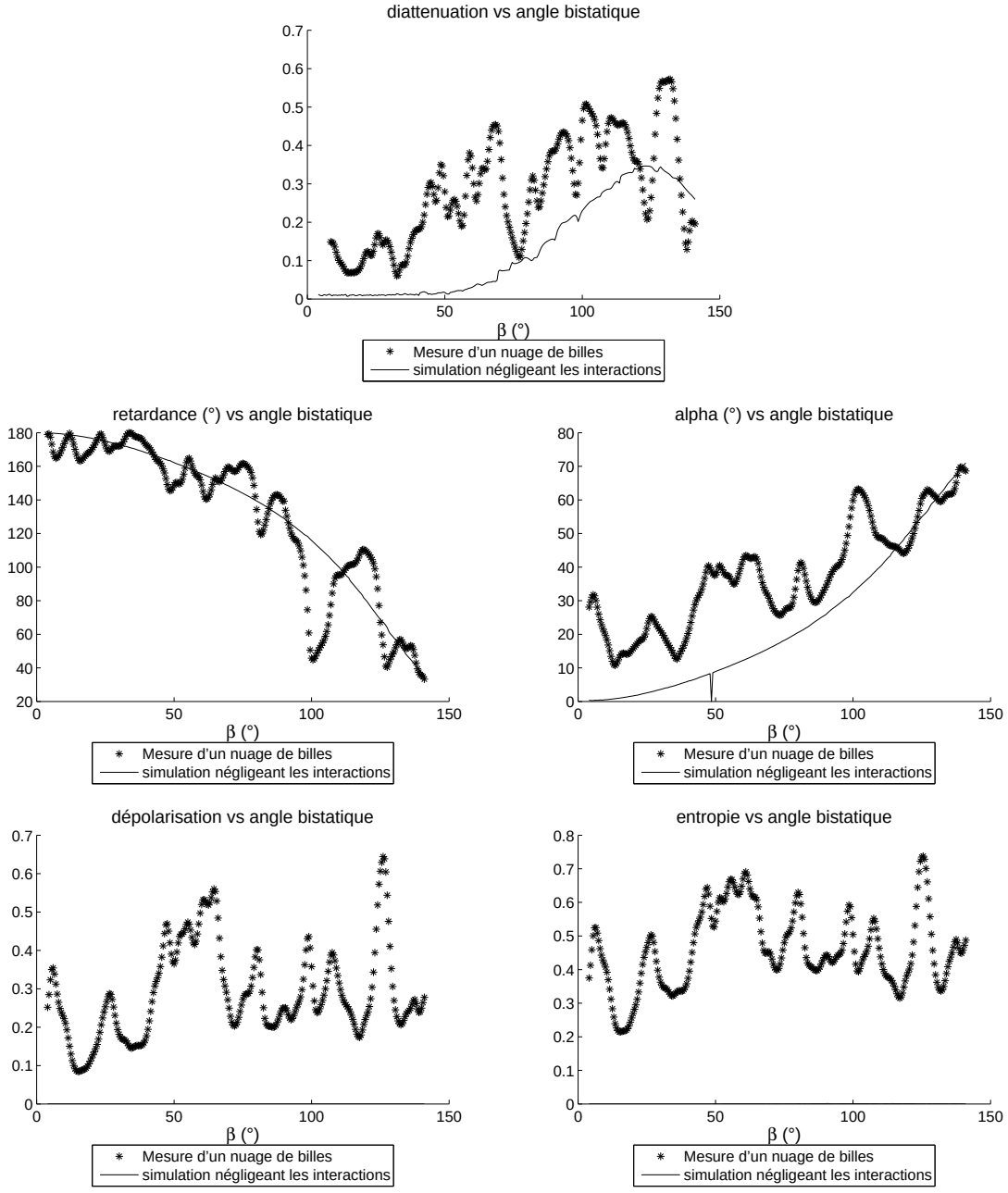


FIGURE 3.4 – Evolution des paramètres polarimétriques classiques en fonction de l'angle de bistatisme pour un nuage de billes dans la zone de résonance

3.2 Analyse et mesure d'un nuage de diffuseurs cylindriques proches de la verticale

3.2.1 Choix et réalisation de la cible

Nous nous intéressons maintenant à une cible ayant à la fois des effets de dépolarisation et des effets d'orientation. Il s'agit d'un nuage de diffuseurs de type cylindrique (donc dichroïque) dont la matrice de diffusion est diagonale réelle lorsque l'axe de symétrie est vertical.

$$S = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \tag{3.6}$$

Nous envisageons alors une distribution de ce diffuseur avec une répartition aléatoire uniforme des positions. En ce qui concerne les orientations de ce diffuseur dans le plan d'onde incident, la distribution envisagée ici est uniforme centrée sur θ_0 et de largeur Δ_θ.

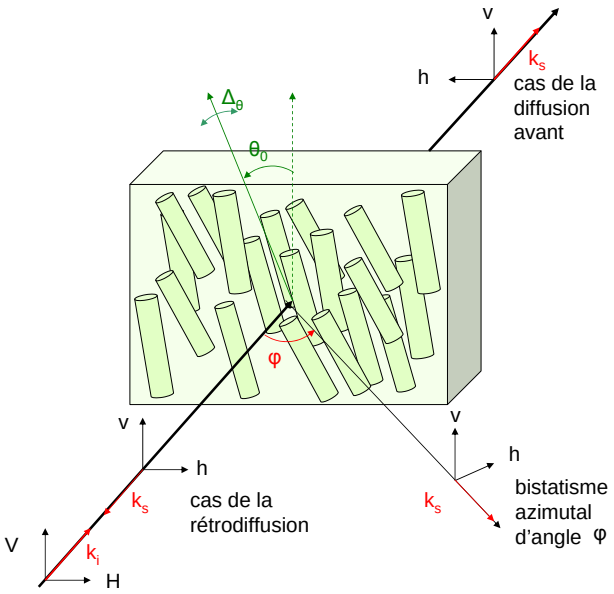


FIGURE 3.5 – Configuration géométrique

Nous essayons alors de tirer les expressions des matrices de cohérence du nuage aléatoire ainsi formé, ainsi que des valeurs propres et des vecteurs propres lorsque c'est possible (Décomposition Cloude et Pottier). Trois cas sont envisagés ; d'abord le cas de la rétrodiffusion, puis celui de la diffusion avant, enfin un cas de bistatisme azimutal d'angle ϕ quelconque. Ces trois cas sont récapitulés sur le schéma de la figure 3.5.

Le cas de la rétrodiffusion correspond à la configuration classique radar monostatique. Nous allons partir de la matrice de Sinclair non orientée, puis en déduire sa matrice de cohérence et enfin appliquer une rotation d'un angle donné à la matrice de cohérence.

Nous partons de l'expression de la matrice de Sinclair :

$$S = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \tag{3.7}$$

Nous en déduisons l'expression de la matrice de cohérence T :

$$T = \begin{pmatrix} (a+b)^2 & a^2-b^2 & 0 & 0 \\ a^2-b^2 & (a-b)^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.8}$$

Considérant toujours un seul cylindre, mais avec une orientation d'un angle θ_0 par rapport à l'axe de la direction de propagation, comme nous l'avons vu précédemment pour le cas de la rétrodiffusion, nous introduisons la matrice $\mathbf{R}$:

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(2\theta_0) & -\sin(2\theta_0) & 0 \\ 0 & \sin(2\theta_0) & \cos(2\theta_0) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.9}$$

Et nous trouvons la matrice de cohérence pour un seul cylindre après rotation :

$$\mathbf{T'} = \mathbf{RTR'} = \begin{pmatrix} (a+b)^2 & (a^2-b^2)\cos(2\theta_0) & (a^2-b^2)\sin(2\theta_0) & 0 \\ (a^2-b^2)\cos(2\theta_0) & (a-b)^2\cos(2\theta_0)^2 & (a-b)^2\sin(2\theta_0)\cos(2\theta_0) & 0 \\ (a^2-b^2)\sin(2\theta_0) & (a-b)^2\sin(2\theta_0)\cos(2\theta_0) & (a-b)^2\sin(2\theta_0)^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.10}$$

Nous allons maintenant simuler notre nuage de cylindres en faisant une moyenne statistique sur un nombre infini de cylindres ayant effectué une rotation avec une densité de probabilité choisie. La moyenne est une somme incohérente car nous négligeons les effets de double rebonds, ainsi nous considérons qu'une onde donnée ne peut interagir qu'avec un seul cylindre du nuage. Nous ne démontrerons pas ici que cette approximation est justifiée, cependant il est facilement compréhensible que l'amplitude relative des rayons diffusés issus des phénomènes de double rebonds sera beaucoup plus faible.

Nous allons donc être amenés à calculer des intégrales du type :

$$\int_{-\pi}^{\pi} P(\theta)\cos(2\theta)d\theta \quad \int_{-\pi}^{\pi} P(\theta)\sin(2\theta)d\theta \quad \int_{-\pi}^{\pi} P(\theta)\cos(2\theta)\sin(2\theta)d\theta \tag{3.11}$$

où $P(\theta)$ est une densité de probabilité quelconque. Dans notre simulation nous avons choisi une densité de probabilité uniforme dans un intervalle centré sur un angle fixé. Ce choix de densité de probabilité permet une résolution littérale, et ne pose pas le problème des valeurs angulaires repliées que l'on peut avoir dans le cas d'une densité gaussienne par exemple. Nous vérifierons plus tard, numériquement, que cette approximation n'entraîne pas une différence trop importante des résultats dans les décompositions matricielles.

Nous posons donc :

$$\mathbf{P}(\theta) = \begin{cases} 1 & \text{si } \theta \in [\theta_0 - \Delta_\theta, \theta_0 + \Delta_\theta] \\ 0 & \text{sinon} \end{cases}$$

C'est ici que nous nous écartons de ce qui avait été fait pour les nuages de cylindres uniformément repartis. En effet dans le cas d'une répartition uniforme les intégrales du type $\cos(\theta_0)\sin(\theta_0)$ ou bien $\sin(\theta_0)$ ou $\cos(\theta_0)$ s'annulent et les intégrales de type $\cos(\theta_0)^2$ et $\sin(\theta_0)^2$ valent $\frac{1}{2}$.

Dans notre cas, nous allons conserver nos deux variables introduites Δ_θ et θ_0 qui définissent la direction privilégiée ainsi que la taille de l'écart autour de cette valeur.

Nous trouvons :

$$\int_{-\pi}^{\pi} P(\theta)\cos(2\theta)d\theta = \cos(2\theta_0)\sin(\Delta_\theta) \tag{3.12}$$

$$\int_{-\pi}^{\pi} P(\theta)\sin(2\theta)d\theta = \sin(2\theta_0)\sin(\Delta_\theta) \tag{3.13}$$

Pour les termes en $\cos(\theta_0)^2$ et $\cos(\theta_0)\sin(\theta_0)$ nous utilisons des formules de trigonométrie qui nous permettent de nous ramener aux deux cas précédents.

Nous trouvons donc une matrice de cohérence moyennée :

$$\mathbf{T}' = \begin{pmatrix} (a+b)^2\Delta_\theta & (a^2-b^2)\cos(2\theta_0)\sin(\Delta_\theta) & (a^2-b^2)\sin(2\theta_0)\sin(\Delta_\theta) & 0 \\ (a^2-b^2)\cos(2\theta_0)\sin(\Delta_\theta) & \frac{(a-b)^2}{2}(\Delta_\theta + \frac{\cos(4\theta_0)\sin(2\Delta_\theta)}{2}) & \frac{(a-b)^2}{4}\sin(4\theta_0)\sin(2\Delta_\theta) & 0 \\ (a^2-b^2)\sin(2\theta_0)\sin(\Delta_\theta) & \frac{(a-b)^2}{4}\sin(4\theta_0)\sin(2\Delta_\theta) & \frac{(a-b)^2}{2}(\Delta_\theta - \frac{\cos(4\theta_0)\sin(2\Delta_\theta)}{2}) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.14}$$

A partir de l'expression de cette matrice dans le cas monostatique, nous pouvons simuler la réponse d'un tel milieu dans une configuration bistatique quelconque. Ceci aboutit à ce que nous appellons par la suite "modèle géométrique simple". Comme pour les sphères, ce modèle permettra de prévoir les tendances qui affecteront les différents paramètres polarimétriques. Comme dans le cas du cylindre seul, il sera confirmé avec l'utilisation de l'approximation du cylindre infini.

3.2.2 Mesures effectuées en chambre anéchoïque

Notre étude sur les nuages de diatténuateurs nous a menés à la réalisation d'une série de mesures en chambre anéchoïque.

Nous avons simulé et réalisé un nuage d'une centaine de cylindres aléatoirement disposés. A l'aide de l'approximation du cylindre infini, nous calculons toutes les matrices de Sinclair relatives à chacun des cylindres auxquels nous avons attribué une orientation statistique. Nous effectuons ensuite une transformation par moyenne pour obtenir les matrices de cohérence pour chaque angle de bistatisme β.

Une fois encore, l'accès à des matériaux diélectriques caractérisés demeurant très coûteux, nous avons réalisé une cible métallique, constituée d'une centaine de clous plantés

dans du polystyrène, avec une inclinaison par rapport à la verticale d'une dizaine de degrés maximum (figure 3.7). Nous nous sommes efforcés de respecter une distribution d'orientation la plus proche d'une distribution quasi verticale ($\theta_0 = 0$, écart angulaire $\Delta_\theta = 10^\circ$). Ainsi nous nous assurons que notre approche simplifiée peut être appliquée. Les clous ont une hauteur de 70 mm, et un diamètre de 3mm.

L'approximation utilisée pour le calcul du cylindre infini ne peut traiter le cas métallique. Pour l'utiliser tout de même, il nous faut choisir une permittivité diélectrique pour nos cylindres ; de préférence avec une grande partie imaginaire. Arbitrairement ici nous fixons une permittivité de 1-30j.

Pour cette étude comme pour le nuage de billes, nous utilisons le rail circulaire de la chambre BABI montré sur la figure 3.6. Une des limites du dispositif de mesure est la situation de diffusion avant. En effet, dans le cas où les antennes se font face, étant donné la taille limitée de notre cible (et de toute les cibles utilisées pour ce type de mesure), une grande partie des ondes passent autour de la cible et donc, par conséquent, arrivent directement dans l'antenne de réception sans avoir été diffusées par la cible. Ainsi la mesure pour des angles de bistatisme proche de $\phi = \pi$ est quasiment inutilisable car noyée dans le bruit du phénomène de "contre jour". Ainsi nous ne présenterons pas les résultats pour ces angles élevés.

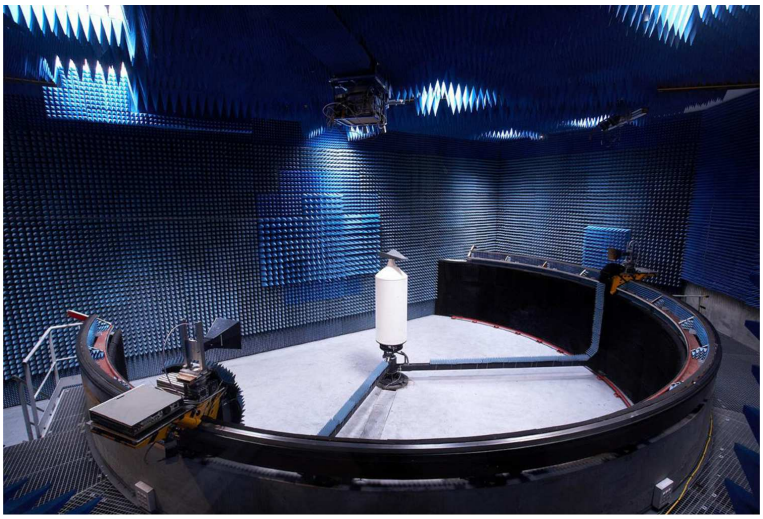


FIGURE 3.6 – La chambre anéchoïque bistatique *BABI* de l'ONERA à Palaiseau où ont été réalisées les mesures

3.2.3 Analyse des résultats obtenus

Les évolutions de paramètres polarimétriques simulés et mesurés pour le nuage de cylindres sont montrés sur la figure 3.8.

Nous y apportons les commentaires suivants :

– Les courbes de retardance font apparaître les mêmes discontinuités que précédemment autour de l'angle de bistatisme de 90 degrés. L'approximation du cylindre

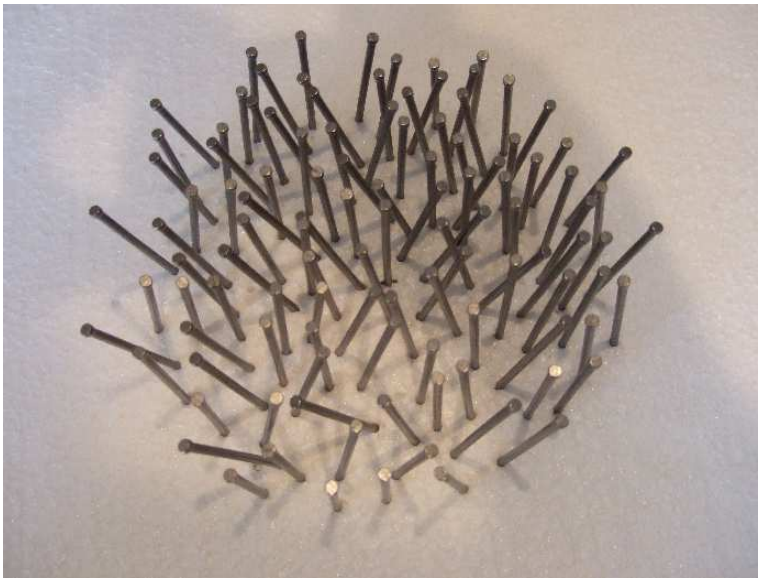


FIGURE 3.7 – Photo de la cible réalisée pour les mesures effectuées en chambre

infini nous a permis de voir que l'amplitude de cette discontinuité dépend de la permittivité du cylindre. Dans le cas diélectrique parfait, les valeurs de retardances seront symétriques par rapport à la valeur de 90˚.

- Dans tous les cas (simulation et mesure), l'angle alpha reste très proche de 45˚ (Valeur classique attribuée au dipôle). Sa variation en fonction de l'angle de bistatisme reste négligeable d'après nos simulations. (Notons que les échelles choisies amplifient les différences entre modèles et simulations. En réalité, les valeurs moyennes trouvées restent très proches)
- Dans tous les cas (simulation et mesure), la diattenuation reste très proche de 1. Là encore, sa variation en fonction de l'angle de bistatisme reste négligeable, devant la valeur déjà très élevée de diatténuation en configuration monostatique.
- On observe, comme pour la sphère, un phénomène de dépolarisation induite autour de l'angle bistatique de 90˚, à la fois pour les mesures et pour la simulation. En dehors de la zone $\beta = 90°$, les valeurs peuvent être assez faibles (autour de 0.2).
- L'entropie simulée est très faible (0.02). L'entropie mesurée est un peu plus élevée (0.2) mais cette valeur reste bien inférieure à celle de la dépolarisation, surtout autour de $\beta = 90°$. Les variations de l'entropie avec l'angle de bistatisme sont négligeables. La valeur moyennement faible de l'entropie peut s'interpréter par le fait que le nuage est très proche d'une orientation déterministe. La simulation sous estime l'entropie. Ceci tendrait à dire que ce sont bien les interactions, et non pas seulement le moyennage fréquentiel, qui crée l'entropie.

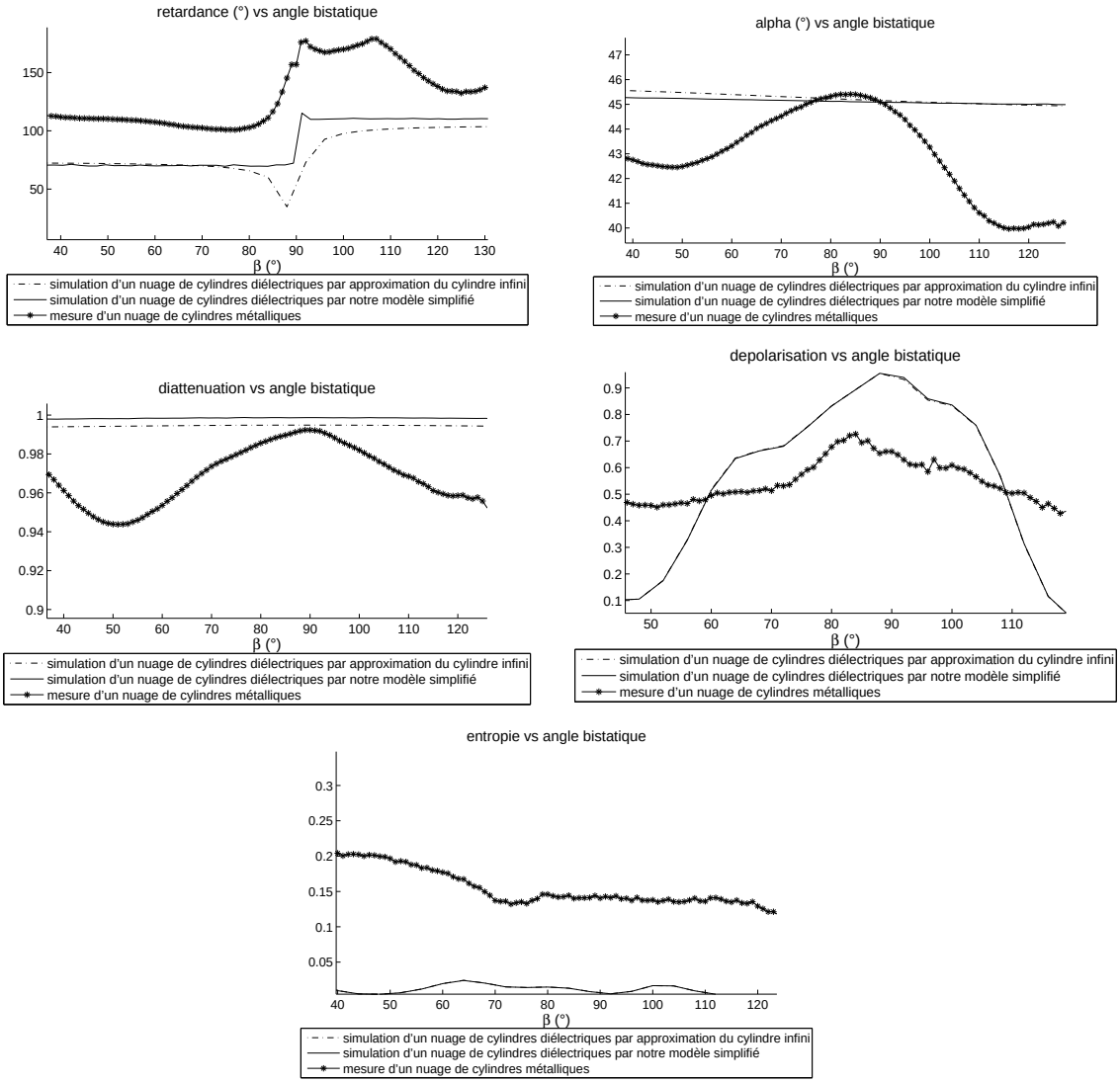


FIGURE 3.8 – Evolution des paramètres polarimétriques classiques en fonction de l'angle de bistatisme pour un nuage de cylindres

Conclusion

Cette partie a été l'occasion de présenter les différentes conventions utilisées pour définir la mesure polarimétrique. Notre principale conclusion porte sur l'importance de mener les analyses polarimétriques en regard de ces choix de convention. Nous préconisons ainsi pour les images radar, de recalculer les matrices de Sinclair mesurées dans la convention du plan de diffusion. Ce calcul nécessite seulement d'estimer deux angles de rotations liés aux repères des antennes. Même si cette nécessité peut sembler évidente dans le monde de l'optique, elle ne l'est pas dans le monde du radar où cette problématique est inexistante pour une configuration monostatique. Or les résultats d'interprétations qui sont alors classiquement déduits (arguments de symétrie, angles d'orientation) dépendent directement de cette convention. Il faut alors prendre également garde à la généralisation de la propriété de "roll invariance" classiquement recherchée en radar. Une fois cette mise en garde effectuée, nous avons pu ensuite tirer des conclusions intéressantes d'un certain nombre de mesures en chambre anéchoïque :

- Dans le plan de diffusion, on note de forts impacts de la géométrie d'acquisition sur la mesure polarimétrique, pour une plage d'angles bistatiques qui s'étend de 90 à 120 degrés. Cet impact se traduit par :
  - une augmentation factice du paramètre de diatténuation pour ces angles, même pour des objets dont la diatténuation devrait être nulle (sphère)
  - un "saut" dans la valeur de retardance dans cette même plage.
  - un impact enfin sur le paramètre de dépolarisation

  Les paramètres issus de la décomposition de Lu et Chipman sont particulièrement sensibles à cette géométrie, peut -être parce que cette décomposition en couches n'est pas la plus adaptée aux cas d'angles bistatiques inférieurs à 90 degrés.
- Au sujet de la nature de la dépolarisation, on note que celle-ci ne peut s'expliquer seulement par la moyenne sur la bande fréquentielle utilisée. Elle pourrait trouver plutôt son origine dans les interactions de groupe. Par ailleurs, ces interactions ne modifient pas de manière sensible le comportement des paramètres déterministes. Cette compréhension du rôle des interactions mériterait d'être étudiée de manière plus approfondie à l'aide de mesures purement optiques, que nous abordons simplement en dernière partie de ce manuscrit.

Deuxième partie

Estimation statistique en radar

Il existe deux grandes stratégies pour exploiter l'information polarimétrique contenue dans une image.

- La première consiste à étudier physiquement l'interaction entre l'onde et les objets observés. Ce processus fait ainsi appel à des étapes de modélisation électromagnétique précises et à une interprétation phénoménologique des comportements polarimétriques des cibles. Ces études aboutissent généralement à des décompositions, projections ou extractions de paramètres qui caractérisent précisément la cible ou le matériau à l'aide de sa réponse polarimétrique. Cette stratégie est naturellement liée à une application donnée, même si il n'est pas rare d'observer des similitudes importantes dans le comportement des cibles d'une application à l'autre. Ainsi ces outils peuvent souvent être exploités dans un cadre qui n'avait pas été envisagé au départ. C'est d'ailleurs le cas lorsqu'on s'intéresse à la fois à des mesures en polarimétrie optique ou en polarimétrie radar où les comportements observés sont régulièrement semblables.
- L'autre stratégie consiste à considérer la mesure polarimétrique comme n'importe quelle mesure d'un signal. Cette information est alors traitée sans chercher à exploiter les raisons physiques qui ont amené à un tel comportement polarimétrique. On recherche uniquement à exhiber une structure ou des éléments particuliers dans l'image en s'appuyant sur le contraste polarimétrique de l'image. Contrairement à la première approche, ces techniques sont universelles, elles peuvent très bien s'appliquer à tout autre type de données multivariées. Elles peuvent donc d'ailleurs aussi bien s'appliquer à la polarimétrie optique que radar, mais aussi à des données non polarimétriques (multispectrales, multicapteurs, temporelles, etc.)

Il faut noter que ces deux approches ne sont pas exclusives et peuvent être judicieusement combinées. Ainsi est-il possible d'extraire un paramètre jugé physiquement pertinent puis d'effectuer une étude statistique sur ce paramètre. Inversement, et ce sera notre approche par la suite, on peut travailler d'abord d'un point de vue statistique sur les données mesurées, avant d'étudier physiquement les données pré-traitées. Ces pré-traitements permettront par exemple de réduire le volume de données en partitionnant l'image en zones homogènes (c'est la segmentation). Ils pourront aussi améliorer les processus d'estimation de paramètres et réduire l'impact du bruit de mesure.

Ce chapitre est dédié aux outils statistiques qui ont pu être développés dans la communauté de la polarimétrie optique et radar. Compte tenu des différences d'approches entre les deux domaines, c'est le monde de la polarimétrie radar qui a développé le plus d'outils dédiés à cette problématique. Nous allons donc d'abord présenter des outils classiques, basés sur le modèle gaussien, utilisé en polarimétrie radar, puis présenter brièvement une alternative, le modèle SIRV, étudiée à l'ONERA et à laquelle j'ai pu contribuer.

Le choix entre un modèle gaussien et non gaussien sera discuté en s'appuyant sur des études sur images réelles haute résolution. Nous proposerons alors une synthèse des différentes distances polarimétriques existantes, gaussiennes et non gaussiennes, ainsi que de nouvelles distances non gaussiennes inédites. Nous proposerons une étude approfondie de leurs performances sur des données réelles et simulées.

Dans un second temps, ces distances polarimétriques nous permettront de construire un processus d'estimation de matrice par croissance de régions contraintes que nous présenterons en détail. Des résultats en interférométrie sur la ville de Toulouse illustreront

une application de cet outil dans la troisième partie.

Enfin, nous aborderons la problématique de la détection polarimétrique dans les milieux gaussiens et non gaussiens. Après un rappel des techniques existantes nous proposerons une nouvelle interprétation géométrique des performances observées. Cette interprétation géométrique nous permettra de proposer une nouvelle famille d'outils de détection ayant des performances très intéressantes dans le cadre de la polarimétrie.

Enfin nos deux algorithmes, la segmentation par croissance de région et la détection hybride, seront combinés dans une application de détection dans des images synthétiques réalisées à l'aide d'un simulateur SAR.

Chapitre 1

Modèles statistiques et distances pour une mesure polarimétrique

1.1 Modélisation du signal polarimétrique : généralités

Pour tout traitement statistique effectué sur une mesure, polarimétrique ou non, il faut avant tout faire une hypothèse sur la nature du signal. Cela se traduit par le choix d'un modèle de distribution statistique qui représente au mieux les observations. En polarimétrie radar, la mesure polarimétrique est représentée par le vecteur de diffusion $\mathbf{k}$. La dimension de ce vecteur peut déjà être sujet à discussion. En effet l'hypothèse de réciprocité amène, dans le cas du monostatisme, la simplification $S_{\mathrm{Hv}} = S_{\mathrm{Vh}}$. Or en polarimétrie complète, les deux grandeurs sont automatiquement mesurées, ce qui nous fournit deux réalisations de bruit différentes.

Classiquement, S_{Hv} et S_{Vh} sont remplacés par la moyenne des deux mesures afin de "réduire" l'impact du bruit. Pour autant, il n'est pas évident qu'il s'agisse de la meilleure méthode.

Nous considérons donc un vecteur de dimension 3 ou 4 à valeurs complexes. Le choix du modèle, dans le cadre d'une mesure radar, va donc être relatif à la distribution suivie par ce vecteur.

Pour décrire une onde parfaitement polarisée, comme expliqué dans la première partie de ce document, seuls les coefficients du premier ordre sont nécessaires. En effet la mesure est alors déterministe, et il suffit d'une seule réalisation pour définir parfaitement le comportement polarimétrique de la cible étudiée.

Il faut cependant distinguer une cible *polarimétriquement déterministe* d'une cible *complètement déterministe*. La nuance vient alors de l'amplitude totale du signal mesuré, qui peut varier dans le premier cas et être constante dans l'autre.

– Le cas de la cible complètement déterministe est décrit par l'équation (1.1) :

$$\mathbf{k_i} = a_0\mathbf{k} + \mathbf{b_i}, \tag{1.1}$$

où a_0 est un terme scalaire d'amplitude déterministe, et $\mathbf{b_i}$ une réalisation de bruit.

– Dans le cas de la cible polarimétriquement déterministe, une variation statistique d'amplitude est considérée, et le vecteur polarimétrique est décrit par l'équation (1.2) :

$$\mathbf{k_i} = a_i \mathbf{k} + \mathbf{b_i} \tag{1.2}$$

où a_i est une réalisation d'amplitude scalaire.

Dans le cadre d'une mesure non déterministe, l'introduction de la matrice de cohérence polarimétrique est également nécessaire. Comme auparavant nous pouvons distinguer deux sous-cas suivant que l'amplitude du signal est elle aussi déterministe ou non.

– Dans le cas ou l'amplitude est constante, la mesure peut encore s'écrire sous la forme (1.3) où $\mathbf{M}$ correspond à la matrice de cohérence, définie dans la première partie comme l'espérance de $\mathbf{k}\mathbf{k}^\dagger$:

$$\mathbf{k_i} = \mathbf{a_0}\sqrt{\mathbf{M}}\mathbf{x_i} + \mathbf{b_i}, \qquad \mathbb{E}(\mathbf{x_i}\mathbf{x_i^\dagger}) = \mathbf{I}. \tag{1.3}$$

– Dans le cas où l'amplitude est elle aussi non déterministe, nous pouvons alors définir la mesure par l'équation (1.4).

$$\mathbf{k_i} = \mathbf{a_i}\sqrt{\mathbf{M}}\mathbf{x_i} + \mathbf{b_i}, \qquad \mathbb{E}(\mathbf{x_i}\mathbf{x_i^\dagger}) = \mathbf{I}. \tag{1.4}$$

L'objectif du traitement de signal polarimétrique en radar au niveau statistique est donc de retrouver à partir des mesures $\mathbf{k_i}$, l'information polarimétrique pertinente $\mathbf{k}$ dans le cas déterministe, et $\mathbf{M}$ dans le cas non déterministe.

Le cas déterministe offre peu de difficultés. En général les cibles ayant des réponses déterministes sont des objets artificiels, avec de très fortes amplitudes rétrodiffusées et donc de bons rapports signal à bruit. C'est la caractérisation des cibles non déterministes qui présente le plus de difficultés et qui suscite donc un nombre conséquent de travaux.

Dans cette section nous nous intéressons aux modèles pouvant être choisis pour décrire le signal polarimétrique. Bien qu'il s'agisse généralement d'un choix motivé par des études statistiques sur des mesures réelles ou des modélisations fines du comportement électromagnétique, le choix reste une hypothèse arbitraire qui ne garantit aucune optimalité dans les traitements déduits de ces hypothèses. Il convient alors de faire attention au sens du mot "optimal". Ce terme est contraint à des hypothèses initiales souvent très restrictives. Il n'existe donc pas de traitement *optimal dans l'absolu*.

Un choix éclairé d'un modèle statistique ne se fait alors véritablement qu'en mettant en oeuvre les différents modèles disponibles sur des échantillons de données. Leurs performances peuvent alors être comparées pour une application donnée.

Dans les sections suivantes, nous présenterons succinctement les distributions employées par leurs densités de probabilité, ainsi qu'éventuellement les estimateurs associés. L'expression de ces densités de probabilité nous permettra de construire, dans la partie suivante, des outils adaptés à ces distributions, comme des distances statistiques, ou des tests de détection.

1.2 Modèle Gaussien

Le modèle le plus fréquemment utilisé pour modéliser le signal polarimétrique est le modèle circulaire gaussien. C'est, de manière générale, le modèle le plus couramment utilisé en traitement du signal pour un spectre d'applications extrêmement variées. L'origine de ce choix provient du théorème central limite (parfois appelé *théorème de la limite centrale*). L'hypothèse faite alors est que le signal mesuré résulte de la somme de signaux diffusés par des éléments indépendants présents dans la même cellule de résolution, ou dans le même secteur angulaire. Si aucun de ces éléments n'a une influence majoritaire et si les réponses polarimétriques de chacun de ces éléments sont indépendantes, alors le signal total mesuré suit une loi de distribution appelée Normale ou Gaussienne.

Le caractère gaussien ou non gaussien de la cible considérée dépend donc très étroitement de la résolution du pixel considéré : suffisamment grande, la cellule de résolution contient un grand nombre de cibles ; trop petite, elle peut contenir une seule cible ponctuelle ; encore plus petite, une cible unique peut elle-même être considérée comme une surface constituée d'un grand nombre d'éléments, et donc à comportement gaussien. C'est ce qu'illustre la figure 1.1.

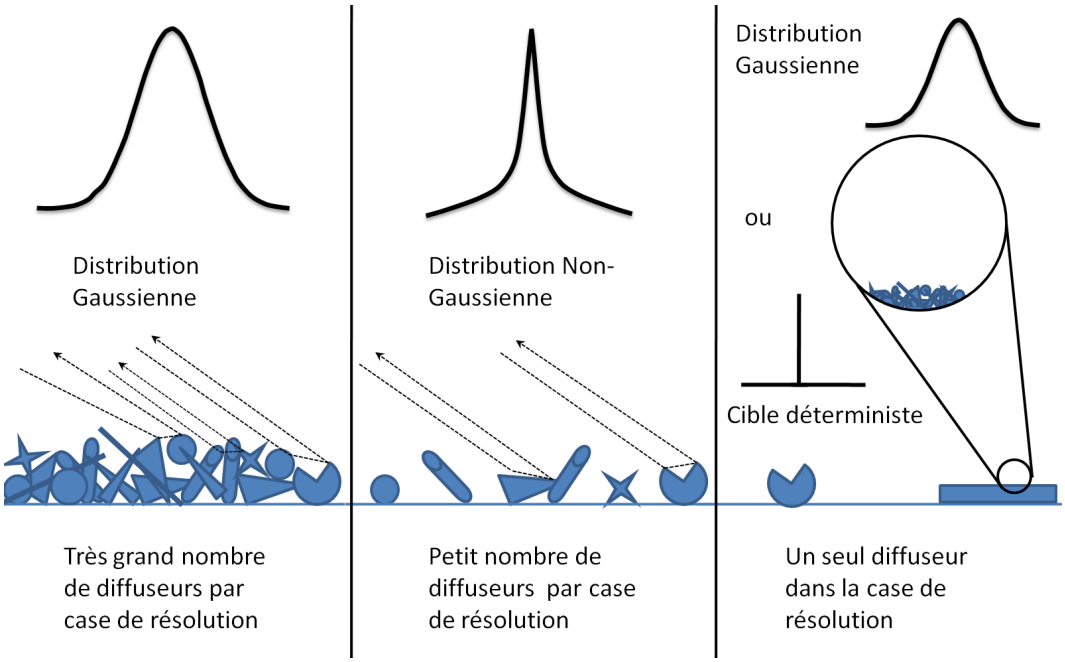


FIGURE 1.1 – Cibles gaussiennes et non gaussiennes : l'influence de la résolution

Le modèle gaussien a donc été très régulièrement utilisé dans des applications radar depuis de nombreuses années [13]. La résolution des systèmes SAR a longtemps été suffisamment grande pour garantir les hypothèses nécessaires à l'application du théorème. Récemment, suite aux progrès importants effectués en terme de résolution, un mouvement s'observe dans la littérature remettant en question la validité de ces hypothèses à haute résolution [14], [15], [16].

1.2.1 Vecteur de diffusion

Dans le cadre d'une distribution gaussienne, on parle alors de la distribution du vecteur de diffusion polarimétrique $\mathbf{k}$, ce qui sous-entend que celui ci n'est pas considéré comme déterministe. Bien entendu cette description inclut le cas particulier des cibles déterministes. Dans ce cadre la distribution du vecteur de diffusion polarimétrique dépend de la matrice de covariance (ou de cohérence) que nous noterons ici $\mathbf{M}$. L'hypothèse de circularité est naturelle, elle suppose que les probabilités de $P(\mathbf{k})$ et $P(e^{j\theta}\mathbf{k})$ sont identiques, et que les phases absolues sont équiprobables. Ainsi la distribution du vecteur $\mathbf{k}$ est décrite par la densité de probabilité suivante d'un vecteur gaussien circulaire de moyenne nulle :

$$P_{\mathbf{M}}(\mathbf{k}) = \frac{1}{\pi^m |\mathbf{M}|} \exp\left(-\mathbf{k}^\dagger \mathbf{M}^{-1} \mathbf{k}\right), \tag{1.5}$$

où m correspond au nombre de dimensions (3 ou 4), $|\mathbf{M}|$ au déterminant de la matrice $\mathbf{M}$ et $\dagger$ à l'opérateur hermitien (transposée conjugué).

1.2.2 Matrice de covariance ou cohérence

Il est possible par un changement de variable de montrer que l'espérance $\mathbb{E}(\mathbf{k}\mathbf{k}^\dagger) = \mathbf{M}$. L'estimateur au sens du maximum de vraisemblance sous hypothèse gaussienne de $\mathbf{M}$ [17] est souvent appelé SCM pour Sample Covariance Matrix et est donné pour les échantillons $\mathbf{k_i}$ par :

$$\hat{\mathbf{M}} = \frac{\mathbf{1}}{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{N}} \mathbf{k_i} \mathbf{k_i}^\dagger. \tag{1.6}$$

La distribution de la matrice estimée $A = \hat{\mathbf{M}}$ à partir de N échantillons $\mathbf{k_i}$, suit quant à elle une distribution de Wishart (déduite de la distribution gaussienne du vecteur de diffusion) :

$$P_{\mathbf{M}}(\hat{\mathbf{A}}) = \frac{|\mathbf{A}|^{N-m} \exp\left(-\mathrm{tr}(\mathbf{M}^{-1}\mathbf{A})\right)}{K(N,m)|\mathbf{M}|^N}, \tag{1.7}$$

où $K(N,m) = \pi^{\frac{1}{2}m(m-1)}\Gamma(N), ..., \Gamma(N-m+1)$, et Γ est la fonction Gamma.

1.3 Modèle Non Gaussien

Dans le contexte des études théoriques sur les distributions statistiques, les modèles non gaussiens ont été étudiés depuis un grand nombre d'années [18]. Leur application en imagerie SAR est par contre sensiblement plus récente [19], [20], [14]. Elle a été stimulée, comme nous l'évoquions plus haut, par l'augmentation des capacités des capteurs à produire des images à très haute résolution. Cette haute résolution remet en cause les hypothèses du théorème central limite, et donc le choix du modèle gaussien. L'existence de comportements non gaussiens est effectivement avérée par certaines campagnes de mesures [21], [22].

Ceci dit, l'incapacité des modèles gaussiens à traiter les données actuelles n'a pas été démontrée pour autant. Beaucoup de travaux actuels adaptent des méthodes existantes à des distributions statistiques de plus en plus complexes, sans avoir encore démontré de manière convaincante un besoin lié à une application donnée.

Avant de décrire plus en détail certains de ces modèles, nous insistons sur les points suivants :

1. Les études statistiques en milieu spatialement hétérogène manquent souvent de rigueur : elles se font à l'aide de fenêtres glissantes, au risque donc de mélanger des populations statistiques appartenant à des types de cibles différents. Il faut différencier des images gaussiennes "par morceaux" dont la distribution globale pourrait donner l'illusion d'un fouillis non gaussien, des zones homogènes ayant des distributions non gaussiennes.
2. Ce n'est pas parce qu'une distribution "complexe" inclut la distribution gaussienne qu'elle permet de traiter efficacement des données gaussiennes. L'estimation des paramètres de lois est nécessairement imparfaite et réduira donc toujours les performances par rapport à un modèle plus adapté. Plus la distribution nécessite l'estimation d'un nombre important de paramètres, plus cette distribution aura des performances dégradées sur des données gaussiennes.
3. Ce n'est pas parce qu'une distribution est effectivement non gaussienne que le traitement adapté aboutira à de meilleures performances. Pour la même raison, l'estimation d'un nombre important de paramètres, qui d'ailleurs peuvent s'avérer être peu contrastés voire constants sur toute une image, peut entraîner des performances moindres que l'emploi d'une distribution plus simple bien que non rigoureusement adaptée.

Ces points seront démontrés et illustrés au cours de la section suivante ainsi qu'au moment de l'application de ces modèles à des données simulées ou réelles.

1.3.1 Modèle Multiplicatif

Comme nous l'avons présenté plus haut, il faut distinguer les variations de l'information polarimétrique pure, des variations d'amplitude. La première correspond à des variations relatives entre les différents canaux polarimétriques, que ce soit en amplitude et/ou en phase. La seconde correspond à une variation relative, de pixel à pixel, de l'amplitude totale des signaux mesurés dans les différents canaux polarimétriques.

Dans l'absolu il n'y a aucune raison pour que ces deux variations soient décorrélées. Au contraire il est fréquent de constater qu'en fonction des mécanismes mis en oeuvre dans la rétrodiffusion, certains phénomènes, comme le double rebond, entraînent à la fois une signature polarimétrique très particulière, mais aussi une augmentation significative de l'énergie rétrodiffusée.

Cependant l'hypothèse d'indépendance entre la mesure polarimétrique et la mesure d'amplitude est souvent utilisée pour les milieux ayant un comportement non déterministe. D'abord parce que seule cette approche permet de développer des techniques et des outils statistiques suffisamment simples. Ensuite parce qu'en pratique, cette hypothèse s'est avérée donner des résultats satisfaisants dans un grand nombre d'applications.

La première approche pour traiter du fouillis non gaussien consiste alors à utiliser une des très nombreuses lois pouvant décrire des données complexes et centrées sur zéro. Une approche assez communément utilisée en polarimétrie suppose plutôt que l'information polarimétrique, contenue dans la matrice de covariance ou de cohérence, constitue le noyau de la distribution. Seule l'énergie totale suit une loi positive, dont la complexité dépend du caractère impulsionnel du fouillis, et qui est décorrélée de la distribution de la matrice polarimétrique.

Dans ce modèle, on suppose donc que la mesure polarimétrique du vecteur de diffusion $\mathbf{k}$ se décompose de la façon suivante :

$$\mathbf{k} = \sqrt{\tau}\ \ \mathbf{x}. \tag{1.8}$$

On suppose donc que le vecteur $\mathbf{x}$ et la variable scalaire positive τ suivent des distributions indépendantes. Le vecteur $\mathbf{x}$ contient l'information polarimétrique et vérifie $\mathbb{E}(\mathbf{x}\mathbf{x}^\dagger) = \mathbf{M}$.

Ici, il n'est pas nécessaire que l'information polarimétrique soit modélisée par une loi gaussienne. Nous allons d'abord présenter quelques modèles à lois connues, qui imposent donc un choix à la fois pour la distribution d'amplitude ou de texture et pour la distribution du vecteur polarimétrique. Dans un second temps nous présenterons un sous produit du modèle multiplicatif qui suppose la distribution du vecteur polarimétrique gaussien et circulaire : le modèle SIRV.

1.3.2 Modèle à lois connues

Loi K

Parmi les modèles multiplicatifs, la loi la plus fréquemment utilisée dans les applications radar pour décrire des données non gaussiennes, et probablement la moins complexe, est la loi K [19]. Il s'agit de la combinaison par produit d'un noyau gaussien et d'une amplitude qui suit une loi Gamma [23] :

$$P_{\mathbf{M},b,\nu}(\mathbf{k}) = \frac{b^{2m}}{\Gamma(\nu)\sqrt{\pi}(2\pi)^m.|\mathbf{M}|}\frac{(b\sqrt{\mathbf{k}^\dagger\mathbf{M^{-1}}\mathbf{k}})^{\nu-m}}{2^{\nu-1}}K_{m-\nu}(b\sqrt{\mathbf{k}^\dagger\mathbf{M^{-1}}\mathbf{k}}) \tag{1.9}$$

où ν est un paramètre de forme de la distribution et b se déduit du paramètre de normalisation de la matrice de covariance avec $\mathrm{tr}(\mathbf{M}) = m$ et $\mathbf{M} = \frac{b^2}{2\nu}\mathbb{E}(\mathbf{k}\mathbf{k}^\dagger)$.

Pour un traitement de détection ou de classification adapté à cette distribution, l'estimation du paramètre de forme ν est nécessaire. Il existe plusieurs méthodes pour cette estimation. Nous donnons ici la méthode du maximum de vraisemblance qui utilise l'équation suivante en passant au logarithme :

$$\ln\nu - \psi^{(0)}(\nu) = \ln\Big(\frac{1}{N}\sum_{i=1}^{N}\mathbf{k_i}^\dagger\mathbf{M^{-1}}\mathbf{k_i}\Big) - \frac{1}{N}\sum_{i=1}^{N}\ln\Big(\mathbf{k_i}^\dagger\mathbf{M^{-1}}\mathbf{k_i}\Big) + \psi^{(0)}(m) - \ln m. \tag{1.10}$$

Le paramètre de forme ν est donc la variable qui permet de modéliser la tendance impulsionnelle de la distribution. Comme nous pouvons le voir sur la figure 1.2, plus ce

paramètre est petit, plus la distribution s'éloigne de la distribution gaussienne. En pratique il est assez courant de rencontrer des fouillis K-distribués [19]. Cependant les paramètres de formes rencontrés sont rarement très faibles (ν entre 3 et 8). Nous montrerons cela sur une étude rapide des fouillis rencontrés sur des données réelles dans une section à venir.

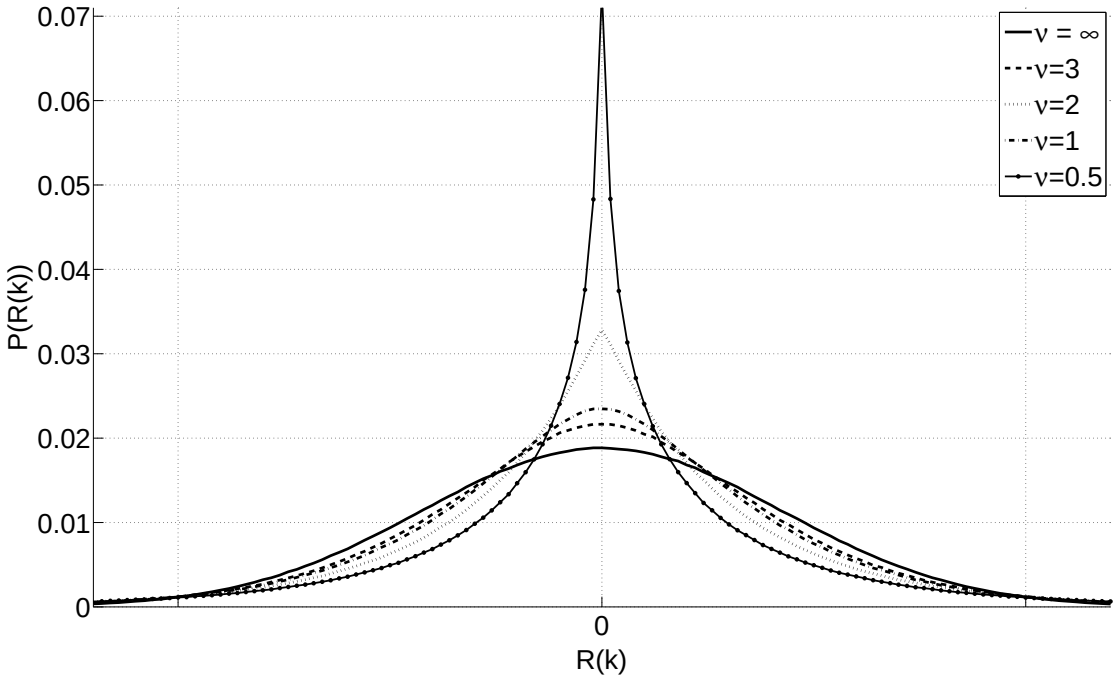


FIGURE 1.2 – Evolution de la densité de probabilité de la partie réelle du vecteur polarimétrique pour différentes valeurs du paramètre de forme ν. Pour $\nu = \infty$ la distribution est gaussienne.

Loi de Fisher (KummerU), et loi Pearson

Un niveau supplémentaire de complexité dans le choix des distributions statistiques est atteint avec la loi de Fisher lorsqu'elle est utilisée pour décrire la variation d'énergie. Cette loi se déduit du quotient de deux distributions du χ^2. Sa combinaison avec un noyau gaussien pour l'information polarimétrique aboutit à une distribution appelée KummerU [20]. A travers mon expérience sur les données actuelles, cette distribution, bien que plus précise, se heurte à de nombreux problèmes d'estimation de paramètres, nécessitant à la fois un très grand nombre de données, et entraînant une perte énorme en résolution et en temps de calcul. Nous nous contenterons ici de donner des références auxquelles le lecteur pourra se reporter pour plus de détails concernant les densités de probabilité et les outils associés à cette distribution. L'étape suivante dans la complexité, qui n'a pas encore véritablement été employée en polarimétrie, consiste à étudier les lois de Pearson d'ordres supérieurs [24]. A mon avis de telles distributions n'auront pas, pour l'instant, de cadre applicatif probant, mais seront probablement étudiées dans les années à venir.

1.3.3 Modèle à lois inconnues : les SIRV

Une approche différente a été considérée et appliquée depuis un certain nombre d'années, notamment à l'ONERA, pour traiter la problématique des données non gaussiennes [25], [26]. Le fondement du modèle est très similaire à celui du modèle multiplicatif. Encore une fois on suppose les variations d'amplitude décorrélées des variations statistiques polarimétriques. Mais ici, le noyau polarimétrique est gaussien. Cette hypothèse provient du théorème de représentation de Yao [18], qui montre que tout processus sphériquement invariant peut s'écrire comme un produit de ce noyau gaussien et d'une variable positive, appelée texture.

On décompose la mesure du vecteur de diffusion $\mathbf{k}$ en un produit d'une texture τ et un vecteur qui cette fois suit une distribution gaussienne circulaire de moyenne nulle $\mathbf{x}$:

$$\mathbf{k} = \sqrt{\tau}\ \ \mathbf{x}.\ \ \&\ \ \mathbf{x} \sim \mathbb{W}(\mathbf{0}, \mathbf{M}) \tag{1.11}$$

Le symbole $\mathbb{W}$ désigne la loi de Wishart. Ce modèle a donc pris le nom de "SIRV" pour *Spherically Invariant Random Vector*. A partir de ce modèle il est possible, comme pour le modèle multiplicatif, de choisir une distribution connue pour la variable de texture. Ces distributions aboutissent à des outils et estimateurs adaptés mais aux mêmes problèmes rencontrés qu'avec les distributions complexes évoquées plus haut.

Dans son utilisation la plus intéressante, le modèle SIRV ne fait pas d'hypothèse sur la distribution du paramètre de texture τ souvent traduit par sa densité de probabilité inconnue $p(\tau)$. La distribution du vecteur de diffusion $\mathbf{k}$ s'écrit alors de manière générale :

$$P_{\mathbf{M},p(\tau)}(\mathbf{k}) = \int_0^{+\infty} \frac{1}{\pi^m \tau^m |\mathbf{M}|} \exp\left(-\frac{\mathbf{k}^\dagger \mathbf{M}^{-1} \mathbf{k}}{\tau}\right) p(\tau) d\tau. \tag{1.12}$$

Comme cette densité de probabilité $p(\tau)$ est inconnue, une approche ne faisant pas d'hypothèse sur cette distribution consiste à supposer que chaque réalisation de la mesure produit une réalisation τ différente et déterministe. Ainsi la densité de probabilité du vecteur de diffusion, qui dépend de ce paramètre inconnu τ, devient :

$$P_{\mathbf{M},\tau}(\mathbf{k}) = \frac{1}{\pi^m \tau^m |\mathbf{M}|} \exp\left(-\frac{\mathbf{k}^\dagger \mathbf{M}^{-1} \mathbf{k}}{\tau}\right). \tag{1.13}$$

Avec cette hypothèse il est possible de calculer un estimateur au sens du maximum de vraisemblance de la texture supposée déterministe et indépendante de pixel à pixel :

$$\hat{\tau}_i = \frac{\mathbf{k_i}^\dagger \mathbf{M^{-1}} \mathbf{k_i}}{m}, \tag{1.14}$$

et également un estimateur de la matrice de cohérence/covariance à un facteur près :

$$\hat{\mathbf{M}}_{\mathrm{FP}} = \frac{m}{N} \sum_{i=1}^{N} \frac{\mathbf{k_i} \mathbf{k_i}^\dagger}{\mathbf{k_i}^\dagger \hat{\mathbf{M}}_{\mathbf{FP}}^{\mathbf{-1}} \mathbf{k_i}}. \tag{1.15}$$

Cet estimateur de la matrice de covariance est souvent appelé par abus de langage *estimateur du point fixe*, car c'est en pratique le processus itératif qui est mis en oeuvre

lors de l'estimation. En effet il a été montré par Pascal dans [26] que cette équation implicite peut être calculée par la méthode du point fixe, de manière itérative quelle que soit l'initialisation.

Ce processus itératif peut être assez coûteux en temps de calcul, puisqu'il implique une inversion matricielle ou une décomposition de Cholesky à chaque étape. Afin de réduire le temps de calcul, il est possible de choisir intelligemment l'initialisation du processus itératif pour accélérer sa convergence. En général c'est l'identité $\mathbf{I_3}$ qui est utilisée pour initialiser le processus. Dans le cadre d'une image, les matrices de covariances/cohérences sont calculées par fenêtres glissantes. L'idée est d'utiliser la matrice résultante de l'estimation du pixel précédent pour initialiser le processus pour le pixel voisin. Ainsi, compte tenu du grand nombre d'échantillons communs dans les deux populations utilisées pour l'estimation, la convergence de l'algorithme est atteinte pour un nombre très réduit d'itérations. Sur une image SAR urbaine par exemple, on peut voir le nombre moyen d'itérations par pixel passer de 7 à 2.

En pratique ce modèle est intéressant en présence de distributions non gaussiennes dont on ne peut prévoir avec précision la distribution a priori. Il faut remarquer que l'estimation de cette matrice est indépendante de l'énergie diffusée et mesurée pour chacun des échantillons. Nous verrons plus tard, au moment de l'étude des performances, que le fait d'ignorer cette information d'énergie n'est profitable que si cette information est véritablement source de perturbations. Bien que le modèle SIRV inclut le modèle gaussien, comme tout modèle non gaussien nous verrons qu'il s'avère bien moins performant que celui-ci dans les cas gaussiens, mais aussi dans les cas faiblement non gaussiens.

1.3.4 Caractérisation des distributions non gaussiennes : Méthode des Log cumulants

Principe

Quand on étudie une image, ou un échantillon d'images, on est souvent amené à s'interroger sur la nature de la distribution rencontrée. Nous venons de voir qu'il existe un grand nombre de lois pouvant, plus ou moins rigoureusement, décrire les comportements statistiques. A l'usage, l'habitude nous permet d'anticiper avec une précision correcte le type de distribution qu'il est souhaitable d'utiliser. Il est toutefois nécessaire, surtout quand on décide d'employer des distributions très complexes, de justifier la nature du signal mesuré.

Une méthode simple consiste à étudier les paramètres de loi "*comme si*" la distribution mesurée suivait la distribution envisagée. Ainsi on peut, si on s'attend à rencontrer un fouillis K-distribué, estimer le paramètre ν de la distribution et, en fonction de la valeur de celui ci, déterminer si les données sont gaussiennes (ν très grand) ou impulsionnelles (ν petit). Il convient ensuite d'utiliser des tests de comparaison (divergence, distance) entre la distribution mesurée, et une distribution tirée à partir des paramètres estimés. Cette méthode n'est toutefois pas idéale et objectivement pas toujours démonstrative.

Un outil introduit par Jean-Marie Nicolas dans [27] permet de caractériser des distributions non gaussiennes en étudiant les Logs-moments et Logs-cumulants de la distribution. L'avantage de cette méthode est qu'elle fournit en plus une représentation graphique.

Celle-ci permet de visualiser dans quelles zones se répartissent les distributions rencontrées dans une image, dans un plan contenant les distributions classiques. Les moments et cumulants d'ordres supérieurs sont caractéristiques des distributions non gaussiennes, il est donc naturel de s'y intéresser afin de les caractériser. L'originalité des travaux concernant les logs moments vient de la propriété des logarithmes à transformer les produits en sommes. Ce qui nous permet, dans le cadre d'un modèle multiplicatif de séparer la contribution de la loi de texture, du noyau gaussien.

Pour toutes les applications où l'on rencontre des produits de distributions, les moments et cumulants dits de deuxième espèce s'avèrent particulièrement élégants : ils allient performance et simplicité. Ce sont des outils qui sont donc régulièrement employés en imagerie SAR afin de caractériser la nature des images rencontrées. Après les travaux de Jean Marie Nicolas, nous notons donc les travaux d'Afinsen qui étend la technique aux *log-det-cumulants* en supposant un modèle produit entre la texture et la matrice de cohérence/covariance polarimétrique [28].

Cette représentation graphique basée sur les *log-cumulants* et *log-det-cumulants* est souvent appelée plan κ_2,κ_3. Elle consiste en effet à représenter les moments de groupe d'échantillons d'une image SAR en fonction de leurs log-cumulants d'ordre 2 et 3.

Ainsi par définition,

$$\tilde{m_1} = \mathbb{E}(\log(x)) \quad \tilde{m_2} = \mathbb{E}(\log(x)^2) \quad \tilde{m_3} = \mathbb{E}(\log(x)^3)$$

$$\begin{aligned} \tilde{\kappa_1} &= \tilde{m_1} \\ \tilde{\kappa_2} &= \tilde{m_2} - \tilde{m_1}^2 \\ \tilde{\kappa_3} &= \tilde{m_3} - 3\tilde{m_1}\tilde{m_2} + 2\tilde{m_1}^3. \end{aligned} \tag{1.16}$$

Dans cette représentation il est ainsi possible de représenter un grand nombre de lois traditionnellement utilisées pour représenter des lois non gaussiennes. Ainsi la figure 1.3 représente dans le plan κ_2,κ_3 les différentes courbes qui correspondent aux positions théoriques des log- cumulants de ces lois ainsi que des moments estimés à l'aide de données simulées à partir de ces mêmes lois.

Problème et limitations

Théoriquement cette méthode et l'utilisation de cette représentation sont effectivement des outils pertinents parce qu'ils permettent de visualiser sur un même plan l'ensemble de ces distributions. Des problèmes surviennent cependant lorsque des conclusions trop optimistes sont déduites de l'utilisation de ce plan. Ou bien lorsque le calcul pratique de ces moments sur des données réelles n'est pas effectué de manière adéquate, et que les interprétations qui en sont faites sont erronées.

1. *Estimation des moments.* Le premier problème qui survient d'après moi, vient de l'estimation des paramètres κ_2 et κ_3.

   Comme pour l'estimation d'une matrice de covariance, ou d'un paramètre d'une loi de distribution, cette estimation nécessite un certain nombre d'échantillons ou d'ob-

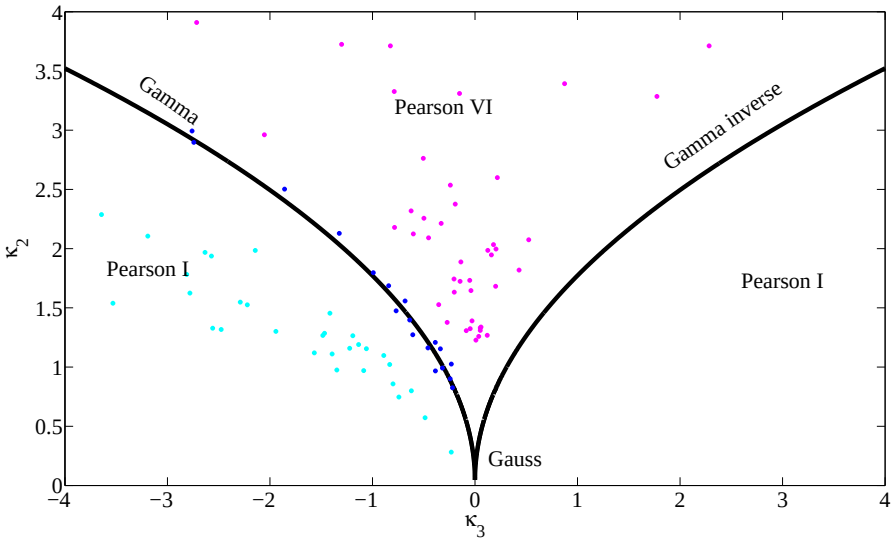


FIGURE 1.3 – Exemple de représentation dans le plan κ_2,κ_3.

servations qui correspondent idéalement au même comportement physique. Classiquement on utilise un moyennage spatial avec une fenêtre glissante sur l'image. Nous allons voir que cette méthode est à proscrire.

L'image est naturellement structurée, et ce de manière encore plus importante dans les images urbaines (qui d'ailleurs sont le plus souvent celles qui motivent l'utilisation de modèles non gaussiens). Cette structure dans l'image fait que, même si toutes les distributions rencontrées sont "gaussiennes" par morceaux ou régions, une fenêtre glissante rencontrera inévitablement un nombre conséquent (voir majoritaire) de zones de transitions. Ces zones de transitions présenteront alors des distributions bi-modales qui donneront alors l'illusion de l'existence de distributions fortement non gaussiennes.

La justification du choix d'une distribution ou de modèle ne peut pas être faite à l'aide d'une estimation par fenêtre glissante. Seule la sélection de zones "homogènes" automatiquement, ou manuellement, avant d'étudier leurs log moments, permet de justifier convenablement le choix de tel ou tel modèle de distribution. Il est préférable à mon avis de ne sélectionner qu'un nombre restreint de zones que l'on va juger caractéristiques de l'image. Nous représentons les effets d'une estimation par fenêtre contre une estimation par région sur la figure 1.5. Sur cette figure, la couleur est reliée à la densité des paramètres κ_2 et κ_3 estimée à l'aide d'une fenêtre glissante sur l'ensemble de l'image présentée sur la figure 1.4. Les points noirs correspondent à l'estimation pour chacune des régions délimitées sur la figure 1.4. L'estimation des paramètres par fenêtre glissante donne l'illusion d'une distribution bien plus impulsionnelle que celle rencontrée réellement. Le centre de la distribution est surévalué par rapport aux distributions réelles rencontrées par région.

2. *Aucun véritable point de repère.* L'autre problème de cette représentation est l'ab-

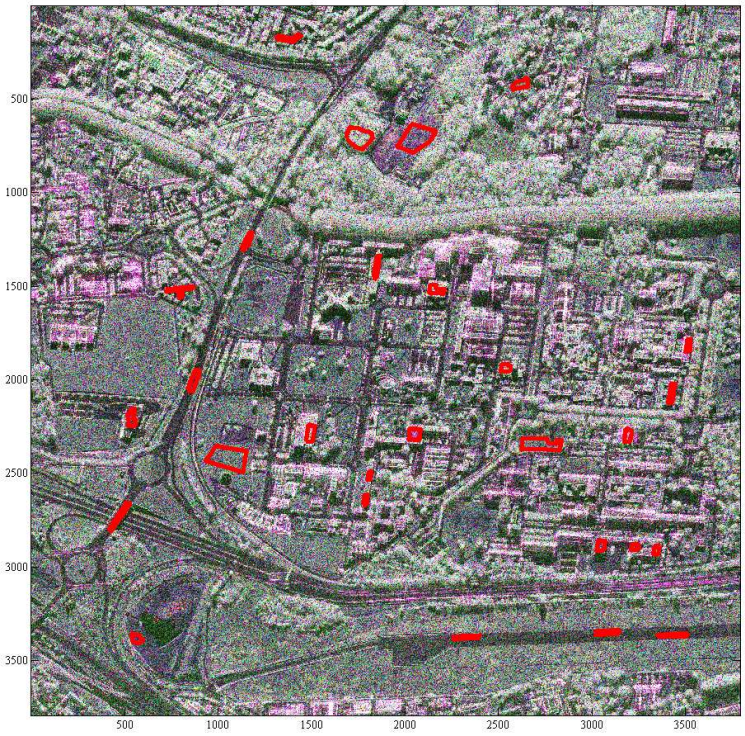


FIGURE 1.4 – Exemple de sélection de zones homogènes, sans frontières, dans une image SAR polarimétrique de la ville de Toulouse (image RAMSES acquise pour le CNES)

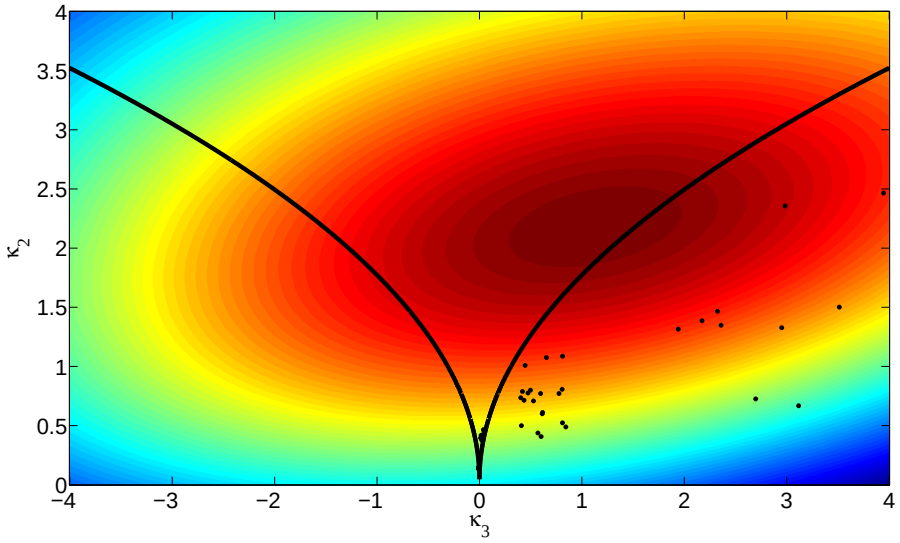


FIGURE 1.5 – Illustration des estimés de paramètres κ_2,κ_3 sur une image haute résolution de la ville de Toulouse (image RAMSES acquise pour le CNES).

sence de point de repère permettant de véritablement choisir entre différents modèles. Les courbes théoriques permettent de délimiter les zones correspondant à certaines distributions (Gamma, Beta, Fisher...). Mais encore une fois, même si les moments ont été parfaitement estimés, la position à un endroit précis du plan ne permet pas à elle seule de préjuger de la nécessité d'utiliser un modèle de distribution particulier. En effet l'intégralité du plan et des courbes de référence forment une figure invariante par changement d'échelle. Ainsi un point situé dans la zone des distributions de Fisher, ne garantit pas que l'utilisation d'outils déduits des distributions de Fisher apporte quoi que ce soit. Ce point peut être extrêmement proche d'une distribution gaussienne ! Cela ne dépendra véritablement que de l'échelle des axes de κ_2 et κ_3. Cette représentation ne permet pas de trancher sur le choix de la distribution tant qu'il n'existe pas de seuil qui nous permette rigoureusement de choisir une distribution.

1.3.5 Exemple sur des images SAR à haute résolution

Brétigny sur Orge

Afin d'obtenir un ordre de grandeur des valeurs de κ_2,κ_3 rencontrées sur des images SAR haute résolution, nous étudions brièvement deux scènes. Nous effectuons une estimation propre de ces paramètres, en évitant donc l'estimation par fenêtre glissante (qui biaise naturellement l'estimation en faveur de la distribution de Fisher). La première scène est une image de la zone de Brétigny sur Orge en région parisienne. L'image a été acquise en bande-X avec une résolution métrique. Nous sélectionnons ici majoritairement des zones naturelles, des champs, des pistes ou des routes. La scène suivante sera plus adaptée à l'estimation des paramètres sur des zones urbaines. Nous pouvons constater que les paramètres se situent majoritairement dans les zones associées à la loi de Pearson I (figure 1.6). Toutefois la valeur de ces paramètres ne nous indique pas si le choix de cette distribution est judicieux ou permettrait vraiment d'apporter de meilleures performances qu'un autre modèle. Il est intéressant aussi de remarquer qu'il semble envisageable d'exploiter les paramètres de texture afin de participer à une classification des différentes zones. Sans toutefois pouvoir quantifier cet apport par rapport à l'information polarimétrique, ou énergétique.

Toulouse

La seconde scène est une image urbaine de Toulouse contenant aussi un certain nombre de parcs, ou de sols nus. Elle a été acquise en 2004 bande X avec une résolution centimétrique, par le capteur RAMSES de l'ONERA, dans le cadre d'une étude pour le CNES. C'est une image acquise en mode interférométrique monopasse. Nous sélectionnons ici à la fois des zones urbaines et des bâtiments, ainsi que quelques zones naturelles, ou non aménagées. Encore une fois les paramètres sont majoritairement situés dans les zones de la loi de Pearson I. (figure 1.7). Rien ne permet de justifier a priori le choix d'une telle distribution. Par contre, contrairement à l'image de Brétigny, les paramètres de texture ne semblent pas permettre de discriminer les différentes zones étudiées. Dans ce cas il semble

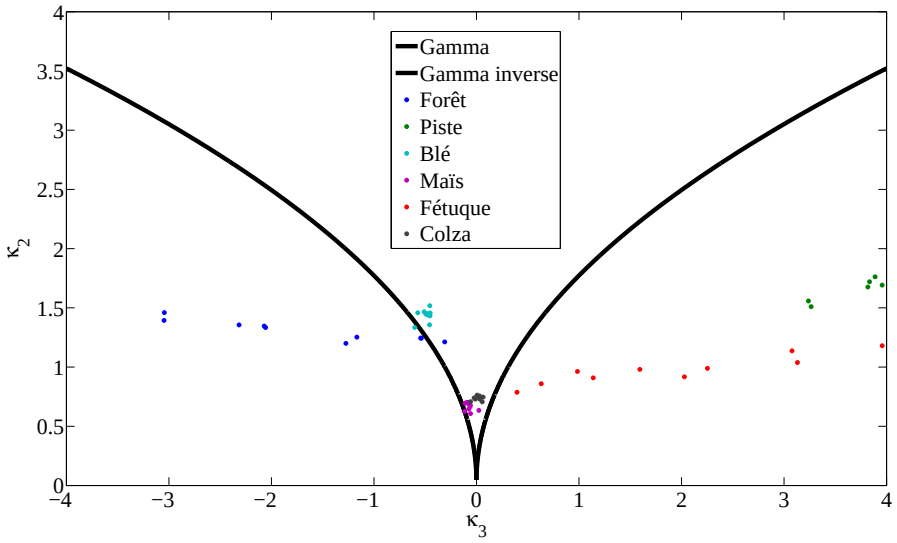


FIGURE 1.6 – Illustration des estimés de paramètres κ_2,κ_3 sur une image haute résolution de Brétigny.

que la texture joue principalement un rôle perturbateur dans l'estimation des paramètres polarimétriques.

Proposition de segmentation du plan κ_2 / κ_3.

L'idée principale proposée ici est un partitionnement de l'espace κ_2 / κ_3, qui permettra de manière objective, de délimiter des zones dans lesquelles il est souhaitable d'utiliser des modèles non gaussiens. Nous illustrons cette idée d'une manière assez simple en comparant uniquement le modèle gaussien et le modèle SIRV. Évidemment l'idée peut être reprise et développée pour tester l'intégralité des modèles rencontrés dans ce plan (Gamma, Beta, etc.), proposant ainsi une segmentation complète du plan et permettant à partir de données réelles (et avec toutes les précautions évoquées plus haut) de choisir immédiatement le modèle le plus adapté.

Nous ne proposons pas ici de méthode permettant d'effectuer cette partition de manière théorique. Je doute d'ailleurs que ce ne soit en pratique réalisable dans la mesure où nous nous intéressons à des performances réelles, et non asymptotiques. Nous cherchons ainsi par exemple à juger de la qualité des classifieurs basés sur deux modèles (ici pour l'exemple Gaussien et SIRV), pour différentes lois de textures, et différents paramètres de lois. La méthode que nous proposons permettra d'étudier les performances non asymptotiques mais fera alors appel à des méthodes de Monte Carlo.

Protocole. La segmentation du plan κ_2 / κ_3 s'effectue de la manière suivante :

1. Nous choisissons un pavage régulier, ou aléatoire du plan κ_2 / κ_3 en choisissant diverses lois, ainsi que différents paramètres pour ces lois. Pour chacun de ces "points", nous effectuons un tirage d'un nombre fini de données, sur lesquelles nous allons alors

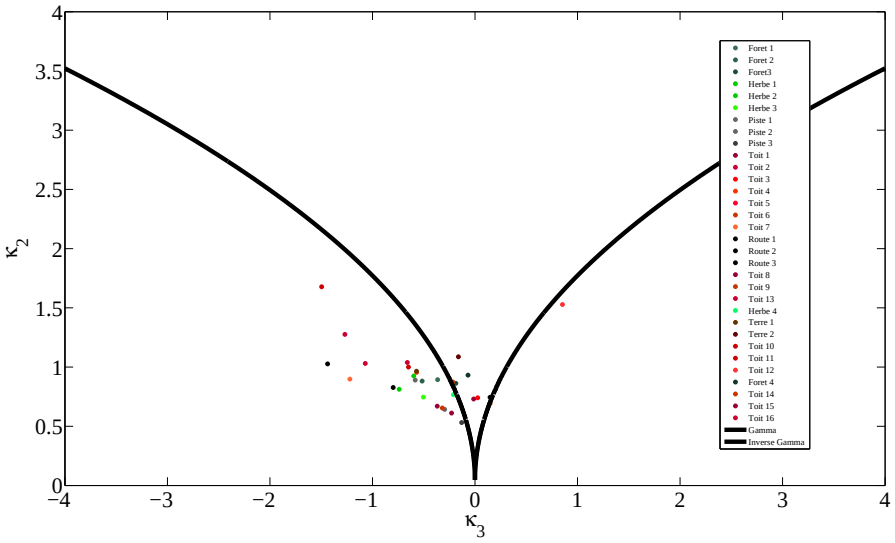


FIGURE 1.7 – Illustration des estimés de paramètres κ_2,κ_3 sur une image haute résolution de la ville de Toulouse.

estimer les paramètres κ_2^i / κ_3^i qui donneront la position estimée du point dans le plan. Ce nombre fini doit être similaire à celui que nous utiliserions si nous estimions ces paramètres sur des données réelles.

2. Pour la loi et les paramètres de loi associés à cette position, nous allons effectuer un très grand nombre de tirages de groupes de données. Pour chaque groupe de données, les tests de classification ou distances statistiques issus des modèles (ici gaussien et SIRV) sont calculés, et leurs performances mesurées. Le protocole exact qui permet de calculer la probabilité de détection pour une fausse alarme donnée est décrit en détail plus loin dans le manuscrit. Le nombre de tirages de groupe de données est alors suffisamment important pour déterminer quel outil propose dans ces conditions non asymptotiques (nombre d'échantillons limités), les meilleures performances (probabilité de bonne classification/segmentation). Ici, nous comparons uniquement deux modèles, mais il serait tout à fait possible de tester tous les modèles et d'enregistrer lequel propose, en cette position du plan, les meilleurs performances.
3. Nous disposons alors d'un nuage de points qui recouvrent une bonne partie du plan κ_2 / κ_3 (figure 1.8). Pour chacun de ces points nous connaissons le modèle qui permettra d'obtenir les meilleures performances, dans ces conditions non asymptotiques. Il reste alors uniquement à chercher là où les frontières où les performances des modèles sont similaires.

Le Coeur de Wishart

Cette frontière, qui permet de séparer les distributions SIRV et gaussiennes dans le plan κ_2 / κ_3 n'a pas du tout la forme circulaire centrée en 0 que nous aurions pu attendre.

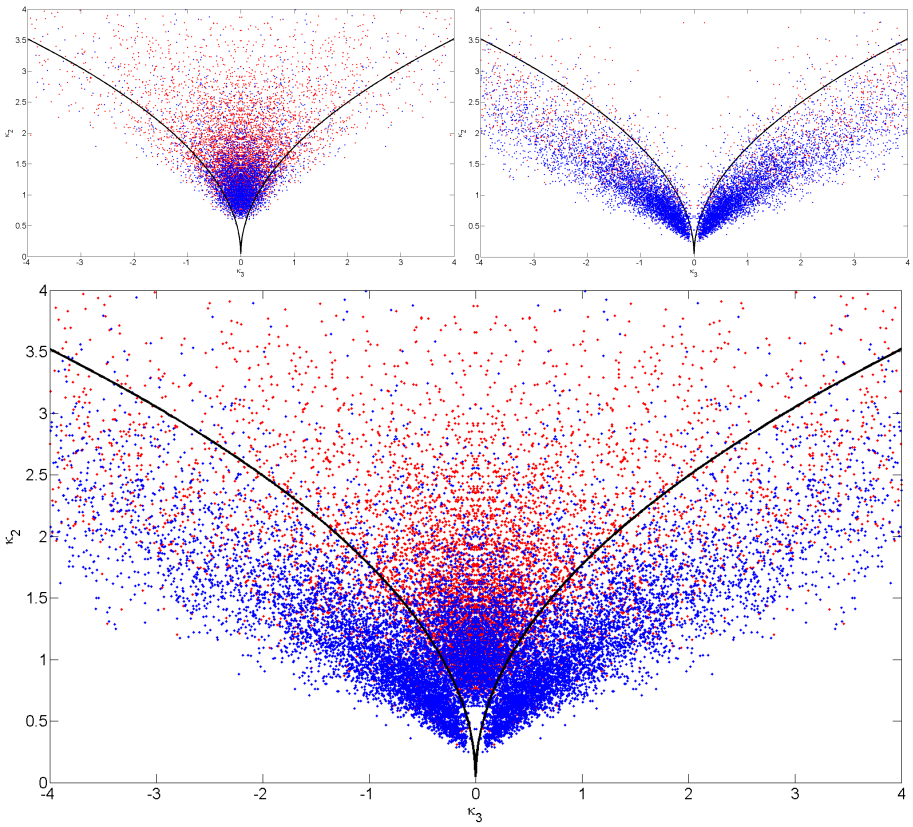


FIGURE 1.8 – Illustration des estimés de paramètres κ_2,κ_3 sur une image haute résolution de la ville de Toulouse. En haut à gauche, les tirages associés à une distribution de Fisher, en haut à droite dans le cas d'une distribution de type Beta (et inverse). En bas, l'ensemble des tirages. Nous ne présentons ici qu'un nombre de tirage restreint.

Cette frontière a une forme générale de coeur (figure 1.9). Elle n'est pas vraiment bien définie au niveau inférieur car en tirant aléatoirement des paramètres de lois Bêta ou Fisher, peu de tirages aboutissent à des log moments situés dans cette zone. On peut toutefois noter que le modèle gaussien reste plus performant que le modèle SIRV jusqu'à des valeurs assez importantes de ces logs-moments $\kappa_2 = 3, \kappa_3 = -3$.

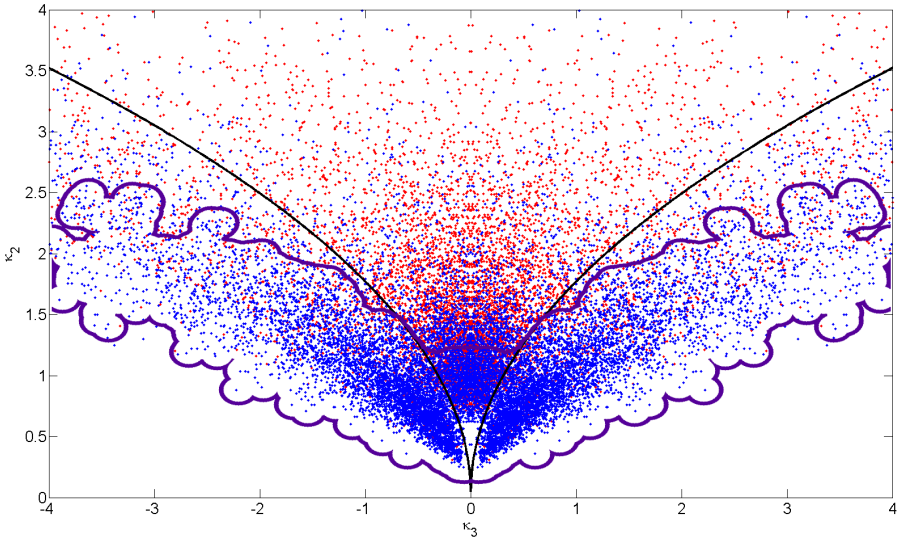


FIGURE 1.9 – Illustration des frontières dans le plan κ_2,κ_3 sur une image haute résolution de la ville de Toulouse.

La frontière montrée ici est très approximative, et nécessiterait un nombre bien plus important de tirages. Ces résultats ne sont qu'un avant-goût d'un projet de segmentation complet du plan en sections correspondant à des outils génériques (SIRV), ou dédiées à des distributions particulières (Fisher, Beta, Gamma). Nous pouvons toutefois conclure, en observant la position des points obtenus sur les données réelles de l'image de Toulouse et de Brétigny, qu'une proportion importante des points est située à l'intérieur de la zone gaussienne, ou à la frontière des deux domaines. Nous pouvons nous attendre alors à observer des performances qui favorisent encore l'utilisation de modèles gaussiens, ou un apport faible des modèles non gaussiens. Nous aurons l'occasion de vérifier ceci dans le chapitre 2.

1.4 Application aux tests statistiques en polarimétrie

Nous venons de présenter des distributions statistiques ainsi que des modèles plus généraux permettant de décrire la nature statistique des signaux mesurés en polarimétrie radar. Nous avons donné pour ces distributions les densités de probabilité, ainsi que des estimateurs de paramètres, obtenus souvent au sens du maximum de vraisemblance. Cela dit, ces densités de probabilité n'ont d'utilité que si elles sont employées à la conception de véritables outils statistiques qui seront à leur tour utilisés dans une application.

Dans cette section nous nous intéressons aux distances statistiques, ou tests statistiques. L'objectif ici est de pouvoir dire si deux ensembles de pixels d'une image polarimétrique correspondent aux mêmes comportements polarimétriques. C'est une problématique rencontrée régulièrement en classification où l'on cherche à regrouper les pixels d'une image en catégories. En classification non supervisée, on cherche uniquement à attribuer une même classe à des pixels ayant un comportement statistique équivalent. En classification supervisée, on identifie d'abord des sites d'entraînement dont on extrait la signature et à partir desquels on recherche ensuite les pixels les plus proches. La segmentation d'une image polarimétrique a besoin du même type d'outils statistiques. Elle s'approche globalement de la classification non supervisée, mais exploite la connexité des pixels dans l'image.

L'ensemble de ces applications utilise des outils statistiques appelés "distances". A l'aide des modèles statistiques on cherche à trouver un critère pertinent et performant pour discriminer les comportements statistiques des pixels mesurés [29]. Le terme "distance" ne doit cependant pas être entendu au sens mathématique du terme. Dans certains cas ces "distances" statistiques ne sont pas symétriques, ce qui peut être gênant pour un certain nombre d'applications, dont la segmentation. Par ailleurs de manière générale, aucune de ces distances ne vérifie l'inégalité triangulaire et nous pouvons parler uniquement de semi-métriques.

1.4.1 Rapport de Vraisemblance

Principe

La première approche que nous présentons est celle du rapport de vraisemblance [30] et plus précisément du rapport de vraisemblance généralisé(e). Cette méthode consiste à poser le problème sous la forme de deux hypothèses complémentaires appelées H_0 et H_1. Chacune des hypothèses attribue arbitrairement une distribution statistique aux observations mesurées (ici les vecteurs de diffusion $\mathbf{k}$). La vraisemblance de cette hypothèse s'exprime alors comme le produit des probabilités de chacune des observations sous cette hypothèse. Ainsi la vraisemblance statistique de l'hypothèse H_0 compte tenu des observations $\mathbf{k_1}, \mathbf{k_2}, \ldots \mathbf{k_N}$, supposées indépendantes, s'écrit :

$$L_{H_0}(\mathbf{k_1}, \mathbf{k_2}, \ldots \mathbf{k_N}) = \prod_{i=1}^{N} P_{H_i}(\mathbf{k_i}). \tag{1.17}$$

Le rapport de la vraisemblance de deux hypothèses permet alors de déterminer, entre deux hypothèses, laquelle est la plus vraisemblable d'après les observations mesurées.

$$Q = \frac{L_{H_0}(\mathbf{k_1}, \mathbf{k_2}, \ldots \mathbf{k_N})}{L_{H_1}(\mathbf{k_1}, \mathbf{k_2}, \ldots \mathbf{k_N})}. \tag{1.18}$$

Si nous choisissons de minimiser la probabilité de non détection ou la probabilité de fausse alarme, nous pourrons déterminer un seuil qui nous permettra de choisir entre les deux hypothèses de départ H_0 et H_1. Pour la détection de cibles en radar, il est classique de se fixer une probabilité de fausse alarme relativement faible, afin de calculer le seuil du rapport de vraisemblance.

Rapport de Vraisemblance Généralisé(e)

Pour un test d'hypothèse simple, le test optimal au sens de Neyman Pearson est fourni par le rapport entre la vraisemblance des deux hypothèses (Eq .1.18) [31]. Malheureusement, la plupart du temps chaque hypothèse prend en compte un certain nombre de paramètres. Il s'agit alors d'un test d'hypothèse composite. L'existence d'un test optimal n'est alors pas garantie. Par exemple dans le cas d'une densité de probabilité gaussienne, la matrice de covariance est un paramètre inconnu qui intervient dans l'expression de la vraisemblance.

La méthode traditionnellement employée, nommée approche "Plug-in", consiste alors à remplacer ces paramètres par des paramètres estimés sur les données testées ou sur un autre jeu de données disponibles. Parmi ces tests "Plug-in", le plus populaire utilise des paramètres estimés au sens du maximum de vraisemblance. C'est le *rapport de vraisemblance généralisé*. De manière générale, il n'est pas optimal au sens UMP (Uniformément plus puissant). Ainsi il est tout à fait possible de concevoir, par l'approche "Plug-in", des tests plus performants. En pratique, le rapport de vraisemblance généralisé offre souvent de très bonnes performances et est très régulièrement utilisé dans la littérature.

Nous faisons cependant remarquer ici qu'il arrive que certains tests de type "Plug-in" soient confondus avec le rapport de vraisemblance généralisé. Au sens du maximum de vraisemblance, les paramètres doivent être estimés sous l'hypothèse associée (H_0 ou H_1), ce qui aboutit la plupart du temps à des estimateurs différents. Cette remarque nous permettra d'ailleurs de calculer de vrais rapports de vraisemblance généralisés dans le cas du modèle SIRV, alors qu'ils n'avaient, à notre connaissance, jamais été calculés exactement dans ce cadre.

Nous aurons par ailleurs l'occasion de construire un certain nombre de tests de type "Plug-in" ayant des performances très intéressantes voir supérieures aux rapports de vraisemblance généralisés.

Test d'égalité de distributions

Il est possible de construire des outils statistiques adaptés aux problématiques de classification et de segmentation à l'aide de la méthode du rapport de vraisemblance. Le but étant de reconnaître les zones ou groupes de pixels ayant le même comportement statistique, il nous suffit alors de choisir une hypothèse adaptée du type :

$$\begin{array}{ll} H_0 : & \text{'les deux groupes d'échantillons suivent la même distribution'} \\ H_1 : & \text{'les deux groupes d'échantillons ne suivent pas la même distribution'} \end{array} \tag{1.19}$$

En pratique ce test va aboutir à des expressions très différentes en fonction des informations à notre disposition a priori, ainsi que de la façon dont on va exprimer mathématiquement le fait que les deux groupes suivent la même distribution. En effet une vraisemblance lie des échantillons à une hypothèse, qui dépend donc d'une loi de distribution et de certains paramètres. Dans notre cas nous ne disposons que de deux groupes d'échantillons que nous notons :

$$(\mathbf{k_1^A}, \mathbf{k_2^A}, \ldots, \mathbf{k_N^A}) \quad \text{et} \quad (\mathbf{k_1^B}, \mathbf{k_2^B}, \ldots, \mathbf{k_P^B}). \tag{1.20}$$

Un test de vraisemblance ne peut pas tester directement la vraisemblance d'un groupe d'échantillons par rapport à un autre. Il est nécessaire de poser arbitrairement une ou plusieurs lois, donc un ou plusieurs paramètres (comme Σ_A ou Σ_B les matrices de cohérences associés aux distributions des familles A et B). Ces paramètres vont nous servir de relais afin de lier les échantillons entre eux dans un test statistique.

Nous allons voir que dans le cas de notre problématique il existe trois manières d'introduire de tels paramètres.

1. *Un seul paramètre.* Nous supposons arbitrairement un seul paramètre connu. Il est associé à la distribution d'une des deux populations. Ici par exemple nous supposons $\mathbf{\Sigma_B}$ connu. Alors nous pouvons exprimer que les deux groupes suivent la même distribution en étudiant la vraisemblance de l'événement suivant : *les échantillons de la famille A suivent cette loi de paramètre* $\mathbf{\Sigma_B}$. Ceci donne le test asymétrique suivant :

$$\begin{array}{ll} H_0 : & \text{'Le groupe d'échantillon de type A suit la loi de paramètre } \mathbf{\Sigma_B}\text{'} \\ H_1 : & \text{'Le groupe d'échantillon de type A ne suit pas la loi de paramètre } \mathbf{\Sigma_B}\text{'} \end{array} \tag{1.21}$$

2. *Deux paramètres.* Nous supposons deux paramètres connus. Chacun des paramètres possède la distribution d'une des deux populations $\mathbf{\Sigma_A}$ et $\mathbf{\Sigma_B}$. Dans cette situation exprimer que les deux familles suivent la même distribution revient alors à dire, que chacune des populations suit la loi associée au paramètre de l'autre. Ce qui donne alors,

$$\begin{array}{ll} H_0 : & \text{'Le groupe d'échantillon de type A suit la loi } \mathbf{\Sigma_B} \text{ et B suit la loi } \mathbf{\Sigma_A}\text{'} \\ H_1 : & \text{'Le groupe d'échantillon de type A ne suit pas la loi } \mathbf{\Sigma_B} \text{ et B ne suit pas la loi } \mathbf{\Sigma_A}\text{'} \end{array} \tag{1.22}$$

3. *Un paramètre commun.* La dernière option est de poser un paramètre commun. il est noté, $\mathbf{\Sigma_C}$, et supposé connu, il correspond dans le cas de l'hypothèse H_0 à la distribution des populations A et B, $(\mathbf{k_1^C}, \mathbf{k_2^C}, \ldots, \mathbf{k_{N+P}^C})$. L'expression du test est aussi cette fois symétrique et s'exprime comme suit :

$$\begin{array}{ll} H_0 : & \text{'Le groupe d'échantillon de type A et de type B suivent la loi } \mathbf{\Sigma_C}\text{'} \\ H_1 : & \text{'Le groupe d'échantillon de type A et de type B ne suivent pas la loi } \mathbf{\Sigma_C}\text{'} \end{array} \tag{1.23}$$

Nous allons voir que les performances relatives de ces trois tests sont extrêmement proches. Cependant le nombre de paramètres à estimer variant, d'un point de vue pratique il existe des différences non négligeables, surtout quand on travaille avec des distributions complexes qui nécessitent alors l'estimation de nombreux paramètres.

Nous allons maintenant traduire ces trois tests en fonction des distributions rencontrées. Nous nous contenterons de rappeler l'expression de ces tests dans le cas d'une distribution gaussienne et de les développer dans le cas d'une distribution SIRV.

Cas Gaussien

Nous allons ici utiliser les expressions présentées dans la partie précédente qui concernait les distributions. Si on pose une loi de distribution gaussienne, centrée, de matrice de covariance $\mathbf{\Sigma}$, alors la vraisemblance de la famille d'observations, supposées indépendantes, $(\mathbf{k_1}, \mathbf{k_2}, \ldots, \mathbf{k_N})$ est donnée par :

$$L_{\mathbf{\Sigma}}(\mathbf{k_1}, \mathbf{k_2}, \ldots \mathbf{k_N}) = \prod_{i=1}^{N} P_{\Sigma}(\mathbf{k_i}) = \prod_{i=1}^{N} \frac{1}{\pi^m |\mathbf{\Sigma}|} \exp\left(-\mathbf{k_i}^{\dagger} \mathbf{\Sigma}^{-1} \mathbf{k_i}\right). \tag{1.24}$$

Il n'est pas possible d'exprimer sous la forme d'une vraisemblance la non-appartenance à une loi. Pour cela sous l'hypothèse H_1 qui correspond ici à la non-appartenance des observations à une loi donnée, nous n'exprimerons qu'une vraisemblance de normalisation qui correspondra à l'expression '*Le groupe d'échantillon de type A est distribué selon une loi* $\mathbf{\Sigma_A}$'. Cette vraisemblance est en effet calculée à partir des données qui serviront à la fois à estimer les paramètres de loi associée $\mathbf{\Sigma_A}$, mais aussi à calculer cette vraisemblance. Cette vraisemblance n'est pas informative et pourrait être appelée "auto vraisemblance", elle peut aussi être assimilée à un terme de normalisation. L'expression de cette vraisemblance s'assimile d'ailleurs à une métrique connue sous le nom de la métrique de Kähler.

Pour une loi gaussienne, en remplaçant la matrice de covariance $\mathbf{\Sigma}$ par son estimateur

au sens du maximum de vraisemblance $\hat{\mathbf{\Sigma}}$, on obtient :

$$\begin{aligned}
L_{H_1}(\mathbf{k_1}, \mathbf{k_2}, \ldots, \mathbf{k_N}) &= \prod_{i=1}^{N} \frac{1}{\pi^m |\sum_{j=1}^{N} \mathbf{k_j}\mathbf{k_j}^\dagger|} \exp\left(-\mathbf{k_i}^\dagger \big(\frac{1}{N}\sum_{j=1}^{N} \mathbf{k_j}\mathbf{k_j}^\dagger\big)^{-1} \mathbf{k_i}\right) \\
&= \frac{1}{\pi^{Nm} |\frac{1}{N}\sum_{j=1}^{N} \mathbf{k_j}\mathbf{k_{=j}}^\dagger|^N} \prod_{i=1}^{N} \exp\left(-\mathbf{k_i}^\dagger \big(\frac{1}{N}\sum_{j=1}^{N} \mathbf{k_j}\mathbf{k_j}^\dagger\big)^{-1} \mathbf{k_i}\right) \\
&= \frac{1}{\pi^{Nm} |\frac{1}{N}\sum_{j=1}^{N} \mathbf{k_j}\mathbf{k_j}^\dagger|^N} \exp\left(\sum_{i=1}^{N} -\mathbf{k_i}^\dagger \big(\frac{1}{N}\sum_{j=1}^{N} \mathbf{k_j}\mathbf{k_j}^\dagger\big)^{-1} \mathbf{k_i}\right) \\
&= \frac{1}{\pi^{Nm} |\frac{1}{N}\sum_{j=1}^{N} \mathbf{k_j}\mathbf{k_j}^\dagger|^N} \exp\left(\sum_{i=1}^{N} -\mathrm{tr}\big(\big(\frac{1}{N}\sum_{j=1}^{N} \mathbf{k_j}\mathbf{k_j}^\dagger\big)^{-1} \mathbf{k_i}\mathbf{k_i}^\dagger\big)\right) \\
&= \frac{1}{\pi^{Nm} |\hat{\mathbf{\Sigma}}|^N} \exp\left(-Nm\right),
\end{aligned} \tag{1.25}$$

avec $\hat{\mathbf{\Sigma}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{k_i}\mathbf{k_i}^\dagger$.

Nous allons maintenant développer les expressions des tests à l'aide des équations (1.24) et (1.25).

1. *Un seul paramètre.* Ici un seul paramètre est supposé connu $\mathbf{\Sigma_B}$. Ce qui donne le rapport de vraisemblance :

$$\begin{aligned}
Q_1 &= \frac{\prod_{i=1}^{N} \frac{1}{\pi^m |\mathbf{\Sigma_B}|} \exp\left(-\mathbf{k_i}^{A\dagger} \mathbf{\Sigma_B}^{-1} \mathbf{k}_i^A\right)}{\frac{1}{\pi^{Nm} |\hat{\mathbf{\Sigma}}_\mathbf{A}|^N} \exp\left(-Nm\right)} \\
&= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\mathbf{\Sigma_B}|^N} \exp\left(\sum_{i=1}^{N} -\mathbf{k_i}^{A\dagger} \mathbf{\Sigma_B}^{-1} \mathbf{k}_i^A + Nm\right) \\
&= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\mathbf{\Sigma_B}|^N} \exp\left(-N\ \mathrm{tr}(\mathbf{\Sigma_B}^{-1} \hat{\mathbf{\Sigma}}_\mathbf{A}) + Nm\right)
\end{aligned} \tag{1.26}$$

Dans les faits $\mathbf{\Sigma_B}$ est rarement, voire jamais connu. Nous pouvons le remplacer par $\hat{\mathbf{\Sigma}}_\mathbf{B} = \frac{1}{P}\sum_{i=1}^{P} \mathbf{k_i^B}\mathbf{k_i}^{\mathbf{B}\dagger}$ ce qui donne :

$$Q_1 = \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\hat{\mathbf{\Sigma}}_\mathbf{B}|^N} \exp\left(-N\ \mathrm{tr}(\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1} \hat{\mathbf{\Sigma}}_\mathbf{A}) + Nm\right). \tag{1.27}$$

Ce test est souvent appelé distance de Wishart dans la littérature. Il est très régulièrement employé dans des applications de classification lorsqu'une des populations (ici B) est bien connue et que l'on cherche à classifier une population plus petite, ici A. L'asymétrie dans ce type d'application n'est pas un problème. Cette distance vérifie certaines propriétés mathématiques d'une véritable distance, mais ne vérifie pas l'inégalité triangulaire. Elle n'est par contre pas applicable dans des problématiques où la symétrie est requise, comme par exemple, en segmentation d'images.

2. *Deux paramètres.* Ici $\mathbf{\Sigma_A}$ et $\mathbf{\Sigma_B}$ sont supposés connus. Le rapport de vraisemblance exposé plus haut devient alors :

$$\begin{aligned} Q_2 &= \frac{\prod_{i=1}^{N} \frac{1}{\pi^m |\mathbf{\Sigma_B}|} \exp\left(-\mathbf{k_i}^{A\dagger} {\mathbf{\Sigma_B}}^{-1} \mathbf{k}_i^A\right) \prod_{i=1}^{P} \frac{1}{\pi^m |\mathbf{\Sigma_A}|} \exp\left(-\mathbf{k_i}^{B\dagger} {\mathbf{\Sigma_A}}^{-1} \mathbf{k}_i^B\right)}{\frac{1}{\pi^{Nm} |\hat{\mathbf{\Sigma}}_\mathbf{A}|^N} \exp\left(-Nm\right) \frac{1}{\pi^{Pm} |\hat{\mathbf{\Sigma}}_\mathbf{B}|^P} \exp\left(-Pm\right)} \\ &= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N |\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\mathbf{\Sigma_B}|^N |\mathbf{\Sigma_A}|^P} \exp\Big(-N\mathrm{tr}({\mathbf{\Sigma_B}}^{-1} \hat{\mathbf{\Sigma}_\mathbf{A}}) - P\mathrm{tr}({\mathbf{\Sigma_A}}^{-1} \hat{\mathbf{\Sigma}}_\mathbf{B}) + (N+P)m\Big) \end{aligned} \tag{1.28}$$

Comme dans le cas précédent, si $\mathbf{\Sigma_A}$ et $\mathbf{\Sigma_B}$ sont remplacés par leurs estimateurs au sens du maximum de vraisemblance, on obtient :

$$Q_2 = \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^{N-P}}{|\hat{\mathbf{\Sigma}}_\mathbf{B}|^{N-P}} \exp\Big(-N\mathrm{tr}(\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1} \hat{\mathbf{\Sigma}}_\mathbf{A}) - P\mathrm{tr}(\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1} \hat{\mathbf{\Sigma}}_\mathbf{B}) + (N+P)m\Big). \tag{1.29}$$

Ce test est souvent utilisé sous l'appellation *distance de Wishart symétrique*, ou bien *distance Wishart révisée* (*Revised Wishart Distance*) [13] [32]. Comme la distance de Wishart, elle ne vérifie pas l'inégalité triangulaire. La symétrie rend cette distance particulièrement intéressante. Dans le cas d'une distribution complexe, elle a l'avantage de ne nécessiter que l'estimation de deux jeux de paramètres, ce qui rend son calcul plus facile.

3. *Un paramètre commun.* Ici nous supposons un paramètre $\mathbf{\Sigma_C}$, correspondant au paramètre de loi, sous l'hypothèse H_0 connu, il est commun aux deux familles. Son estimateur au sens du maximum de vraisemblance est donné par : $\hat{\mathbf{\Sigma}}_\mathbf{C} = 1/(N+P) \sum_{i=1}^{N+P} \mathbf{k}_i^{(A,B)} {\mathbf{k_i}}^{(A,B)\dagger}$. Nous pouvons remarquer, et cela sera vrai *uniquement* dans le cas gaussien, que nous pouvons écrire : $\hat{\mathbf{\Sigma}}_\mathbf{C} = 1/(N+P)\Big(N\hat{\mathbf{\Sigma}}_\mathbf{A} + P\hat{\mathbf{\Sigma}}_\mathbf{B}\Big)$.

Le rapport de vraisemblance devient alors :

$$\begin{aligned} Q_3 &= \frac{\prod_{i=1}^{N} \frac{1}{\pi^m |\Sigma_C|} \exp\left(-\mathbf{k_i}^{A\dagger} \mathbf{\Sigma_C}^{-1} \mathbf{k}_i^A\right)}{\frac{1}{\pi^{Nm} |\hat{\mathbf{\Sigma}}_\mathbf{A}|^N} \exp\left(-Nm\right)} \frac{\prod_{i=1}^{P} \frac{1}{\pi^m |\mathbf{\Sigma_C}|} \exp\left(-\mathbf{k_i}^{B\dagger} \mathbf{\Sigma_C}^{-1} \mathbf{k}_i^B\right)}{\frac{1}{\pi^{Pm} |\hat{\mathbf{\Sigma}}_\mathbf{B}|^P} \exp\left(-Pm\right)} \\ &= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N |\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\mathbf{\Sigma_C}|^{N+P}} \exp\left(-N\mathrm{tr}(\mathbf{\Sigma_C}^{-1} \hat{\mathbf{\Sigma}}_\mathbf{A}) - P\mathrm{tr}(\mathbf{\Sigma_C}^{-1} \hat{\mathbf{\Sigma}}_\mathbf{B}) + (N+P)m\right) \\ &= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N |\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\mathbf{\Sigma_C}|^{N+P}} \exp\left(-\mathrm{tr}(\mathbf{\Sigma_C}^{-1}(N\hat{\mathbf{\Sigma}}_\mathbf{A} + P\hat{\mathbf{\Sigma}}_\mathbf{B})) + (N+P)m\right) \end{aligned} \tag{1.30}$$

Si on remplace $\mathbf{\Sigma_C}$ par son estimateur au sens du maximum de vraisemblance on obtient alors :

$$Q_3 = \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N |\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\hat{\mathbf{\Sigma}}_\mathbf{C}|^{N+P}}. \tag{1.31}$$

Cette semi métrique est connue sous le nom de distance de Bartlett [33], [34]. Elle est couramment utilisée dans des applications de classification ou de segmentation. Nous verrons par contre qu'avec des modèles non gaussiens son expression devient plus complexe.

Cas du modèle SIRV

Nous reprenons ici le modèle SIRV avec l'approximation de la texture déterministe et l'estimateur de texture τ_i et de matrice de covariance du point fixe M_{FP}. Dans ce cadre nous développons les trois rapports de vraisemblance précédents.

La vraisemblance des observations $(\mathbf{k_1}, \mathbf{k_2}, \ldots, \mathbf{k_N})$ dépend des paramètres $\Sigma, \tau_1, \tau_2, \ldots, \tau_N$ et s'exprime par :

$$L_{\Sigma,\tau_1,\tau_2,\ldots,\tau_N}(\mathbf{k_1}, \mathbf{k_2}, \ldots \mathbf{k_N}) = \prod_{i=1}^{N} P_{\Sigma,\tau_i}(\mathbf{k_i}) = \prod_{i=1}^{N} \frac{1}{\pi^m \tau_i^m |\Sigma|} \exp\left(\frac{-\mathbf{k_i}^\dagger \Sigma^{-1} \mathbf{k_i}}{\tau_i}\right). \tag{1.32}$$

Comme pour le cas gaussien il convient d'étudier la vraisemblance des données sous l'hypothèse H_1 dans le cas SIRV. On va donc étudier la vraisemblance d'un jeu de données en remplaçant tous les paramètres par leurs estimateurs au sens du maximum de vraisemblance avec ce même jeu de données. Nous rappelons que la construction de ces estimateurs s'effectue en deux étapes. D'abord nous estimons la texture, en chaque échantillon i, par l'estimateur : $\tau_i = \frac{\mathbf{k_i}^\dagger \Sigma^{-1} \mathbf{k_i}}{m}$. On en déduit alors l'estimateur de la matrice de covariance, par la méthode du point fixe. Dans l'hypothèse H_1 nous supposons que les données sont distribuées suivant la distribution associée à leur propre matrice de cohérence estimée $\hat{\mathbf{\Sigma}}_\mathbf{A}$ ou $\hat{\mathbf{\Sigma}}_\mathbf{B}$. L'estimateur de la texture s'écrit alors, par exemple pour l'échantillon de la famille A : $\hat{\tau_i^A} = \frac{\mathbf{k_i}^{A\dagger} \hat{\mathbf{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A}}{m}$.

Ainsi nous avons :

$$\begin{aligned} L_{H_1}(\mathbf{k_1}, \mathbf{k_2}, \ldots \mathbf{k_N}) &= \prod_{i=1}^{N} \frac{1}{\pi^m \tau_i^m |\hat{\Sigma}|} \exp\left(\frac{-\mathbf{k_i}^\dagger \hat{\Sigma}^{-1} \mathbf{k_i}}{\tau_i}\right) \\ &= \prod_{i=1}^{N} \frac{1}{\pi^m \left(\frac{\mathbf{k_i}^\dagger \hat{\mathbf{\Sigma}}^{-1} \mathbf{k_i}}{m}\right)^m |\hat{\mathbf{\Sigma}}|} \exp\left(\frac{-\mathbf{k_i}^\dagger \hat{\mathbf{\Sigma}}^{-1} \mathbf{k_i}}{\frac{\mathbf{k_i}^\dagger \hat{\mathbf{\Sigma}}^{-1} \mathbf{k_i}}{m}}\right) \\ &= \frac{m^{Nm}}{\pi^{Nm} \prod_{i=1}^{N} \left(\mathbf{k_i}^\dagger \hat{\mathbf{\Sigma}}^{-1} \mathbf{k_i}\right)^m |\hat{\mathbf{\Sigma}}|^N} \exp\left(-Nm\right) \end{aligned} \tag{1.33}$$

1. *Un seul paramètre.* Nous supposons le paramètre $\mathbf{\Sigma_B}$ connu. Dans un second temps nous pourrons le remplacer par son estimateur au sens du maximum de vraisemblance, qui correspond à l'estimateur du point fixe. Dans l'expression de la vraisemblance sous l'hypothèse H_0 interviennent à la fois les échantillons $\mathbf{k_i^A}$ mais aussi le coefficient de texture τ_i^A que nous allons devoir estimer. Il convient de faire particulièrement attention à l'estimation de ce paramètre.

   En toute rigueur, si nous voulons vraiment obtenir le rapport de vraisemblance généralisé, il convient d'estimer ce paramètre au sens du maximum de vraisemblance sous l'hypothèse H_0. Sous cette hypothèse, les échantillons $\mathbf{k_i^A}$ suivent alors une distribution SIRV, de matrice de cohérence $\mathbf{\Sigma_B}$ (supposée connue). L'estimateur au sens du maximum de vraisemblance est donc $\hat{\tau_i^A} = \frac{\mathbf{k_i}^{A\dagger} \mathbf{\Sigma_B}^{-1} \mathbf{k_i^A}}{m}$. Celui-ci diffère cette fois de l'estimation du paramètre sous l'hypothèse H_1. Ainsi, les tests basés sur le modèle SIRV qui utilisent le même estimateur de τ pour les deux hypothèses ne sont que des rapports de vraisemblance "Plug-in". C'est le cas par exemple dans [25]. Naturellement a priori il n'y a pas de raison de privilégier une méthode par rapport à une autre, c'est pourquoi nous présentons ici les deux rapports de vraisemblances. Notons que les expressions que nous obtenons pour les "vrais" rapports de vraisemblance généralisés sont inédites.

   – *Rapport de vraisemblance "Plug in"*, $\hat{\tau_i^A}|H_0 = \frac{\mathbf{k_i}^{A\dagger} \hat{\mathbf{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A}}{m}$,

$$\begin{aligned} Q_1 &= \frac{\prod_{i=1}^{N} \frac{m^m}{\pi^m (\mathbf{k_i}^{A\dagger} \hat{\mathbf{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A})^m |\mathbf{\Sigma_B}|} \exp\left(-m \frac{\mathbf{k_i}^{A\dagger} \mathbf{\Sigma_B}^{-1} \mathbf{k_i}^A}{\mathbf{k_i}^{A\dagger} \hat{\mathbf{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A}}\right)}{\frac{m^{Nm}}{\pi^{Nm} \prod_{i=1}^{N} \left(\mathbf{k_i}^{A\dagger} \mathbf{\Sigma_A}^{-1} \mathbf{k_i^A}\right)^m |\hat{\mathbf{\Sigma}}_\mathbf{A}|^N} \exp\left(-Nm\right)} \\ &= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\mathbf{\Sigma_B}|^N} \prod_{i=1}^{N} \exp\left(-m \frac{\mathbf{k_i}^{A\dagger} \mathbf{\Sigma_B}^{-1} \mathbf{k_i}^A}{\mathbf{k_i}^{A\dagger} \hat{\mathbf{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A}} + Nm\right) \\ &= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\mathbf{\Sigma_B}|^N} \exp\left(\sum_{i=1}^{N} -m \frac{\mathrm{tr}(\mathbf{\Sigma_B}^{-1} \mathbf{k_i}^A \mathbf{k_i}^{A\dagger})}{\mathbf{k_i}^{A\dagger} \hat{\mathbf{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A}} + Nm\right) \\ &= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\mathbf{\Sigma_B}|^N} \exp\left(-N \mathrm{tr}(\mathbf{\Sigma_B}^{-1} \hat{\mathbf{\Sigma}}_\mathbf{A}) + Nm\right). \end{aligned} \tag{1.34}$$

A la troisième ligne nous faisons apparaître l'expression de l'estimateur du point fixe, qui nous permet d'obtenir une simplification. On retrouve alors la même expression que dans le cas gaussien, seule l'estimation des matrices est différente et utilise la méthode du point fixe.
Si nous remplaçons maintenant $\mathbf{\Sigma_B}$ par son estimateur au sens du maximum du vraisemblance nous obtenons alors :

$$Q_1 = \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\hat{\mathbf{\Sigma}}_\mathbf{B}|^N} \exp\left(-N\mathrm{tr}(\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\hat{\mathbf{\Sigma}}_\mathbf{A}) + Nm\right). \tag{1.35}$$

– *Rapport de vraisemblance Généralisé,* $\hat{\tau_i^A}|H_0 = \frac{\mathbf{k_i}^{A\dagger}{\mathbf{\Sigma_B}}^{-1}\mathbf{k_i^A}}{m}$

$$\begin{aligned} Q_1 &= \frac{\prod_{i=1}^N \frac{m^m}{\pi^m(\mathbf{k_i}^{A\dagger}{\mathbf{\Sigma_B}}^{-1}\mathbf{k_i^A})^m|\mathbf{\Sigma_B}|} \exp\left(-m\frac{\mathbf{k_i}^{A\dagger}{\mathbf{\Sigma_B}}^{-1}\mathbf{k_i}^A}{\mathbf{k_i}^{A\dagger}{\mathbf{\Sigma_B}}^{-1}\mathbf{k_i^A}}\right)}{\frac{m^{Nm}}{\pi^{Nm}\prod_{i=1}^N\left(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A}\right)^m|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}\exp(-Nm)} \\ &= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\mathbf{\Sigma_B}|^N}\prod_{i=1}^N\frac{(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}{\mathbf{\Sigma_B}}^{-1}\mathbf{k_i^A})^m}. \end{aligned} \tag{1.36}$$

L'expression de ce test est donc inédite. Le test est asymétrique et il n'existe d'après moi pas de solution pour simplifier son expression. Son calcul nécessite alors l'ensemble des données. Cela dit, la forme quadratique peut se calculer facilement de manière vectorielle. De la même manière si on remplace $\mathbf{\Sigma_B}$ par son estimateur au sens du maximum du vraisemblance nous obtenons alors :

$$Q_1 = \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}{|\hat{\mathbf{\Sigma}}_\mathbf{B}|^N}\prod_{i=1}^N\frac{(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^A})^m}. \tag{1.37}$$

2. *Deux paramètres.* Ici $\mathbf{\Sigma_A}$ et $\mathbf{\Sigma_B}$ sont supposés connus. Sous l'hypothèse H_0 nous pouvons encore une fois distinguer un test "Plug in", où les estimateurs de $\hat{\tau_i^A}$ et $\hat{\tau_i^B}$ sont les mêmes que sous l'hypothèse H_1, du rapport de vraisemblance généralisé où les estimateurs sont $\hat{\tau_i^A}|H_0 = \frac{\mathbf{k_i}^{A\dagger}{\mathbf{\Sigma_B}}^{-1}\mathbf{k_i^A}}{m}$ et $\hat{\tau_i^B}|H_0 = \frac{\mathbf{k_i}^{B\dagger}{\mathbf{\Sigma_A}}^{-1}\mathbf{k_i^B}}{m}$.

   – *Rapport de vraisemblance "Plug in",* $\hat{\tau_i^A}|H_0 = \frac{\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A}}{m}$ et $\hat{\tau_i^B}|H_0 = \frac{\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B}}{m}$.

$$
\begin{aligned}
Q_2 &= \frac{\prod_{i=1}^{N} \frac{m^m}{\pi^m(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A})^m|\mathbf{\Sigma_B}|}\exp\left(-m\frac{\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k}_i^A}{\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A}}\right)}{\frac{m^{Nm}}{\pi^{Nm}\prod_{i=1}^{N}\left(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A}\right)^m|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}\exp\left(-Nm\right)} \\
&\quad \cdot \frac{\prod_{i=1}^{P} \frac{m^m}{\pi^m(\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B})^m|\mathbf{\Sigma_A}|}\exp\left(-m\frac{\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k}_i^B}{\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B}}\right)}{\frac{m^{Pm}}{\pi^{Pm}\prod_{i=1}^{P}\left(\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B}\right)^m|\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}\exp\left(-Pm\right)} \\
&= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N|\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\mathbf{\Sigma_B}|^N|\mathbf{\Sigma_A}|^P}\exp\Big(-N\mathrm{tr}(\mathbf{\Sigma_B}^{-1}\hat{\mathbf{\Sigma}}_\mathbf{A}) - P\mathrm{tr}(\mathbf{\Sigma_A}^{-1}\hat{\mathbf{\Sigma}}_\mathbf{B}) + (N+P)m\Big)
\end{aligned}
\tag{1.38}
$$

Encore une fois on retrouve la même expression que dans le cas gaussien. De la même manière si on remplace $\mathbf{\Sigma_A}$ et $\mathbf{\Sigma_B}$ par leurs estimateurs au sens du maximum du vraisemblance nous obtenons alors :

$$
Q_2 = \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^{N-P}}{|\hat{\mathbf{\Sigma}}_\mathbf{B}|^{N-P}}\exp\Big(-N\mathrm{tr}(\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\hat{\mathbf{\Sigma}}_\mathbf{A}) - P\mathrm{tr}(\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\hat{\mathbf{\Sigma}}_\mathbf{B}) + (N+P)m\Big)
\tag{1.39}
$$

– *Rapport de vraisemblance Généralisé,* $\hat{\tau_i^A}|H_0 = \frac{\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k_i^A}}{m}$ et $\hat{\tau_i^B}|H_0 = \frac{\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k_i^B}}{m}$.

$$
\begin{aligned}
Q_2 &= \frac{\prod_{i=1}^{N} \frac{m^m}{\pi^m(\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k_i^A})^m|\mathbf{\Sigma_B}|}\exp\left(-m\frac{\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k}_i^A}{\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k_i^A}}\right)}{\frac{m^{Nm}}{\pi^{Nm}\prod_{i=1}^{N}\left(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A}\right)^m|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}\exp\left(-Nm\right)} \\
&\quad \cdot \frac{\prod_{i=1}^{P} \frac{m^m}{\pi^m(\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k_i^B})^m|\mathbf{\Sigma_A}|}\exp\left(-m\frac{\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k}_i^B}{\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k_i^B}}\right)}{\frac{m^{Pm}}{\pi^{Pm}\prod_{i=1}^{P}\left(\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B}\right)^m|\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}\exp\left(-Pm\right)} \\
&= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N|\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\mathbf{\Sigma_B}|^N|\mathbf{\Sigma_A}|^P}\prod_{i=1}^{N}\frac{(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k_i^A})^m}\prod_{i=1}^{P}\frac{(\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B})^m}{(\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k_i^B})^m}
\end{aligned}
\tag{1.40}
$$

Comme pour tous les cas précédents nous remplaçons $\mathbf{\Sigma_A}$ et $\mathbf{\Sigma_B}$ par leurs esti-

mateurs,

$$Q_2 = \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^{N-P}}{|\hat{\mathbf{\Sigma}}_\mathbf{B}|^{N-P}} \prod_{i=1}^{N} \frac{(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}_\mathbf{B}}^{-1}\mathbf{k_i^A})^m} \prod_{i=1}^{P} \frac{(\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B})^m}{(\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^B})^m} \tag{1.41}$$

De la même manière nous obtenons un test inédit, qui dépend encore une fois de l'ensemble des données.

3. *Troisième paramètre.* Ici comme dans le cas gaussien nous supposons $\mathbf{\Sigma_C}$ connu. Par la suite son estimateur au sens du maximum de vraisemblance est calculé par la méthode du point fixe sur l'ensemble des données $\mathbf{k_i^A}$ et $\mathbf{k_i^B}$. *Contrairement au cas gaussien cet estimé n'est pas égal à la moyenne pondérée des deux estimateurs de* $\mathbf{\Sigma_A}$ *et* $\mathbf{\Sigma_B}$. Comme dans les deux cas précédents nous distinguons l'approche "Plug in" de l'approche *rapport de vraisemblance généralisé*,
   - *Rapport de vraisemblance "Plug in"*, $\hat{\tau_i^A}|H_0 = \frac{\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A}}{m}$ et $\hat{\tau_i^B}|H_0 = \frac{\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B}}{m}$.

$$\begin{aligned}
Q_3 &= \frac{\prod_{i=1}^{N} \frac{1}{\pi^m(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A})^m|\mathbf{\Sigma_C}|} \exp\left(-m\frac{\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_C}^{-1}\mathbf{k}_i^A}{\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A}}\right)}{\frac{1}{\pi^{Nm}(\mathbf{k_i}^{A\dagger}\hat{\mathbf{\Sigma}}_\mathbf{A}^{-1}\mathbf{k_i^A})^m|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N}\exp(-Nm)} \\
&\quad \cdot \frac{\prod_{i=1}^{P} \frac{1}{\pi^m(\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B})^m|\mathbf{\Sigma_C}|} \exp\left(-m\frac{-\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_C}^{-1}\mathbf{k}_i^B}{\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B}}\right)}{\frac{1}{\pi^{Pm}(\mathbf{k_i}^{B\dagger}\hat{\mathbf{\Sigma}}_\mathbf{B}^{-1}\mathbf{k_i^B})^m|\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}\exp(-Pm)} \\
&= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N|\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\mathbf{\Sigma_C}|^{N+P}} \exp\Big(-N\mathrm{tr}(\mathbf{\Sigma_C}^{-1}\hat{\mathbf{\Sigma}}_\mathbf{A}) - P\mathrm{tr}(\mathbf{\Sigma_C}^{-1}\hat{\mathbf{\Sigma}}_\mathbf{B}) + (N+P)m\Big) \\
&= \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N|\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\mathbf{\Sigma_C}|^{N+P}} \exp\Big(-\mathrm{tr}(\mathbf{\Sigma_C}^{-1}(N\hat{\mathbf{\Sigma}}_\mathbf{A} + P\hat{\mathbf{\Sigma}}_\mathbf{B})) + (N+P)m\Big)
\end{aligned} \tag{1.42}$$

Si on remplace alors $\mathbf{\Sigma_C}$ par son estimateur,

$$Q_3 = \frac{|\hat{\mathbf{\Sigma}}_\mathbf{A}|^N|\hat{\mathbf{\Sigma}}_\mathbf{B}|^P}{|\hat{\mathbf{\Sigma}_\mathbf{C}}|^{N+P}} \exp\Big(-\mathrm{tr}(\hat{\mathbf{\Sigma}_\mathbf{C}}^{-1}(N\hat{\mathbf{\Sigma}}_\mathbf{A} + P\hat{\mathbf{\Sigma}}_\mathbf{B})) + (N+P)m\Big) \tag{1.43}$$

L'expression du test est alors assez proche de celle du cas gaussien à l'exception de la dernière simplification qui n'est plus possible car $\hat{\mathbf{\Sigma}_\mathbf{C}} \neq N\hat{\mathbf{\Sigma}}_\mathbf{A} + P\hat{\mathbf{\Sigma}}_\mathbf{B}$.
   - *Rapport de vraisemblance Généralisé*, $\hat{\tau_i^A}|H_0 = \frac{\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_C}^{-1}\mathbf{k_i^A}}{m}$ et $\hat{\tau_i^B}|H_0 = \frac{\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_C}^{-1}\mathbf{k_i^B}}{m}$.

$$
\begin{aligned}
Q_3 &= \frac{\prod_{i=1}^{N} \frac{1}{\pi^m (\mathbf{k_i}^{A\dagger} \boldsymbol{\Sigma_C}^{-1} \mathbf{k_i^A})^m |\boldsymbol{\Sigma_C}|} \exp\left(-m \frac{\mathbf{k_i}^{A\dagger} \boldsymbol{\Sigma_C}^{-1} \mathbf{k}_i^A}{\mathbf{k_i}^{A\dagger} \boldsymbol{\Sigma_C}^{-1} \mathbf{k_i^A}}\right)}{\frac{1}{\pi^{Nm} (\mathbf{k_i}^{A\dagger} \hat{\boldsymbol{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A})^m |\hat{\boldsymbol{\Sigma}}_\mathbf{A}|^N} \exp(-Nm)} \\
&\quad \cdot \frac{\prod_{i=1}^{P} \frac{1}{\pi^m (\mathbf{k_i}^{B\dagger} \boldsymbol{\Sigma_C}^{-1} \mathbf{k_i^B})^m |\boldsymbol{\Sigma_C}|} \exp\left(-m \frac{-\mathbf{k_i}^{B\dagger} \boldsymbol{\Sigma_C}^{-1} \mathbf{k}_i^B}{\mathbf{k_i}^{B\dagger} \boldsymbol{\Sigma_C}^{-1} \mathbf{k_i^B}}\right)}{\frac{1}{\pi^{Pm} (\mathbf{k_i}^{B\dagger} \hat{\boldsymbol{\Sigma}}_\mathbf{B}^{-1} \mathbf{k_i^B})^m |\hat{\boldsymbol{\Sigma}}_\mathbf{B}|^P} \exp(-Pm)} \\
&= \frac{|\hat{\boldsymbol{\Sigma}}_\mathbf{A}|^N |\hat{\boldsymbol{\Sigma}}_\mathbf{B}|^P}{|\boldsymbol{\Sigma_C}|^{N+P}} \prod_{i=1}^{N} \frac{(\mathbf{k_i}^{A\dagger} \hat{\boldsymbol{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger} \boldsymbol{\Sigma_C}^{-1} \mathbf{k_i^A})^m} \prod_{i=1}^{P} \frac{(\mathbf{k_i}^{B\dagger} \hat{\boldsymbol{\Sigma}}_\mathbf{B}^{-1} \mathbf{k_i^B})^m}{(\mathbf{k_i}^{B\dagger} \boldsymbol{\Sigma_C}^{-1} \mathbf{k_i^B})^m}
\end{aligned}
\tag{1.44}
$$

Si on remplace alors $\boldsymbol{\Sigma_C}$ par son estimateur,

$$
Q_3 = \frac{|\hat{\boldsymbol{\Sigma}}_\mathbf{A}|^N |\hat{\boldsymbol{\Sigma}}_\mathbf{B}|^P}{|\hat{\boldsymbol{\Sigma}}_\mathbf{C}|^{N+P}} \prod_{i=1}^{N} \frac{(\mathbf{k_i}^{A\dagger} \hat{\boldsymbol{\Sigma}}_\mathbf{A}^{-1} \mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger} \hat{\boldsymbol{\Sigma}}_\mathbf{C}^{-1} \mathbf{k_i^A})^m} \prod_{i=1}^{P} \frac{(\mathbf{k_i}^{B\dagger} \hat{\boldsymbol{\Sigma}}_\mathbf{B}^{-1} \mathbf{k_i^B})^m}{(\mathbf{k_i}^{B\dagger} \hat{\boldsymbol{\Sigma}}_\mathbf{C}^{-1} \mathbf{k_i^B})^m}
\tag{1.45}
$$

Comme les deux autres, ce test est inédit et nécessite l'ensemble des données pour être calculé.

Par la suite, comme tous les paramètres sont estimés, nous ne distinguerons plus $\boldsymbol{\Sigma}$ et $\hat{\boldsymbol{\Sigma}}$ et présenterons uniquement les expressions des tests en fonction des matrices. Nous distinguerons uniquement l'estimateur qui sera différent, dans le cas des tests associés au modèle SIRV.

Tableau récapitulatif des tests statistiques dans le cas gaussien et dans le cas SIRV.

Dans le tableau 1.1, nous donnons un récapitulatif des différents tests proposés à l'aide du principe du rapport de vraisemblance. Dans le cas gaussien les estimateurs de $\boldsymbol{\Sigma_A}$, $\boldsymbol{\Sigma_B}$ et $\boldsymbol{\Sigma_C}$ sont les estimateurs classiques par moyenne statistique (SCM) pour les données k_i^A, k_i^B et $k_i^A \bigcup k_i^B$. Dans le cas SIRV ceux-ci correspondent à l'estimateur du point fixe appliqué aux observations respectives. Nous distinguons les tests Plug-in des rapports de vraisemblances (RVG).

Modèle	#	Expression	Sym	Nom
Gaussien	1	$\frac{\lvert\mathbf{\Sigma_A}\rvert^N}{\lvert\mathbf{\Sigma_B}\rvert^N}\exp\left(-N\mathrm{tr}(\mathbf{\Sigma_B}^{-1}\mathbf{\Sigma_A})+Nm\right)$	Non	Wishart
Gaussien	2	$\frac{\lvert\mathbf{\Sigma_A}\rvert^{N-P}}{\lvert\mathbf{\Sigma_B}\rvert^{N-P}}\exp\left(-N\mathrm{tr}(\mathbf{\Sigma_B}^{-1}\mathbf{\Sigma_A})-P\mathrm{tr}(\mathbf{\Sigma_A}^{-1}\mathbf{\Sigma_B})+(N+P)m\right)$	Oui	Wishart révisée ou symétrique
Gaussien	3	$\frac{\lvert\mathbf{\Sigma_A}\rvert^N\lvert\mathbf{\Sigma_B}\rvert^P}{\lvert\mathbf{\Sigma_C}\rvert^{N+P}}$	Oui	Bartlett
SIRV Plug-in	1	$\frac{\lvert\mathbf{\Sigma_A}\rvert^N}{\lvert\mathbf{\Sigma_B}\rvert^N}\exp\left(-N\mathrm{tr}(\mathbf{\Sigma_B}^{-1}\mathbf{\Sigma_A})+Nm\right)$	Non	Wishart+PF
SIRV Plug-in	2	$\frac{\lvert\mathbf{\Sigma_A}\rvert^{N-P}}{\lvert\mathbf{\Sigma_B}\rvert^{N-P}}\exp\left(-N\mathrm{tr}(\mathbf{\Sigma_B}^{-1}\mathbf{\Sigma_A})-P\mathrm{tr}(\mathbf{\Sigma_A}^{-1}\mathbf{\Sigma_B})+(N+P)m\right)$	Oui	Wishart révisée ou symétrique +PF
SIRV Plug-in	3	$\frac{\lvert\mathbf{\Sigma_A}\rvert^N\lvert\mathbf{\Sigma_B}\rvert^P}{\lvert\mathbf{\Sigma_C}\rvert^{N+P}}\exp\left(-\mathrm{tr}(\mathbf{\Sigma_C}^{-1}(N\mathbf{\Sigma_A}+P\mathbf{\Sigma_B}))+(N+P)m\right)$	Oui	?
SIRV RVG	1	$\frac{\lvert\mathbf{\Sigma_A}\rvert^N}{\lvert\mathbf{\Sigma_B}\rvert^N}\prod_{i=1}^{N}\frac{(\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k_i^A})^m}$	Non	?
SIRV RVG	2	$\frac{\lvert\mathbf{\Sigma_A}\rvert^{N-P}}{\lvert\mathbf{\Sigma_B}\rvert^{N-P}}\prod_{i=1}^{N}\frac{(\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k_i^A})^m}\prod_{1:P}\frac{(\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k_i^B})^m}{(\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k_i^B})^m}$	Oui	?
SIRV RVG	2	$\frac{\lvert\mathbf{\Sigma_A}\rvert^N\lvert\mathbf{\Sigma_B}\rvert^P}{\lvert\mathbf{\Sigma_C}\rvert^{N+P}}\prod_{i=1}^{N}\frac{(\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_A}^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}\mathbf{\Sigma_C}^{-1}\mathbf{k_i^A})^m}\prod_{1:P}\frac{(\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_B}^{-1}\mathbf{k_i^B})^m}{(\mathbf{k_i}^{B\dagger}\mathbf{\Sigma_C}^{-1}\mathbf{k_i^B})^m}$	Oui	?

TABLE 1.1 – Résumé des tests statistiques développés à l'aide du principe du rapport de vraisemblance

1.4.2 Effets liés aux nombres d'observations

Jusqu'à maintenant, nous avons étudié les tests dans leur forme la plus générale, et nous avons vu les nombres d'observations N et P de chaque famille intervenir régulièrement dans les expressions des tests statistiques.

La vraisemblance prend en compte la confiance qu'elle a dans les estimateurs des différents paramètres dans le calcul du test. Ainsi, plus le nombre d'observations est important, plus la vraisemblance est grande. Il n'y a donc pas de problème inhérent à l'utilisation du test, dans un cas simple, lorsqu'il s'agit de prendre une seule décision, si le seuil a correctement été calculé, théoriquement ou avec des données sous l'hypothèse H_0.

Un problème survient cependant lorsque ces tests statistiques sont utilisés comme des distances. Le plus souvent la décision n'est plus de choisir entre les hypothèses H_0 et H_1 mais plutôt de choisir, parmi plusieurs populations, laquelle est la plus proche de la population étudiée. Nous ne sommes alors plus amenés à comparer entre deux populations de N et de P observations, mais à choisir entre la population de P échantillons et celle de Q observations. Le problème vient alors de l'impact de ce nombre d'échantillons sur la mesure de distance polarimétrique. Dans la plupart des tests décrits plus haut, une asymétrie du nombre d'échantillons tend à biaiser la vraisemblance de l'hypothèse H_0.

Pour illustrer cet effet nous allons d'abord regarder un cas concret très simple. On fixe le nombre d'observations de la population A à $N = 30$. Le nombre d'observation des deux populations testées va, lui, varier de 9 à 125 : $P \in [9; 125]$. On mesure la distance de Bartlett (voir plus haut), d'abord entre la population A et une population tirée à partir de la même matrice de cohérence (H_0), avec une distribution gaussienne. Puis entre la même population A et une autre population tirée à l'aide d'une matrice de cohérence complètement différente (H_1). Le logarithme du rapport de vraisemblance en fonction du nombre d'échantillons P est tracé sur la figure 1.10.

Nous pouvons observer les choses suivantes :

1. Pour un nombre d'échantillons similaire, la distance de Bartlett permet de façon très claire de distinguer entre les deux hypothèses. La distance entre les deux familles d'observations qui suivent la même distribution est bien plus faible qu'entre deux familles qui suivent des distributions différentes.

2. Par contre, supposons que nous ayons plus d'échantillons disponibles pour une population qui suit une distribution différente, ne serait-ce que 20 observations de plus, nous mesurons alors une distance entre les deux populations qui sera bien plus faible qu'entre les deux populations qui suivraient pourtant une distribution identique ! Ici, avoir plus d'échantillons disponibles crée un biais qui entraîne une confusion entre les deux hypothèses.

En conclusion, *il n'est pas possible d'utiliser directement les tests issus de rapports de vraisemblances comme des outils de distance*. Lorsque ces tests sont utilisés comme outils dans des algorithmes de classification ou de segmentation, les effets liés à la taille des populations sont très importants. Dans une classification par exemple, on va mesurer la distance entre une population test, et plusieurs populations de référence. Si la taille des populations de référence varie fortement, les distances mesurées seront alors souvent plus liées aux nombres d'observations qu'à la signature polarimétrique. On observera ainsi

en pratique les populations sur-représentées dans les bases d'apprentissage l'emporter en moyenne sur les petites populations.

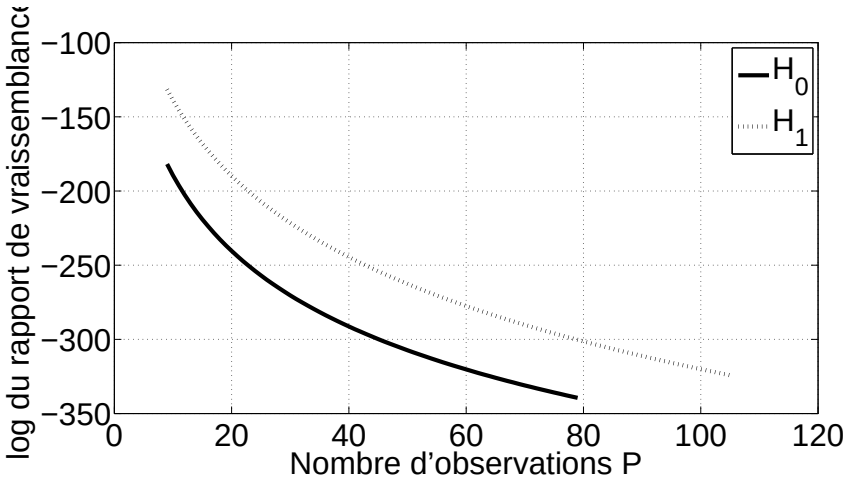


FIGURE 1.10 – Effet du nombre d'observations sur le calcul de distances polarimétriques (distance de Bartlett)

Nous pouvons observer cet effet sur la segmentation d'images polarimétriques. Ici aussi nous prenons une décision fondée sur une distance statistique, qui tient compte de la taille des différentes zones. Dans [35] et [36] le problème a été soulevé dans le cadre de la segmentation d'images optiques en couleur, et les problèmes liés aux nombres d'observations ont aussi été soulevés. Calderero appelle ainsi les distances qui prennent en compte les tailles des populations des distances "weighted", que l'on peut traduire par "pondérées". Une distance non "pondérée" correspond alors à l'expression du test statistique où N et P sont égaux. Nous pouvons même aller plus loin et remplacer alors N et P par 1 puisque dans le cas $N = P$, les expressions de ces tests sont égales à un facteur près.

Cette pondération des distances n'est toutefois pas toujours préjudiciable. C'est pourquoi ce problème a été peu, voire pas du tout, soulevé dans le cadre des applications aux images SAR polarimétriques. Cette pondération entraîne le plus souvent, dans le cadre d'une segmentation ou d'une classification, une fusion des classes ou des régions les plus petites. La plupart du temps, ces classes correspondent en fait à des petits groupes de pixels, souvent liés à des détails infimes, ou à des points brillants singuliers qu'il n'est pas toujours souhaitable de conserver comme une classe indépendante. Le problème peut toutefois survenir en fonction de la topologie de l'image que l'on étudie, et en fonction de ce que nous voulons conserver comme niveau de détail. Avec une distance pondérée nous n'avons aucun contrôle sur cette gestion des régions ou classes de petite taille. Les outils résultants produisent ainsi des partitions ou classifications de la scène en zones ayant globalement la même taille.

Nous pouvons ainsi exhiber facilement le type de cas où une distance pondérée peut entraîner des problèmes dans le cas d'une segmentation. Nous avons utilisé pour notre exemple une mire à la topologie particulière qui va mettre en relief l'impact des distances

pondérées. Sur la figure 1.11 à gauche nous présentons une segmentation de l'image réalisée à l'aide de la distance de Bartlett pondérée. Sur la droite nous avons la même image, segmentée avec un algorithme rigoureusement identique à l'exception de la distance de Bartlett qui a ici été utilisée dans sa version non pondérée. Les deux segmentations ont été appliquées jusqu'à atteindre le nombre exact de régions qui correspondent à la vérité connue de la segmentation "idéale".

Dans cet exemple, les carrés ont un comportement polarimétrique très contrasté avec la zone de fond, cependant la distance pondérée cherche à produire une partition en zones de tailles homogènes, ce qui ne correspond pas à la topologie de l'image. Ainsi on peut observer sur l'image de gauche que le fond a été fusionné par endroits de façon erronée avec des carrés ayant des comportements polarimétriques très différents. Inversement, la zone de fond, très grande, a été segmentée en plusieurs zones, bien que les signatures polarimétriques de ces sous-zones soient parfaitement identiques. A l'aide de la même distance non pondérée, à droite, tous ces problèmes disparaissent.

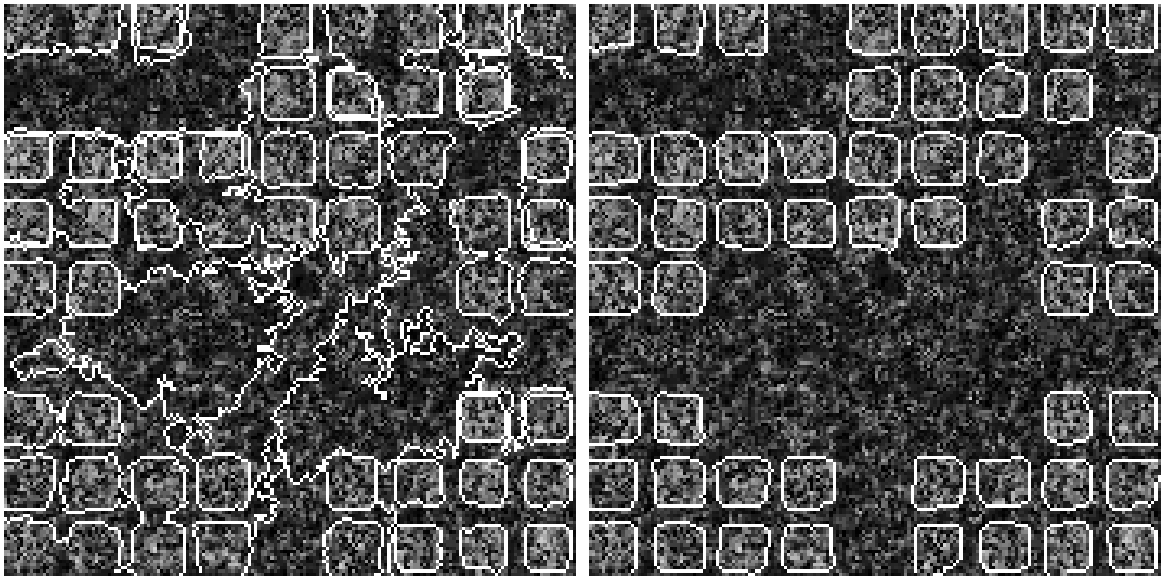


FIGURE 1.11 – Illustration des effets liés à la pondération des distances polarimétriques sur des résultats de segmentation d'image synthétique. La topologie de l'image est choisie pour mettre en relief les problèmes induits dans l'utilisation d'une distance pondérée (à gauche). Ici le contraste polarimétrique est très élevé et toutes les zones suivent des distributions parfaitement gaussiennes.

Nous pouvons donc conclure que l'utilisation des distances statistiques doit être faite à l'aide des distances non pondérées. Le problème des petits détails et des points aberrants reste toutefois à résoudre dans ce cas-là. Calderero propose une méthode dans ses articles. Nous proposerons une autre solution applicable dans le cadre de la segmentation d'images polarimétriques dans les sections à venir.

1.4.3 Distances issues du monde de la théorie de l'information

La seconde approche, régulièrement employée en traitement d'images, consiste à utiliser une des mesures classiques de distance entre des distributions provenant de la théorie de l'information. Ces mesures classiques ont été introduites il y a des dizaines d'années. Elles font toujours l'object d'études théoriques et appliquées. Les études théoriques concernent surtout les comportements asymptotiques qui nous intéressent ici assez peu. Les études appliquées notamment sur des images classiques optiques couleurs s'avèrent plus proches de nos besoins. Nous remarquons ici encore le travail de Calderero [35] [36]. En polarimétrie ces méthodes ont été étudiée par Jérome Morio, François Goudail et Philippe Réfrégier dans, [37], [38] et [39]. Le choix parmi ces différentes distances sera pour nous encore une fois plus motivé par les performances simulées dans un contexte similaire à celui de notre étude (nombre d'observations restreint, faible dimension, fouillis gaussien ou modérément non gaussien). Nous étudierons aussi les performances sur des données réelles. Parmi ces très nombreuses distances, quatre distances sont couramment utilisées dans des applications proches des nôtres :

1. *la mesure de Kullback Leibler :* c'est une mesure asymétrique qui s'exprime en fonction des deux probabilités $P_A(X)$ et $P_B(X)$ [40, 41]. Elle est aussi parfois nommée Information de Kullback Leibler. Nous la noterons $I(\mathbf{A}, \mathbf{B})$ et elle est définie par :

$$I(\mathbf{A}, \mathbf{B}) = \int P_A(X).\log\left(\frac{P_A(X)}{P_B(X)}\right) dX. \tag{1.46}$$

   Ici X correspond à la variable mesurée d'intérêt, donc dans notre cas il s'agira, comme toujours, du vecteur de diffusion $\mathbf{k}$.

2. *La divergence de Kullback Leibler* : il s'agit de la distance symétrique, issue de la mesure de Kullback Leibler [40, 41]. Elle est parfois nommé Entropie Croisée. Nous la noterons $J(\mathbf{A}, \mathbf{B})$

$$J(\mathbf{A}, \mathbf{B}) = \int (P_A(X) - P_B(X)).\log\left(\frac{P_A(X)}{P_B(X)}\right) dX. \tag{1.47}$$

3. *La distance de Chernoff* est définie par :

$$C(\mathbf{A}, \mathbf{B}) = \max_{0<t<1} \Big(-\log\Big[\int P_A(X)^t P_B(X)^{1-t} dX\Big].\Big) \tag{1.48}$$

4. *La distance de Bhattacharyya :* c'est une distance symétrique [42], issue de la distance de Chernoff [43]. Le calcul du maximum en fonction de t est parfois coûteux en temps, la distance de Bhattacharyya est en fait une version simplifiée qui consiste à prendre $t = 1/2$ et on a alors par définition :

$$B(\mathbf{A}, \mathbf{B}) = -\log\Big[\int P_A(X)^{1/2} P_B(X)^{1/2} dX\Big]. \tag{1.49}$$

Comme pour les rapports de vraisemblance, nous allons d'abord exprimer et retrouver les expressions de ces distances dans le cas d'une distribution de fouillis gaussien, avec des démonstrations plus simples que la plupart qui sont présentées classiquement. Dans un second temps nous calculerons ces distances dans le cas de distribution de fouillis SIRV.

Cas Gaussien

Avant de commencer nous rappelons des propriétés sur les densités de probabilité qui nous serviront par la suite. $P_{\mathbf{A}}(\mathbf{k})$, correspond à une distribution gaussienne de matrice de covariance $\mathbf{A}$. On a donc, par définition d'une densité de probabilité :

$$\int P_{\mathbf{A}}(\mathbf{k})\mathbf{dk} = 1. \tag{1.50}$$

Par définition de la matrice de covariance/cohérence :

$$\int \mathbf{k}\mathbf{k}^{\dagger}\ P_{\mathbf{A}}(\mathbf{k})\mathbf{dk} = \mathbb{E}_{\mathbf{A}}(\mathbf{k}\mathbf{k}^{\dagger}) = \mathbf{A}. \tag{1.51}$$

Dans l'ensemble des calculs suivants nous chercherons à faire apparaître ces expressions afin d'effectuer des simplifications à partir des expressions sous forme d'intégrales.

1. *la mesure de Kullback Leibler*, nous remplaçons pour commencer uniquement les densités de probabilité dans le logarithme,

$$\begin{aligned}
I(\mathbf{A},\mathbf{B}) &= \int P_{\mathbf{A}}(\mathbf{k}).\log\left(\frac{P_{\mathbf{A}}(\mathbf{k})}{P_{\mathbf{B}}(\mathbf{k})}\right)\mathbf{dk} \\
&= \int P_{\mathbf{A}}(\mathbf{k}).\log\Big(\frac{\frac{1}{\pi^m|\mathbf{A}|}\exp\left(-\mathbf{k}^{\dagger}\mathbf{A}^{-1}\mathbf{k}\right)}{\frac{1}{\pi^m|\mathbf{B}|}\exp\left(-\mathbf{k}^{\dagger}\mathbf{B}^{-1}\mathbf{k}\right)}\Big)\mathbf{dk} \\
&= \int P_{\mathbf{A}}(\mathbf{k}).\Big(\log\left(\frac{|\mathbf{B}|}{|\mathbf{A}|}\right) + \mathbf{k}^{\dagger}\mathbf{B}^{-1}\mathbf{k} - \mathbf{k}^{\dagger}\mathbf{A}^{-1}\mathbf{k}\Big)\mathbf{dk} \\
&= \log\left(\frac{|\mathbf{B}|}{|\mathbf{A}|}\right)\int P_{\mathbf{A}}(\mathbf{k})\mathbf{dk} + \int Tr\Big(\mathbf{B}^{-1}\mathbf{k}\mathbf{k}^{\dagger} - \mathbf{A}^{-1}\mathbf{k}\mathbf{k}^{\dagger}\Big)P_{\mathbf{A}}(\mathbf{k})\mathbf{dk} \\
&= \log\left(\frac{|\mathbf{B}|}{|\mathbf{A}|}\right) + \mathrm{tr}\Big((\mathbf{B}^{-1} - \mathbf{A}^{-1})\int \mathbf{k}\mathbf{k}^{\dagger}P_{\mathbf{A}}(\mathbf{k})\mathbf{dk}\Big) \\
&= \log\left(\frac{|\mathbf{B}|}{|\mathbf{A}|}\right) + \mathrm{tr}\Big(\mathbf{B}^{-1}\mathbf{A} - \mathbf{I}\Big).
\end{aligned} \tag{1.52}$$

L'expression obtenue est exactement similaire à celle que nous avions trouvée pour le rapport de vraisemblance dans lequel nous avions introduit un seul paramètre $\mathbf{\Sigma_B}$. Dans le cas gaussien, la mesure de Kullback Leibler est donc équivalente à la distance de Wishart. Nous noterons qu'ici, nous avons par contre une expression qui ne dépend pas d'estimateurs de matrices mais que ceux-ci seront de toute façon nécessaires, exactement comme dans le cas du rapport de vraisemblance. Par contre, nous constatons aussi que le nombre d'échantillons dans les populations n'a pas d'impact ici. Pour obtenir exactement les mêmes expressions que dans les rapports de vraisemblance, nous devrions étudier la mesure de Kullback Leibler entre deux distributions de Wishart et utiliser la densité de probabilité de la loi de la matrice et non du vecteur. La démonstration est par contre beaucoup moins simple et aboutit, à l'exception de l'impact du nombre d'échantillons, à la même expression.

2. *La divergence de Kullback Leibler :* elle s'exprime de la manière suivante :

$$J(\mathbf{A},\mathbf{B}) = I(\mathbf{A},\mathbf{B}) + I(\mathbf{B},\mathbf{A}) = \mathrm{tr}\Big(\mathbf{B}^{-1}\mathbf{A} + \mathbf{A}^{-1}\mathbf{B} - 2\mathbf{I}\Big). \tag{1.53}$$

De la même manière on retrouve une expression qui est similaire à un test issu du rapport de vraisemblance dans lequel nous avions introduit deux paramètres A et B. Cette divergence est identique à la distance de Wishart révisée. On retrouverait aussi la même expression, bien que plus difficilement, à partir des densités matricielles. Celles-ci dépendraient aussi, comme pour le rapport de vraisemblance, du nombre d'observations.

3. *La distance de Bhattacharyya :* le calcul est légèrement plus complexe que pour les deux autres. Nous commençons par écrire :

$$\begin{aligned} B(\mathbf{A},\mathbf{B}) &= -\log\Big[\int P_{\mathbf{A}}(\mathbf{k})^{1/2} P_{\mathbf{B}}(\mathbf{k})^{1/2}\mathbf{dk}\Big], \\ &= -\log\Big[\int \frac{1}{\pi^{\frac{m}{2}}|\mathbf{A}|^{\frac{1}{2}}}\exp\left(-\frac{\mathbf{k}^\dagger\mathbf{A}^{-1}\mathbf{k}}{2}\right)\frac{1}{\pi^{\frac{m}{2}}|\mathbf{B}|^{\frac{1}{2}}}\exp\left(-\frac{\mathbf{k}^\dagger\mathbf{B}^{-1}\mathbf{k}}{2}\right)\mathbf{dk}\Big], \\ &= -\log\Big[\frac{1}{|\mathbf{A}|^{\frac{1}{2}}|\mathbf{B}|^{\frac{1}{2}}}\int\frac{1}{\pi^m}\exp\left(-\mathbf{k}^\dagger\frac{(\mathbf{B}^{-1}+\mathbf{A}^{-1})}{2}\mathbf{k}\right)\mathbf{dk}\Big], \end{aligned} \tag{1.54}$$

Il convient alors de poser un troisième paramètre intermédiaire, défini comme $\mathbf{C} = \Big(\frac{(\mathbf{B^{-1}}+\mathbf{A^{-1}})}{\mathbf{2}}\Big)^{\mathbf{-1}}$,

$$\begin{aligned} B(\mathbf{A},\mathbf{B}) &= -\log\Big[\frac{1}{|\mathbf{A}|^{\frac{1}{2}}|\mathbf{B}|^{\frac{1}{2}}}\int\frac{|\mathbf{C}|}{\pi^m|\mathbf{C}|}\exp\big(-\mathbf{k}^\dagger C^{-1}\mathbf{k}\big)\mathbf{dk}\Big], \\ &= -\log\Big[\frac{|\mathbf{C}|}{|\mathbf{A}|^{\frac{1}{2}}|\mathbf{B}|^{\frac{1}{2}}}\int P_{\mathbf{C}}(\mathbf{k})\mathbf{dk}\Big], \\ &= -\log\Big[\frac{|\mathbf{A}|^{\frac{1}{2}}|\mathbf{B}|^{\frac{1}{2}}}{|\frac{(\mathbf{A}^{-1}+\mathbf{B}^{-1})}{2}||\mathbf{A}||\mathbf{B}|}\Big], \\ &= -\log\Big[\frac{|\mathbf{A}|^{\frac{1}{2}}|\mathbf{B}|^{\frac{1}{2}}}{|\frac{(\mathbf{A}+\mathbf{B})}{2}|}\Big]. \end{aligned} \tag{1.55}$$

La distance de Bhattacharyya correspond, elle, au rapport de vraisemblance utilisant un paramètre $\mathbf{\Sigma_C}$ dans le cas gaussien. Ce n'est pas surprenant puisque ce paramètre est justement introduit pendant son calcul.

4. *La distance de Chernoff* : nous pouvons, par un calcul similaire trouver :

$$C(\mathbf{A},\mathbf{B}) = \max_{0<t<1} -\log\Big[\frac{|\mathbf{A}|^{1-t}|\mathbf{B}|^t}{|(t\mathbf{A}+(1-t)\mathbf{B})|}\Big]. \tag{1.56}$$

Dans le cas SIRV

Dans le cas SIRV nous rappelons que nous décomposons le vecteur $\mathbf{k}$ sous la forme $\hat{\mathbf{k}} = \sqrt{\tau}\,\hat{\mathbf{z}}$, donc par définition de la matrice de covariance/cohérence nous pourrons utiliser la propriété suivante :

$$\int \frac{\mathbf{k}\mathbf{k}^\dagger}{\tau}\; P_{\mathbf{A}}(\mathbf{k})\mathbf{dk} = \mathbb{E}_{\mathbf{A}}(\hat{\mathbf{z}}\hat{\mathbf{z}}^\dagger) = \mathbf{A}. \tag{1.57}$$

Nous allons refaire rapidement les démonstrations dans le cas d'un vecteur suivant une loi SIRV "déterministe", c'est à dire où la texture τ est supposée déterministe. Nous allons voir que nous retrouvons ici les mêmes expressions pour les tests que dans le cas gaussien.

1. *la mesure de Kullback Leibler :* de la même manière que dans le cas gaussien, on écrit,

$$\begin{aligned} I(\mathbf{A},\mathbf{B}) &= \int P_{\mathbf{A}}(\mathbf{k}).\log\left(\frac{P_{\mathbf{A}}(\mathbf{k})}{P_{\mathbf{B}}(\mathbf{k})}\right)\mathbf{dk}, \\ &= \int P_{\mathbf{A}}(\mathbf{k}).\log\Big(\frac{\frac{1}{\pi^m\tau^m|\mathbf{A}|}\exp\left(-\frac{\mathbf{k}^\dagger\mathbf{A}^{-1}\mathbf{k}}{\tau}\right)}{\frac{1}{\pi^m\tau^m|\mathbf{B}|}\exp\left(-\frac{\mathbf{k}^\dagger\mathbf{B}^{-1}\mathbf{k}}{\tau}\right)}\Big)\mathbf{dk} \\ &= \int P_{\mathbf{A}}(\mathbf{k}).\Big(\log(\frac{|\mathbf{B}|}{|\mathbf{A}|}) + \frac{\mathbf{k}^\dagger\mathbf{B}^{-1}\mathbf{k} - \mathbf{k}^\dagger\mathbf{A}^{-1}\mathbf{k}}{\tau}\Big)\mathbf{dk} \\ &= \log(\frac{|\mathbf{B}|}{|\mathbf{A}|}) + \mathrm{tr}\Big((\mathbf{B}^{-1}-\mathbf{A}^{-1})\int\frac{\mathbf{k}^\dagger\mathbf{k}}{\tau}P_{\mathbf{A}}(\mathbf{k})\mathbf{dk}\Big) \\ &= \log(\frac{|\mathbf{B}|}{|\mathbf{A}|}) + \mathrm{tr}\Big(\mathbf{B}^{-1}\mathbf{A}-\mathbf{I}\Big). \end{aligned} \tag{1.58}$$

2. *La divergence de Kullback Leibler :* de même qu'en gaussien, on a

$$J(\mathbf{A},\mathbf{B}) = I(\mathbf{A},\mathbf{B}) + I(\mathbf{B},\mathbf{A}) = \mathrm{tr}\Big(\mathbf{B}^{-1}\mathbf{A} + \mathbf{A}^{-1}\mathbf{B} - 2\mathbf{I}\Big). \tag{1.59}$$

3. *La distance de Bhattacharyya :* comme pour le cas gaussien, on fait apparaître un paramètre intermédiaire $\mathbf{C}$.

$$\begin{aligned} B(\mathbf{A},\mathbf{B}) &= -\log\Big[\int P_{\mathbf{A}}(\mathbf{k})^{1/2}P_{\mathbf{B}}(\mathbf{k})^{1/2}\mathbf{dk}\Big], \\ &= -\log\Big[\frac{1}{|\mathbf{A}|^{\frac{1}{2}}|\mathbf{B}|^{\frac{1}{2}}}\int\frac{1}{\pi^m\tau^m}\exp\left(-\mathbf{k}^\dagger\frac{(\mathbf{B}^{-1}+\mathbf{A}^{-1})}{2\tau}\mathbf{k}\right)\mathbf{dk}\Big], \\ &= -\log\Big[\frac{|\mathbf{A}|^{\frac{1}{2}}|\mathbf{B}|^{\frac{1}{2}}}{|\frac{(\mathbf{A}+\mathbf{B})}{2}|}\Big]. \end{aligned} \tag{1.60}$$

4. *La distance de Chernoff :* de la même manière, il est possible de montrer que l'expression de la distance de Chernoff est la même que dans le cas de distributions gaussiennes.

Nom référence	Acronyme	Expression
Kullback Leibler Mesure	KLM	$\log(\frac{\lvert\Sigma_{\mathbf{B}}\rvert}{\lvert\Sigma_{\mathbf{A}}\rvert}) + \mathrm{tr}(\Sigma_{\mathbf{B}}^{-1}\Sigma_{\mathbf{A}}) - m$
Kullback Leibler Divergence	KLD	$\mathrm{tr}(\Sigma_{\mathbf{B}}^{-1}\Sigma_{\mathbf{A}} + \Sigma_{\mathbf{A}}^{-1}\Sigma_{\mathbf{B}}) - 2m$
Chernoff Distance	Che	$\max_{0<t<1} -\log\Big[\frac{\lvert\Sigma_{\mathbf{A}}\rvert^{1-t}\lvert\Sigma_{\mathbf{B}}\rvert^{t}}{\lvert(t\Sigma_{\mathbf{A}} + (1-t)\Sigma_{\mathbf{B}})\rvert}\Big]$
Bhattacharyya Distance	Bha	$-\log\Big[\frac{\lvert\Sigma_{\mathbf{A}}\rvert^{1/2}\lvert\Sigma_{\mathbf{B}}\rvert^{1/2}}{\lvert\frac{N\Sigma_{\mathbf{A}}+P\Sigma_{\mathbf{B}}}{N+P}\rvert}\Big]$

TABLE 1.2 – Expressions des tests ou distances dans le cadre d'une modélisation gaussienne du fouillis.

Les quatres distances sur lesquelles nous nous sommes concentrés sont, à l'exception de la distance de Chernoff, équivalentes à des tests statistiques issus de rapports de vraisemblance. Ce résultat est assez confortant et laisse présager, a priori, de bonnes performances pour ces tests. Il convient cependant de faire les remarques suivantes :

- Contrairement aux distances issues de rapport de vraisemblance, les distances issues de la théorie de l'information sont naturellement indépendantes du nombre d'observations. Nous venions de voir en quoi le nombre d'observations pouvait influencer négativement les performances dans certains types d'applications. Nous n'aurons pas ce problème ici.
- Dans le cas de distributions SIRV, les distances sont rigoureusement identiques et seuls les estimateurs de matrices utilisés seront différents. On observe donc une divergence entre les tests issus du rapport de vraisemblance et ces distances. Pour la distance de Wishart et la distance de Wishart Révisée, qui correspondent respectivement à la Mesure de Kullback Leibler et la Divergence de Kullback Leibler, on conserve la même expression dans le cas SIRV. C'est entre la distance de Bartlett et de Bhattacharyya, identiques dans le cas gaussien, qu'on obtient une expression différente.

Compte tenu de la ressemblance très importante des distances issues des deux mondes, nous n'aurons à étudier qu'un nombre restreint de distances. Nous allons récapituler les distances dont nous allons étudier les performances ainsi que les noms auxquels nous ferons référence par la suite.

Bien évidemment nous testons ces distances à la fois dans des fouillis gaussiens (table 1.2) mais aussi des fouillis non gaussiens (table 1.3). Dans le premier cas $\mathbf{\Sigma_A}$ et $\mathbf{\Sigma_B}$ correspondent aux estimées des matrices de covariance ou de cohérence à l'aide de l'estimateur

classique, souvent appelé SCM. Ainsi $\mathbf{\Sigma} = \sum_{\mathbf{n=1}}^{\mathbf{M}} \mathbf{k_n}\hat{\mathbf{k}}_\mathbf{n}^\dagger$. Nous ne faisons pas intervenir $\mathbf{\Sigma_C}$ dans la mesure où ce paramètre s'exprime en fonction des deux précédents. Cela ne sera plus valable dans le cas non gaussien. Nous préférons utiliser autant que possible les noms associés aux distances provenant de la théorie de l'information puisque les noms associés aux rapports de vraisemblance seront utilisés dans le cas non gaussien, où les expressions diffèrent.

Dans le tableau 1.3, $\Sigma_\mathbf{A}$ et $\Sigma_\mathbf{B}$ correspondent aux estimées des matrices de covariance ou de cohérence à l'aide de l'estimateur du point fixe. Ainsi $\Sigma = \sum_{n=1}^{M} \frac{\mathbf{k_n}\hat{\mathbf{k}}_\mathbf{n}^\dagger}{\mathbf{k}^\dagger\Sigma^{-1}\mathbf{k}}$. Ici $\Sigma_\mathbf{C}$ correspond à l'estimée de la matrice de covariance sur l'ensemble des deux populations A et B. Par souci de cohérence, les distances quasi identiques aux distances gaussiennes, à l'exception du processus d'estimation, gardent le même nom (avec un + PF).

1.4.4 Comparaison des performances sur données simulées

Les distances construites dans la section précédente ne présentent aucune garantie d'optimalité puisqu'il ne s'agit que de tests approchant le rapport de vraisemblance pour le test d'hypothèse simple où tous les paramètres seraient connus. Seule l'étude du comportement de ces tests sur des données nous permet d'étudier leurs performances respectives dans un cadre applicatif. On s'attend à avoir des performances assez similaires pour toutes ces distances qui approchent toutes plus ou moins le même rapport de vraisemblance.

Dans cette section on va s'intéresser aux performances sur des tirages de données suivant la méthode de Monte Carlo. Nous effectuerons donc des tirages de données gaussiennes, mais aussi de données non gaussiennes pour étudier les performances relatives des outils SIRV et des outils Gaussiens sur des données ayant différents degrés d'impulsionnalité. Pour simuler l'effet d'un fouillis non gaussien, nous utiliserons simplement une loi de distribution K. Nous ajusterons le paramètre de forme ν pour faire évoluer progressivement la distribution à partir de la distribution gaussienne.

Protocole de simulation

Afin d'étudier les performances de ces différents tests nous procédons de la manière suivante : l'ensemble du problème est posé de la même manière que le test d'hypothèse qui a servi à construire ces tests statistiques. L'hypothèse H_0 correspond à l'hypothèse '*les deux groupes d'échantillons suivent la même distribution*' et l'hypothèse H_1 '*les deux groupes d'échantillons ne suivent pas la même distribution*'.

1. Nous choisissons deux matrices distinctes $\mathbf{M_A}$ et $\mathbf{M_B}$ que nous distinguerons de $\mathbf{\Sigma_A}$ et $\mathbf{\Sigma_B}$ parce qu'elles correspondent aux 'vraies' matrices de covariance.
2. Pour calculer le test sous l'hypothèse H_0 nous tirons deux groupes d'observations, deux familles d'échantillons, qui auront tous deux la même matrice de covariance $\mathbf{M_A}$, et éventuellement des réalisations de texture totalement indépendantes pour chacun des échantillons. Ainsi, nous allons pouvoir appliquer l'ensemble de ces tests sur ces deux familles et calculer la statistique de ces tests sous l'hypothèse H_0.

Nom référence	Acronyme	Expression
Kullback Leibler Mesure et Point Fixe	KLM+FP	$\log(\frac{\lvert\Sigma_{\mathbf{B}}\rvert}{\lvert\Sigma_{\mathbf{A}}\rvert}) + \mathrm{tr}(\Sigma_{\mathbf{B}}^{-1}\Sigma_{\mathbf{A}}) - m$
Kullback Leibler Divergence et Point Fixe	KLD+FP	$\mathrm{tr}(\Sigma_{\mathbf{B}}^{-1}\Sigma_{\mathbf{A}} + \Sigma_{\mathbf{A}}^{-1}\Sigma_{\mathbf{B}}) - 2m$
Chernoff Distance et Point Fixe	Che+FP	$\max_{0<t<1} -\log\Big[\frac{\lvert\Sigma_{\mathbf{A}}\rvert^{1-t}\lvert\Sigma_{\mathbf{B}}\rvert^{t}}{\lvert(t\Sigma_{\mathbf{A}} + (1-t)\Sigma_{\mathbf{B}})\rvert}\Big]$
Bhattacharyya Distance et Point Fixe	Bha+FP	$-\log\Big[\frac{\lvert\Sigma_{\mathbf{A}}\rvert^{1/2}\lvert\Sigma_{\mathbf{B}}\rvert^{1/2}}{\lvert\frac{N\Sigma_{\mathbf{A}}+P\Sigma_{\mathbf{B}}}{N+P}\rvert}\Big]$
Bartlett SIRV Distance	BS	$-\log\Big[\frac{\lvert\Sigma_{\mathbf{A}}\rvert^{1/2}\lvert\Sigma_{\mathbf{B}}\rvert^{1/2}}{\lvert\Sigma_{\mathbf{C}}\rvert}\Big] + \mathrm{tr}(\Sigma_{\mathbf{C}}^{-1}(\frac{N\Sigma_{\mathbf{A}}+P\Sigma_{\mathbf{B}}}{N+P})) - m$
Rapport de Vraisemblance Généralisé 1	RVG1	$\frac{\lvert\Sigma_A\rvert}{\lvert\Sigma_B\rvert}\prod_{i=1}^{N}\frac{(\mathbf{k_i}^{A\dagger}\Sigma_A^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}\Sigma_B^{-1}\mathbf{k_i^A})^m}$
Rapport de Vraisemblance Généralisé 2	RVG2	$\prod_{i=1}^{N}\frac{(\mathbf{k_i}^{A\dagger}\Sigma_A^{-1}\mathbf{k_i^A})^m}{(\mathbf{k_i}^{A\dagger}\Sigma_B^{-1}\hat{\mathbf{k}}_\mathbf{i}^\mathbf{A})^m}\prod_{i=1}^{P}\frac{(\mathbf{k_i}^{B\dagger}\Sigma_B^{-1}\mathbf{k_i^B})^m}{(\mathbf{k_i}^{B\dagger}\Sigma_A^{-1}\mathbf{k_i^B})^m}$
Rapport de Vraisemblance Généralisé 3	RVG3	$\frac{\lvert\Sigma_A\rvert\lvert\Sigma_B\rvert}{\lvert\Sigma_C\rvert^2}\prod_{i=1}^{N}\frac{(\mathbf{k_i}^{A\dagger}\Sigma_A^{-1}\mathbf{k_i^A})^m}{(\hat{\mathbf{k}}_\mathbf{i}^{A\dagger}\Sigma_C^{-1}\mathbf{k_i^A})^m}\prod_{i=1}^{P}\frac{(\mathbf{k_i}^{B\dagger}\Sigma_B^{-1}\mathbf{k_i^B})^m}{(\mathbf{k_i}^{B\dagger}\Sigma_C^{-1}\mathbf{k_i^B})^m}$

TABLE 1.3 – Expressions des tests ou distances dans le cadre d'une modélisation gaussienne du fouillis non gaussien.

3. Pour calculer le test sous l'hypothèse H_1 nous tirons deux groupes d'observations, deux familles d'échantillons, qui auront cette fois pour matrice de covariance M_A et M_B respectivement, et encore une fois des réalisations de texture totalement indépendantes pour chacun des échantillons. Nous appliquons aussi l'ensemble des tests statistiques sur ces deux familles afin de générer la statistique sous l'hypothèse H_1.
4. La fausse alarme est définie sous l'hypothèse H_0 et correspond à une décision amenant à conclure que les deux familles ont des distributions différentes alors que ce n'est pas le cas. En pratique nous fixons une probabilité de fausse alarme assez faible ($P_{FA} = 10^{-3}$ par exemple), et déterminons un seuil, sous l'hypothèse H_0 pour que cette probabilité de fausse alarme soit effectivement mesurée sur les données simulées.
5. La probabilité de bonne détection sous l'hypothèse H_1 correspond à la conclusion que les deux familles ont des distributions différentes quand c'est effectivement le cas. Cette probabilité est mesurée à l'aide du seuil préalablement calculé pour garantir la probabilité de fausse alarme sous l'hypothèse H_0.
6. Les courbes présentées par la suite ont été obtenues à l'aide de $3 * 10^6$ tirages pour chaque couples de matrices de cohérence. 500 couples de matrices de cohérence ont été utilisé pour obtenir des courbes moyennes. Ainsi pour chaque position de l'axe des abscisses des figures suivante $1.5 * 10^9$ tirages ont été effectués.

Le choix des paramètres, notamment des matrices de covariance $\mathbf{M_A}$ et $\mathbf{M_B}$ mais aussi la probabilité de fausse alarme et enfin le nombre d'observations dans chacune des familles, a une grande importance. En effet si on choisit des matrices très différentes, alors il sera très facile de distinguer l'hypothèse H_0 de l'hypothèse H_1 et tous les tests statistiques donneront des probabilités de bonne décision égales à 1. Inversement, si les matrices sont trop proches, on risque de voir les tests uniformément incapables de les distinguer. L'objectif est alors d'étudier les performances dans un domaine où les tests ont tous des performances moyennes, où la bonne décision est donc possible mais non triviale. Nous allons donc étudier les performances des tests sur des plages de paramètres qui contiennent ces domaines pertinents pour la comparaison.

1. Les performances des tests dépendent du choix de $\mathbf{M_A}$ et $\mathbf{M_B}$. Afin d'étudier des performances moyennes indépendantes de ce choix, nous allons calculer des performances sur un très grand nombre de couples de matrices de covariance tirées aléatoirement. Pour un couple de matrices donné, on s'intéressera aussi aux performances lorsque les deux matrices sont de plus en plus différentes. Ainsi on étudiera la matrice pondérée $t\mathbf{M_A} + (1-t)\mathbf{M_B}$ avec $t \in [0, 1]$. Pour $t = 1$ il est impossible de distinguer les deux matrices. On doit alors normalement retrouver une probabilité de bonne détection égale à la probabilité de fausse alarme choisie. Pour $t = 0$ la différence entre les deux matrices utilisées pour générer les échantillons est maximum. Ici t joue le même rôle qu'un rapport signal à bruit dans le cadre de la détection.
2. Les nombres d'observations disponibles, que nous avions appelés N et P, ont eux aussi un effet sur les performances des tests statistiques. Nous aurons par la suite, une section consacrée à l'influence d'un nombre différent d'observations pour les

Nom référence	$\nu = +\infty$	$\nu = 10$	$\nu = 7$	$\nu = 4$	$\nu = 3$	$\nu = 2$	$\nu = 1$	$\nu = 0.5$
KLM	0.50	0.45	0.43	0.38	0.30	0.24	0.10	0.01
KLD	0.748	0.713	0.712	0.699	0.663	0.613	0.497	0.374
Che	0.41	0.36	0.35	0.34	0.29	0.25	0.19	0.12
Bha	0.746	0.711	0.709	0.697	0.660	0.608	0.487	0.374
KLM+FP	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
KLD+FP	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56
Che+FP	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
Bha+FP	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54
BS	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54
RVG1	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34
RVG2	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
RVG3	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56

TABLE 1.4 – Tableau récapitulatif des performances (probabilité de bonne décision) en fonction du paramètre de forme du fouillis K distribué ν, pour $P_{FA} = 10^{-3}$ et $N_{sample} = 50$.

deux populations. Ici nous nous restreindrons au cas $N = P$. Nous allons toutefois faire évoluer ce paramètre afin d'étudier les variations de performances de ces tests quand le nombre d'échantillon est petit ou grand. L'ordre de grandeur est choisi en fonction de la gamme des tailles de fenêtres de moyennage couramment utilisée. Ainsi nous ferons varier le nombre d'échantillon de 9 (3× 3) à 81 (9×9).

3. Enfin, la fausse alarme, la probabilité que nous fixons a priori, de supposer que deux familles d'échantillons ont des distributions différentes alors que ce n'est pas le cas, détermine le calcul du seuil. Ce seuil permet de calculer la probabilité de detection. Nous étudierons donc la relation entre la probabilité de fausse alarme et la probabilité de détection.

Impact de la distribution

Nous nous intéressons ici à l'évolution des performances des différents tests ou distances statistiques en fonction du caractère impulsionnel du fouillis. Nous nous plaçons dans une configuration moyenne, une probabilité de fausse alarme de $P_{FA} = 10^{-3}$ et un nombre d'échantillons $N_{sample} = 50$. Nous étudions en utilisant le protocole que nous venons de présenter les performances des distances pour un fouillis K-distribué avec différentes valeurs du paramètre de forme ν. Pour rappel un paramètre de forme $\nu = +\infty$ correspond à une distribution gaussienne. Plus ce paramètre est petit, et plus la distribution est non gaussienne.

L'ensemble des performances sont présentées sur le tableau 1.4.

Nous ne discuterons en détail des performances relatives de chacune des distances qu'après avoir montré l'ensemble des courbes de performance. Cela dit, ce tableau nous permet de regarder l'évolution de ces distances en fonction du paramètre ν. Si on observe par exemple les performances de la divergence de Kullback Leibler, avec modèle gaussien

et modèle SIRV, on se rend compte qu'il faut atteindre un degré d'impulsionnalité très important $\nu = 1$ pour exhiber de meilleures performances avec le modèle SIRV. Et encore, ce gain de performance ne devient important que pour une distribution encore nettement plus impulsionnelle avec $\nu = 0.5$.

Pour nos courbes de performances nous choisissons alors d'étudier à la fois le cas gaussien, $\nu = +\infty$ et un cas très fortement impusionnel $\nu = 1$.

Evolution en fonction de la probabilité de fausse alarme

Nous allons d'abord observer l'évolution des performances par la probabilité de détection (ou de bonne classification), en fonction de la probabilité de fausse alarme (mauvaise classification). Nous fixons donc le nombre d'échantillons à $N_{sample} = 50$ et observons les deux distributions envisagées sur les figures 1.12 et 1.13. Nous pouvons constater une évolution importante concernant les distances asymétriques (Chernoff, Mesure de Kullback Leibler avec ou sans modèle SIRV). Ces distances semblent gagner très nettement en performance pour des probabilités de fausse alarme élevées. Par ailleurs, on remarque certaines variations qui concernent les performances relatives de certaines distances. Ainsi à faible probabilité de fausse alarme, dans le cas d'une distribution non gaussienne, la divergence de Kullback Leibler SIRV offre de meilleures performances que le troisième rapport de vraisemblance généralisé associé au modèle SIRV. Cette tendance s'inverse quand la probabilité de fausse alarme augmente. En fouillis gaussien on remarque de la même manière que l'avance de la divergence de Kullback Leibler vis à vis de la distance de Bhattacharyya s'estompe avec l'augmentation de la probabilité de fausse alarme.

Evolution en fonction du nombre d'échantillons utilisés dans l'estimation des matrices de cohérences

Nous fixons maintenant la probabilité de fausse alarme à $P_{FA} = 10^{-2}$. Nous étudions l'évolution des performances en fonction du nombre d'échantillons utilisés dans l'estimation des paramètres et des distances. Nous faisons varier ce nombre de 9 à 81, valeurs qui correspondent à des tailles de fenêtre crédibles pour une estimation par fenêtres glissantes de la taille (3×3) à la taille(9×9). Comme précédemment, nous considérons le cas gaussien sur la figure 1.14 et le cas fortement non gaussien sur la figure 1.15. Comme pour l'evolution de la fausse alarme, nous notons une forte variation des distances asymétriques, qui gagnent proportionnellement en performances vis à vis des autres distances avec l'augmentation du nombre d'échantillons. Nous remarquons les mêmes phénomènes pour les performances relatives de la divergence de Kullback Leibler et la distance de Bhattacharyya en gaussien, ainsi que la distance de Kullback Leibler SIRV et le RVG3 en non gaussien.

Nous remarquons de manière globale que même si les distances associées à la distribution SIRV ont de meilleures performances en non gaussien, c'est de moins en moins le cas quand le nombre d'échantillons devient très petit. Ainsi, même pour un fouillis aussi impulsionnel qu'une K distribution de paramètres de forme $\nu = 1$, si le nombre d'échantillons est très faible il est encore préférable d'utiliser un modèle gaussien.

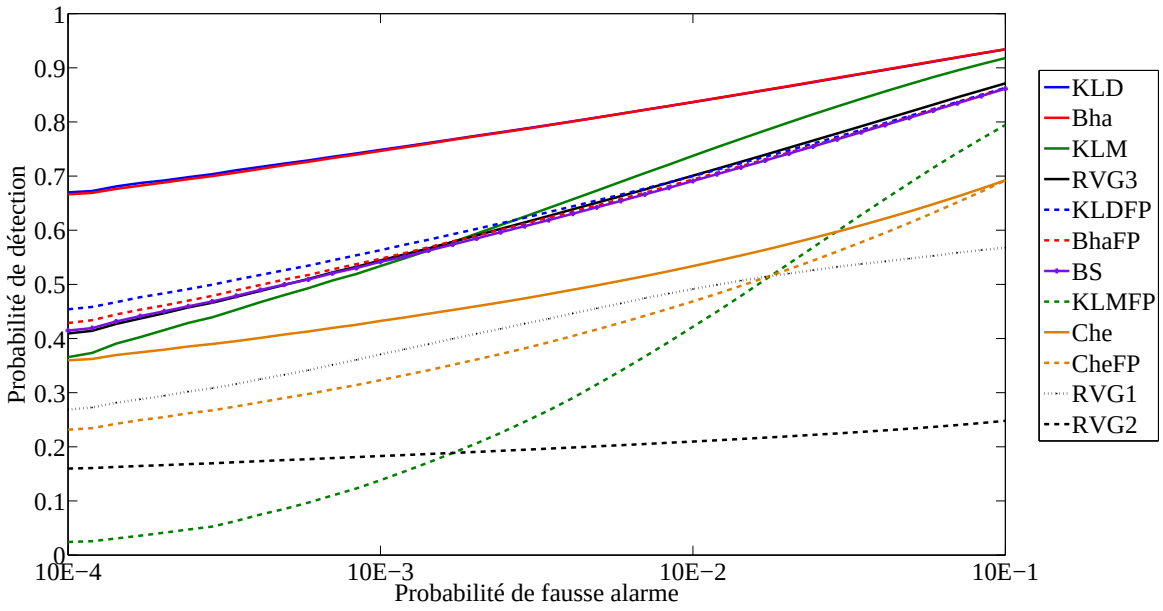


FIGURE 1.12 – Probabilité de détection (bonne décision) en fonction de la probabilité de fausse alarme pour l'ensemble des distances étudiées dans le cas d'un fouillis gaussien, $N_{sample} = 50$.

Bilan sur les performances des différentes distances

Pour les distances associées au modèle gaussien.

1. La divergence de Kullback Leibler, c'est la distance qui offre les meilleures performances en fouillis gaussien. Le gain par rapport à la distance de Bhattacharyya est cela dit extrêmement léger et n'est véritablement sensible qu'à très faible probabilité de fausse alarme ou très faible nombre d'échantillons. En non gaussien, elle reste la meilleure des distances gaussiennes et arrive même à surpasser les distances SIRV pour de faibles nombres d'échantillons.
2. La distance de Bhattacharyya offre des performances quasiment identiques à la divergence de Kullback Leibler. En non gaussien, elle arrive aussi à surpasser certaines distances SIRV pour les faibles nombres d'échantillons. Dans une utilisation réelle il est peu probable qu'on observe une différence entre la distance de Bhattacharyya et la divergence de Kullback Leibler.
3. La mesure de Kullback Leibler, par son asymétrie, offre des performances bien plus faibles que les distances symétriques. C'est une remarque importante car cela veut dire que non seulement l'asymétrie est problématique au niveau de certains algorithmes, comme la segmentation d'image, mais elle est aussi problématique quand le paramètre supposé connu Σ_B est mal estimé ; les performances sont alors très mauvaises.

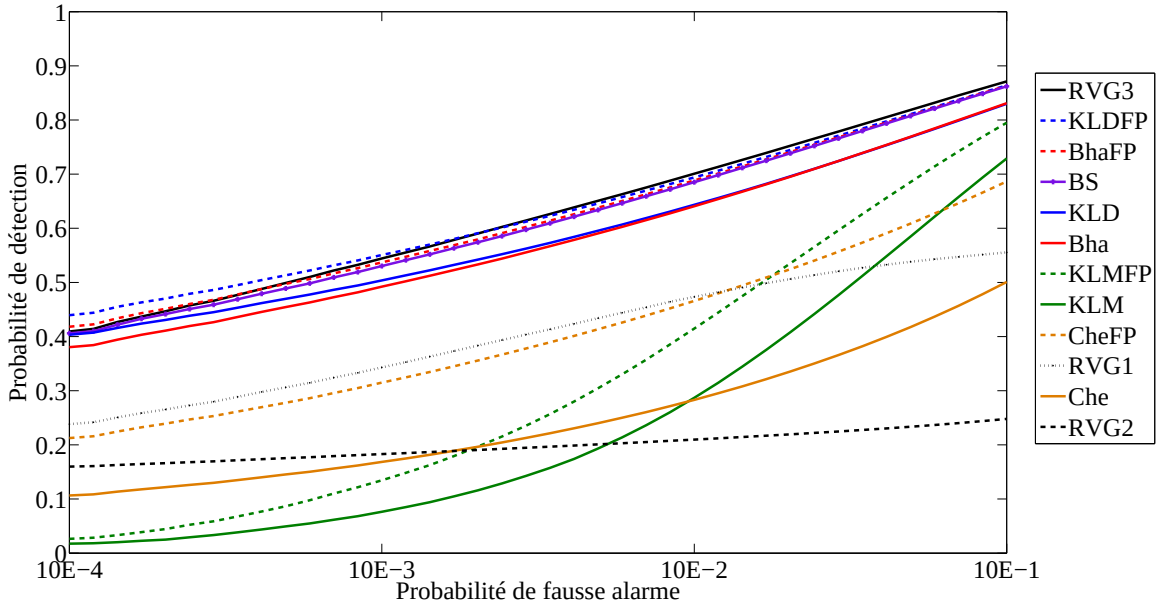


FIGURE 1.13 – Probabilité de détection (bonne décision) en fonction de la probabilité de fausse alarme pour l'ensemble des distances étudiées dans le cas d'un fouillis K distribué,$\nu = 1$, $N_{sample} = 50$.

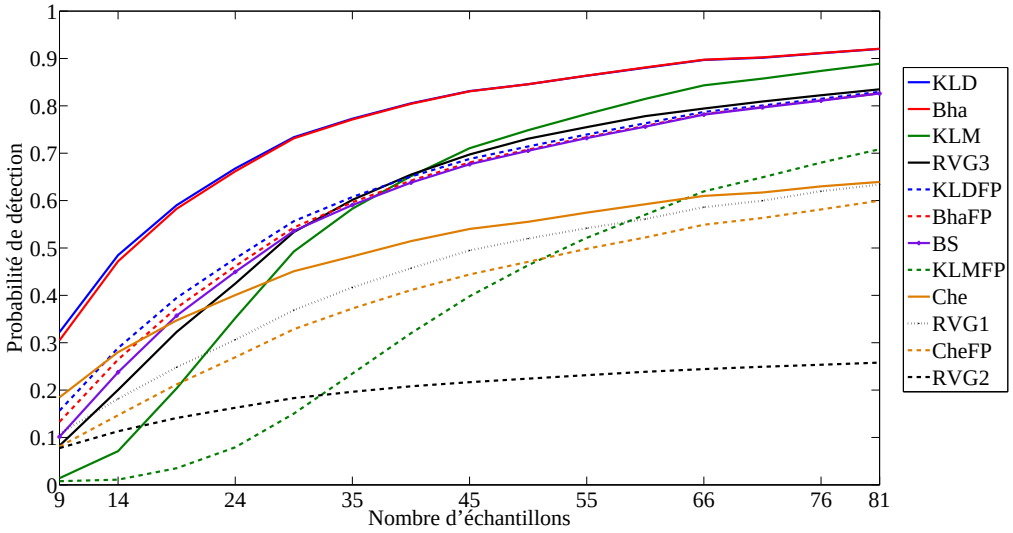


FIGURE 1.14 – Probabilité de détection en fonction du nombre d'échantillon utilisé dans l'estimation des matrices de cohérence pour l'ensemble des distances étudiées dans le cas d'un fouillis gaussien, $P_{FA} = 10^{-2}$.

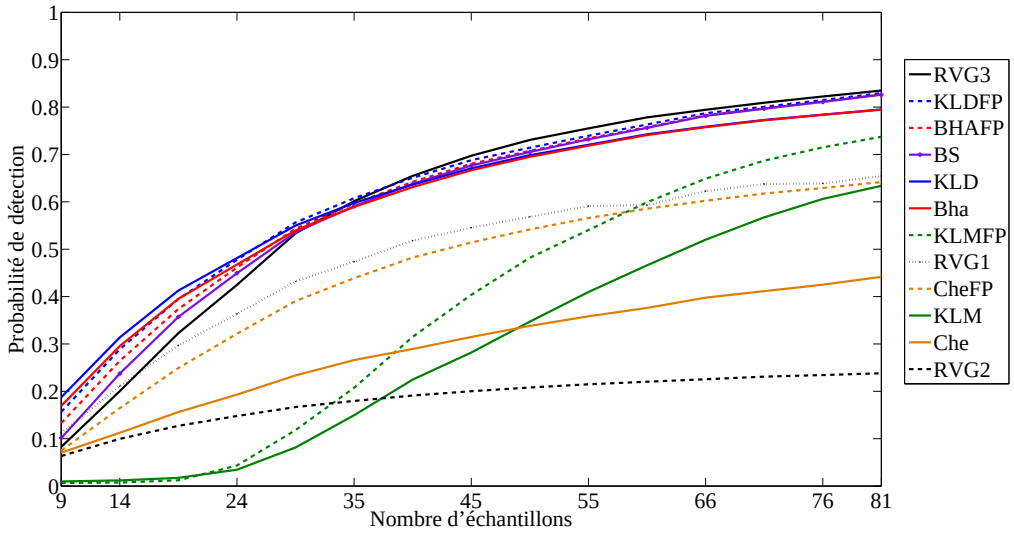


FIGURE 1.15 – Probabilité de détection en fonction du nombre d'échantillon utilisé dans l'estimation des matrices de cohérence pour l'ensemble des distances étudiées dans le cas d'un fouillis K distribué,$\nu = 1$, $P_{FA} = 10^{-2}$.

4. La distance de Chernoff a des performances qui nous ont surpris. D'après la littérature nous nous attendions à observer des performances supérieures à la distance de Bhattacharyya. Nous attribuons cette sous-performance à son asymétrie. En effet le paramètre t qui maximise l'expression s'éloigne très souvent de la valeur moyenne $t = 1/2$. Ainsi les performances, comme pour la Mesure de Kullback Leibler, sont médiocres dans une situation symétrique.

Pour les distances associées au modèle SIRV.

1. La divergence de Kullback Leibler, dans le cas d'un fouillis fortement non gaussien offre de très bonnes performances. Elle est encore une fois talonnée par la distance de Bhattacharyya.
2. La distance de Bhattacharyya, comme en gaussien, a des performances très proches de la divergence de Kullback Leibler.
3. La distance de Barlett 'SIRV', dont l'expression est très proche de la distance de Bhattacharyya a des performances extrêmement proches de cette dernière, tout en étant très légèrement inférieures.
4. La mesure de Kullback Leibler, comme toutes les distances asymétriques, n'offre pas de sursaut dans le cas du modèle SIRV : les performances restent bien plus faibles que pour les équivalents symétriques.
5. La distance de Chernoff, comme la mesure de Kullback Leibler, présente des performances qui restent médiocres.

6. Le rapport de vraisemblance généralisé 3, offre, lui, les meilleures performances globalement, même si à très faibles nombres d'échantillons ou probabilités de fausse alarme, d'autres distances sont faiblement plus performantes.
7. Les rapports de vraisemblance généralisés 1 et 2, offrent, eux, des performances très décevantes. Pour le premier, nous pouvons attribuer ce fait au même problème d'asymétrie, pour le second nous n'avons pas d'explications.

Conclusion sur les distances polarimétriques

Au travers de notre étude sur les distances polarimétriques et distances statistiques appliquées à la polarimétrie nous avons pu établir les points suivants :

- Les distances du monde de la théorie de l'information et celles obtenues par rapport de vraisemblance ont souvent des expressions similaires, et des performances presque identiques.
- Les rapports de vraisemblance généralisés et rapports de vraisemblance de type "Plug-in" peuvent chacun à leur tour être plus performants suivant les gammes de probabilité de fausse alarme. Nous confirmons le rapport de vraisemblance généralisé n'offre aucune garantie d'optimalité.
- Nous avons proposé un calcul de ces distances dans le cas du modèle SIRV, pour les distances de la théorie de l'information, Kullback Leibler et Bhattacharyya ainsi que pour les rapports de vraisemblance généralisés. Les expressions obtenues pour ce dernier sont inédites. Nous avons également démontré que les expressions pour Kullback Leibler et Bhattacharyya sont les mêmes qu'en milieu gaussien.
- Nous avons montré par simulation de type Monte Carlo, que le modèle SIRV n'apporte un gain de performance que pour des fouillis fortement non gaussiens, et qu'il n'est pas souhaitable d'utiliser ce modèle pour des images partiellement gaussiennes compte tenu des bien plus faibles performances dans ce type de fouillis.
- La divergence de Kullback Leibler offre les meilleures performances en fouillis gaussien, mais la distance Bhattacharyya offre des performances tellement proches que nous doutons pouvoir exhiber une quelconque différence dans une utilisation réelle.
- En fouillis non gaussien, c'est le rapport de vraisemblance généralisé "3" associé au modèle SIRV qui offre les meilleures performances. Il est cela dit suivi de très près par la divergence de Kullback Leibler et la distance Bhattacharyya utilisées avec l'estimateur du point fixe. Étant donné que le calcul du rapport de vraisemblance peut être très contraignant, elles offrent donc de bonnes alternatives.

Chapitre 2

La segmentation

2.1 Application des outils statistiques pour la segmentation d'images polarimétriques

Cette section est dédiée à la segmentation d'images polarimétriques.

Cette étude a été motivée par les tentatives de représentations de paramètres polarimétriques en milieu urbain. A cause de la très forte dynamique de ce type d'images, les méthodes classiques d'estimation spatiales des matrices de cohérence polarimétriques conduisent à de mauvaises estimations des ces paramètres. Par exemple, de très nombreux diffuseurs ont des réponses contenant des lobes secondaires significatifs qui viennent polluer les signatures polarimétriques des zones voisines. Par ailleurs, même si nous parvenons à extraire une information intéressante de ces paramètres, l'image résultante est difficile à interpréter visuellement, la structure et les différents objets présents dans l'image ayant été très affectés par l'effet de convolution induit par la fenêtre glissante.

A l'origine, l'objectif était donc de disposer d'un outil capable d'améliorer cette estimation. La partition de l'image en régions homogènes est une solution naturellement envisageable. C'est ainsi l'objectif de la thèse [38].

Dans un premier temps nous avons donc cherché à construire un outil de segmentation rapide, permettant d'estimer alors les paramètres polarimétriques sur les régions trouvées. Nous n'avions pas initialement pour objectif le caractère innovant de l'algorithme, mais nous souhaitions simplement disposer d'un outil utilisable pour le travail de thèse.

Au fur et à mesure du développement de cet outil et de son utilisation, nous avons eu l'occasion de l'améliorer, tout en l'adaptant aux problèmes rencontrés sur les images SAR de milieux urbains. Aujourd'hui, fort de toutes ces évolutions, l'outil produit des résultats tout à fait intéressants. Ce chapitre lui est donc consacré.

2.1.1 La segmentation d'image

La segmentation est un processus de partitionnement d'une image en de multiples segments, ou régions. Ces segments sont des groupes de pixels dont la réunion couvre l'intégralité de l'image. Ils peuvent aussi correspondre à un certain nombre de frontières

extraites de l'image. Dans [44], les outils de segmentation sont rangés selon trois catégories :

- Les algorithmes de classification, fondés sur le clustering (ou agglomération) de pixels ayant des caractéristiques proches ou identiques ;
- La détection de frontières ;
- La croissance de régions.

Les algorithmes de classification ou d'agglomération n'exploitent pas d'information spatiale (comme la connexité des différents pixels). Ils utilisent uniquement des critères statistiques. Ainsi, bien qu'en traitant les données de l'image dans son ensemble, ces algorithmes sont insensibles à une permutation aléatoire des pixels dans l'image.

Inversement, la détection de bords ou de contours s'intéresse majoritairement à l'information spatiale mais ne permet pas, à elle seule, de construire des régions ou segments fermés.

Enfin, la croissance de régions permet de construire des régions ou segments fermés tout en exploitant une information spatiale "locale", en s'appuyant sur des critères similaires à ceux utilisés en classification.

Il existe bien sûr des méthodes bien plus complexes qui sortent de ce cadre, ou qui combinent plusieurs approches. Ainsi la thèse de Jérôme Morio [38] combine une croissance de régions et des méthodes de contours actifs et permet de construire des régions fermées tout en exploitant une information spatiale "globale". Ces travaux sont d'ailleurs tout à fait remarquables et ont servi de base de travail pour l'ensemble des travaux que j'ai pu effectuer par la suite.

Comme je l'ai expliqué plus haut, mon objectif initial étant de construire un algorithme relativement simple, j'ai opté pour une méthode de croissance de régions. Nous verrons par la suite que cette méthode a été enrichie progressivement et permet d'obtenir des résultats d'une qualité qui approche maintenant les résultats de méthodes plus complexes.

2.1.2 Principe

Nous expliquons le principe de la croissance de régions, ou croissance de régions hiérarchiques sur la figure 2.1. C'est un algorithme itératif, qui à chaque étape, permet de passer d'une partition ou segmentation donnée de l'image, à une partition moins fine, en fusionnant deux régions connexes, sur des critères de ressemblances statistiques, ou aussi, dans notre cas de formes. Ce processus est alors répété de façon itérative jusqu'à atteindre le partitionnement souhaité. Il convient alors naturellement d'initialiser ce processus itératif par une partition de départ qui peut être un quadrillage d'une finesse équivalente à la résolution, ou choisie plus grande si on cherche à réduire le temps de calcul. Dans notre application nous cherchons à conserver le maximum de finesse au niveau des résultats de segmentation ; ainsi nous conserverons toujours une partition initiale d'une finesse égale à la résolution.

2.1.3 Processus détaillé

Le seuil d'arrêt est une problématique qui peut être délicate en fonction des applications, et qui a donné lieu à plusieurs travaux comme dans [45]. Il peut être en effet

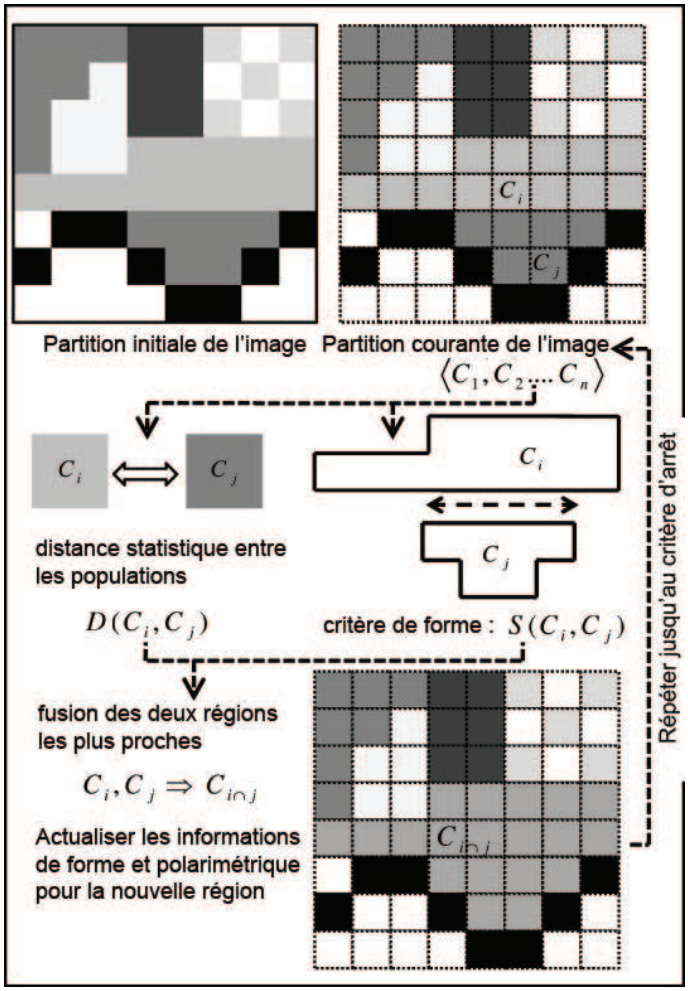


FIGURE 2.1 – Principe hiérarchique de la méthode de croissance de régions

souhaitable de connaître avec précision le nombre de zones différentes dans une image de manière objective. C'est d'ailleurs une problématique communément rencontrée en classification où il est souhaitable d'obtenir un nombre de classes "optimal". Ici l'objectif n'est pas de dénombrer le nombre de régions homogènes mais d'obtenir des régions de taille suffisamment importante en moyenne pour à la fois réduire l'impact du speckle et améliorer l'aspect visuel, et permettre une bonne estimation des matrices de covariance polarimétrique au sein de chacune des régions. C'est pourquoi nous n'insisterons pas sur le problème du critère d'arrêt et nous nous contenterons d'arrêter le processus itératif une fois une taille moyenne de région atteinte.

Nous avons étudié de façon très détaillée un grand nombre de distances statistiques dans la section précédente. Dans cette section nous nous contenterons de choisir parmi les différentes distances en justifiant au mieux notre choix. Outre un choix motivé de distances statistiques, l'originalité du travail provient de l'utilisation de critères de formes, qui vont contraindre la croissance des régions. L'utilisation de tels critères n'est pas nouvelle, mais notre approche ainsi que les critères choisis permettent à notre outil de se démarquer, notamment sur les images urbaines.

A priori cette méthode est simple. Par contre sa mise en oeuvre pratique nécessite une certaine réflexion pour éviter des temps de calculs prohibitifs. Un algorithme simple consisterait à calculer à chaque étape toutes les zones connexes et toutes les distances statistiques associées, puis à répéter cette opération après avoir fusionné les deux régions les plus proches. Mais son application deviendrait impossible tant le nombre de calculs redondants est alors important.

Nous proposons ici une méthode originale qui permet d'obtenir des temps de calcul tout à fait raisonnables. L'objectif est double : on souhaite à la fois éviter de recalculer des paramètres plusieurs fois (optimisation du temps de calcul), mais aussi éviter de stocker des informations inutiles, redondantes, ou très rapides à calculer (optimisation de la mémoire). En outre il est souhaitable d'éviter au maximum des allocations ou désallocations de mémoire en cours de segmentation, même s'il sera impossible d'éviter d'y avoir recours dans la mesure où la géométrie du problème varie grandement entre le début du processus (grand nombre de petites régions) et la fin du processus (petit nombre de grandes régions).

Formatage des données

Nous allons décrire ici les données stockées et les formats, bien que nous ne rentrerons pas finement dans les détails.

1. *Les segments.* Chaque segment S_i est indexé par i. Pour chaque segment il est nécessaire de stocker : sa matrice de cohérence ou de covariance M_i et le nombre de pixels qui le composent. Il est possible de manière facultative de conserver aussi des paramètres utiles si on désire les utiliser dans des critères de formes par exemple (nous verrons ceci plus en détail ulterieurement). On pourra ainsi par exemple conserver le périmètre, le nombre de régions voisines, ou tout autre paramètre que l'on pourra introduire plus tard comme l'orientation principale, ainsi que le degré d'orientation. Par ailleurs, comme certaines distances statistiques font régulièrement intervenir l'inverse de la matrice de covariance M_i^{-1}, il est possible aussi de stocker l'inverse de

la matrice afin de réduire sensiblement le temps de calcul de la distance. Ce choix est alors un compromis temps/mémoire.

2. *Les frontières.* A chaque couple de segments S_i et S_j connexes correspond une frontière F_{ij}. Il est alors nécessaire de stocker : les deux indices des deux segments et la distance entre les deux segments (distance au sens statistique uniquement ou bien en incluant l'influence de la forme des différentes régions). Parmi les paramètres facultatifs, il est possible de stocker des paramètres utiles au calcul de critères de formes comme la longueur de la frontière.
3. *L'image.* Il n'est pas très efficace de stocker une liste de positions de pixels pour chaque segment. Au contraire il est assez facile de conserver une image d'indices des segments. Celle-ci sera mise à jour à chaque fusion de région.

Fusion de deux régions

A partir d'un partitionnement donné, nous allons fusionner les deux régions les plus proches au sens de la distance choisie. Il convient alors de procéder aux étapes suivantes :

1. *Déterminer les deux indices.* Le choix des deux indices se fait en parcourant l'ensemble des frontières F_{ij}. Il peut être judicieux de stocker celles-ci de manière ordonnée afin d'accélérer la recherche de minimum. Les deux indices de la frontière ayant la distance la plus petite seront fusionnés.
2. *Mettre à jour l'image.* L'image des indices peut être mise à jour immédiatement. Plutôt que de créer un nouvel indice pour la nouvelle région, il est plus simple de réutiliser un des deux indices existants. Arbitrairement nous utilisons l'indice le plus faible pour représenter la nouvelle région ($i < j$). Ainsi il suffit de remplacer toutes les occurrences de j par i dans l'image.
3. *Mettre à jour le nouveau segment.* Le nouveau segment reprend donc l'indice le plus petit i. Il faut alors calculer la matrice de cohérence de la nouvelle région (ainsi qu'éventuellement son inverse). Si on travaille avec un modèle gaussien, ceci peut se déduire très simplement des deux matrices précédentes ainsi que de la taille des deux régions. Dans le cas d'un modèle SIRV, il faut réutiliser les données et recalculer le point fixe à l'aide de l'image des indices. De manière aussi immédiate on met à jour la taille du segment. Pour des paramètres plus complexes qui pourraient avoir été introduits comme le nombre de voisins, il faudra exploiter des informations que nous ferons apparaître lors de la mise à jour de la carte de frontières. Le périmètre de la nouvelle région se calcule très simplement à partir des deux périmètres et la longueur de l'ancienne frontière entre les deux régions.
4. *Mettre à jour la carte des frontières.*
   - *Supprimer les anciennes frontières.* Toutes les anciennes frontières liant les régions i et j doivent être supprimées (bien que les informations associées doivent être momentanément conservées pour générer les nouvelles frontières). Afin d'éviter de la réallocation dynamique de mémoire qui entraînerait des coûts de calcul très importants, il est plutôt souhaitable de neutraliser la frontière (en lui attribuant une distance infinie). Un nettoyage des frontières neutralisées, entraînant une

réallocation de mémoire peut être réalisée à intervalles réguliers (toutes les 10,000 itérations par exemple) afin de réduire l'utilisation de mémoire sans augmenter le temps de calcul de façon significative.
- *Nouveaux voisins.* Les voisins de la nouvelle région se déduisent des voisins des deux segments précédents. On distingue l'union distincte des deux ensembles de voisins, de l'intersection. Pour les voisins uniques (qui étaient connexes à un des deux segments uniquement), la longueur de la frontière est la même que pour la frontière précédente. Pour les voisins communs aux deux anciens segments, la longueur de la nouvelle frontière est naturellement la somme des deux anciennes frontières. A partir de l'union des indices des voisins des deux régions précédentes on obtient les indices des segments voisins à la nouvelle région.
- *Calcul des distances.* A partir de la liste des nouveaux voisins il suffit de calculer les distances entre les régions (au sens statistique et ou de forme en utilisant les paramètres du nouveau segment ainsi que des régions voisines).

5. *Supprimer l'ancien segment.* Comme pour la suppression des frontières, il est préférable de neutraliser le segment et d'effectuer un nettoyage régulier afin d'éviter trop d'allocation dynamique de mémoire. Ici, le segment est naturellement neutralisé car il n'est plus lié aux autres segments. Nous fixons cependant sa taille à zéro afin de pouvoir repérer les segments inutiles lors du nettoyage de la mémoire.

2.1.4 Utilisation des distances statistiques polarimétriques

Les performances théoriques des distances statistiques, amplement décrites dans la première section de ce chapitre, montrent que la plupart des distances sont globalement équivalentes. Le choix de l'utilisation du modèle SIRV dans notre cas, y compris pour des données urbaines, s'est effectué à la fois sur l'étude statistique présentée mais aussi sur une exploitation directe sur des images. En pratique, même dans des données urbaines, l'utilisation du point fixe ne montre pas de résultats suffisamment différents des méthodes d'estimations classiques pour justifier son utilisation. En outre, nous avons montré sur une étude par Monte Carlo que même dans le cadre des données urbaines, le modèle SIRV n'apporte un gain que pour une faible fraction des zones de l'image, les performances étant réduites pour la plupart des autres zones de l'image.

Le traitement de données hautes résolution entraîne naturellement un besoin d'optimisation de temps de calcul compte tenu du volume très important de données. Bien que nous proposions une structure d'algorithme évitant au maximum de répéter des opérations redondantes, le calcul des distances statistiques joue un rôle important dans le processus de segmentation et peut impacter grandement le temps de calcul. La possibilité de réaliser ces calculs de distance de manière vectorielle ou non aura notamment une influence importante sur les temps de segmentation. Parmi les distances donnant de bonnes performances théoriques, la divergence de Kullback Leibler est la plus rapide à calculer si les inverses des matrices de cohérence de chaque région sont conservées. En outre son calcul est complètement vectorisable ce qui la rend particulièrement adaptée à un processus de segmentation comme le notre.

2.2 Critères images

L'information spatiale et plus particulièrement de forme peut jouer un rôle très important dans un processus de segmentation. Au départ, une segmentation par croissance de région exploite une information spatiale assez restreinte. La "visibilité" reste en effet très locale et les fusions n'ont lieu qu'en fonction de comportements statistiques locaux n'ayant aucun moyen d'exploiter une structure plus globale de l'image. En général, cela se traduit par une segmentation globalement efficace mais qui produit dans le cas de zones de transition ou dégradées des formes que l'on peut qualifier de fractales et dont les frontières sont tentaculaires. Par exemple une ligne interrompue sera complètement ignorée, et l'ensemble des frontières des segments conservera un aspect rugueux.

La structure globale est pourtant une information cruciale qui est d'ailleurs naturellement utilisée par le couple oeil-cerveau humain. Ainsi par exemple sur une image urbaine, même une personne n'ayant pas particulièrement l'habitude d'observer des images SAR peut aisément reconnaître routes ou bâtiments. Un oeil expert peut même améliorer son interprétation des images en observant les lignes de double-rebond et les ombres pour en déduire la forme ou la taille de ces bâtiments. Notre idée est donc d'utiliser justement cet a priori concernant la forme ou la structure de l'image, afin d'améliorer les résultats de segmentation.

L'exploitation d'informations géométriques plus globales se fait alors par des contraintes sur la forme des régions, qui vont se traduire par une certaine pondération des distances statistiques utilisées. Dans le cadre d'applications visant des milieux naturels, il est évident que seul très peu d'a priori pourront être exploités. Dans un milieu urbain par contre nous pouvons espérer exploiter quantité d'a priori concernant la forme et la structure des objets rencontrés. Ce type de segmentation est alors appelé croissance de région contrainte. Dans le cadre de la segmentation d'image SAR polarimétriques, Beaulieu et Touzi ont déja utilisé deux très simples critères de forme afin d'encourager leur segmentation à produire des régions compactes [46] et [47]. Dans notre outil, nous proposons de pousser l'approche de la croissance de région contrainte beaucoup plus loin, avec pour objectif de pouvoir traiter efficacement des images polarimétriques urbaines. Ces travaux pourront par la suite s'appliquer à tout scénario où il est possible de produire une connaissance a priori sur la forme ou la taille des objets rencontrés.

Tant que les règles restent simples ou peu nombreuses comme dans [46], la gestion de ces contraintes est relativement facile. Cependant, quand on désire augmenter le nombre de règles et complexifier celles-ci, leur gestion conjointe devient difficile et nécessite un cadre plus rigoureux. Dans mes expériences, j'ai ainsi pu constater l'apparition de certaines interférences, lorsque plusieurs règles se contredisent, ou bien de dominance inattendue d'une règle dans certains cas particuliers. Le juste équilibre entre toutes les règles contraignant la forme, tout en conservant le rôle important des distances statistiques est difficile à atteindre. Deux solutions ont été envisagées et testées :

1. Apprentissage des règles par réseaux de neurones. Plutôt que de définir des règles avec cette méthode nous pourrions permettre à l'outil de segmentation de les apprendre par lui-même avec une image de référence. L'idée était alors de stocker un certain nombre de paramètres (tailles des deux régions, périmètres des deux régions, nombre de voisins exclusifs, commun, longueur de la frontière), et de renseigner à l'aide d'une image dont

on connaît la segmentation "idéale", si la fusion de deux régions est un bon choix, ou non. Le problème de cette approche provient de la génération de la base d'apprentissage. En pratique les règles basées sur la forme n'auront une influence importante qu'à quelques moments clés dans le processus de segmentation. La majorité des fusions qui ont lieu au sein d'une zone homogène ne présentent pas de problème et ne stimulent donc pas l'apprentissage des règles qui pourront faire la différence aux moments difficiles. Même en sélectionnant uniquement les cas critiques (les fusions qui ont lieu alors qu'elles n'auraient pas dû), il est plus difficile de repérer, même avec une base d'apprentissage, des fusions qui auraient dû avoir lieu. Au final, avec un réseau de neurones comportant un très grand nombre de couches de neurones, et des bases d'apprentissage de plusieurs millions de cas, je n'ai pas réussi à faire émerger automatiquement des règles de décision adaptées en me basant sur les paramètres de formes.

2. Spécification humaine des règles et optimisation automatique de paramètres. Ayant constaté l'incapacité des réseaux à "apprendre" naturellement des règles de segmentation basées sur la forme, la seconde solution qui a finalement été choisie consiste à créer chacune des règles que nous jugeons nécessaires pour obtenir de bons résultats. L'équilibre entre ces différentes règles, qui pourront a priori interagir de façon destructive, sera alors trouvé à l'aide d'un processus de fusion de décision. J'ai choisi d'utiliser la logique floue pour sa simplicité. La fusion se fera par différents poids et paramètres qui pourront être optimisés à l'aide des bases d'apprentissage que nous avions utilisées dans le cadre des réseaux de neurones. Ces poids détermineront les interactions entre les différentes règles et pourront, s'ils sont calculés correctement, éviter les interactions préjudiciables entre les différentes règles. La logique floue est un outil qui apporte surtout par sa simplicité d'utilisation et de compréhension, ainsi que par la facilité à transformer des règles exprimées verbalement par un opérateur, en opérations mathématiques. La seule limitation de cette méthode provient alors de la créativité et de la pertinence de l'opérateur qui définit les lois et les contraintes sur la forme des régions.

2.2.1 Principe

L'utilisation de la logique floue pour la création de critères de formes est résumée sur la figure 2.2. Elle comporte les principales étapes suivantes :

1. *Choix des paramètres.* Les paramètres que nous allons utiliser comme critères de formes.
2. *Définition des sous ensembles flous* associés aux paramètres, ainsi que des fonctions d'appartenance aux sous-espaces flous.
3. *Définition des règles de formes* et traduction en logique floue avec des règles de la forme "SI-ALORS".
4. *Fusion des décisions.* Chacune des règles ayant abouti à une décision, définie comme un degré d'appartenance à un sous-ensemble flou, les décisions sont fusionnées puis transformées en variables réelles de décision.

Nous allons expliquer l'ensemble de ces étapes en détail dans cette section. Les travaux présentés ne sont qu'une application simple de la logique floue et ont été intégralement

inspirés du livre de Bernadette Bouchon-Meunier [48], ainsi que des documents du cours qu'elle dispense dans le cadre des formations CNRS.

Choix des paramètres.

Le choix des paramètres est bien entendu lié aux paramètres dont on aura besoin pour définir des règles de décisions et des critères sur la forme des différentes régions. J'ai déjà évoqué quelques paramètres simples qui interviennent naturellement dans plusieurs règles de formes. Il s'agit par exemple du périmètre, de l'aire des régions ou de la longueur des frontières. Dans la littérature qui concerne moins spécifiquement la segmentation d'image en général, il existe un nombre impressionnant de paramètres à notre disposition pour qualifier la forme d'une région. Naturellement ces paramètres s'orientent principalement sur des applications où des a priori de formes peuvent être utilisés, comme les objets marqués par l'influence de l'homme. Parmi ces nombreux paramètres, il existe un très grand spectre de complexité, au sens calculatoire, de paramètres.

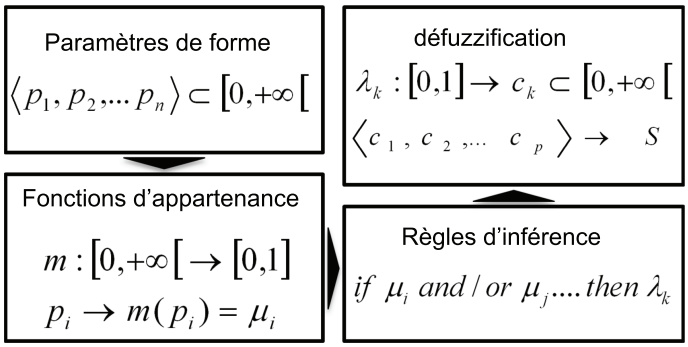


FIGURE 2.2 – Contraintes de forme gérées par la logique floue

Par exemple, une des spécificités des objets, bâtiments, voire même des champs ou portions de forêt délimitées par l'homme est la compacité. Elle différencie grandement les zones purement naturelles des zones sous l'influence humaine. Il existe plusieurs manières de calculer la compacité d'une région dans une image. La plus simple peut être assimilée à un rapport entre le périmètre et l'aire d'une région, paramètres rapidement déductibles d'anciens paramètres lors de la fusion de deux régions. Mais il est tout à fait possible d'envisager des méthodes bien plus complexes (mais aussi plus précises), comme le rapport entre le périmètre de la région et le périmètre de son enveloppe convexe, ou de son enveloppe pseudo convexe comme dans [47]. Cependant, l'estimation de ces paramètres comme l'enveloppe convexe nécessite des calculs très longs à chaque fusion, pour estimer le nouveau paramètre. Ou, pire, pour estimer tous les paramètres des potentielles formes créées par les potentielles fusions !

Les campagnes de mesure polarimétriques radar à haute résolution produisent de très grands volumes de données. Compte tenu de la taille des images à segmenter et des temps de calcul potentiellement longs, j'ai choisi de restreindre volontairement la complexité des paramètres utilisés pour les critères de forme. La limite que je fixe est de se contraindre aux paramètres déductibles des paramètres précédents. Ainsi il n'est pas nécessaire de

retravailler à partir des données initiales, ou de l'ensemble des positions des pixels qui composent les régions. Un exemple simple est le périmètre. Celui-ci se déduit très facilement des périmètres des deux régions ainsi que de la longueur de la frontière qui les sépare. De cette manière, la gestion des critères de forme n'apporte pas de contrainte supplémentaire pouvant faire augmenter les temps de calcul. Au contraire, il n'est pas rare de remarquer qu'en produisant des formes compactes tout le long de la segmentation on réduit grandement le nombre de petites frontières et donc le nombre de calculs de distances. Ainsi, des critères de formes adéquats accélèrent la segmentation d'image.

Nous distinguons trois familles de paramètres :

- *Les paramètres régions.* Ceux-ci sont uniquement dépendants de la région concernée. Dans nos règles, nous utiliserons donc le périmètre, le nombre de pixels qui composent la région (ce qui correspond aussi à l'aire), le nombre de voisins, l'angle d'orientation principale (plus de détails seront donnés dans la suite sur ce point) et le degré d'orientation.
- *Les paramètres frontières.* Ces paramètres sont relatifs à un couple de régions. Ici nous nous restreignons à la longueur de la frontière entre les deux régions. Cependant, si nous avions décidé de sacrifier du temps de calcul il existe de nombreux paramètres très intéressants qui pourraient servir à renseigner la frontière (linéarité, rugosité), ou bien à décrire la forme qui serait obtenue en cas de fusion des deux régions. Beaulieu compare, par exemple, les plus petits rectangles contenant chacune des régions (paramètres régions), avec le plus petit rectangle contenant les deux régions (paramètre frontière).
- *Les paramètres globaux.* Ces paramètres ne sont pas liés à une région ni à un couple de régions, ils servent en quelque sorte de commande extérieure. Ils permettent d'influencer les règles de fusion, en tenant compte de l'état global de segmentation de l'image. Un exemple simple que nous aurons l'occasion de détailler plus tard concerne la taille des régions. En effet un jugement porté sur la taille d'une région, par exemple la qualifier de "petite", prend un sens numérique très différent au cours de l'avancement du processus de segmentation. Ces paramètres globaux nous permettent d'ajuster les paramètres utilisés dans l'élaboration des contraintes de forme de façon différente en fonction des différentes étapes dans la segmentation. Un problème peut cependant survenir si on abuse des paramètres globaux. En effet, à chaque étape de fusion, nous calculons les paramètres de la nouvelle région, et des nouvelles frontières, ainsi que les distances avec les voisins de la nouvelle région. Ce dernier calcul est assez rapide car le nombre de distances à calculer est relativement restreint. Par contre la modification d'un paramètre global nécessite de mettre à jour l'intégralité de la carte des distances existantes. Il convient donc de ne pas multiplier inutilement les variations de paramètres globaux pour éviter une inflation des temps de calculs.

Les paramètres sont des valeurs numériques. Ils forment ainsi ce qu'on appelle un domaine de discours, ou un univers de discours (Universe of discourse en anglais). A partir de cet univers, nous allons définir des sous-espaces flous (Fuzzy Subspaces), à l'aide d'une variable linguistique (mot ou expression décrivant une qualité informative sur ces valeurs). A partir de l'ensemble des valeurs rencontrées pour la taille des régions, on pourra définir des sous-espaces à partir d'une description linguistique qui concerne la taille.

Nous pourrons donc définir les sous espaces flous des "petites", "moyennes" et "grandes" régions. Pour chaque valeur numérique de taille, l'appartenance à un sous-espace flou est caractérisé par une fonction d'appartenance (Membership function).

2.2.2 Définition des sous-ensembles flous, et fonction d'appartenance

Un sous-ensemble flou est lié à une variable linguistique (petit, grand, pointu, etc.). Plus précisément il est lié à un domaine de discours, que nous noterons U, ensemble des valeurs numériques rencontrées par la variable concernée. Pour définir un sous ensemble flou (noté V) sur ce domaine, il est nécessaire de définir une fonction de U, à valeur dans $[0, 1]$. Ainsi pour toute valeur numérique $x \in U$, le degré d'appartenance de x est noté $m(x)$. Si l'élément x de U est complètement inclus dans V, son degré d'appartenance est $m(x) = 1$. Inversement si x est complètement exclu de V son degré d'appartenance est $m(x) = 0$. Enfin, et c'est la particularité des sous-ensembles flous, si x appartient partiellement à V alors $m(x)$ prend une valeur comprise entre 0 et 1. Plus $m(x)$ est proche de un et plus x "appartient à V".

Les fonctions d'appartenance sont le plus souvent construites avec des fonctions linéaires par morceaux. Dans notre application, nous préférons des fonctions linéaires pour éviter des décrochements brutaux dans la segmentation. Parmi les fonctions aux dérivées continues, les plus populaires sont les fonctions sigmoïdes. Nous nous contenterons alors de travailler avec des fonctions d'appartenance sigmoïdes ou sigmoïdes composées (produit de sigmoïdes).

Pour reprendre l'exemple décrit plus haut, nous pouvons décrire la taille d'une région avec les variables linguistiques "petite", "moyenne" et "grande". A l'aide d'une sigmoïde nous pouvons, en fonction de la taille n d'une région (en pixel), calculer le degré d'appartenance d'une taille de régions, à chacun des sous espaces flous associés à ces variables linguistiques. Pour cet exemple nous montrons ce que pourrait donner l'expression de ces fonctions d'appartenance sur la figure 2.3. Comme nous l'avons évoqué plus haut, ces sigmoïdes dépendent aussi d'un paramètre global (qui, lui, par contre prendra des valeurs discrètes qui ne changeront donc qu'à quelques moments clés dans la segmentation). Une région que nous pouvons qualifier de taille "moyenne" en début de segmentation, pourra appartenir au sous-espace "petite" en fin de segmentation.

Les paramètres utilisés pour les critères de formes vont être qualifiés par une variable linguistique par l'intermédiaire des fonctions d'appartenances. Ainsi on pourra par exemple dire d'une région qu'elle est petite (avec un degré 0.8), avec un grand périmètre (degré 0.6), et très fortement orientée (degré 0.9). Cet exemple peut être caractéristique d'une route ou d'une ligne de chemin de fer. Si on observe les paramètres de sa frontière avec une autre région ayant des paramètres similaires, on pourrait aussi par exemple noter que leur frontière est fine (0.8), et que leurs orientations principales sont très proches (1). Nous pouvons alors avoir à faire à deux segments de route alignés, dont nous allons encourager la fusion. C'est à l'aide de règles de logique floue, aussi appelées règles d'inférence,

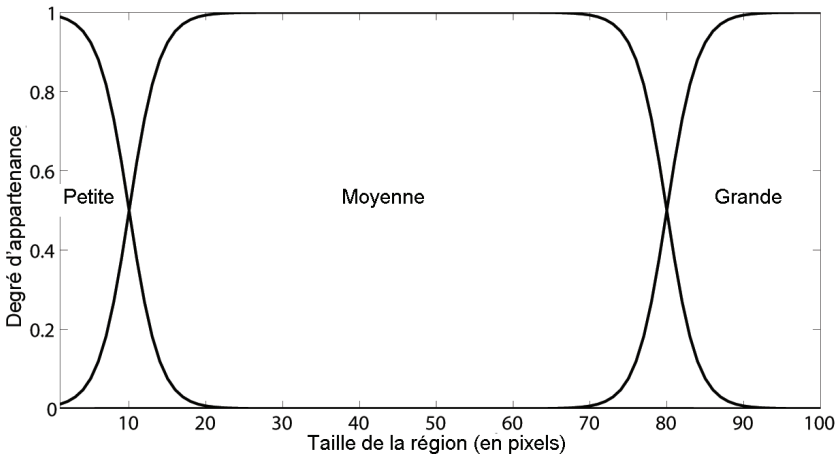


FIGURE 2.3 – Exemple de fonction d'appartenance : taille d'un segment pour les premières étapes de segmentation.

que nous allons traduire ce type de raisonnement en contraintes de forme.

2.2.3 Règles d'inférence

Les règles d'inférence floue traduisent des règles sur les variables linguistiques de la forme "si, alors" en opérations logiques et numériques sur les degrés d'appartenance des différents paramètres. Une règle exprimée par un expert ou un utilisateur s'exprime en français sous une forme du type : "si une des régions est 'X', que l'autre est 'Y', alors il faut encourager la fusion avec un critère 'Z'". Évidemment le nombre de conditions peut être bien plus complexe. La condition peut aussi comprendre des opérations logiques comme le *ou*, le *ou exclusif*, le *non*, etc. On notera μ_i les degrés d'appartenance et $\lambda_i = f_i(\mu_1, \mu_2, ...\mu_n) \in [0, 1]$, la règle d'inférence. La valeur numérique λ_i donnée par le calcul de la règle d'inférence est aussi un degré d'appartenance à un sous espace flou. Ainsi par exemple si une règle cherche à favoriser la fusion de deux régions, la sortie sera un degré d'appartenance au sous-espace "fusion encouragée" de l'univers de discours des décisions de fusion. Nous pouvons donc donner naturellement un poids relatif aux différentes règles si celles ci encouragent ou découragent plus ou moins la fusion des régions. Après calcul des degrés d'appartenance aux sous-espaces de décision de toutes les règles, nous obtiendrons une liste de degrés et de sous espaces qui pourront ressembler à : fusion encouragée à 0.9, fusion très fortement encouragée à 0.5, et fusion découragée à 0.1. Nous expliquerons plus tard comment fusionner l'ensemble des degrés d'appartenance produits par les différentes règles en une seule décision.

Il existe plusieurs façons de définir les opérations logiques floues. Certaines font appel à des opérations de min, de max ou des pondérations plus complexes. J'ai opté pour une solution arithmétique qui permet de simplifier grandement les calculs pour des règles complexes, et qui exprime ces règles sous forme de produits et de sommes de degrés d'appartenance. Les principales règles logiques et leurs expressions en fonction des degrés d'appartenance sont exprimées ici :

$$\begin{aligned} \neg A &\rightarrow \mu_{\neg A} = (1 - \mu_A), \\ A \wedge B &\rightarrow \mu_{A \wedge B} = \mu_A . \mu_B, \\ A \vee B &\rightarrow \mu_{A \vee B} = \mu_A + \mu_B - \mu_A . \mu_B, \\ A \veebar B &\rightarrow \mu_{A \veebar B} = \mu_A + \mu_B - 2\mu_A . \mu_B. \end{aligned} \tag{2.1}$$

Les seules limitations pour la création de règles sont les paramètres disponibles, et les temps de calcul associés à leur éventuelle estimation. En effet le calcul des règles en lui-même est trivial, et peut facilement se faire de façon vectorielle et parallèle. Après, seule la créativité de l'utilisateur et la pertinence des choix font l'efficacité de la méthode qui reste au final extrêmement simple. Les règles que je présente dans le manuscrit sont construites avec l'application des images SAR urbaines en vue. Mais comme il s'agit d'un système « à la carte », il est tout à fait possible d'exploiter des a priori sur d'autres types d'images. Je ne présenterai pas de manière exhaustive toutes les règles que j'ai pu mettre en oeuvre. Un premier exemple est expliqué en détail mais je passerai plus vite sur les règles suivantes, en résumant surtout l'idée sous-jacente.

Exemple détaillé : gestion du speckle.

Le speckle (ou chatoiement) est un des problèmes récurrents de l'imagerie polarimétrique radar. En segmentation ou classification, l'effet du speckle est en général la persistance de certains pixels, ou groupes de pixels, dont l'interférence constructive a produit une amplitude bien plus élevée que la moyenne locale ou un comportement polarimétrique très contrasté. Ces pixels représentent en quelque sorte les queues de distribution, et sont inévitables compte tenu du caractère aléatoire des réalisations statistiques. Même si un éventuel léger pré filtrage des données est possible, nous cherchons avant tout à éviter de perdre la résolution par un moyennage trop important et nous devons donc disposer d'une méthode pour traiter ces réalisations rares. Dans la section concernant les distances statistiques, j'avais expliqué l'impact de la taille des régions sur le calcul des distances. Cet effet est en moyenne préjudiciable et biaise la segmentation en faveur des couples de régions de tailles très différentes. En pratique, les distances pondérées sont moins sensibles au speckle à cause de ce biais qui favorise la fusion entre une région de grande taille, et une toute petite région produite par un point brillant. La solution que nous proposons ici consiste à travailler sur ces points de speckle de manière plus fine grâce à la mesure d'autres paramètres que la taille. L'expression de la règle sous forme linguistique est la suivante :

'Si la région A est très petite, et la région B est grande,
et le rapport frontière périmètre pour la région A est très grand,
alors la fusion doit être très fortement encouragée'. (2.2)

Nous noterons donc $\mu_{tp}(A)$ le degré d'appartenance de la région A au sous espace flou "très petite région", $\mu_g(A)$, pour "grande", et $\mu_{fp}(A, A \cup B)$ le sous-espace flou "grand"

pour le rapport entre la frontière entre les régions A et B, et le périmètre de la région A. La traduction de la règle exprimée ci-dessus en logique floue s'écrit alors :

$$\lambda = (\mu_{tp}(A) \wedge \mu_g(B)) \wedge \mu_{fp}(A, A \cup B). \tag{2.3}$$

En pratique la règle doit être symétrisée (c'est le cas pour toutes les règles d'ailleurs), car il n'y a pas de différence a priori entre les régions A et B. Par contre il faut distinguer les règles qui peuvent être symétrisées par un "*ou*", de celles à symétriser par un "*ou exclusif*". Dans cet exemple nous ne pouvons pas imaginer avoir une situation symétrique. Il convient donc d'utiliser un ou exclusif ce qui nous donne finalement :

$$\lambda = ((\mu_{tp}(A) \wedge \mu_g(B)) \wedge \mu_{fp}(A, A \cup B)) \veebar ((\mu_{tp}(B) \wedge \mu_g(A)) \wedge \mu_{fp}(B, B \cup A)) \tag{2.4}$$

L'effet de cette règle peut être résumé sur la figure 2.4.

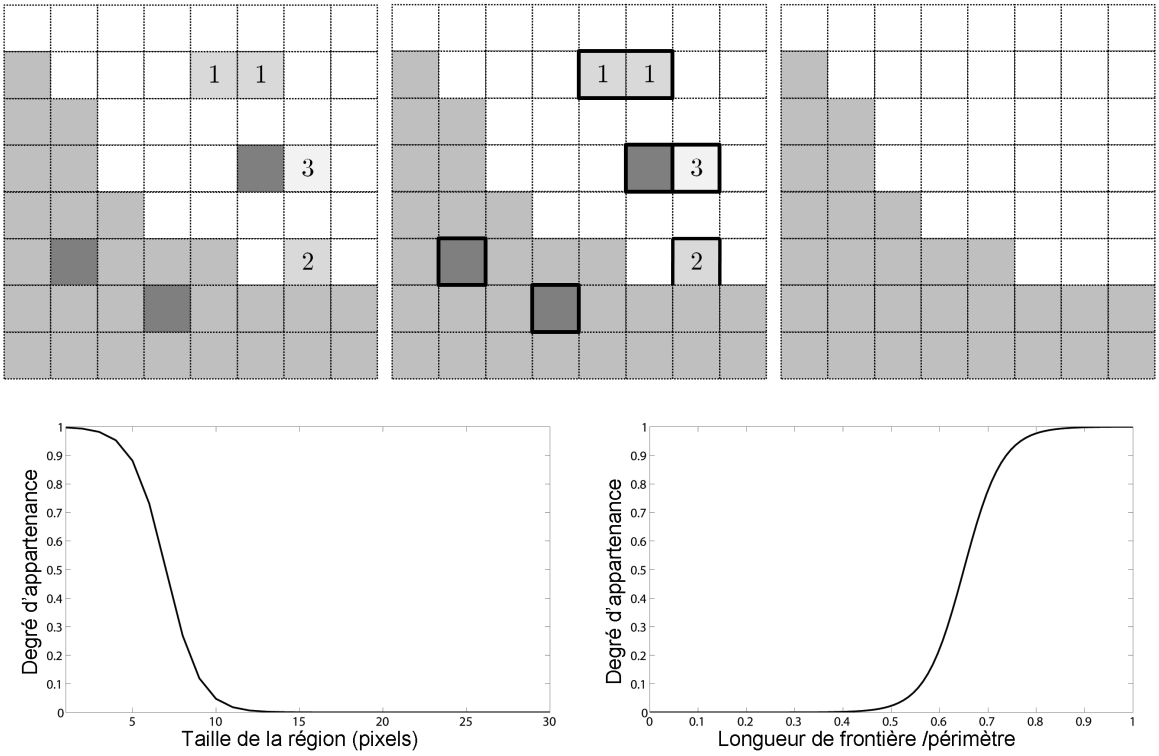


FIGURE 2.4 – Les points brillants, ou très contrastés, petits et entourés par une région plus grande sont encouragés à être fusionnés avec cette région.

Exemple complémentaire

L'utilisation de la règle précédente permet de régler de façon élégante la gestion des pixels de petite taille. Les paramètres des fonctions d'appartenance étant réglés à l'aide de base d'apprentissage, les résultats sont très bons. Toutefois, en observant les résultats

de segmentation sur certains exemples, on remarque qu'il existe dans certains cas, une persistance de certains points de speckle au sein de zones homogènes. Nous nous sommes rendu compte que ces points de speckle sont alors souvent des points doubles, composés de deux régions, ayant des tailles assez proches et des frontières communes assez étendues. Comme la frontière entre ces deux petites régions est assez importante, la règle précédente ne s'applique pas. Bien que le couple de régions soit bien entouré par la grande région, prise indépendamment chacune des deux régions ne l'est pas !

Une règle complémentaire à la règle précédente peut être ajoutée. Elle s'exprime sous la forme linguistique de la façon suivante :

'*Si la région A est très petite, et la région B est très petite*
et la région A a exactement 2 régions voisines, et la région B a exactement 2 régions voisines
alors la fusion doit être très fortement encouragée'. (2.5)

Si les deux régions ont chacune exactement deux régions voisines, nécessairement elles sont chacune comprise dans une des deux régions voisines, ce qui laisse alors une unique région qui circonscrit les régions A et B. La fusion de ces deux régions, A et B, met immédiatement la nouvelle région résultante dans les conditions de la règle précédente. Et si la région qui les entoure est suffisamment grande, la fusion aura alors lieu à l'étape suivante. Il faut noter que la fonction d'appartenance associée à l'expression "exactement deux régions voisines", prend la valeur 0 pour tout autre nombre de voisins que 2. Il s'agit donc d'un sous-ensemble classique, mais qui est naturellement inclus dans l'ensemble des sous-ensembles flous.

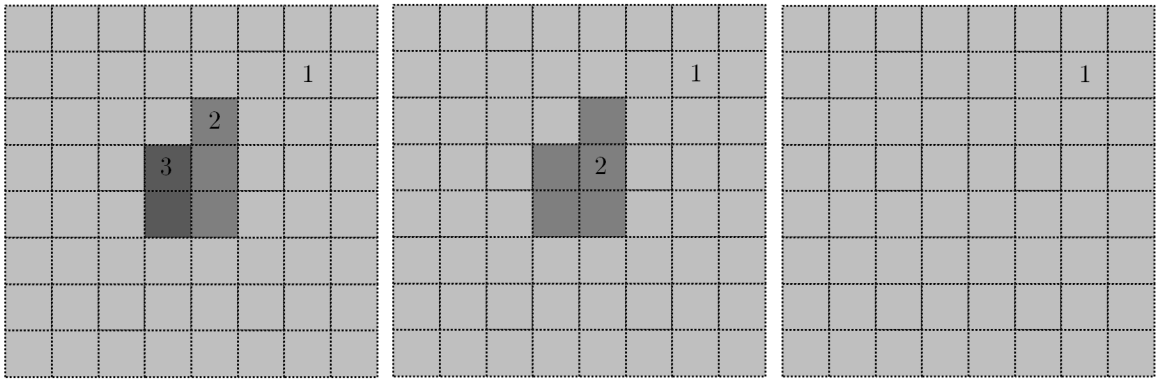


FIGURE 2.5 – Principe du fonctionnement de règle d'inférence floue, pour les paires de petites régions. Un exemple de situation d'application : les régions 2 et 3 sont petites, mais pas suffisamment entourées par la région 1 pour que la règle précédente s'applique. Mais comme elles ont chacune exactement deux voisins, nous les fusionnons. A l'étape suivante, la région deux est toujours petite, et est complètement entourée par la région 1 : elle est à son tour fusionnée.

Gestion des grandes régions

Il est assez courant de rencontrer des cas de figures où de grandes sections de l'image ont des comportements statistiques proches mais sont séparées par une région plus fine. Cela peut être le cas avec deux champs, deux sections de parking divisées par une route ou deux bâtiments sensiblement proches. Dans ce genre de situation, si cette ligne de séparation est trop fine, voire interrompue, il arrive que ces deux grandes sections se retrouvent être connexes à un moment ou un autre dans le processus de segmentation. Si ces deux régions ont un comportement polarimétrique proche, l'algorithme aurait naturellement tendance à les fusionner. Par la suite il n'est alors pas rare de voir l'intégralité de la frontière être fusionnée progressivement dans la grande région qui est alors formée. Nous pouvons éviter cette fusion à l'aide d'une règle sur la longueur de la frontière, la taille et le périmètre. Elle peut s'exprimer sous la forme suivante :

'Si les régions A et B sont très grandes
et que leur frontière est très petite devant leurs périmètres (2.6)
alors la fusion doit être fortement découragée'.

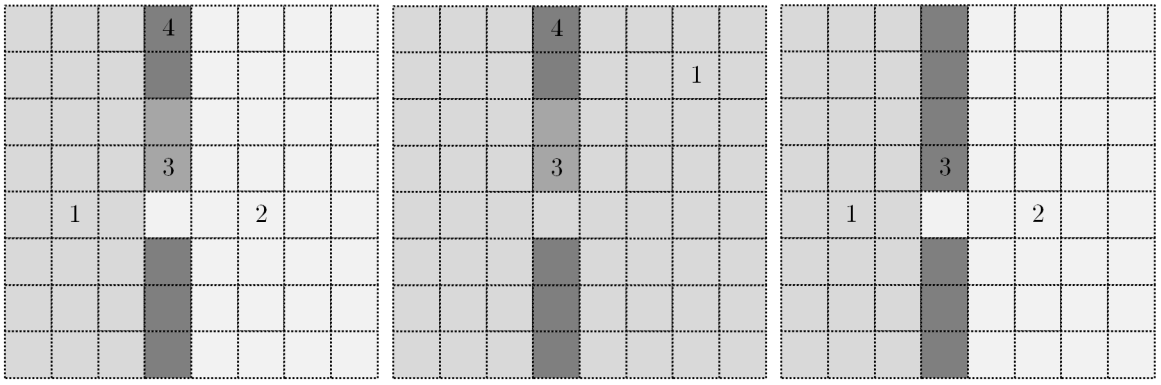


FIGURE 2.6 – Dans l'image de gauche, la situation initiale. Deux grandes régions sont séparées par une ligne brisée. Au milieu, la fusion qui a lieu en l'absence de la règle d'inférence, les deux régions sont fusionnées. A droite la règle empêche la fusion de ces deux régions, au profit d'une fusion mieux choisie.

2.2.4 Lignes et bords

Le processus de segmentation par croissance de régions est un processus local, qui n'a donc aucune vision d'ensemble sur la structure de l'image. Ainsi nous pouvons rencontrer un ensemble de pixels formant une frontière ou une ligne, et c'est d'ailleurs fréquemment le cas dans des images SAR en zones urbaines. Un objet de forme linéaire aura naturellement des difficultés à être préservé dans la segmentation en présence de speckle. En effet, même si pour la majorité des pixels qui le compose, le comportement statistique sera suffisamment proche pour encourager la fusion des différents éléments dont il est formé, il suffit d'un ou quelques pixels aberrants, ou issus de queues de distributions pour

interrompre la fusion et créer ainsi une ligne brisée. Nous avons présenté une règle qui vise à limiter les dégâts induits par les frontières brisées, nous proposons ici une règle permettant de les prévenir.

L'idée générale est d'introduire un paramètre qui indiquera, à chaque pixel et à chaque région, sa potentielle appartenance à un élément linéaire ou une frontière. Nous procédons donc de la manière suivante :

- Nous effectuons une convolution de l'image par un filtre linéaire tournant, pour différentes valeurs d'angle θ_i.
- Cette mesure de contraste, en fonction de l'angle θ_i nous permet de stocker pour chaque pixel, l'angle d'orientation qui maximise le contraste θ_i, ainsi que la valeur de ce contraste permettant de déterminer si nous sommes ou non en présence d'une ligne.
- Lorsque nous fusionnons deux régions, les nouveaux paramètres sont calculés par moyenne pondérée des paramètres précédents (sachant par contre que les valeurs de θ sont discrètes).

Il faut noter que ces paramètres perdent leurs sens au fur et à mesure des fusions et de la création de régions de grandes tailles. Ainsi nous nous restreindrons en pratique aux régions de petites tailles, qui seront majoritairement présentes en début de segmentation. A partir de ces paramètres nous pouvons donc établir la règle suivante :

'Si la région A et la région B sont petites et leurs angles d'orientation sont très proches et si leur coefficient de contraste est élevé alors la fusion doit être très fortement encouragée'. (2.7)

Nous illustrons sur la figure 2.7, le principe de la règle d'inférence. Sur la figure 2.8 nous présentons les résultats d'estimation de l'angle d'orientation ainsi que du contraste pour des sections d'images SAR. Sur cet exemple, on observe que dans le champ nous sommes capables d'établir une différenciation importante entre les sillons. Ces détails seront alors conservés après segmentation.

Gestion des très petites et très grandes régions

Les règles exposées plus haut proposent une solution fine, pour la gestion des grandes et petites régions, et permettent d'éviter la majorité des problèmes liés à la taille des régions. Il n'est cependant pas rare de conserver un certain nombre de régions très petites qui n'auront pas rencontré une des règles précédentes, et auront un comportement polarimétrique très différent de leurs voisines. Inversement, il n'est pas impossible d'observer la construction de très grandes zones, qui bien que correspondant à la topologie de l'image peuvent nuire au temps de calcul et n'apportent plus d'intérêt au niveau de la qualité d'estimation des matrices de covariance (une région ayant un très grand nombre de voisins nécessite beaucoup plus de calculs à chacune de ces fusions). Nous nous contraignons donc à établir deux règles qui permettront, dans ces cas là, et uniquement à des étapes avancées de segmentation, de restreindre la création de trop grandes zones, et de limiter les régions trop petites.

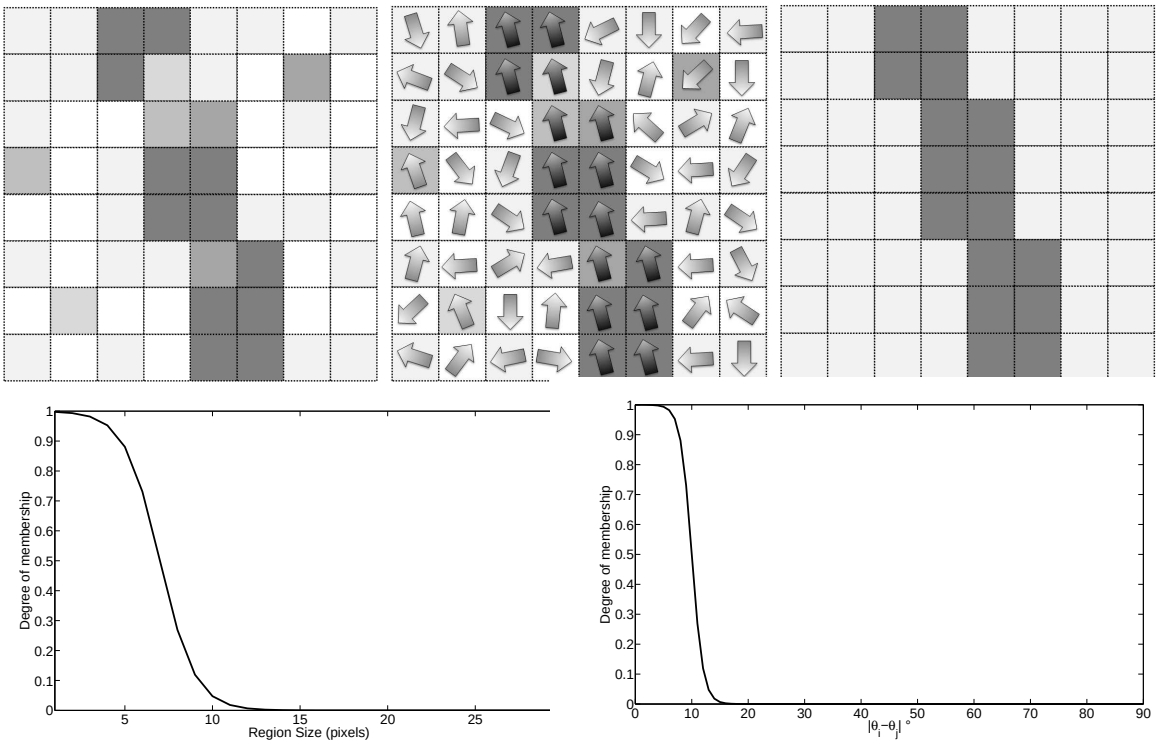


FIGURE 2.7 – Exemple d'application de la règle d'inférence : à chaque pixel est attribué une orientation qui correspond à une structure plus globale qu'à l'échelle du pixel. En fonction des similitudes d'orientation et de contraste, les pixels puis les régions de petite taille sont fusionnés de façon privilégiée.

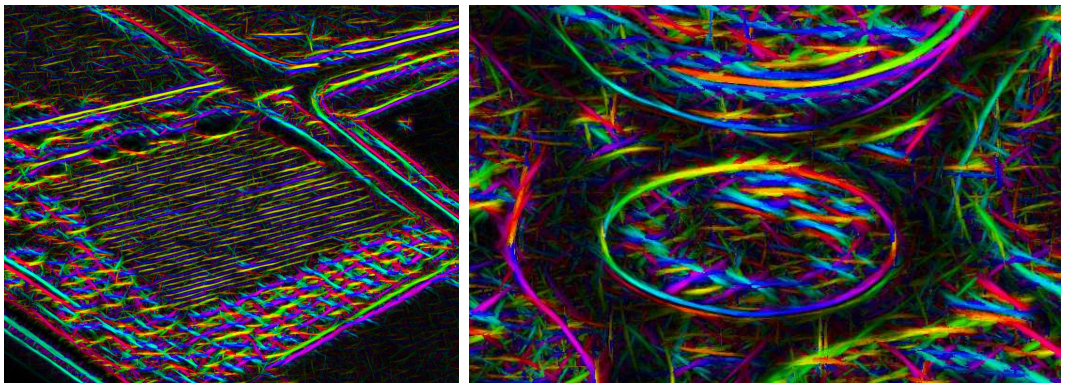


FIGURE 2.8 – Illustration des paramètres d'orientations pour un extrait d'image SAR. En composition colorée nous représentons l'angle d'orientation maximale du pixel, et en intensité la valeur du contraste.

'*Si la région A ou la région B est extrêmement petite alors la fusion doit être très fortement encouragée*'. (2.8)

'*Si la région A et la région B sont très grandes alors la fusion doit être très fortement découragée*'. (2.9)

Autre règles

Nous avons présenté dans ce document les règles les plus importantes, celles dont l'impact est le plus visible car elles interviendront très régulièrement au fur et à mesure du processus de croissance de régions. Nous avons cependant un certain nombre de règles, à l'impact très mineur, pouvant n'être appliquées qu'un nombre minime de fois, ou en ayant un impact extrêmement léger sur les résultats. Nous ne les présenterons pas en détail ici.

2.2.5 Fusion des différentes règles et décision.

Une fois le degré d'appartenance au sous-espace flou associé à chaque règle d'inférence, nous disposons d'un nombre conséquent de degrés d'appartenance. La fusion des régions A et B peut être "fortement encouragée" avec un degré d'appartenance de 0.4, "encouragée" avec un degré d'appartenance de 0.7, et "faiblement découragée" avec un degré d'appartenance de 0. L'objectif de la fusion est de prendre une décision, et de retrouver un paramètre scalaire, que nous utiliserons pour moduler la distance statistique polarimétrique. Ainsi un scalaire supérieur à 1 décourage la fusion, et inversement un scalaire inférieur à 1 l'encourage.

Les règles sur la forme restent pour nous un complément à un processus de croissance de région basé sur une information polarimétrique qui après tout fonctionne correctement toute seule. Nous décidons alors de restreindre l'application de règles de forme uniquement dans le cas ou celles-ci donnent des résultats cohérents. Nous effectuons une opération de *ou exclusif* entre l'ensemble des règles encourageants la fusion et celles qui la découragent. Dans le cas où une des règles de forme indiquerait une fusion à la fois "fortement encouragée" et "faiblement découragée", nous préférons ne prendre aucun risque et le paramètre de forme scalaire sera alors égal à 1.

Dans le cas où l'ensemble des degrés d'appartenance sont cohérents, le calcul du paramètre scalaire se fait alors de la façon suivante :

1. Chaque sous-espace flou associé à une variable linguistique telle que "fortement encouragée" est associé à une valeur scalaire numérique S_i.
2. Pour chaque sous espace flou nous avons calculé le degré d'appartenance $\mu_i \in [0, 1]$.
3. Le scalaire de décision sur la forme S est alors calculé par moyenne pondérée des scalaires, par les degrés d'appartenance, $S = \frac{\sum_i \mu_i S_i}{\sum_i \mu_i}$.

2.2.6 Gestion des grands volumes de données

Une conséquence importante des systèmes radar très haute résolution est l'augmentation importante du volume de données produites. Pour donner un ordre de grandeur, l'image SAR de Toulouse sur laquelle j'ai pu travailler fait 108 Millions de pixels, celle de Salon de Provence plus de 500 Millions de pixels. La plupart des traitements développés et publiés dans la littérature ne s'intéressent que peu à l'optimisation des temps de calcul, même pour les algorithmes utilisant des distributions très complexes. Compte tenu de la chance que j'avais d'avoir à ma disposition de telles images, il aurait été dommage de ne pouvoir les traiter que partiellement. La limitation principale de l'algorithme de croissance de région vient de la recherche des deux régions les plus proches, parmi l'ensemble des couples de régions de l'image. Une idée assez naturelle est de procéder à des segmentations partielles de plusieurs zones de l'image, avant de traiter l'image dans sa globalité. La recherche du minimum se fait alors par minimum local, parmi un nombre bien plus restreint de couples, et est grandement accélérée. En outre les traitements de chacune des sous images sont des opérations indépendantes, qui peuvent donc aisément s'effectuer de façon parallèle sur un processeur multi-coeurs. Lorsque nous travaillons sur de grandes images, nous procédons donc de la façon suivante :

- L'image est découpée en sous images, selon la même puissance de deux $2^{N_{part}}$ selon chaque axe, aboutissant donc à une puissance de 4 de sous images. Les données sont simplement transformées en des tableaux de trois dimensions au lieu des deux précédentes. L'espace mémoire utilisé est donc identique.
- A chaque étape, l'algorithme est appliqué sur chaque sous-image, produisant une segmentation partielle de cette zone. Nous notons par exemple un taux de compression qui correspond entre le nombre initial de régions et le nombre final de régions par un facteur C compris entre 0 et 1. Chaque zone est alors segmentée par un coeur de processeur différent. Ainsi en fonction de la machine disponible les zones peuvent être segmentées en 8 par 8 ou en 12 par 12 images. Comme ces sous images sont de petites tailles, cette segmentation est bien plus rapide puisque la recherche du minimum se fait aisément parmi un nombre bien plus faible de couples de régions.
- Ensuite, quatre régions connexes sont fusionnées en une sous image quatre fois plus grande laquelle subira à son tour une segmentation toujours partielle d'un facteur C. Cette opération est répétée à autant de reprises que la puissance de deux choisie comme subdivision de l'image.
- L'image est maintenant à nouveau complète mais a été compressée (ou segmentée) d'un facteur $C^{N_{part}}$. Nous pouvons alors terminer le processus de segmentation sur cette image de manière linéaire.

Cette méthode nous permet de bénéficier au maximum des processeurs multi-coeur très répandus de nos jours. Nous observons ainsi un gain en temps de calcul d'un facteur 7 pour un processeur Nehalem à 8 coeurs. Cependant, même si chacune des sous régions est calculée de manière linéaire à l'aide d'un processeur monocoeur, cette méthode réduit tout de même de façon extrêmement importante les temps de calcul en réduisant le nombre de couples candidats à chaque étape.

2.3 Résultats et synthèse

Nous ne proposons ici qu'une fraction des données sur lesquelles nous avons eu l'occasion d'appliquer notre algorithme de croissance de région. De très grandes images segmentées seront présentées dans la troisième partie consacrée aux applications. L'exemple présenté sur les figures 2.9, est une portion de l'image de Toulouse, avec une estimation classique, par moyenne spatiale de l'information polarimétrique, et avec une estimation basée sur les résultats de segmentation.

La composition colorée est celle de Pauli (HH+VV en rouge HH-VV en bleu, et 2HV en vert). Nous pouvons voir une réduction importante de l'impact du speckle tout en conservant les caractéristiques spatiale essentielles de l'image (frontières, bords de bâtiments). Un des avantages de travailler sur une image segmentée, pour l'application ultérieure de traitements polarimétriques est aussi la compression importante de l'image. Ainsi une image comme celle ci est compressée a plus de 98%.

Dans la dernière partie des applications, nous présentons des résultats obtenus sur une image de forêt, avec ou sans information polarimétrique, puis en désactivant une majorité des critères de formes. Cependant nous ne pouvons désactiver l'intégralité des critères à moins d'utiliser une distance qui dépend de la taille des régions (problème évoqué dans le chapitre 1).

Nous avons pu confronter aussi les distances issues du modèle SIRV et celles issues du modèle gaussien sur des données réelles. Nous n'avons pas pu constater de différences déterminantes au niveau des résultats (d'où le choix de ne pas les présenter). Nous pouvions nous y attendre compte tenu de l'étude statistique réalisée dans le premier chapitre qui situait la majorité des secteurs de la zone de Toulouse au voisinage de la frontière de Wishart, où les performances des deux modèles sont très proches.

Dans notre algorithme les critères de formes jouent donc un rôle nettement plus important que le choix du modèle (SIRV ou Gaussien).

Il est de toute façon malheureusement très difficile de comparer objectivement deux processus de segmentation ayant des performances proches. A long terme il serait souhaitable de proposer, dans le cadre de la polarimétrie radar, d'une base de données commune ainsi qu'une segmentation idéale de référence. Méthode très répandue dans la communauté du traitement d'image « classique » comme celle de Berkeley.

A défaut de pouvoir quantifier les performances de l'algorithme, je propose ici de quantifier le temps d'exécution de notre algorithme. Sur la figure 2.10 je représente ainsi le temps de calcul pour une image d'un million de pixels sur un processeur Intel Nehalem (serveur datant de l'année 2009). Ainsi en utilisant les huit processeurs disponibles sur cette configuration, nous pouvons obtenir une compression de 90% de l'image en environ 180 secondes. Pour la même image l'execution d'une décomposition matricielle polarimétrique en produit comme celle de Lu et Chipman prend environ 250 secondes.

Dans ce travail, l'objectif était simplement d'adapter une méthode existante au cadre difficile des milieux urbains. La logique floue, utilisée pour la gestion de critères de forme, est dans ce cadre une solution simple et efficace. La plupart des règles de formes que nous avons proposées sont très performantes pour une charge calculatoire minime. Elles seront, je l'espère, utilisée dans la conception d'algorithmes de segmentation plus ambitieux.

Nous présentons également, avec la même composition colorée, des résultats sur une

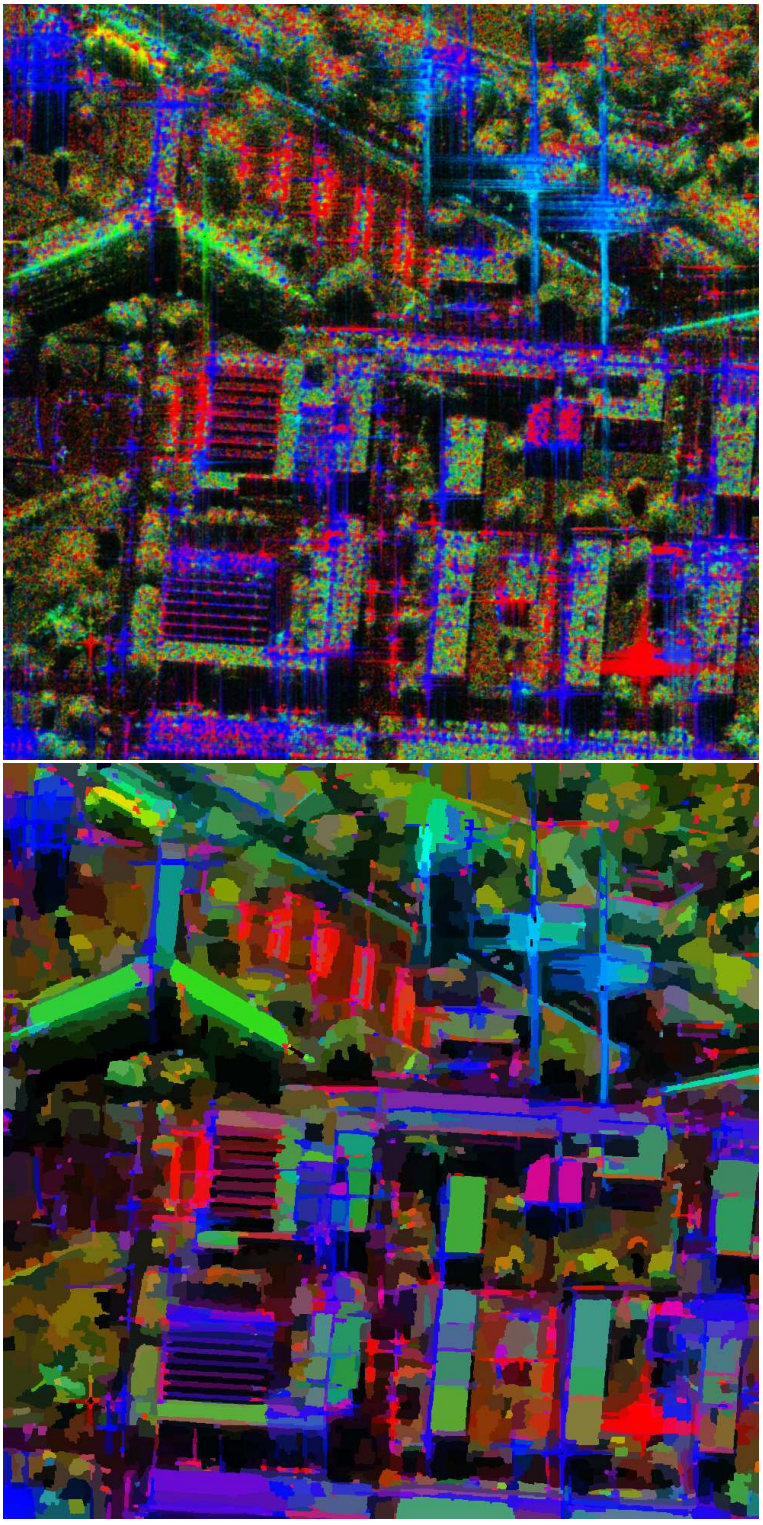


FIGURE 2.9 – Image PolSAR en bande X acquise par RAMSES sur Toulouse. Représentation de Pauli de la matrice de cohérence estimée sur l'image initiale par un filtrage sur une fenêtre carrée de 5x5 pixels en haut, et estimée sur les segments de l'image en bas.

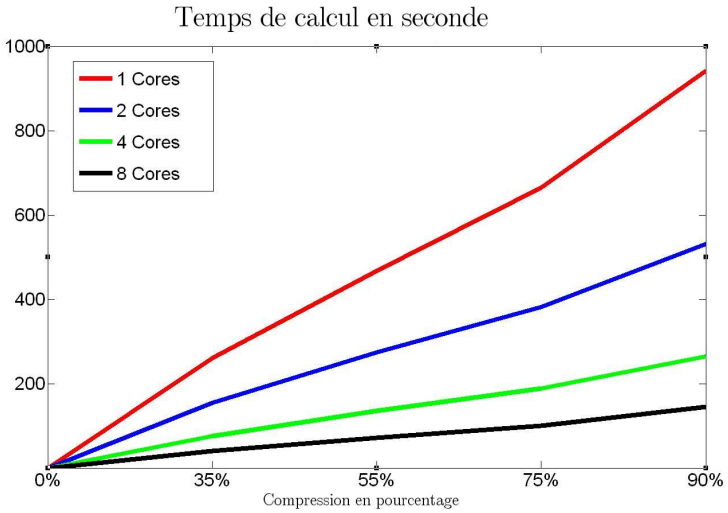


FIGURE 2.10 – Temps de calcul en seconde de l'algorithme de segmentation appliqué sur une image polarimétrique de taille $1000 * 1000$ en fonction du nombre de processeur utilisés.

image obtenue par imagerie de Mueller d'une tumeur cancéreuse. Nous aurons l'occasion d'étudier plus en détail ce type d'image dans la troisième partie. La même représentation colorée est utilisé ici, et nous pouvons voir sur la figure 2.11 l'image avant et après application du processus de croissance de régions. Dans le cadre de l'application sur une image dans un contexte très différents de celui initialement prévu, nous désactivons naturellement certaines règles spécifiques aux milieux urbains. Nous pouvons constater que l'algorithme donne des résultats tout à fait corrects.

2.4 Perspectives

A long terme, si nous étions amenés à perfectionner cet algorithme de croissance de régions, nous proposerions d'exploiter l'information interférométrique lorsqu'elle est disponible. Nous pourrions l'exploiter à la fois comme une information statistique supplémentaire (déjà présentée par Jérome Morio), mais en utilisant aussi des a priori concernant l'évolution des hauteurs qui pourraient intervenir de manière élégante dans les critères de formes. Avec un avant-goût sur le chapitre suivant, il serait aussi possible de proposer des méthodes hybrides, comme nous l'avons fait en détection, permettant d'améliorer les performances des distances non gaussiennes dans des environnements faiblement non gaussiens. Nous avons effectué des tentatives dans ce sens, dont les résultats sont prometteurs, mais dont les justifications théoriques manquent encore aujourd'hui de rigueur pour mériter d'être présentées.

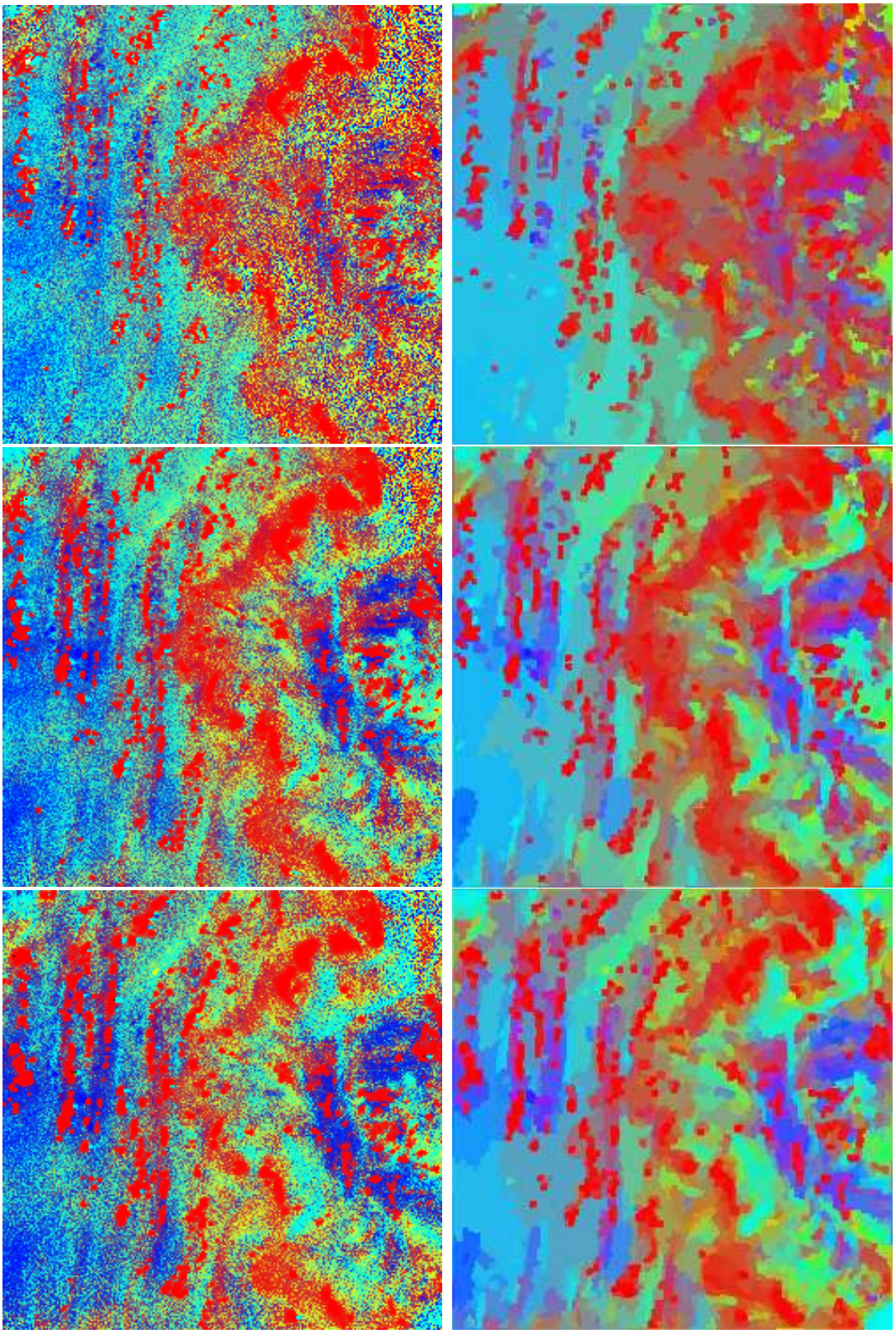


FIGURE 2.11 – Application du processus de croissance de régions sur des images en polarimétrie de Mueller, à gauche l'image avant segmentation, à droite l'image segmentée. De haut en bas des images prises à 500nm, 600nm, et 700nm .

Chapitre 3

La détection

Nous venons d'étudier l'utilisation de la modélisation statistique du signal appliqué aux distances polarimétriques. Ces distances sont utiles dans un grand nombre d'applications qui concernent principalement la classification ou la segmentation d'images. Toutes ces applications ont comme point commun de s'intéresser à des sections importantes des images, des zones, des groupes conséquents de pixels, que nous cherchons à caractériser. Une autre application très répandue consiste à rechercher, dans une image, un comportement bien particulier dont nous aurions connaissance a priori. Cette application s'appelle la détection. La différence entre les deux applications est toutefois assez faible. Il existe par exemple des processus de classification "monoclasse" qui peuvent s'apparenter à de la détection ; inversement des procédés de détection itératifs peuvent aboutir à des classifications d'images complètes.

Dans cette section nous allons nous concentrer sur l'approche la plus standard mise en oeuvre pour la détection, appliquée à la détection d'objets ou de cibles dans des images polarimétriques.

Ces travaux ont donné lieu à un rapport technique ONERA, ainsi qu'à plusieurs publications qui pourront permettre au lecteur d'avoir plus de détails, notamment sur les démonstrations de théorèmes que nous ne présenterons que brièvement ici [49], [50].

La problématique précédente concernait la comparaison de deux groupes de fouillis. Ici nous allons nous intéresser à une problématique qui concerne une seule distribution de fouillis. L'objet de l'étude est alors de déterminer la présence ou l'absence d'une cible ayant une réponse polarimétrique déterministe préalablement connue dans ce fouillis. Comme dans les travaux sur la classification, la modélisation de ce fouillis est aussi essentielle dans l'élaboration des tests de détection. Pour chaque modèle de distribution statistique du fouillis polarimétrique, il existe une stratégie de détection déduite de ces modèles.

Comme pour l'application aux tests statistiques ainsi qu'aux distances polarimétriques, nous nous restreindrons volontairement au modèle classique gaussien et au modèle multiplicatif SIRV avec l'hypothèse de texture déterministe.

3.1 Détecteurs Gaussiens et amplitude inconnue

L'approche la plus courante dans les applications de détection est encore une fois la méthode du rapport de vraisemblance et plus précisément du rapport de vraisemblance généralisé [51] [31]. Comme pour la classification ou l'élaboration de distances statistiques, il convient de poser des hypothèses qui concernent la distribution d'un certain nombre d'observations à notre disposition.

Ici nous supposons que nous avons une ou plusieurs observations d'une position de l'image où nous allons chercher à établir l'existence d'une cible ayant une réponse polarimétrique donnée.

A cela nous ajoutons un certain nombre d'observations voisines à cette position de l'image. Il peut s'agir par exemple d'une couronne circulaire entourant la zone d'étude. Nous pourrons éventuellement garder un certain intervalle de garde afin d'éviter qu'une cible étendue vienne perturber ces observations de l'environnement de la cible.

Les données secondaires que nous noterons $(\hat{\mathbf{k}}_\mathbf{1}, \hat{\mathbf{k}}_\mathbf{2}, \ldots, \hat{\mathbf{k}}_\mathbf{N})$ servent alors à estimer les paramètres de la loi de distribution choisie pour modéliser le clutter. Ici encore, le choix de la loi va déterminer les paramètres à estimer ainsi que des estimateurs pour les paramètres inconnus. Le test du rapport de vraisemblance va alors concerner l'observation "principale" que nous noterons $\hat{\mathbf{y}}$ dans la position où nous allons tester l'existence de la cible.

Le test d'hypothèse peut se poser ainsi : sous l'hypothèse H_0, l'échantillon central $\hat{\mathbf{y}} = \hat{\mathbf{k}}_\mathbf{0}$ suit la même distribution que les échantillons secondaires. Sous l'hypothèse H_1, l'échantillon central $\hat{\mathbf{y}}$ est une somme cohérente d'une réalisation suivant la même distribution que les échantillons secondaires et du produit du vecteur de diffusion recherché $\hat{\mathbf{s}}$ par une amplitude complexe inconnue A.

Nous pouvons poser dans le cas d'une distribution quelconque :

$$\begin{array}{ll} H_0 : & \text{"}\hat{\mathbf{y}} = \hat{\mathbf{k}}_\mathbf{0} \text{ } et \\ & (\hat{\mathbf{k}}_\mathbf{0}, \hat{\mathbf{k}}_\mathbf{1}, \hat{\mathbf{k}}_\mathbf{2}, \ldots, \hat{\mathbf{k}}_\mathbf{N}) \text{ } suivent\ la\ même\ distribution.\text{"} \\ H_1 : & \text{"}\hat{\mathbf{y}} = A\hat{\mathbf{s}} + \hat{\mathbf{k}}_\mathbf{0} \text{ } et \\ & (\hat{\mathbf{k}}_\mathbf{0}, \hat{\mathbf{k}}_\mathbf{1}, \hat{\mathbf{k}}_\mathbf{2}, \ldots, \hat{\mathbf{k}}_\mathbf{N}) \text{ } suivent\ la\ même\ distribution.\text{"} \end{array} \tag{3.1}$$

Cas du modèle gaussien

Comme dans la partie précédente nous noterons Σ la matrice de covariance polarimétrique. *Nous supposons dans l'élaboration des tests que cette matrice est connue* ; Nous l'estimerons dans un second temps à partir des données secondaires. La dimension sera quasiment exclusivement égale à 3 pour les applications monostatiques sans interférométrie. Dans la problématique de la détection en modèle gaussien, l'estimateur de la matrice de covariance polarimétrique reste classiquement la moyenne des $\hat{\mathbf{k}}_\mathbf{i}\hat{\mathbf{k}}_\mathbf{i}^\dagger$. Dans le test du rapport de vraisemblance il convient alors de remplacer l'amplitude A (seul autre paramètre inconnu) par son estimateur au sens du maximum de vraisemblance. Nous obtenons alors le test suivant, qui est donc un rapport de vraisemblance généralisé à comparer à un seuil τ_G :

$$\frac{|\hat{\mathbf{s}}^\dagger \Sigma^{-1} \hat{\mathbf{y}}|^2}{\hat{\mathbf{s}}^\dagger \Sigma^{-1} \hat{\mathbf{s}}} \underset{H_0}{\overset{H_1}{\gtrless}} \tau_G. \tag{3.2}$$

Le calcul du seuil τ_G peut se faire théoriquement. Dans une utilisation classique radar, il est calculé empiriquement de manière à garantir un nombre fixé de fausses alarmes sur l'ensemble de la scène traitée.

Ce test de détection est notamment appelé OGD, pour Optimal Gaussian Detector dans [52]. Dans nos articles, nous avons préféré l'appeler GLRT-GAUSS, (pour Generalized Likelihood Ratio Test) car le terme optimal est malheureusement souvent trompeur en traitement du signal et la notation ne se généralise pas à d'autres rapports de vraisemblance généralisés obtenus avec d'autres modèles.

Cas du modèle SIRV

Dans la problématique de la détection en modèle SIRV l'estimateur de la matrice de covariance polarimétrique est l'estimateur du point fixe calculé sur les données secondaires. Dans le test du rapport de vraisemblance il convient alors de remplacer l'amplitude A et la texture par leurs estimés par maximum de vraisemblance [53]. Nous obtenons alors le test suivant, qui est donc un rapport de vraisemblance généralisé à comparer au seuil τ_S :

$$\frac{|\hat{\mathbf{s}}^\dagger \Sigma^{-1} \hat{\mathbf{y}}|^2}{(\hat{\mathbf{s}}^\dagger \Sigma^{-1} \hat{\mathbf{s}})(\hat{\mathbf{y}}^\dagger \Sigma^{-1} \hat{\mathbf{y}})} \underset{H_0}{\overset{H_1}{\gtrless}} \tau_S. \tag{3.3}$$

Ce test de détection a été obtenu de différentes manières. Il possède donc de nombreux noms : GLRT-LQ dans [54], Asymptotic BORD dans [52], NMF dans [55] et ACE dans [56] et [57]. Par cohérence avec les différents modèles statistiques, nous avons choisi d'y faire référence sous le nom de GLRT-SIRV. Le calcul du seuil théorique pour une fausse alarme donnée se fait avec la formule $P_{FA} = (1 - \tau_S)^{m-1}$. Cependant comme pour le cas gaussien il est rare de retrouver en pratique la fausse alarme désirée à l'aide du seuil théorique. Ainsi dans son utilisation pratique, il est toujours souhaitable d'ajuster ce seuil pour obtenir exactement le nombre de fausses alarmes voulu. C'est d'ailleurs important pour comparer de manière probante différents tests de détection sur le même jeu de données. Contrairement au GLRT-GAUSS, dont le seuil de détection dépend de la loi de texture, ce test de détection est naturellement *CFAR* ou TFAC (Constant False Alarm Rate, Taux de Fausses Alarmes Constant) [58].

Comparaison rapide des performances

Avant de rentrer plus en détail sur l'étude de ces deux tests de détection, nous réalisons rapidement une simulation par Monte Carlo, dans le cadre d'un fouillis K-distribué présenté sur la figure 3.3. Le rapport signal à bruit (SNR) est calculé en sortie de filtre adapté. Il prend donc en compte la matrice de covariance, et le gain d'intégration :

$$\text{SNR} = |A|^2 \frac{\hat{\mathbf{s}}^\dagger \Sigma^{-1} \hat{\mathbf{s}}}{\hat{\mathbf{k}}^\dagger \Sigma^{-1} \hat{\mathbf{k}}}, \tag{3.4}$$

Nous pouvions nous en douter suite à l'étude préalable sur les distances, mais ici encore, même dans un fouillis non gaussien, le test de détection SIRV n'est pas nécessairement

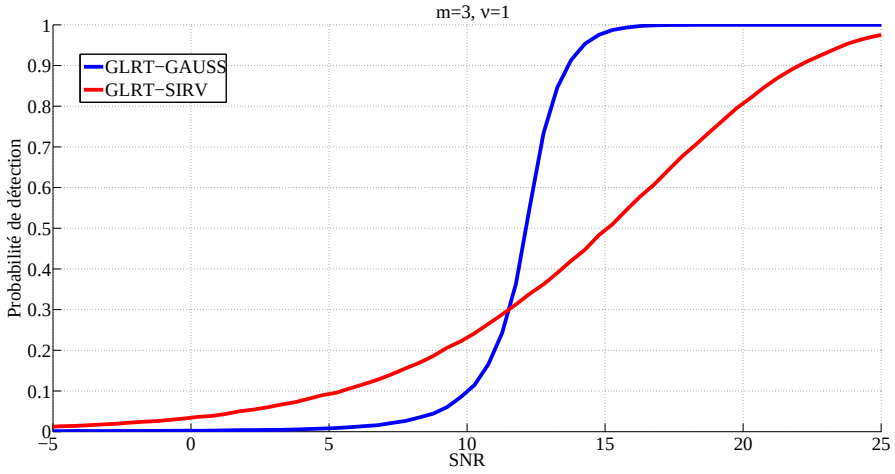


FIGURE 3.1 – Comparaison des performances du GLRT-GAUSS et du GLRT-SIRV dans un environnement non gaussien (Bruit K distribué $\nu = 1$)

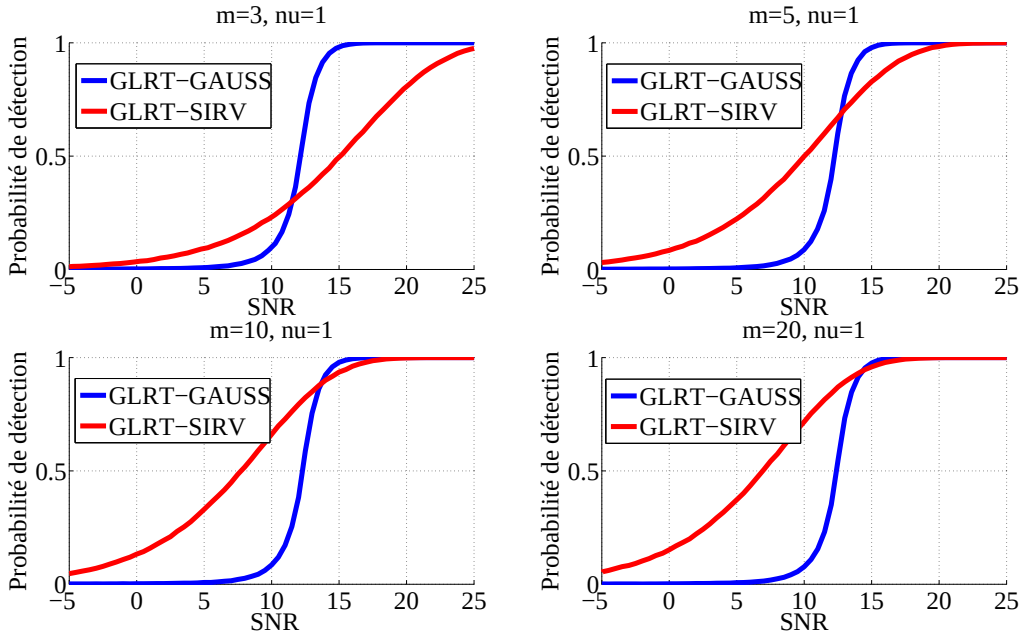


FIGURE 3.2 – Comparaison des performances du GLRT-GAUSS et du GLRT-SIRV dans un environnement non gaussien (Bruit K distribué $\nu = 1$ pour plusieurs dimensions)

meilleur. Ici nous représentons un ν fixé et une dimension polarimétrique classique. Nous observons qu'il semble y avoir un SNR à partir duquel le GLRT-GAUSS devient plus performant que le GLRT-SIRV, même dans ce fouillis non gaussien. Dans le cadre des distances statistiques nous n'avons pas observé ce comportement. En effet il existait un seuil à partir duquel le fouillis devenait suffisamment impulsionnel pour justifier l'intérêt du modèle SIRV de manière uniforme. Ici, dans tous les cas de figure, et même pour les distributions les plus impulsionnelles, il semblerait qu'il existe un SNR à partir duquel le GLRT-GAUSS devient plus performant. Cette remarque a été le point de départ de l'ensemble des travaux que nous avons pu effectuer sur la détection dans du fouillis non gaussien.

La deuxième remarque, plus relative à la polarimétrie, concerne les performances observées dans notre cas $m = 3$. En nous plaçant dans ce cas à faible dimension, nous avons aussi exhibé un comportement du GLRT-SIRV qui n'avait à notre connaissance pas vraiment été mis en relief jusqu'à présent. Si on étudie les performances pour différentes dimensions comme sur la figure 3.2, on observe que plus la dimension est faible, plus les performances du GLRT-SIRV se dégradent par rapport à celles du GLRT-GAUSS. Dans le cadre de notre étude plus générale sur les détecteurs, nous avons expliqué plus en détail les effets de la dimension du signal. Ici nous nous sommes naturellement restreints aux faibles dimensions ; les plus faibles performances du GLRT-SIRV dans ce domaine sont alors malheureusement inévitables. Il est assez facile de comprendre pourquoi. La normalisation par l'amplitude de la mesure correspond au sacrifice d'un degré de liberté, dans le cadre d'une mesure en contenant très peu. A forte dimension, ce sacrifice devient proportionnellement négligeable.

Un argument en faveur du modèle SIRV, souvent mis en avant, est sa capacité à inclure le modèle gaussien : la texture suit alors une distribution de Dirac en 1. Il est souvent sous-entendu que le GLRT-SIRV doit pouvoir aussi fonctionner, bien qu'avec des performances évidemment plus faibles, dans un fouillis gaussien. Cependant, l'amplitude de cette baisse de performance est assez conséquente, comme nous pouvons le voir sur la figure 3.3. Nous faisons tendre le paramètre de forme de la K-distribution vers l'infini, et nous nous approchons d'une distribution gaussienne. Dans ce domaine, les performances du GLRT-SIRV se dégradent très nettement par rapport à celles du GLRT-GAUSS. Il faut noter que cette dégradation est d'autant plus forte que la dimension est faible. Il est donc important de noter que bien que le GLRT-SIRV fonctionne sur du bruit gaussien, les performances sont très nettement inférieures à un détecteur adapté. Le choix du détecteur à utiliser dans une image comprenant à la fois des zones gaussiennes et non gaussiennes est délicat.

La recherche d'une explication à cette intersection des deux courbes nous a permis de mettre en relief un nombre important de propriétés sur le GLRT-SIRV qui n'avaient pas ou peu été étudiées dans la littérature. Cette interprétation concernant les performances a pu être faite par le biais d'une interprétation géométrique que je vais présenter brièvement.

Interprétation géométrique.

Les tests de détection peuvent s'interpréter géométriquement si on considère la forme quadratique qui intervient à de nombreuses reprises dans ces tests : $\hat{\mathbf{a}}^{\dagger}\Sigma^{-1}\hat{\mathbf{b}}$. Cette forme

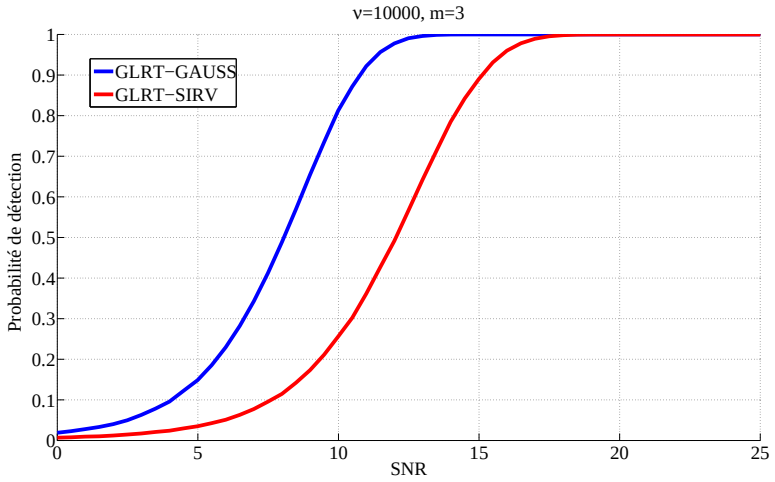


FIGURE 3.3 – Comparaison des performances entre le GLRT-GAUSS et le GLRT-SIRV dans un environnement gaussien (Bruit K distribué $\nu = 10000$ quasiment gaussien)

quadratique définit un produit scalaire, ce qui permet de proposer l'interprétation géométrique suivante :

1. Le GLRT-GAUSS consiste en la projection du vecteur mesuré $\hat{\mathbf{y}}$, sur l'espace signal défini par le vecteur signal $\hat{\mathbf{s}}$ (connu et recherché). Il s'agit donc de la mesure de la norme de la projection de ce vecteur dans l'espace signal. Le seuil de détection τ_G définit alors une partition de l'espace en deux. Au delà du seuil, le vecteur mesuré est accepté sous l'hypothèse H_1. En dessous du seuil, il est rejeté. Ainsi, dans le cas réel, deux hyperplans situés en τ_G et $-\tau_G$ permettent de représenter la géométrie de ce détecteur dans l'espace sur la figure 3.4. L'espace d'acceptation du détecteur GLRT-GAUSS est noté Φ_G ainsi que sa zone complémentaire $\overline{\Phi_G}$.
2. Le GLRT-SIRV peut s'interpréter comme la projection du vecteur $\hat{\mathbf{y}}$ sur $\hat{\mathbf{s}}$, normalisée à la fois par la norme du vecteur signal, mais aussi par la norme du vecteur mesuré. Il s'agit alors, d'après la relation de Cauchy-Schwarz, du cosinus au carré de l'angle formé par les vecteurs $\hat{\mathbf{y}}$ et $\hat{\mathbf{s}}$. Ainsi la zone d'acceptation du test de détection est, dans le cas réel, un double cône d'angle θ_S tel que $\cos(\theta_S)^2 = \tau_S$. Nous notons Φ_S et $\overline{\Phi_S}$ la zone d'acceptation et la zone complémentaire de rejet du test de détection GLRT-SIRV. Elles sont aussi représentées sur la figure 3.4 dans le cas réel.

Cette approche nous permet d'interpréter géométriquement les comportements observés sur les performances de détection. Ainsi les meilleures probabilités de détection à faible niveau de signal à bruit pour le GLRT-SIRV s'expliquent par l'acceptation des données de très faibles énergies situées dans la pointe du cône, comprise dans la zone $\Phi_S \bigcap \overline{\Phi_G}$. Il suffit que la direction du vecteur mesuré soit suffisamment proche pour qu'il soit accepté. Inversement, à fort rapport signal à bruit cette zone d'exclusivité est surpassée par la zone $\overline{\Phi_S} \bigcap \Phi_G$ située à l'extérieur du cone. L'angle d'ouverture du cône dépend du seuil du GLRT-SIRV lié à la dimension du signal comme nous l'avons vu dans sa formule théorique. Ainsi plus la dimension est faible, plus le cône d'acceptation est fin, réduisant

d'autant plus les performances.

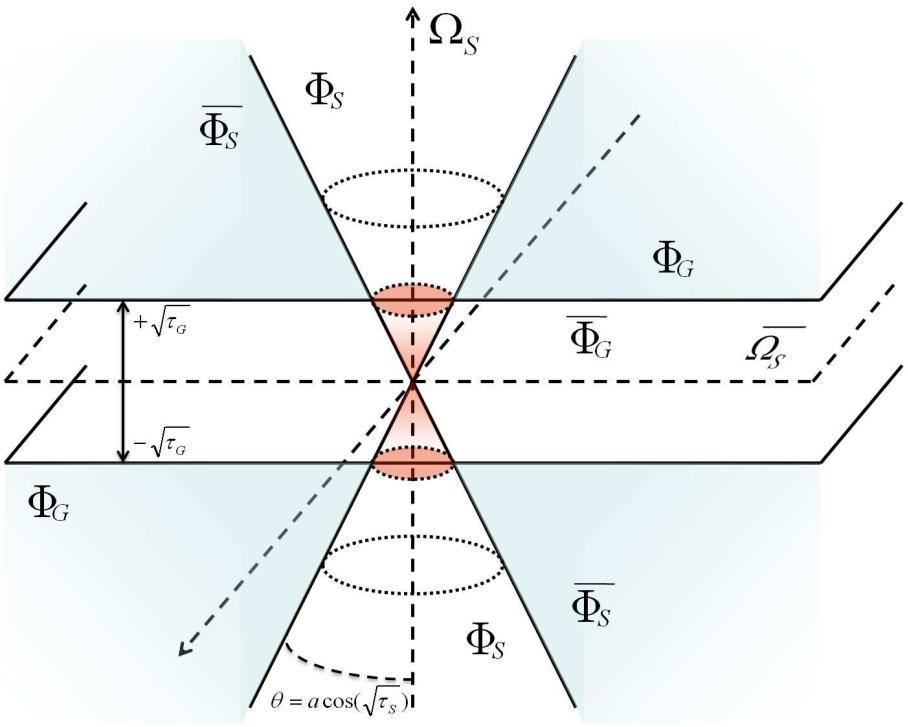


FIGURE 3.4 – Comparaison de la géométrie entre le GLRT-GAUSS et du GLRT-SIRV dans le cas réel. En rouge, la zone d'acceptation du GLRT-SIRV uniquement, en bleu celle du GLRT-GAUSS uniquement.

Influence du Rapport Signal à bruit, SNR

A l'aide de cette interprétation géométrique, nous avons pu démontrer que le comportement observé se vérifie quelle que soit la distribution de texture :

Théorème 1. *Si la distribution de fouillis est sphériquement invariante (modèle produit), alors, pour une fausse alarme fixée $P_{FA} > 0$, il existe un rapport signal à bruit SNR $\gamma_{P_{FA}}$ tel que,*

$$\forall \gamma \geq \gamma_{P_{FA}}, P_D^G(P_{FA}, \gamma) \geq P_D^S(P_{FA}, \gamma).$$

De plus, pour toute distribution non bornée, l'inégalité est stricte.

La démonstration de ce théorème est détaillée dans un article soumis dans la revue IEEE Signal Processing [49]. Nous ne la présenterons pas en détail ici. Elle consiste à construire une image par rotation de la région $\Phi_S \bigcap \overline{\Phi_G}$ autour du vecteur signal (multiplié par une amplitude A bien choisie). Ainsi en exploitant l'invariance sphérique de la distribution on montre qu'au delà de cette amplitude, la probabilité sous l'hypothèse H_1 que l'observation appartienne à l'espace $\Phi_S \bigcap \overline{\Phi_G}$ est plus faible que la probabilité que l'obervation appartienne à l'espace $\overline{\Phi_S} \bigcap \Phi_G$.

Influence du vecteur signal "cible".

Une difficulté qui apparaît immédiatement lors de la mise en oeuvre de ces tests de détection, est la définition du vecteur cible. C'est une problématique qui a été abordée dans [59]. En effet nous cherchons un comportement polarimétrique préalablement connu $\hat{\mathbf{s}}$. Cela suppose d'avoir naturellement un a priori sur la signature des objets d'intérêt. Cet a priori est tout à fait raisonnable, puisque par exemple les cibles humaines, ou véhicules ont des comportements polarimétriques bien particulier, comme par exemple les effets de double rebond. Cependant cette connaissance a priori n'est qu'imparfaite, il est plus probable que nous connaissions la réponse de la cible avec une certaine incertitude, et il est même sûr qu'au travers des effets comme la géométrie d'acquisition ou de rotation de la cible elle-même la signature réelle de la cible diffère de la signature recherchée. C'est une problématique qui a aussi été abordée hors du cadre de la polarimétrie dans [60] [61] et [62].

C'est une problématique générale qui ne concerne pas particulièrement le choix du modèle statistique utilisé dans l'élaboration des tests de détection. Mais en simulant une imprécision sur le signal recherché, nous avons pu constater une différence de robustesse entre les deux processus de détection. Notons $\hat{\mathbf{s}}_\mathbf{b}$ le vrai signal, diffusé par la cible et mesuré par le radar, ou le polarimètre, qui serait mesuré en cas d'absence total de fouillis ou de bruit de mesure. $\hat{\mathbf{s}}$ reste le signal "attendu" par l'opérateur, suite à une modélisation imparfaite de la cible ou de la chaîne de mesure. Nous définissons alors l'angle

$$\alpha = \arccos \sqrt{\frac{|\hat{\mathbf{s}}^\dagger \boldsymbol{\Sigma}^{-1} \hat{\mathbf{s}}_\mathbf{b}|^2}{(\hat{\mathbf{s}}^\dagger \boldsymbol{\Sigma}^{-1} \hat{\mathbf{s}})(\hat{\mathbf{s}}_\mathbf{b}^\dagger \boldsymbol{\Sigma}^{-1} \hat{\mathbf{s}}_\mathbf{b})}}. \tag{3.5}$$

Cet angle correspond à l'erreur faite sur le signal cible, et nous permet de mesurer et de quantifier cette erreur. Naturellement cet angle est compris entre 0 et $\pi/2$, cette dernière valeur correspondant à un vecteur signal supposé parfaitement orthogonal au vrai signal.

Nous appliquons, dans les mêmes conditions, les deux tests de détection en dimension $m = 3$, pour différentes valeurs de α, et représentons les résultats sur la figure 3.5. Nous pouvons observer qu'une valeur croissante de l'angle d'erreur sur le signal réel entraîne des baisses de performances totalement différentes pour le GLRT-GAUSS et le GLRT-SIRV. Pour le GLRT-GAUSS cette perte de performance est très légère (au point qu'elle est à peine visible sur la figure à cette échelle). Pour le GLRT-SIRV cette perte de performance est très rapide, et pour un α suffisamment grand, fait tendre la probabilité de détection vers 0 quand le SNR augmente.

L'interprétation géométrique expliquée plus haut nous permet de comprendre ce phénomène. L'hyperplan du GLRT-GAUSS est une partition de l'espace qui garantit qu'à partir d'un certain rapport signal à bruit, le vecteur de mesure dépasse le niveau de l'hyperplan. Une erreur sur le signal, traduite par un angle, correspond à une rotation de cet hyperplan. Seules les réalisations et mesures les plus extrêmes et donc les moins probables sont les plus affectées. Ceci explique la faible influence de l'erreur sur le vecteur signal.

Pour le GLRT-SIRV, par contre, la zone d'acceptation en forme de cône est uniquement sensible à l'information angulaire. Si le vecteur $\hat{\mathbf{s}}_\mathbf{b}$ n'est pas inclus dans le cône, l'augmentation du rapport signal à bruit n'a aucune chance de ramener le signal mesuré

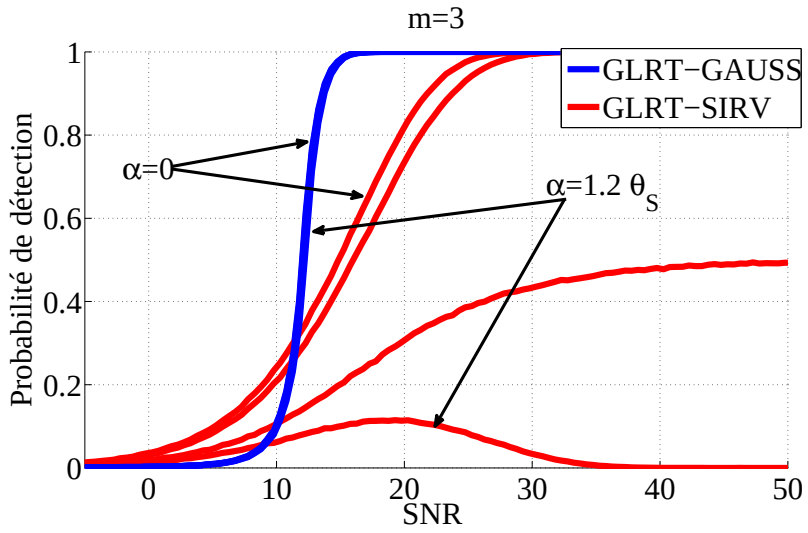


FIGURE 3.5 – $\alpha = \{0; \quad \theta_S/2; \quad \theta_S; \quad 1.2\theta_S\}$

dans le cône. C'est même l'effet contraire ! A fort rapport signal à bruit, la distribution du signal mesuré s'éloigne de plus en plus du cône. Cela explique pourquoi sur la figure 3.5, on voit d'abord la probabilité de détection augmenter, puis décroître. Ce manque de robustesse est d'autant plus flagrant à faible dimension car l'angle d'ouverture du cône est très faible.

Robustesse

On notera $P_D^G(\gamma, \alpha)$ et $P_D^S(\gamma, \alpha)$ les probabilités de détection du GLRT-GAUSS et du GLRT-SIRV respectivement, pour un rapport signal à bruit γ et une erreur d'angle α sur le vecteur signal réel. Ce constat sur l'influence du vecteur signal et la robustesse de ces deux détecteurs nous ont amené à proposer la définition de la robustesse d'un détecteur suivante :

Définition. *Un détecteur est α-robuste si* $\lim_{\gamma \to +\infty} P_D(\gamma, \alpha) = 1$.

Avec cette définition, il est possible de démontrer que le GLRT-GAUSS est robuste pour tout $\alpha \in [0, \pi/2[$, et que le GLRT-SIRV n'est robuste que pour $\alpha \in [0, \theta_S[$. De plus la probabilité de détection tend vers 0 pour tout angle $\alpha > \theta_S$ (les démonstrations sont aussi dans les articles [50] et [49]). Pour le GLRT-SIRV, la robustesse dépend donc de la dimension dans le sens où elle dépend du seuil, par l'intermédiaire de θ_S. Plus la dimension est grande, plus le détecteur est robuste. Naturellement en polarimétrie monostatique la robustesse est donc malheureusement assez faible.

Premier bilan

Il s'avère donc, comme pour les distances statistiques, que le choix entre approche gaussienne ou non gaussienne n'est pas évident, même dans du fouillis aussi impulsionnel

qu'une K-distribution avec un paramètre $\nu = 1$. Nous avons pu faire les constatations suivantes :

1. Le GLRT-GAUSS
   - donne de très bonnes performances en fouillis gaussien.
   - donne de bonnes performances à fort SNR même en fouillis non gaussien.
   - est robuste.
2. Le GLRT-SIRV
   - donne de moins bonnes performances que le GLRT Gauss en fouillis gaussien, particulièrement à faible dimension.
   - donne de bonnes performances à faible et moyen SNR en fouillis non gaussien.
   - est peu robuste.

Ces constatations ainsi que l'interprétation géométrique montrent que dans le cadre de la détection polarimétrique, aucun des deux outils n'est adapté aux distributions modérément impulsionnelles ou aux images présentant à la fois des sections gaussiennes et non gaussiennes. Notre idée est de construire des processus de détection géométriquement à l'aide du produit scalaire évoqué plus haut. Il s'agit d'exploiter les propriétés à faible et moyen SNR du GLRT-SIRV, ainsi que la qualité à fort SNR et la robustesse du GLRT-GAUSS. Nous avons ainsi proposé une famille de détecteurs que nous pouvons qualifier d'hybrides, qui exploitent à la fois l'information angulaire (donc exclusivement polarimétrique), et l'information de puissance, ou de norme projetée. La section suivante leur est consacrée.

3.2 Les détecteurs hybrides

.

Dans cette section nous allons proposer une famille de détecteurs fondée sur une logique géométrique inspirée des deux processus de détection que nous venons d'étudier. L'objectif était initialement d'obtenir un compromis entre les performances sur fouillis gaussiens et non gaussiens, ainsi que d'améliorer la robustesse en fouillis non gaussien. Les résultats dépassent nos espérances et permettent de concevoir des détecteurs robustes et plus performants que les deux méthodes pour des gammes importantes de rapport signal à bruit en fouillis non gaussien.

Notons $\hat{\mathbf{y}}_S = (\hat{\mathbf{s}}^\dagger \Sigma^{-1} \hat{\mathbf{y}})/(\hat{\mathbf{s}}^\dagger \Sigma^{-1} \hat{\mathbf{s}})$ la projection de $\hat{\mathbf{y}}$ sur l'espace signal Ω_S et $\hat{\mathbf{y}}_N = \hat{\mathbf{y}} - \hat{\mathbf{y}}_S$ la projection de $\hat{\mathbf{y}}$ sur l'espace bruit. Nous définissons la famille de détecteur hybride $\mathcal{D}$ de la façon suivante :

Définition. *Un détecteur appartient à la famille $\mathcal{D}$ si sa région de décision Ω_D peut s'écrire :*

$$\Omega_D = \{\hat{\mathbf{y}} : \|\hat{\mathbf{y}}_S\| \underset{H_0}{\overset{H_1}{\gtrless}} \sqrt{\tau_D} f(\|\hat{\mathbf{y}}_N\|), \tag{3.6}$$

où τ_D est un seuil calculé pour assurer la fausse alarme désirée et $f(x)$ une application de R^+ dans R vérifiant les propriétés suivantes :

Efficacité $\forall x \in \mathbb{R}^+, \frac{df}{dx}(x) \geq 0$, *i.e.* $f(x)$ *est croissante sur* $\mathbb{R}^+$;

Robustesse $\exists M \in \mathbb{R}^+ : \forall x \in \mathbb{R}^+, f(x) \leq M$, *i.e.* $f(x)$ *est bornée sur* $\mathbb{R}^+$.

La propriété d'eficacité et de robustesse a été choisie afin de garantir que cette classe contienne majoritairement des "bons" détecteurs. En effet une fonction f quelconque peut être utilisée, mais ne produira pas nécessairement un détecteur intéressant. Plus on s'éloigne de l'espace signal angulairement, plus l'amplitude nécessaire pour être acceptée est importante. L'objectif étant d'accepter, comme le GLRT-SIRV, les mesures de faible amplitude qui sont très proches du signal, et comme le GLRT-GAUSS d'accepter les mesures plus éloignées de l'espace signal, si elles ont une amplitude suffisante.

La propriété de robustesse consiste à imposer qu'asymptotiquement, le comportement du détecteur va suivre celui du GLRT-GAUSS et va permettre ainsi une plus grande robustesse à une erreur sur le signal recherché. Il est en fait possible de démontrer le théorème suivant :

Théorème 2. *(Robustesse de la famille* $\mathcal{D}$*) Tout détecteur appartenant à la famille* $\mathcal{D}$ *est* α*-robuste pour tout* $0 \leq \alpha < \pi/2$.

La démonstration est fournie dans [49]. Cette définition permet de regrouper ces différentes approches et aussi de construire plus simplement des détecteurs encore plus efficaces. Nous allons maintenant présenter les détecteurs pouvant ou non appartenir à la classe définie ci-dessus.

Exemples de détecteurs hybrides, et appartenance à la classe $\mathcal{D}$.

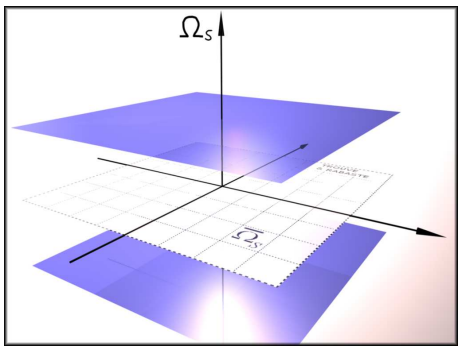


FIGURE 3.6 – Zone d'acceptation du détecteur GLRT-GAUSS

1. *Le GLRT-GAUSS* est un cas particulier des détecteurs de classe $\mathcal{D}$. Il correspond à la fonction $f(x) = 1$. Il vérifie bien les deux propriétés d'efficacité et de robustesse. La partition de l'espace de ce détecteur est représentée sur la figure 3.6.

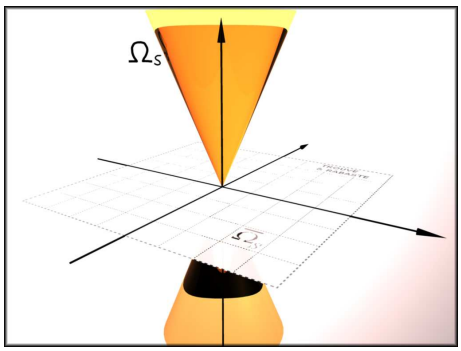


FIGURE 3.7 – Zone d'acceptation du GLRT-SIRV

2. *Le GLRT-SIRV*, bien que définissable par la fonction $f(x) = x$, n'est pas borné. Il vérifie uniquement la propriété d'efficacité, mais comme nous l'avons vu il n'est pas robuste, et n'appartient donc pas à la classe $\mathcal{D}$. La partition de l'espace de ce détecteur est représenté sur la figure 3.7.

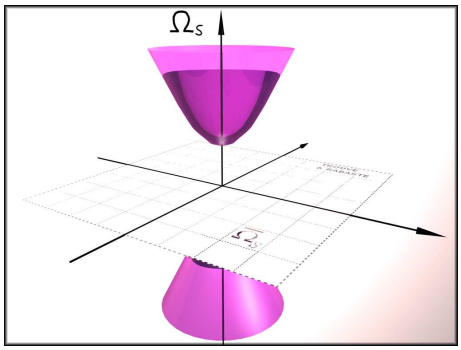


FIGURE 3.8 – Zone d'acceptation du détecteur de Kelly

3. *Le Détecteur de Kelly*, défini dans [63], par le test

$$\frac{|\hat{\mathbf{s}}^{\dagger}\Sigma^{-1}\hat{\mathbf{y}}|^2}{(\hat{\mathbf{s}}^{\dagger}\Sigma^{-1}\hat{\mathbf{s}})(1+\hat{\mathbf{y}}^{\dagger}\Sigma^{-1}\hat{\mathbf{y}})} \underset{H_0}{\overset{H_1}{\gtrless}} \tau_K, \tag{3.7}$$

correspond à la fonction $f(x) = \sqrt{1+x^2}$. Comme le GLRT-SIRV, il vérifie la propriété d'efficacité, mais n'est pas robuste. La partition de l'espace de ce détecteur est représentée sur la figure 3.8.

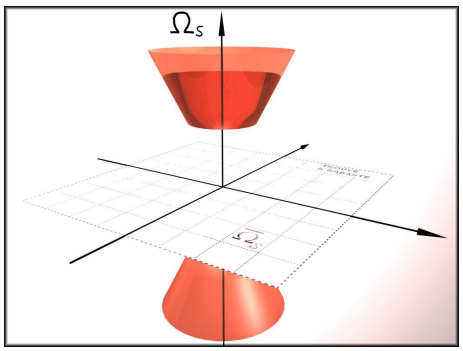


FIGURE 3.9 – Zone d'acceptation du détecteur ASB

4. *Le détecteur ASB* (Pour Adaptive Sidelobe Blanker), proposé par Richmond dans [64]. Richmond est le premier à avoir proposé un processus de détection hybride, inspiré des deux détecteurs GLRT-GAUSS et GLRT-SIRV. Il effectue ainsi une opération logique de type ET entre le GLRT-GAUSS et le GLRT-SIRV. Il peut être décrit pas l'application suivante :

$$f(x) = \begin{cases} 1 & \text{pour} \quad x \in [0, 1/\sqrt{\tau_{ASB}}], \\ \sqrt{\tau_{ASB}}x & \text{pour} \quad x \in [1/\sqrt{\tau_{ASB}}, +\infty[, \end{cases} \tag{3.8}$$

Comme le précédent, ce détecteur est efficace, mais non robuste car il a un comportement asymptotiquement équivalent à un GLRT-SIRV. La partition de l'espace de ce détecteur est représentée sur la figure 3.9.

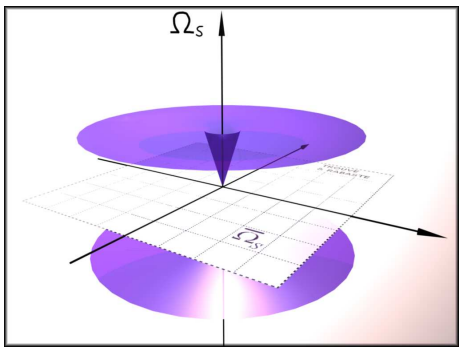


FIGURE 3.10 – Zone d'acceptation du détecteur OU

5. *La famille OU :* dans la logique précédente, il est possible de construire un détecteur en utilisant l'opération logique OU. Nous l'avons ainsi proposé dans [50]. Il peut être défini par la fonction,

$$f(x) = \begin{cases} \sqrt{\tau_{OR}}x & \text{pour} \quad x \in [0, 1/\sqrt{\tau_{OR}}], \\ 1 & \text{pour} \quad x \in [1/\sqrt{\tau_{OR}}, +\infty[. \end{cases} \tag{3.9}$$

Ce détecteur vérifie les propriétés d'efficacité et de robustesse. Il est intéressant car il permet d'accepter les signaux de faible amplitude suffisamment proches du signal, mais permet aussi d'accepter les signaux de forte amplitude même s'ils sont moins

proches du signal recherché. La partition de l'espace de ce détecteur est représentée sur la figure 3.10.

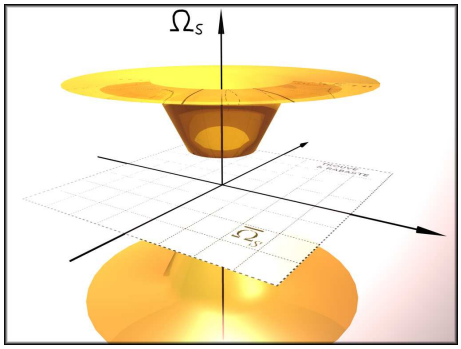


FIGURE 3.11 – Zone d'acceptation du détecteur PWL3

6. *Le détecteur PWL3*, (Pour PieceWise Linear, ie linéaire par morceau). Cette famille de détecteur peut être vue comme l'amélioration logique des deux détecteurs précédents. Il est défini par une fonction f linéaire par morceaux à 3 paramètres (2 fixés, 1 déterminé par la fausse alarme). Il peut aussi être interprété comme l'opération logique OU, appliquée entre le détecteur ASB, et le détecteur GLRT-GAUSS. Ainsi il peut s'écrire :

$$f(x) = \begin{cases} 1 & \text{pour} \quad x \in \left[0, \frac{1}{\sqrt{\tau_{PWL,1}}}\right], \\ \sqrt{\tau_{PWL,1}}x & \text{pour} \quad x \in \left[\frac{1}{\sqrt{\tau_{PWL,1}}}, \frac{\sqrt{\tau_{PWL,2}}}{\sqrt{\tau_{PWL,1}}}\right], \\ \sqrt{\tau_{PWL,2}} & \text{pour} \quad x \in \left[\frac{\sqrt{\tau_{PWL,2}}}{\sqrt{\tau_{PWL,1}}}, +\infty\right[. \end{cases} \tag{3.10}$$

Par rapport au détecteur OU, ce détecteur permet de gagner en performance à rapport signal à bruit intermédiaire, ce qui correspond à des plages intéressantes de probabilités de détection (de 0.5 à 0.9). Il sacrifie par contre les performances à très faible rapport signal à bruit, mais où les performances sont faibles (Probabilité de détection <0.3). Il appartient à la famille $\mathcal{D}$ car il est à la fois efficace et robuste. La partition de l'espace de ce détecteur est représentée sur la figure 3.11. Dans la même idée il est possible de généraliser cette approche avec des détecteurs du type PWLN, linéaires par morceau avec N degrés de liberté.

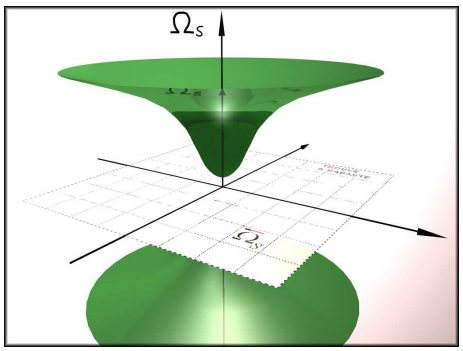


FIGURE 3.12 – Zone d'acceptation du détecteur BELL

7. *Le détecteur BELL*, pour "cloche" (En pratique nous proposons ainsi une frontière en forme de gaussienne, mais utiliser à nouveau cette appellation porterait très fortement à confusion !). L'idée est ici de proposer un détecteur ayant un comportement plus continu et progressif, tout en approchant les performances globales du détecteur PWL3. Il peut ainsi être défini par la fonction :

$$f(x) = 1 + \sqrt{\tau_{BELL,1}}\left(1 - \exp(-\tau_{BELL,2}x^2)\right). \tag{3.11}$$

   La partition de l'espace de ce détecteur est représentée sur la figure 3.12.

Nous avons ainsi montré que de nombreux processus de détection adaptatifs présentés dans la littérature peuvent s'écrire sous la forme proposée. Toutefois, seule une partie d'entre eux vérifient la propriété de robustesse que nous désirons. Les trois nouveaux détecteurs proposés ici, exemples parmi de nombreuses possibilités, ont été choisis pour leur propriété de robustesse ainsi que leurs bonnes performances à la fois en bruit gaussien et en bruit non gaussien. Les propriétés de ces détecteurs sont résumées sur le tableau 3.1. Nous allons proposer dans un premier temps des résultats d'application de ces détecteurs sur des données simulées, puis dans le cadre d'un scénario radar polarimétrique synthétique.

3.3 Performances

3.3.1 Performances théoriques.

Nous avons ici appliqué, comme précédemment, les différents détecteurs sur du fouillis synthétique, gaussien, non gaussien ainsi qu'avec une erreur sur le signal. Sur la figure 3.13, nous présentons les performances des détecteurs OU, $PWL3$ et $BELL$ ainsi que des deux détecteurs classiques dans notre fouillis polarimétrique fortement non gaussien K distribué de paramètre de forme $\nu = 1$. Nous pouvons donc observer que le détecteur $PWL3$ et, encore plus le détecteur $BELL$ permettent d'obtenir des performances qui sont, dans la gamme de $[8 - 15dB]$, meilleures que le GLRT-SIRV et le GLRT-GAUSS.

Sur la figure 3.14, nous présentons les performances des mêmes détecteurs dans le cas d'un fouillis polarimétrique gaussien. Nous constatons que l'ensemble des détecteurs hybrides ont des performances extrêmement proches du GLRT-GAUSS, ce qui n'est pas le cas pour le GLRT-SIRV. Les détecteurs que nous avons proposés sont donc à la fois

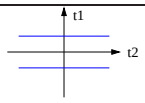
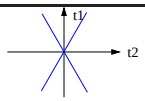
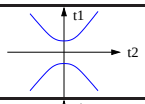
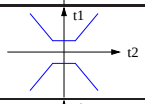
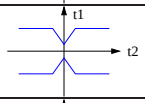
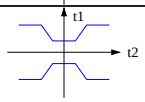
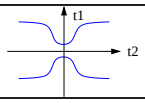
Détecteurs	Fonctions	$\mathcal{D}$	α-Robustesse	Région de décision
GLRT-GAUSS	$f(x) = 1$	oui	$\left[0, \frac{\pi}{2}\right[$	
GLRT-SIRV	$f(x) = x$	non	$[0, \theta_S[$	
Kelly	$f(x) = \sqrt{1+x^2}$	non	$[0, \theta_K[$	
ASB	$f(x) = \begin{cases} 1 & \text{pour} \quad x \in [0, 1/\tau_1] \\ \tau_1 x & \text{pour} \quad x \in [1/\tau_1, +\infty[\end{cases}$	non	$[0, \theta_{ASB}[$	
OU	$f(x) = \begin{cases} \tau_1 x & \text{pour} \quad x \in [0, 1/\tau_1] \\ 1 & \text{pour} \quad x \in [1/\tau_1, +\infty[\end{cases}$	oui	$\left[0, \frac{\pi}{2}\right[$	
PWL3	$f(x) = \begin{cases} 1 & \text{pour} \quad x \in [0, 1/\tau_1] \\ \tau_1 x & \text{pour} \quad x \in [1/\tau_1, \tau_2/\tau_1] \\ \tau_2 & \text{pour} \quad x \in [\tau_2/\tau_1, +\infty[\end{cases}$	oui	$\left[0, \frac{\pi}{2}\right[$	
BELL	$f(x) = 1 + \tau_1 \left(1 - \exp(-\tau_2 x^2)\right)$	oui	$\left[0, \frac{\pi}{2}\right[$	

TABLE 3.1 – Resumé des propriétés des différents détecteurs.

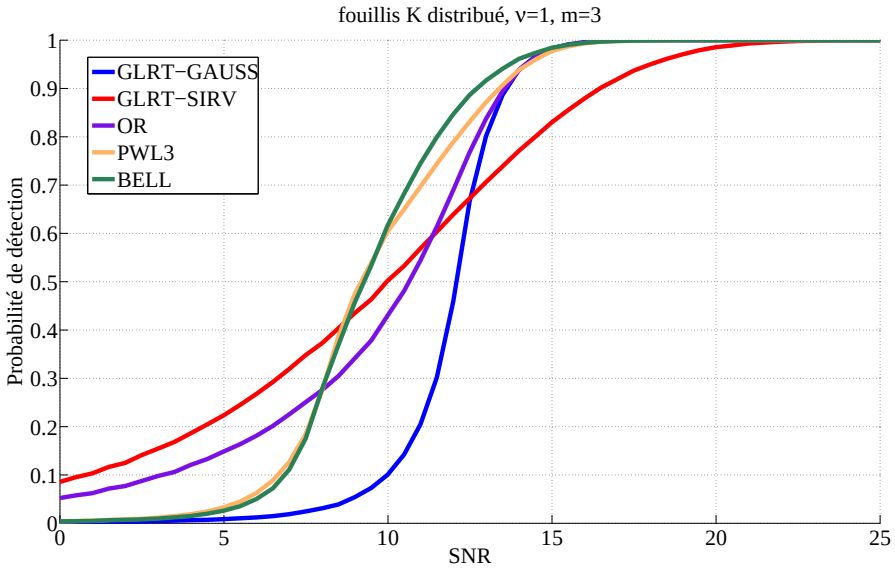


FIGURE 3.13 – Performances dans du fouillis polarimétrique K distribué pour les trois détecteurs proposés, OU, PWL3 et BELL, ainsi que les deux détecteurs classiques GLRT-GAUSS et GLRT-SIRV.

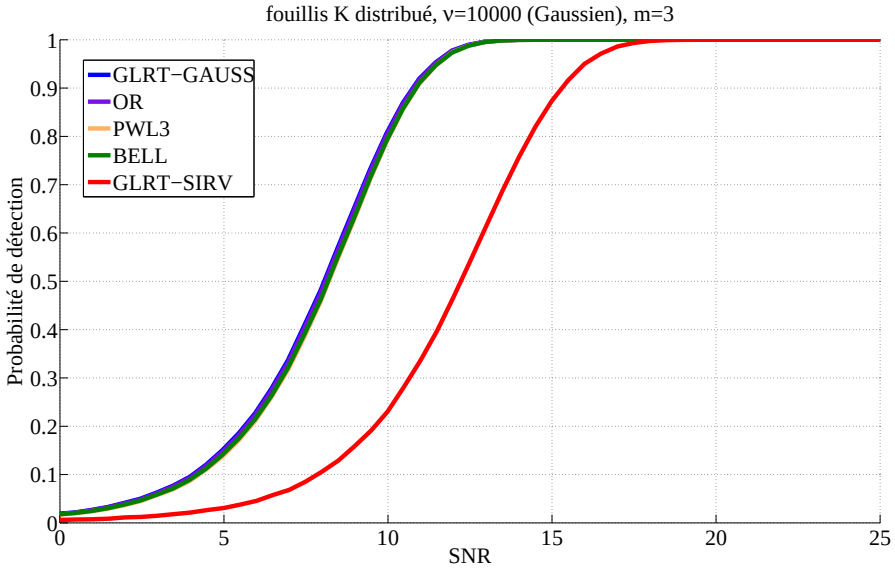


FIGURE 3.14 – Performances dans du fouillis polarimétrique gaussien pour les trois détecteurs proposés, OU, PWL3 et BELL, ainsi que les deux détecteurs classiques GLRT-GAUSS et GLRT-SIRV ; $P_{FA} = 10^{-3}$.

très performants dans des situations fortement non gaussiennes, mais offrent aussi de très bonnes performances dans un fouillis gaussien.

Sur la figure 3.15 nous présentons les mêmes détecteurs pour le même fouillis non gaussien, mais en présence d'une erreur sur le signal recherché. Nous montrons ainsi que les détecteurs hybrides ont des performances, certes dégradées, mais bien moins affectées que le GLRT-SIRV.

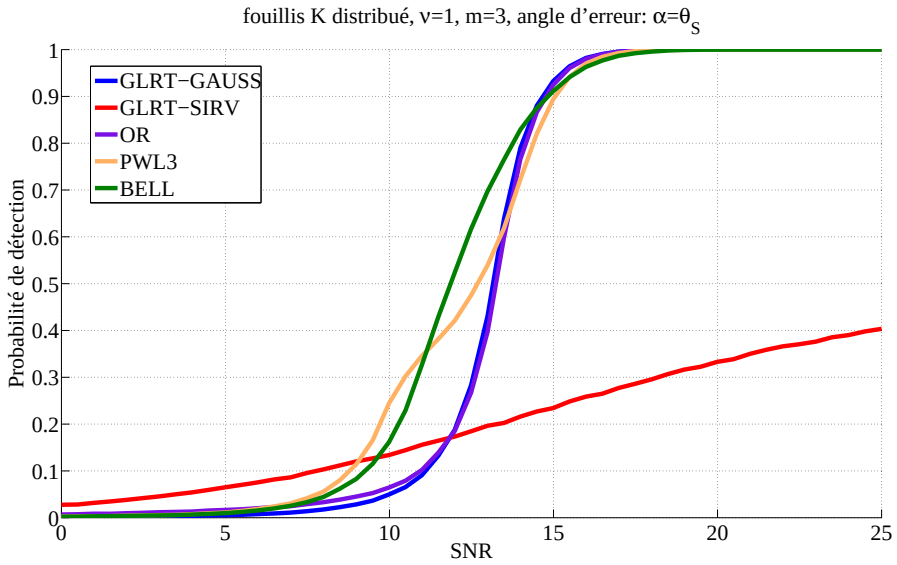


FIGURE 3.15 – Performances dans du fouillis polarimétrique K-distribué avec une erreur sur le signal recherché pour les trois détecteurs proposés, OU, PWL3 et BELL, ainsi que les deux détecteurs classiques GLRT-GAUSS et GLRT-SIRV ; $P_{FA} = 10^{-3}$.

3.3.2 Performances sur une image SAR synthétique

Les performances théoriques sont donc très intéressantes. Cela dit, nous voudrions montrer que cette approche hybride, ainsi que le calcul des seuils, qui peut paraître plus contraignant que pour les schémas de détection classiques, est tout à fait réalisable dans un cas pratique. Nous avons choisi d'utiliser un simulateur de scènes polarimétriques SAR nommé MOCEM, que nous allons présenter brièvement. Ce simulateur présente une étape séduisante de validation des performances entre les calculs théoriques sur simulation de type Monte Carlo et les applications sur données réelles. En effet l'application d'un processus de détection sur une image réelle est rarement convaincante à elle seule car le nombre de cibles connues est très faible, et il est impossible d'exhiber un gain de performances de quelques dBs ou d'une petite variation de probabilité de détection. Un simulateur SAR, parce qu'il permet de produire de nombreuses réalisations de la même scene, permet de tracer des courbes comparables aux courbes de performances théoriques dans un contexte qui est très proche de celui rencontré dans une image réelle.

MOCEM

MOCEM est un logiciel développé par la DGA et Alyotech Technologies qui permet de générer des images SAR à partir de modèles 3D habillés électromagnétiquement par des matériaux. Il prend à la fois en compte le calcul de RCS (surface équivalente radar), les multiples réflexions et les interactions avec l'environnement [65].

MOCEM prend en entrée un modèle 3D de la scène et des objets que l'on veut imager. Nous en voyons un exemple sur la figure 3.16 et 3.17, pour lesquels nous avons généré un bâtiment en 3D, de dimensions réalistes (Les dimensions réelles du bâtiment modélisé ont été gracieusement fournies par le CNES.) Il faut ensuite définir la zone capturée, l'angle de gisement et d'incidence et le matériau de chaque composante de nos objets. Ce choix induit naturellement un a priori statistique sur la réponse polarimétrique et radiométrique de chacun des matériaux, en fonction de l'angle d'incidence. Les matériaux pré-définis par le logiciel sont tout à fait suffisants pour produire des scénarios réalistes pour notre application. Compte tenu de la rapidité du logiciel nous pouvons exploiter de façon statistique de multiples réalisations de la même image, afin de produire des courbes de performances.

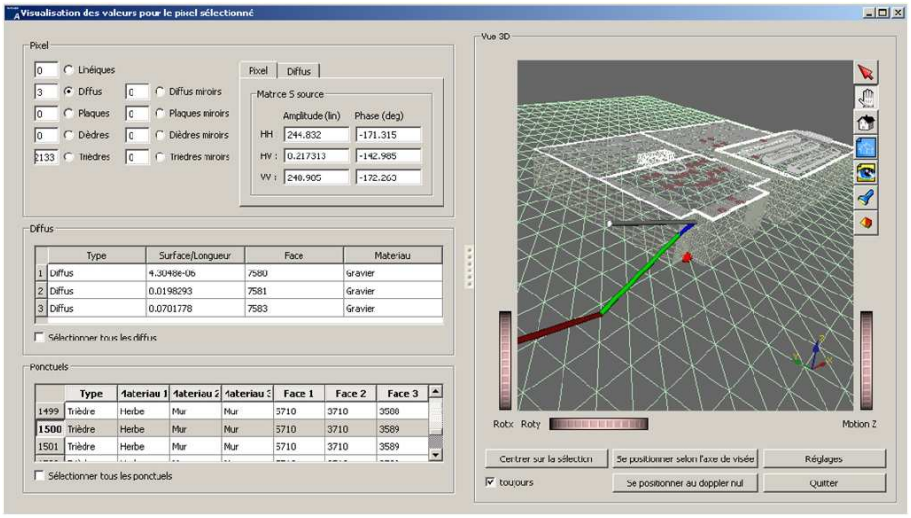


FIGURE 3.16 – Représentation d'une fenêtre de MOCEM. Celle-ci affiche le modèle 3D étudié et les rayons associés au calcul de la contribution d'un ponctuel sélectionné

La scène

Nous avons construit une scène semi urbaine en 3 dimensions à l'aide d'un logiciel dédié. Nous représentons cette scène en vue optique sur la figure 3.18, les couleurs et textures de la vue optique sont représentatives des matériaux que nous choisissons lors de l'importation de la scène sous MOCEM. Nous choisissons un tel scénario car il correspond aux types de données réelles sur lesquelles nous avons pu travailler. La scène comporte ainsi un nombre conséquent d'éléments urbains, bâtiments, une station service, une tour et un pont métallique, mais aussi des zones naturelles et sensiblement moins structurées. La

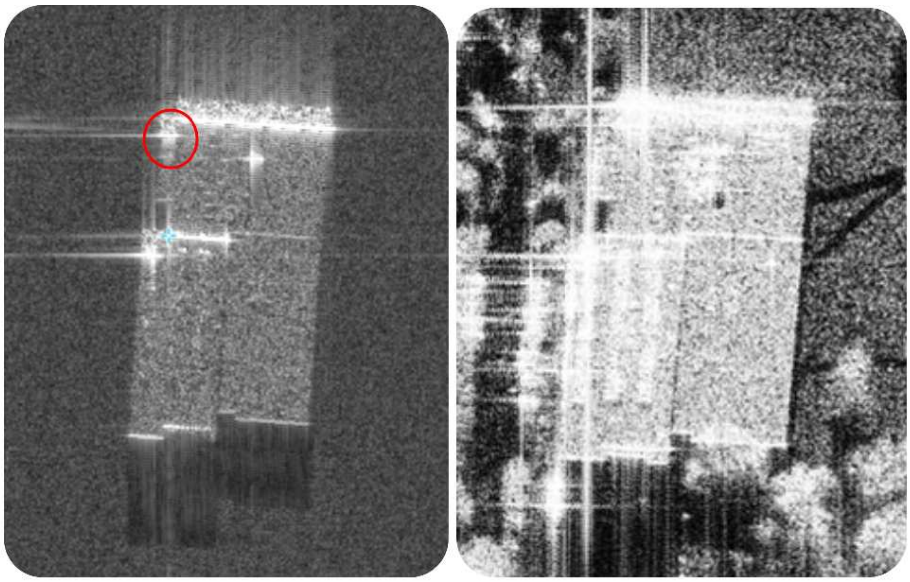


FIGURE 3.17 – Image générée par MOCEM à gauche - A droite, image SAR réelle de référence pour le bâtiment concerné.

figure 3.19 représente l'image résultante de la synthèse sous MOCEM pour la polarisation horizontale.

La légère différence par rapport aux tests théoriques est la gestion de la fausse alarme. En pratique, dans les courbes précédentes, nous n'étions pas concernés par l'estimation du fouillis "voisin", ou des matrices de cohérence polarimétrique qui interviennent dans les tests de détection. Dans un premier temps nous allons utiliser le principe classique de la moyenne mobile pour estimer les matrices de cohérence ainsi que le niveau de bruit moyen environnant. La matrice de cohérence intervient à la fois sur la projection sur l'espace signal, et sur la projection sur l'espace bruit. Concernant la projection sur l'espace signal il convient de remarquer que celle ci va dépendre du niveau de bruit moyen ambiant. Naturellement le détecteur GLRT-GAUSS n'est pas un outil "CFAR" ou "TFAC" (Constant False Alarm Rate, ou Taux de fausse alarme constant). La méthode classique pour obtenir un test TFAC est donc de normaliser le test par un niveau de bruit moyen, estimé sur un voisinage. Ainsi le seuil global du test pour l'ensemble de l'image est le même. Nous verrons plus tard que cette méthode souffre des effets de bords. Le GLRT-SIRV est par contre lui naturellement et théoriquement TFAC, et seul l'estimation de la matrice de cohérence est nécessaire.

Nous utilisons donc la même fenêtre de voisinage, circulaire pour l'estimation des matrices de cohérence, ainsi que pour l'estimation du niveau de bruit moyen. L'estimation des matrices de cohérence, est celle qui est naturellement adaptée au modèle qui a donné naissance au test de détection. Nous utilisons ainsi la méthode du point fixe pour le GLRT-SIRV et l'estimateur classique de la moyenne pour le GLRT-GAUSS. Dans la scène nous ne considérons que les positions valides de l'image, qui sont donc en ligne de vue direct du radar, afin d'éviter de biaiser les courbes de performances. Par contre, l'ensemble des

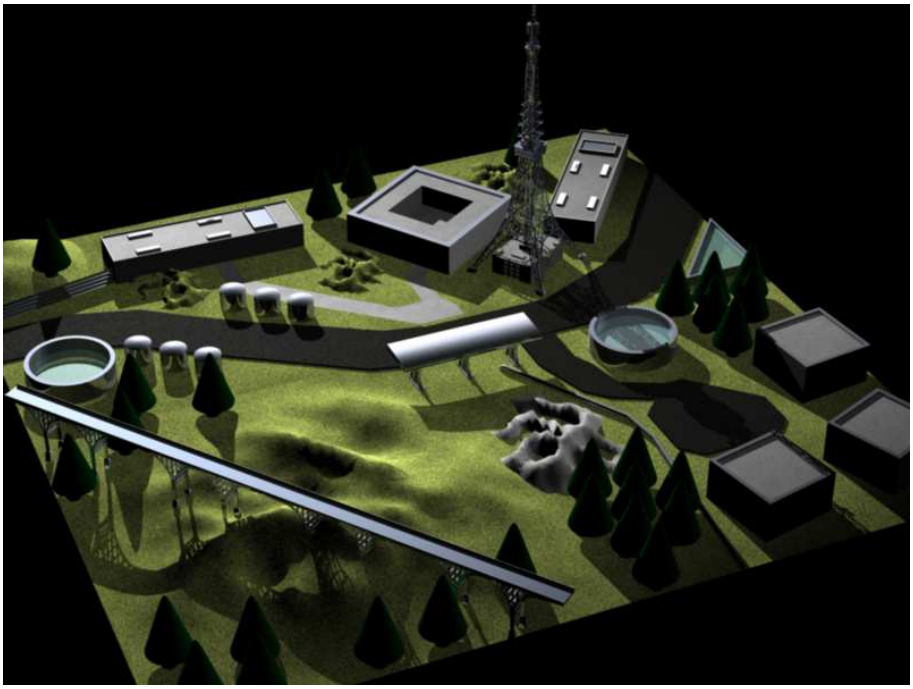


FIGURE 3.18 – Image de la scène synthétique péri-urbaine

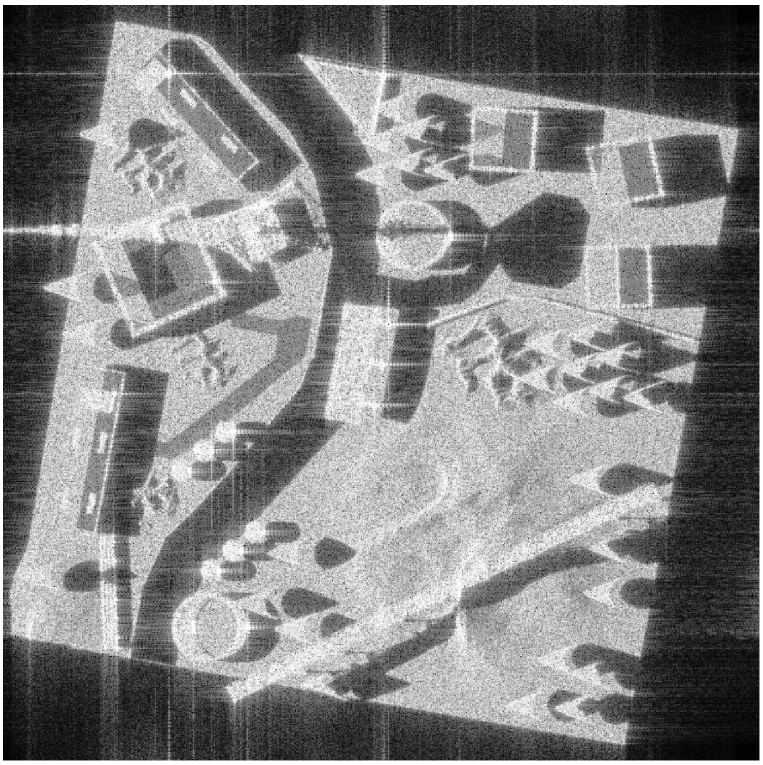


FIGURE 3.19 – Exemple de rendu de la scène par MOCEM, polarisation horizontale

positions de l'images en ligne de vue sont testées. Les courbes de performances obtenues ont des structures relatives au scénario. Par ailleurs, nous faisons remarquer qu'un vrai calcul de rapport signal a bruit n'a pas de sens compte tenu des divers niveaux de fouillis rencontrés dans l'image. Nous représentons alors uniquement l'amplitude de la cible. Il convient alors de s'intéresser uniquement aux positions relatives des différentes courbes.

Sur la figure 3.20 nous testons la présence d'une cible dont le vecteur polarimétrique est parfaitement connu, pour les détecteurs habituels sur l'image synthétique. Plusieurs réalisations sont utilisées afin d'obtenir des courbes non bruitées.

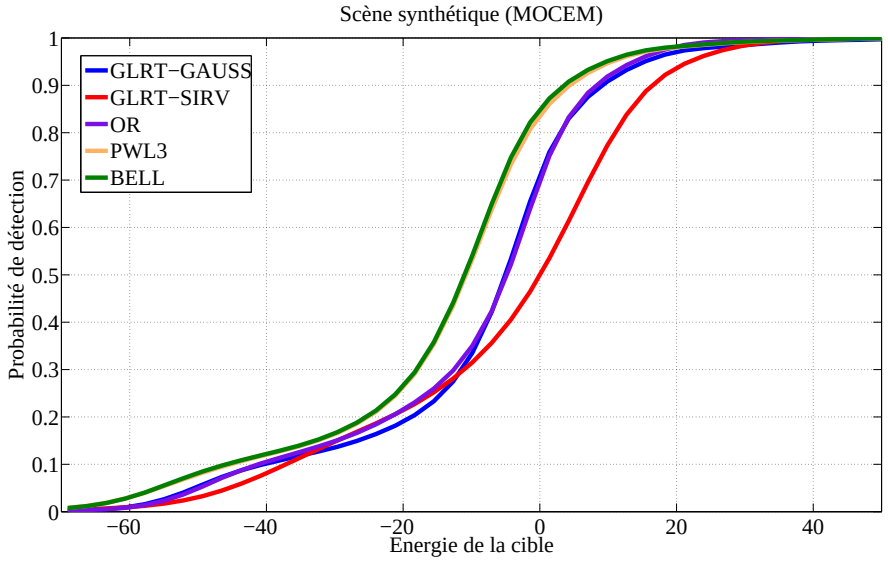


FIGURE 3.20 – Probabilité de détection dans la scène synthétique pour les différents détecteurs, probabilité de fausse alarme $P_{FA} = 10^{-3}$.

Sur la figure 3.21 nous représentons les mêmes résultats à l'exception d'une erreur supposée sur la signature polarimétrique recherchée.

Nous pouvons donc confirmer la mise en oeuvre pratique des outils de détection hybride ainsi que le calcul pratique des différents seuils de détection. Nous constatons, comme théoriquement, un gain important de l'approche hybride par rapport aux schémas de détection classiques qu'étaient le GLRT-GAUSS et le GLRT-SIRV. L'effet d'une erreur sur la modélisation de la signature polarimétrique est elle aussi conforme à nos observations sur les courbes théoriques.

3.3.3 Utilisation de la segmentation d'image pour l'amélioration des processus de détection.

Nous venons de mettre en oeuvre de façon pratique les nouveaux détecteurs que nous avons pu développer. Nous avons toutefois soulevé le problème de l'estimation du fouillis voisin, ou du choix des données secondaires, nécessaires à l'estimation des paramètres tel que la matrice de cohérence ou le niveau de bruit moyen. Le choix classique, comme pour

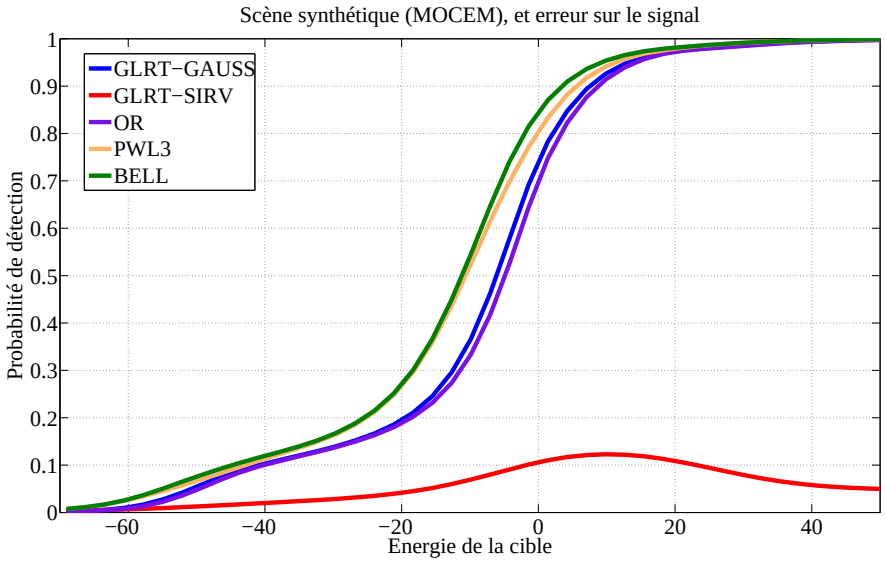


FIGURE 3.21 – Probabilité de détection dans la scène synthétique pour les différents détecteurs avec une erreur sur le signal, probabilité de fausse alarme $P_{FA} = 10^{-3}$.

toutes les problématiques en polarimétrie, consiste à utiliser le voisinage spatial, avec des fenêtres plus ou moins finement choisies (circulaires, carrées, avec ou sans intervalle de garde, etc.). Nous avons par ailleurs développé un outil de partitionnement de l'image, sous la forme d'un algorithme de croissance de région. Une idée naturelle est de croiser les deux outils afin d'exploiter cette segmentation et d'effectuer un processus de détection qui s'adapte à la structure de l'image. Ainsi, nous pouvons estimer les matrices de cohérence ainsi que le niveau de bruit "par région", au lieu de l'estimer par voisinage. L'amélioration espérée se ferait alors majoritairement ressentir à proximité des fortes variations spatiales de l'image qui nuisent à l'estimation locale.

Principe

Le processus est donc très simple, nous appliquons d'abord l'algorithme de segmentation à l'image, puis effectuons les mêmes schémas de détection présentés plus haut. Seule l'estimation des paramètres se fait en fonction de la région d'appartenance du pixel testé au lieu d'utiliser une fenêtre de voisinage.

Nous présentons sur la figure 3.22, les résultats de la segmentation sur l'image synthétique. Nous notons que nous avons désactivé certaines règles sur la forme qui visaient à obtenir des frontières bien nettes. Dans une application de détection ceux ci n'étaient pas nécessaires et risquaient au contraire d'augmenter les fausses alarmes.

Nous représentons sur la figure 3.23, la carte de fausse alarme du GLRT-GAUSS (qui correspond au résultat du test en l'absence de cible), avant et après segmentation. Nous pouvons donc observer qu'à l'aide de ce partitionnement de l'image, nous homogénéisons le fouillis polarimétrique, et donc la carte de fausse alarme. La segmentation n'a par contre que peu d'effet sur la carte de fausse alarme du GLRT-SIRV. Sur la figure 3.24

FIGURE 3.22 – Segmentation de la scène synthétique. La couleur des frontières est calculée par contraste avec la région voisine (Blanc sur Noir / Noir sur Blanc).

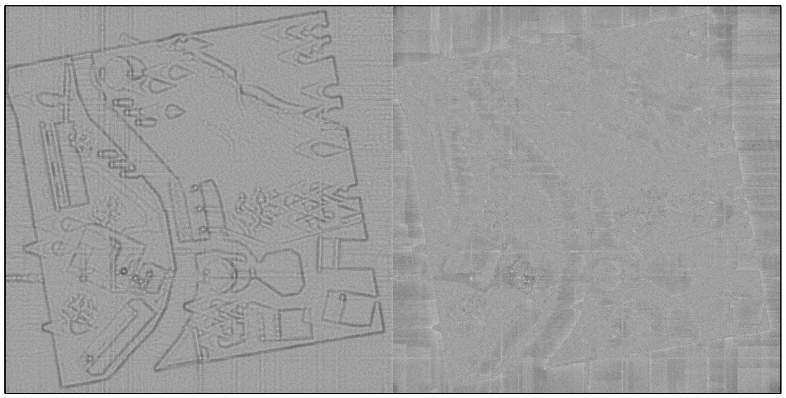


FIGURE 3.23 – Carte de fausse alarme pour le GLRT-GAUSS, à gauche avec estimation locale, à droite avec estimation par régions.

nous présentons les performances calculées sur de multiples réalisations de la scène, pour les détecteurs hybrides ainsi que les détecteurs classiques, en utilisant ou non le résultat de segmentation.

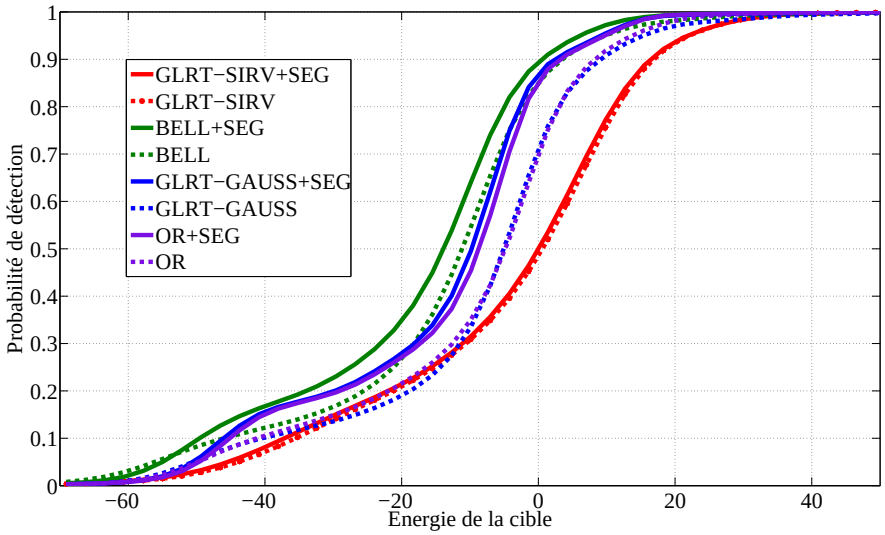


FIGURE 3.24 – Probabilité de détection dans la scène synthétique pour les différents détecteurs, probabilité de fausse alarme $P_{FA} = 10^{-3}$. Avec ou sans segmentation préalable de la scène.

Nous pouvons donc faire les constatations suivantes :

- La segmentation apporte grandement au détecteur classique gaussien, le GLRT-GAUSS, elle améliore grandement les priorité TFAC de ce détecteur. Par contre elle n'a aucun, ou presque aucun effet sur le GLRT-SIRV. Il semblerait donc qu'une erreur sur l'estimation de la matrice de cohérence polarimétrique ait beaucoup moins d'influence sur les tests de détection qu'une erreur sur l'estimation du niveau moyen de bruit.
- Pour les détecteurs hybrides, bien qu'une partie de l'information qu'ils utilisent n'ait pas bénéficié de la segmentation, le gain sur l'information de puissance est suffisant pour apporter un gain important au niveau des performances. Ainsi, le détecteur *BELL* couplé à la segmentation apporte un gain très important par rapport aux méthodes classiques que sont le GLRT-GAUSS et le GLRT-SIRV. Le gain apporté par la segmentation est du même ordre de grandeur que les gains apportés par l'approche hybride.

3.3.4 Conclusion

Dans ce chapitre nous avons eu l'occasion de compléter notre analyse des modèles non gaussiens appliqués aux images SAR. Dans le premier et second chapitre nous avions pu constater la difficulté d'application des modèles non gaussiens dans le cadre de la

polarimétrie radar. Ici en détection nous avons pu constater les même difficultés, et même interpréter géométriquement l'origine de ces limitations. Les modèles non gaussiens, en polarimétrie, ne peuvent être utilisés tels quels, que dans le cas de fouillis très fortement non gaussiens, et cela est dû à la faible dimension du vecteur signal polarimétrique. De tels fouillis sont très rarement présents, même à haute résolution, même en milieux urbains, comme nous avons pu le voir dans le premier chapitre et le vérifier en appliquant ces outils sur des données synthétiques.

Pour autant, l'information angulaire, qui correspond à une information polarimétrique "pure", indépendante de l'énergie, qui est produite par les détecteurs non gaussiens, apporte clairement une information importante et pertinente. Celle-ci permet lorsqu'elle est combinée de manière fine avec l'approche gaussienne classique, de gagner de façon très significative en probabilité de détection. La famille de détecteurs hybrides proposée a permis de mettre en oeuvre cette combinaison et nous avons proposé différents détecteurs hybrides ayant différents degrés de complexité, mais aussi de gains de performance. Les détecteurs "souples", comme le détecteur BELL, montrent des gains très importants de performances.

Cette utilisation conjointe de l'énergie et de l'information polarimétrique est une démonstration claire de l'apport de la polarimétrie en détection dans les images SAR. Nous avons pu améliorer la robustesse à une erreur sur la signature polarimétrique recherchée, mais la détection polarimétrique restera toujours dépendante de notre capacité à anticiper la signature polarimétrique des objets recherchés.

La combinaison des outils de segmentation et de détection est aussi une avancée importante pour les milieux urbains, qui permet à nouveau de gagner en probabilité de détection. Le gain est du même ordre de grandeur que l'utilisation des outils hybrides. Cela dit le coût en calculs est quand même plus élevé. La combinaison des deux approches nous permet d'exhiber au final des gains très importants par rapport aux méthodes classiques sur données synthétiques.

3.3.5 Perspectives

Les travaux sur la détection hybride sont aujourd'hui assez aboutis. Il est toujours possible de construire des détecteurs aux formes encore plus complexes, mais nous avons pu constater qu'au delà du niveau de complexité du détecteur BELL, les gains de performances deviennent de plus en plus négligeables. Des travaux théoriques, qui permettrait de calculer les probabilité de détection théorique des détecteurs "souples" seraient aussi intéressant (nous ne pouvons actuellement exprimer les probabilités de détection que des détecteurs linéaire par morceaux). Il serait aussi à long terme intéressant de tester ces outils de détection dans des cadres différents, comme la détection sous feuillage, ou la détection d'embarcation en mer. Encore une fois, obtenir des courbes de performances nécessiterait un outil adapté de simulation pour ces situations. Enfin, comme pour le chapitre précédent, nous pourrions étendre notre étude à la polarimétrie interférométrie, l'augmentation de la dimension du vecteur ayant un impact important sur l'ensemble des performances relatives des différents détecteurs.

Conclusion

Dans cette partie nous avons d'abord étudié avec précision les outils statistiques existant en polarimétrie radar. Nous nous sommes intéressés aux problématiques d'estimation des matrice de cohérence polarimétrique en radar et à l'utilisation de distances statistiques sur ces matrices. Dans un second temps nous avons abordé la segmentation d'image polarimétrique, et la détection dans des images polarimétriques. Après avoir fait des rappels sur les méthodes existantes nous avons pu contribuer à chacun de ces thèmes :

- Le choix d'une distance polarimétrique, dans une application de détection ou de classification n'a pas un impact majeur sur les performances. Les différences de performances entre de nombreuses distances ont été étudiés et montrent que les distances du monde la théorie de l'information comme la divergence de Kullback Leibler et la distance de Bhattacharyya offrent ainsi des performances très proches. Bien que sans aucune garantie d'optimalité, les test issus du rapport de vraisemblance généralisé sont aussi très compétitifs. Nous avons eu l'occasion de dériver ces tests, dans le cadre d'un modèle non gaussien, ce qui nous a permis de trouver des expressions inédites et dont les performances surpassent largement les tests classiques en milieux fortement non gaussiens.
- L'utilisation de modèles non gaussiens ne doit pas être automatique en haute résolution ou dans un scénario urbain. Il convient d'estimer avec rigueur, en évitant la méthode de la fenêtre glissante, la statistique de l'image. En polarimétrie, seule la présence d'un fouillis très impulsionnel peut motiver l'utilisation de modèles non gaussiens pour la gestion ou l'estimation de l'information polarimétrique. Notre étude suppose surtout que cette caractéristique, non gaussienne, est éventuellement source de perturbation dans l'estimation des matrices polarimétriques. Il est tout à fait possible que des paramètres de textures puissent être utile à la discrimination de différents objets ou zones dans l'image. Nos conclusions ne portent que sur l'application à la polarimétrie. Une idée de segmentation du plan κ_2,κ_3 à été proposée, nous aurons l'occasion d'affiner cette méthode dans l'avenir.
- Dans le cas ou les fouillis rencontrés sont faiblement impulsionnels, des modèles non gaussiens qui exploitent uniquement l'information polarimétrique peuvent quand même être utilisés mais d'une manière plus fine. Nous l'avons surtout démontré dans le cadre de la détection. Des résultats préliminaires indiquent que cela sera aussi le cas pour les distances polarimétriques. L'approche hybride semble produire des meilleurs performances pour des fouillis modérément impulsionnels, utilisant ainsi à la fois une information d'énergie projetée, mais aussi une information polarimétrique pure.
- Cette approche hybride, et la nouvelle famille d'outils de détection proposée sont aussi inédites et permettent d'obtenir des gains de performances importants. Ces méthodes auront, nous l'espérons, un cadre d'application bien plus large que la polarimétrie. Nous avons pu aussi aborder la problématique de la robustesse à une erreur sur le signal polarimétrique recherché et proposé un critère de robustesse.
- Un algorithme de croissance de régions contrainte a été proposé, et utilisé pour améliorer l'estimation des matrices de cohérence dans ce contexte. La logique floue a été proposée et exploitée de manière rapide et efficace afin d'exploiter une connais-

sance a priori sur la topologie des milieux urbains. L'outil proposé est donc rapide, et assez performant pour nos besoin. Il permet de réduire visiblement l'impact du speckle et contribue à préserver les détails et frontières de l'image. Enfin il permet une compression importante des images, qui accélère grandement les traitements ultérieurs.

Troisième partie

Les décompositions polarimétriques

Ce chapitre est dédié aux applications de la polarimétrie optique et radar. Nous supposons à partir de maintenant que nous traitons des données calibrées, corrigées éventuellement de l'influence de la géométrie, et éventuellement pré traitées de manière statistique. Nous travaillerons ainsi par exemple sur des images déjà segmentées et compressées. Nous nous intéressons donc à l'étape finale, l'extraction de l'information utile contenue dans la mesure polarimétrique.

Les applications peuvent faire intervenir des outils purement statistiques, comme de la classification, des outils inspirés de la physique comme l'extraction de paramètres, ou une combinaison des deux approches. Dans ce document nous ne proposons pas un inventaire exhaustif des applications de la polarimétrie, mais nous présentons les applications que nous avons pu mettre en oeuvre pendant la thèse.

Nous proposons un rappel de l'état de l'art concernant les méthodes d'extraction de paramètres ou décompositions polarimétriques utilisées en polarimétrie radar ou en polarimétrie optique. Nous appliquerons ensuite ces décompositions à des exemples d'images radar (images SAR) et d'images optique. Ces applications concerneront d'abord l'étude des milieux dépolarisants, comme les forêts à l'échelle radar, ou des tissus cancéreux à l'échelle optique. Dans un second temps nous nous intéresserons aux images SAR urbaines. Nous exploiterons encore les effets de dépolarisation afin de différentier les objets artificiels des zones naturelles, ou détecter des objets ou véhicules cachés dans des zones naturelles. Ensuite nous montrerons comment exploiter nos résultats de segmentation dans le cadre d'images PolInSAR (Polarimétrie et Interférométrie SAR) afin d'extraire une information de hauteur et de reconstruire des zones urbaines en trois dimensions.

Nous insistons sur la multiplicité des décompositions et des données sur lesquelles nous avons pu travailler. Ainsi nous ne croiserons pas tous les traitements avec toutes nos données, mais nous choisirons au mieux les combinaisons à présenter.

Chapitre 1

État de l'art sur les décompositions polarimétriques

Une décomposition polarimétrique est un outil qui permet d'interpréter physiquement le comportement d'une cible mesurée en la décomposant en éléments plus simples. Les cibles que nous sommes amenés à mesurer sont plus compliquées que des cibles élémentaires simples (retardateur, diatténuateur, dépolariseur, surface, dièdre, etc). En effet, une cible peut comporter plusieurs éléments de natures polarimétriques différentes. Ces différents effets polarimétriques n'étant pas forcément directement accessibles, l'analyse de la cible peut être facilitée par la décomposition de la matrice de Mueller ou la matrice de diffusion expérimentale en éléments simples.

Cette analyse se fait souvent à l'aide une réduction du nombre de paramètres ou du nombre de degrés de liberté. La plupart des décompositions polarimétriques aboutissent à trois paramètres. Ce nombre correspond naturellement au système de représentation colorée RGB (ou RVB). L'objectif est alors de proposer à l'opérateur une représentation colorée significative de l'image polarimétrique. Nous appellerons cette approche 'sous-paramétrage'. Dans un second temps une analyse plus fine peut être envisagée, avec des méthodes adaptées à la problématique du scénario.

Les décompositions polarimétriques dites "exactes" produisent un résultat qui permet de reconstruire la matrice polarimétrique initiale. Nous pourrions alors qualifier cette approche de re-paramétrage. L'objectif est de proposer des paramètres plus faciles à interpréter que les mesures brutes, mais sans simplification ni réduction des degrés de libertés. Nous rappellerons cette distinction, de 're' ou 'sous' paramétrage, au fur et à mesure de notre présentation des différentes décompositions.

Quelle que soit la longueur d'onde utilisée, les propriétés de la cible sont à analyser à l'échelle de la longueur d'onde employée. La notion de rugosité, par exemple, dépend de l'échelle à laquelle nous la définissons. La notion d'éléments élémentaires également. Ainsi, il apparaît que les interprétations et les analyses physiques de cible peuvent se rejoindre entre les différents domaines de fréquences utilisés.

Prenons l'exemple d'une forêt de troncs regardée en imagerie radar basse fréquence (longueur d'onde de l'ordre du mètre). Son comportement peut être assez bien rendu par la description sous forme d'un ensemble de cylindres fins quasiment verticaux. Un nuage de clous, analysé avec une longueur d'onde proche de la longueur de ce clou (de

l'ordre de la dizaine ce centimètres), présentera les mêmes comportements de diffusion de l'onde électromagnétique. Ainsi, une telle cible pourra être mesurée et analysée en chambre. Enfin, à l'échelle optique, une croissance de nanotubes de carbone sur un substrat présentera des similitudes de composition de structure (figure 2.16). Il apparaît donc que les décompositions polarimétriques traditionnellement dévolues au domaine optique ont de grandes chances de pouvoir aider l'interprétation dans le domaine radar, et vice-versa.

Parmi les décompositions polarimétriques, on distingue plusieurs catégories :

- Les décompositions cohérentes et les décompositions incohérentes. Les premières s'appliquent à la matrice de Sinclair ou matrice de Jones. Leur application à la matrice de Mueller est également possible, mais dans le cas où la matrice de Mueller est une matrice de "Mueller-Jones", c'est à dire lorsque la matrice de Mueller est de rang 1 et qu'il n'y a alors pas de dépolarisation.
- Les décompositions en sommes et les décompositions en produits. Les premières correspondent à une vision de la cible comme "superposition" de plusieurs éléments de comportements polarimétriques différents. Les deuxièmes correspondent à une représentation de la cible comme cascade d'éléments différents, successivement traversés par l'onde électromagnétique. (voir figure 1.1) Traditionnellement, les radaristes ont travaillé sur les décompositions en sommes, tandis que les opticiens ont travaillé sur les décompositions en couches. Pourtant, cette dernière représentation nous semble tout à fait intéressante pour un certain nombre d'applications radar, où les modélisations de cibles se font elles-même en différentes strates ou couches (forêts, glaciers, etc.).

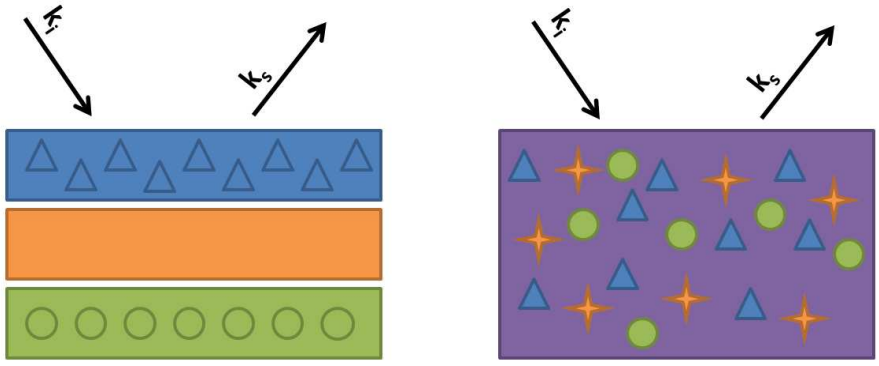


FIGURE 1.1 – La cible peut être vue comme une succession de divers éléments (décompositions multiplicatives, à gauche) ou bien comme un mélange de différents diffuseurs (décompositions additives, à droite)

Parmi les comportements de cibles élémentaires, et donc parmi les paramètres polarimétriques extraits de décomposition, on recherchera par exemple :

- Des propriétés simples de forme : cylindre fin, dièdre, trièdre, hélice, etc. Ces propriétés se déduiront principalement de décompositions additives.
- des propriétés d'anisotropie de structures : ce sera le cas par exemple pour des éléments canoniques de type retardeur, diatténuateur, etc. Ces propriétés se déduiront

principalement des décompositions multiplicatives.

1.1 Valeurs propres et vecteurs propres de cibles dépolarisantes

Parmi les décompositions polarimétriques additives utilisées en radar, un certain nombre s'appuie sur la diagonalisation de la matrice de cohérence polarimétrique. Celle-ci est hermitienne par construction. Elle est donc décomposable sur une base orthogonale de trois (ou quatre en bistatique) vecteurs propres, et les valeurs propres associées sont réelles positives. Ainsi on peut écrire :

$$\mathbf{T} = \sum_{i=1..3} \lambda_i \mathbf{T_i} \tag{1.1}$$

où les matrices $\mathbf{T_i}$ sont des matrices de rang 1 construites à partir des vecteurs propres $\mathbf{v_i}$ par $\mathbf{T_i} = \mathbf{v_i}\mathbf{v_i}^\dagger$. Les décompositions utilisant ces valeurs propres et vecteurs propres proposent l'interprétation physique suivante : dans la cellule de résolution sont présents les trois types de diffuseurs orthogonaux dont les vecteurs de diffusion polarimétriques sont représentés par les vecteurs propres. Les valeurs propres associées nous renseignent alors sur l'amplitude relative de chacun de ces types de diffuseurs. Ainsi dans le cas particulier où toutes les valeurs propres sont égales, les trois diffuseurs élémentaires ont le même poids. Dans l'autre cas limite où deux valeurs propres sont nulles, alors un seul mécanisme de diffusion a lieu dans la cellule de résolution. La décomposition en valeurs propres vecteurs propres est naturellement un re-paramétrage des données qui n'empêche aucunement de reconstruire le signal initial.

1.1.1 Cloude Pottier

Historiquement, la première décomposition polarimétrique à exploiter cette propriété mathématique est la décomposition de Cloude-Pottier. Dans [5], S. Cloude et E. Pottier proposent trois paramètres déduits de cette décomposition. Nous rappelons ici brièvement ces paramètres que nous avons eu l'occasion d'utiliser plus tôt dans ce document. Ce sont

1. L'entropie H. Elle représente le caractère aléatoire des mécanismes de diffusion présents dans les pixels dont on a calculé la moyenne. Si les λ_i sont les valeurs propres de la matrice de cohérence, l'entropie est définie par :

$$H = -\sum_{i=1}^{3} p_i \ \log_3 \ p_i, \tag{1.2}$$

   où les p_i sont les valeurs propres pondérées : $p_i = \lambda_i / \sum \lambda_i$. Cette valeur p_i peut être considérée comme étant la probabilité que le mécanisme de diffusion représenté par λ_i soit présent dans le groupe de pixels considérés.

2. Le paramètre α représente le mécanisme de diffusion présent dans les données. Il est défini à partir des premières composantes de chacun des vecteurs propres normalisés. Ces vecteurs propres sont paramétrés de la manière suivante :

$$v = \begin{vmatrix} \cos\alpha \\ \sin\alpha\cos\beta e^{j\delta} \\ \sin\alpha\sin\beta e^{j\gamma} \end{vmatrix} \tag{1.3}$$

On associe les valeurs de alpha aux mécanismes de diffusion dominants présents dans la cible. Cet angle est compris entre 0° et 90°. La valeur 0° correspond à la surface idéale ou une diffusion par réflexion impaire. La valeur 45° représente une diffusion par dipôle et la valeur 90° représente une diffusion par dièdre ou réflexion paire. Malgré ces définitions, l'interprétation physique de l'angle alpha demeure difficile, car la plupart des diffuseurs naturels ne s'insèrent pas directement dans ces trois catégories.

3. L'anisotropie A décrit l'importance des mécanismes de diffusion secondaires. Elle est définie à partir des deux valeurs propres les plus faibles par $A = (\lambda_2 - \lambda_3)/(\lambda_2 + \lambda_3)$. Elle indique l'intensité relative des deuxième et troisième mécanismes de diffusion.

A ce jour c'est encore la décomposition la plus utilisée en polarimétrie radar. Elle a donné lieu à de nombreuses études concernant l'estimation des paramètres, et applications en classification ou segmentation. Le plan des paramètres $H\alpha$ a aussi été utilisé pour proposer des classifications d'images polarimétriques en fonction des comportements physiques [66]. Pour nous, les qualités essentielles de cette décomposition sont la rigueur mathématique sur laquelle elle repose, et la quantité d'information contenue dans les paramètres H et α. Nous pouvons toutefois lui reprocher de n'utiliser qu'une partie seulement de l'information polarimétrique existante puisqu'il s'agit d'un sous-paramétrage. En outre, son calcul peut être long dans le cadre d'images très volumineuses comme en très haute résolution.

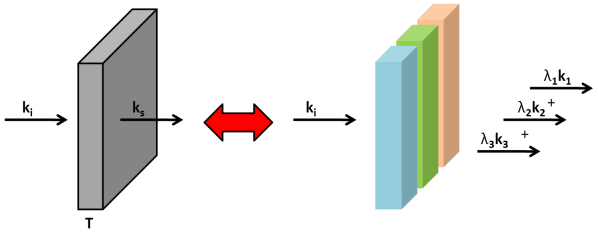


FIGURE 1.2 – Principe des décompositions en valeur propres et vecteurs propres.

Par ailleurs, seule la définition de l'entropie s'étend facilement au cas de la configuration bistatique. Il a été montré que l'interprétation de α ne s'étend pas facilement à des configurations bistatiques [67], [68], à partir du moment où même la notion de rebond devient délicate (qu'appelle-t-on double rebond dans une situation avec un angle bistatique azimutal de $\pi/2$?)

1.1.2 Alternative

Une décomposition alternative a été proposée par J. Praks et E. Koeniguer dans [69]. Elle consiste à s'affranchir de l'étape de la décomposition en valeurs propres et vecteurs

propres, tout en proposant des paramètres alternatifs à H et α. L'avantage est alors une bien plus grande simplicité et rapidité dans le calcul de ces paramètres, tout en conservant l'ensemble des interprétations physiques de ces paramètres et des statistiques très proches. Là encore, il s'agit d'un sous-paramétrage.

1.1.3 Touzi

Nous n'avons pas beaucoup eu l'occasion de travailler sur la décomposition de R. Touzi [70], nous la mentionnons ici toutefois car elle exploite aussi la décomposition en valeurs propres et vecteurs propres. A la différence de la décomposition de Cloude Pottier, la paramétrisation des vecteurs propres s'appuie sur un modèle cohérent et "roll-invariant" de chacune des cibles élémentaires représentées. Cette nouvelle paramétrisation s'appuie sur la décomposition en valeurs singulières de chacune des matrices de diffusion associées aux vecteurs propres trouvés. C'est une décomposition plus complète dans le sens où si l'ensemble des paramètres sont conservés, elle permet de reconstruire exactement la matrice originale. L'interprétation de paramètres globaux pour les matrices à forte entropie est toutefois délicate. Si on considère l'ensemble des paramètres produits pour tous les vecteurs propres, cette décomposition est en fait même un 'sur-paramétrage' des données polarimétrique, puisque le nombre de paramètres déduits dépasse le nombre de degrés de liberté.

1.2 Décompositions appuyées sur des modèles.

La deuxième grande famille de décompositions utilisées en radar est aussi formée de décompositions additives, mais qui ne s'appuient pas sur une diagonalisation de la matrice de cohérence. Ces décompositions s'appliquent à projeter les matrices mesurées sur des matrices canoniques de référence, caractéristiques de comportements polarimétriques attendus. Ce sont par exemple les matrices de cohérence d'une surface avec une orientation quelconque, ou d'un volume de particules aléatoirement réparties.

L'avantage de cette approche est que ce type de décompositions peut s'adapter facilement à une problématique précise. Si les comportements polarimétriques rencontrés peuvent être modélisés a priori, il est possible de construire un ensemble de matrices qui servira d'une certaine manière de base sur laquelle nous allons projeter nos matrices initiales. Il convient d'insister sur la non-exactitude mathématique de ces décompositions qui fait qu'en pratique les matrices canoniques utilisées ne forment absolument pas une vraie base. En pratique, la somme de tous les éléments obtenus n'est donc jamais égale à la matrice initiale. Cette somme produit des matrices 'residues' non hermitiennes. Par ailleurs, il n'est pas rare de voir apparaître des puissances négatives dans la décomposition, difficiles à interpréter physiquement. Toutes ces décompositions sont donc des sous-paramétrages de l'information polarimétrique.

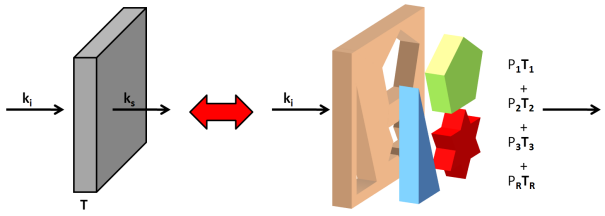


FIGURE 1.3 – Principe de la décomposition de Freeman Durden ou Yamaguchi.

1.2.1 Freeman Durden

La première décomposition basée sur ce principe est la décomposition d'Anthony Freeman dans [71]. Elle s'appuie sur une modélisation simple d'un volume aléatoire, d'une cible de type dièdre, siège de double-rebond, et d'une cible ayant une diffusion de type surface. L'algorithme consiste alors à estimer des paramètres de puissances (P_d, P_v, P_s) affectées à chacune de ces matrices modèles :

- Le modèle de diffusion volumique est basé sur des diffuseurs cylindriques étroits, dont la position et l'orientation suivent une distribution aléatoire uniforme.
- Le modèle de diffusion à double réflexion est basé sur des réflecteurs dièdres, dont les surfaces sont faites de matériaux diélectriques différents.
- la diffusion surfacique est un modèle de diffusion de Bragg.

La somme pondérée de ces matrices approche au mieux la matrice mesurée. Indépendamment de l'existante d'un résidu, cette décomposition permet d'obtenir rapidement des résultats dont l'interprétation est satisfaisante dans des situations classiques. Comme la décomposition de Cloude et Pottier, cette décomposition a aussi été utilisée dans des méthodes de classification d'images polarimétriques [72].

$$T = f_s \begin{pmatrix} |\beta|^2 & 0 & \beta \\ 0 & 0 & 0 \\ \beta^\dagger & 0 & 0 \end{pmatrix} + f_d \begin{pmatrix} |\alpha|^2 & 0 & \alpha \\ 0 & 0 & 0 \\ \alpha^\dagger & 0 & 0 \end{pmatrix} + \frac{f_v}{8} \begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix} \tag{1.4}$$

1.2.2 Yamaguchi

La décomposition de Yamaguchi est une extension de la décomposition de Freeman Durden. Elle consiste à ajouter une nouvelle matrice au modèle de Freeman, décrivant le comportement de type "hélice". Comme sa grande soeur, cette décomposition estime quatre paramètres de puissances (P_d, P_v, P_s, P_h). En pratique cette décomposition est plus durement affectée que la précédente par des problèmes de résidus. Ceux-ci s'accompagnent très souvent de paramètres de puissance négatifs. Son auteur a toutefois proposé récemment des ajustements visant à réduire l'occurrence de ces problèmes, mais l'interprétation physique de ces solutions n'est pour nous pas encore très claire.

1.2.3 Limitations

Les limitations sont principalement liées aux puissances négatives et aux résidus qui peuvent apparaître. L'interprétation physique n'est alors plus possible. Les résidus eux-mêmes n'ont aucune interprétation physique puisqu'ils ne correspondent plus à des matrices hermitiennes positives (propriété conservée dans le cas d'une somme, mais pas dans le cas d'une soustraction). Bien qu'il s'agisse de sous paramétrage ces décompositions restent toutefois de bon outils d'analyse rapide, et donnent des représentations colorées d'images satisfaisantes.

1.3 Lu Chipman et Décomposition Symétrique d'Ossikovski

En optique les décompositions sont le plus souvent multiplicatives. La représentation en couches correspond naturellement aux changements successifs de polarisation qui peuvent être induits le long d'un axe optique ; c'est donc une représentation naturelle pour un opticien.

1.3.1 La décomposition polaire des milieux non dépolarisants

Nous devons d'abord distinguer l'appellation "matrice de Mueller-Jones" de l'appellation générale "matrice de Mueller". La première s'applique aux milieux non dépolarisants. Dans ce cas, à une matrice de Mueller correspond une matrice de Jones. Par contre, si ce n'est pas le cas, le milieu est dépolarisant. La matrice de Mueller Jones est l'analogue optique des matrices de cohérence de rang 1 qui, en radar, peuvent être associées à un unique vecteur de diffusion et donc une unique matrice de Sinclair.

En optique, il existe un algorithme de décomposition qui s'applique aux milieux non dépolarisants [6]. Celui-ci utilise la décomposition polaire appliquée à la matrice de Jones.

Toute matrice de Jones peut ainsi être décomposée en un produit de deux matrices de Jones homogènes caractéristiques d'éléments polarimétriques dits élémentaires : une matrice hermitienne qui peut être associée à un diatténuateur elliptique homogène, et une matrice unitaire qui peut être associée à un retardateur elliptique homogène.

Une approche similaire a été proposée en radar, avec comme support théorique celui des quaternions [73]. Le principe de décomposition polaire sous-jacent est le même.

1.3.2 Lu Chipman

Dans le cadre des cibles dépolarisantes, Lu et Chipman ont montré dans [11] qu'il est possible de décomposer toute matrice de Mueller en trois éléments permettant d'obtenir le dichroïsme, le déphasage et la dépolarisation introduits par l'échantillon. Un élément dichroïque ou diatténuateur présente une anisotropie d'absorption. Il modifie uniquement les amplitudes des composantes du champ électrique. Un déphaseur ou retardateur retarde la phase des états propres de polarisation l'un par rapport à l'autre.

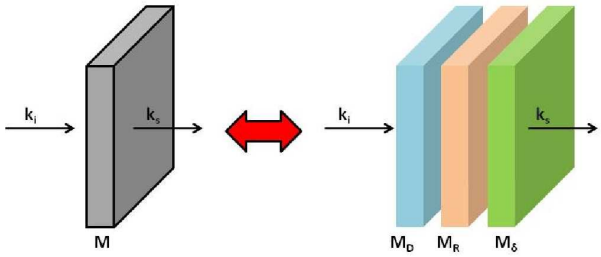


FIGURE 1.4 – La décomposition de Lu et Chipman, décomposition de la matrice de Mueller en un produit d'un diatténuateur, d'un retardeur et d'un dépolariseur.

La matrice de Mueller $\mathbf{M}$ peut alors s'écrire à l'aide d'une décomposition polaire comme le produit de trois matrices : un élément dichroique $\mathbf{M_D}$ suivi par un déphaseur $\mathbf{M_R}$ puis par un dépolariseur $\mathbf{M}_\delta$:

$$\mathbf{M} = \mathbf{M_D}\mathbf{M_R}\mathbf{M}_\delta \tag{1.5}$$

La matrice M_D est une matrice de la forme suivante :

$$\mathbf{M_D} = \begin{pmatrix} 1 & \mathbf{D^t} \\ \mathbf{D} & \mathbf{m_D} \end{pmatrix} \tag{1.6}$$

$\mathbf{D}$ est le vecteur dichroisme. On note $\hat{\mathbf{D}}$ ce même vecteur normalisé, et D sa norme. La sous-matrice $\mathbf{m_D}$ prend la forme $\mathbf{m_D} = \sqrt{1-D^2}\mathbf{I} + (1-\sqrt{1-D^2}\hat{\mathbf{D}}\hat{\mathbf{D}}^\mathbf{t})$.

La matrice $\mathbf{M_R}$ est une matrice de la forme :

$$\mathbf{M_R} = \begin{pmatrix} 1 & \mathbf{0^t} \\ \mathbf{0} & \mathbf{m_R} \end{pmatrix} \tag{1.7}$$

où $\mathbf{m_R}$ est une matrice de rotation.

La matrice $\mathbf{M}_\delta$ est une matrice de la forme suivante :

$$\mathbf{M}_\delta = \begin{pmatrix} 1 & \mathbf{0^t} \\ \mathbf{P} & \mathbf{m}_\delta \end{pmatrix} \tag{1.8}$$

où $\mathbf{P}$ est le vecteur polarisance qui représente la capacité d'un système à polariser une onde incidente dépolarisée. $\mathbf{m}_\delta$ est une matrice 3x3 symétrique qui contient les termes dépolarisants du système.

La décomposition de Lu et Chipman est exacte dans le sens où il s'agit d'un re-paramétrage, et qu'il est possible de retrouver la matrice de Mueller originale. Cela dit, il est possible d'obtenir des valeurs d'un nombre restreint de paramètres comme la diatténuation, le retard, ou la dépolarisation à partir des différentes matrices de la décomposition. La décomposition de Lu et Chipman propose donc à la fois un re-paramétrage et un sous paramétrage qui permet aussi des représentations colorées.

1.3.3 Décomposition "reverse"

La décomposition proposée par Lu et Chipman est aujourd'hui fréquemment utilisée mais comme pour la décomposition polaire d'un milieu non dépolarisant, se pose le problème de la non commutativité des matrices. Ce problème a été abordé dans [74]. En fonction de l'ordre des matrices employées dans une telle décomposition, six arrangements possibles peuvent être obtenus. Ils se répartissent en deux familles suivant la position du diatténuateur vis-à-vis du dépolariseur. Chaque famille possède des valeurs de diatténuation, retardance et dépolarisation différentes. Seuls changent les états propres des matrices élémentaires. Cependant, les deux familles ont des propriétés différentes.

En effet, il est possible d'obtenir des matrices de Mueller non "physiques" dans le cas où le diatténuateur précède le dépolariseur, notamment lorsque diatténuation et dépolarisation sont importants tous les deux. Pour résoudre ces problèmes survenant dans cette deuxième famille de décomposition, [75] propose de définir l'élément dépolarisant de la manière suivante :

$$\mathbf{M}_\delta = \begin{pmatrix} 1 & \mathbf{D^t} \\ \mathbf{0} & \mathbf{m}_\delta \end{pmatrix} \tag{1.9}$$

caracatérisé par un vecteur polarisance nul et une diatténuation non nulle.

1.3.4 Décomposition symétrique d'Ossikovski

Une des limitations de la décomposition de Lu et Chipman est qu'elle produit dans certains cas, des matrices de dépolarisation non diagonales. En outre, comme nous venons de le soulever, l'ordre dans la décomposition est souvent sujet à discussion. R. Ossikovski propose une solution alternative dans [76] qui consiste à décomposer la matrice en Mueller en un produit "symétrique" de cinq matrices :

$$\mathbf{M} = \mathbf{M_{D1}}\mathbf{M_{R1}}\mathbf{M}_\delta\mathbf{M_{R2}}\mathbf{M_{D2}} \tag{1.10}$$

La matrice de dépolarisation $\mathbf{M}_\delta$ est alors systématiquement diagonale. Elle est entourée des deux couples formés chacun d'un diatténuateur et d'un retardeur. Comme la décomposition de Lu et Chipman, il s'agit à la fois du re-paramétrage et d'un sous-paramétrage éventuel de l'information polarimétrique.

L'avantage de cette décomposition est qu'elle permet de tenir compte d'un aller-retour dans un milieu, ce qui peut être particulièrement adapté à ce qui se passe physiquement dans le cas de la rétrodiffusion, ou dans tout cas où l'angle bistatique est inférieur à π. Cela sera d'autant plus judicieux si il y a véritablement pénétration de l'onde dans les milieux, comme ça peut être le cas dans des applications telles que l'étude des forêts ou des terrains agricoles dans des bandes bien choisies. Par ailleurs, cette décomposition permet de produire deux couples de retardeur/diatténuateurs différents pour l'aller et pour le retour de l'onde. Dans le cas monostatique, les deux couples produits (R_1, D_1) et (R_2, D_2) sont identiques.

Deux limitations ont été rencontrées lors de son utilisation, notamment sur des mesures en chambre anéchoïque sur des cibles proches de cibles canoniques. La première apparaît lorsque la matrice décomposée par Lu et Chipman donne déjà une matrice de

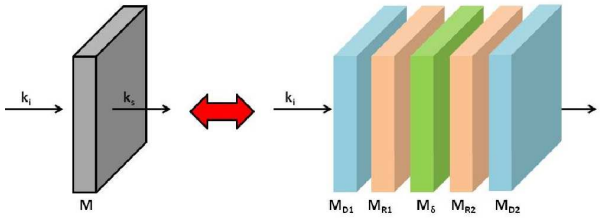


FIGURE 1.5 – La décomposition de la matrice de Mueller comprend ici cinq éléments canoniques : deux paires de diatténuateurs/retardateurs, entourant une matrice de dépolarisation diagonale.

dépolarisation diagonale. C'est notamment le cas lorsque le vecteur diatténuation associé à plusieurs composantes nulles. En pratique il convient alors d'abord de "tester" la décomposition de Lu et Chipman afin de voir si celle-ci produit une matrice de dépolarisation diagonale. Si c'est le cas, la matrice de dépolarisation commute avec les deux autres matrices et il est donc possible de l'écrire sous la forme symétrique avec une matrice de retard nul et de diatténuation nulle à gauche ou à droite de la décomposition. Dans ce cas la décomposition symétrique est inutile. Dans le cas contraire alors la décomposition symétrique s'applique correctement.

Le second problème vient de la non-unicité du couple formé par les deux retardeurs. En effet une rotation d'un angle $+\theta$ du premier retardeur est équivalente à la rotation du second retardeur d'un angle $-\theta$. Cette incertitude peut toutefois se résoudre simplement par le principe du minimum de retardance qui consiste à fixer arbitrairement un des deux retards afin de minimiser la retardance totale.

Dans l'ensemble, cette décomposition nous semble très prometteuse pour les futures campagnes de mesures radar polarimétriques bistatiques.

1.4 Visualisation et compositions colorée

Nous avons vu que nous pouvons utiliser trois paramètres pour une visualisation colorée de l'image. Chacun de ces paramètres est affecté à un canal de couleur : Rouge, Vert ou Bleu. Cependant, le résultat de ces représentations colorées dépendent d'un certain nombre de choix :

- Il est possible de prendre en compte le span ou intensité totale comme facteur multiplicatif de chacun des canaux de couleur. Celui-ci vient alors pondérer l'intensité du pixel. Si le span n'est pas pris en compte, alors l'image présente peu de contraste. Si le span est pris en compte, alors le résultat est plus satisfaisant de manière qualitative.
- Dans tous les cas, les paramètres doivent être bornés par des limites communes. Cela peut poser problème notamment pour les décompositions qui amènent à trouver des paramètres non bornés, comme la décomposition de Yamaguchi. On peut être amené à seuiller les paramètres pour éviter les valeurs non significatives, mais ce seuillage reste pour le moment arbitraire.

– La dynamique de ces paramètres produit parfois des résultats visuellement décevants si cette dynamique n'est pas modifiée. On peut donc choisir ou non de faire une normalisation d'histogramme, c'est à dire de transformer l'image de telle manière à ce que chaque valeur soit présente en quantité égale. La raison pour laquelle on réalise une normalisation d'histogramme est d'utiliser au maximum la plage de valeurs disponibles. Une image peu constrastée peut présenter une foule de détails après une égalisation.

En fonction de ces différents choix, les résultats produits peuvent être vraiment très différents. Une représentation colorée de 3 paramètres polarimétriques ne suffit donc pas à elle seule à démontrer la pertinence des paramètres choisis. Dans la suite, nous choisissons donc de ne donner que la représentation colorée des puissances calculées dans la base de Pauli, avec utilisation du span et égalisation des histogrammes, de manière à montrer au maximum l'apport de l'information polarimétrique. Les autres paramètres seront montrés en niveaux de gris et simplement normalisés entre 0 et 1.

Chapitre 2

Etude des milieux dépolarisants

L'étude des milieux dépolarisants est une problématique commune au monde de la polarimétrie optique et de la polarimétrie radar. Dans les milieux complexes, au fur et à mesure que l'onde pénètre dans les matériaux, une perte de degré de polarisation est observée. L'effet de dépolarisation permet de caractériser des paramètres importants sur le milieu, comme la densité des éléments constituants, leur structure ou leur taille.

En optique, le domaine d'application principal est le domaine médical, principalement les tissus cellulaires. Il a notamment été observé que la dépolarisation des tissus cancéreux est différente de celle des tissus sains, et évolue par ailleurs en fonction de l'avancement de la maladie.

En radar, les deux domaines d'applications principaux de la polarimétrie sont l'étude des forêts et des terres agricoles dont la dépolarisation permet de caractériser la densité, la hauteur, le type d'arbres ou de cultures.

Dans cette partie nous présentons des résultats de décompositions polarimétriques sur des données optique et radar.

2.1 Forêts et Agriculture

Dans cette section nous nous concentrons sur les applications des milieux dépolarisants dans le monde du radar. A l'échelle humaine ces études concernent donc les milieux naturels, les végétaux par exemple. En fonction de l'application concernée, il est possible de rencontrer de multiples comportements polarimétriques différents :

- Les effets liés au feuillage et aux branches de petites tailles créent de la dépolarisation par diffusion multiples.
- Les effets de clairières produisent des réponses polarimétriques semblables à des surfaces rugueuses.
- Des terrains agricoles structurés comme des rangées de vignes voient alterner des comportements polarimétriques différents.
- Dans le cadre des forêts, il n'est pas rare de rencontrer des effets de type "double rebond" qui sont liés aux troncs des arbres formant des angles droits avec le sol. Ces effets d'orientation verticale et de double-rebond se retrouvent également dans les cultures.

2.1.1 Objectifs

Pour les applications agricoles, les objectifs sont majoritairement civils ; l'application majeure consiste à classifier les zones agricoles en différents types de cultures. Dans le cas de données multi temporelles, il est aussi possible de suivre la croissance des cultures au cours du temps.

Pour les forêts, les intérêts sont multiples. Parmi les applications civiles on note principalement les études sur l'évolution de la biomasse. C'est cette application qui a d'ailleurs motivé le lancement de nombreux satellites d'observation en modes polarimétriques partiels ou complets. La polarimétrie utilisée seule permet d'apporter des informations intéressantes. La proportion entre les zones dépolarisantes et les zones de clairières nous renseigne par exemple sur l'espace moyen entre les arbres. L'intensité de la dépolarisation peut nous renseigner sur leur taille ou leur densité.

Cependant c'est surtout lorsque la polarimétrie est couplée à l'interférométrie qu'elle trouve son maximum de potentiel pour l'estimation de la biomasse. Dans cette problématique, l'approche est d'utiliser des modèles basés sur des considérations électromagnétiques poussées [77] ou plus simples [78], et d'inverser ces modèles afin d'estimer un grand nombre de paramètres.

Les décompositions ne permettent pas de retrouver des paramètres physiques à elles seules. Cependant, elles permettent une visualisation colorée rapide et une interprétation au premier ordre des images. Ces paramètres peuvent aussi être utilisés comme vecteurs d'états dans des méthodes de classification pour améliorer leurs performances.

2.1.2 Exemple : Banlieue de Brétigny, image monostatique en bande X

L'image de Brétigny permet d'observer un panel assez complet des différents types de zones que nous pouvons rencontrer en imagerie polarimétrique radar. Elle a été acquise en configuration monostatique en bande X, avec une résolution de 1.3 m. Elle représente une zone de banlieue ou se côtoient des bâtiments, des champs et les pistes d'un aérodrome. Les champs sont composés de cultures très diverses, comme des champs de blé, de maïs, de fétuque et de colza.

Sur cette image, nous avons calculé les paramètres issus de la décomposition de Cloude Pottier et ceux issus de la décomposition de Lu et Chipman. Nous utilisons une version segmentée de l'image pour accélérer le calcul des différentes décompositions et améliorer le rendu visuel.

La représentation colorée des puissances calculées dans la base de Pauli après normalisation des histogrammes et utilisation du span est donnée sur la figure 2.1. Nous pouvons distinguer, directement sur la représentation de Pauli différentes zones :

– En rouge, nous observons des terrain nus ou des cultures très jeunes de très faible hauteur. Pour ce type de terrain il s'agit donc principalement d'une diffusion de type surface, dont l'amplitude (span) dépend de la rugosité du terrain. Ainsi le comportement polarimétrique des routes et des pistes correspond aussi à une diffusion de surface ; mais la rugosité de ces éléments étant très faibles, l'amplitude du signal rétrodiffusé est également très faible car il s'agit d'une réflexion spéculaire.

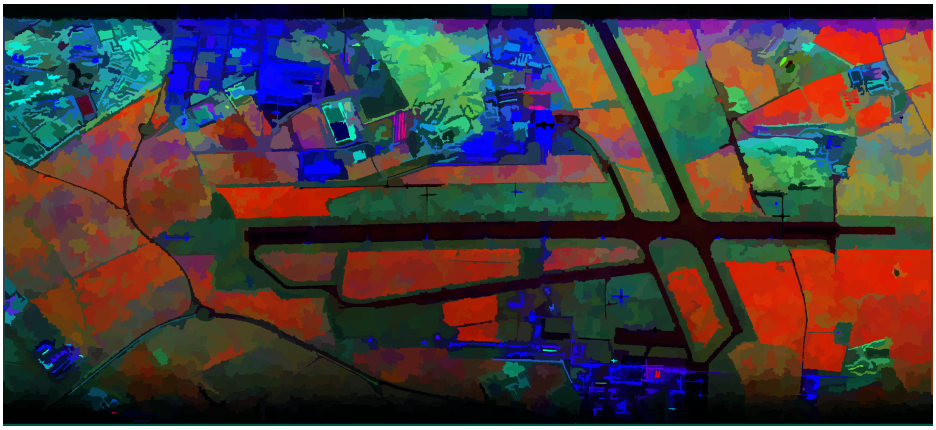


FIGURE 2.1 – Représentation de l'image de Brétigny dans la base de Pauli, en rouge $HH + VV$, en bleu $HH - VV$, en vert HV (après normalisation des histogrammes)

- En vert nous observons des éléments plus aléatoires, des diffusions de type volumique. Ainsi la plupart des champs qui apparaissent en vert sont des cultures dont la hauteur et la densité est suffisante pour que les effets de volume surpassent la diffusion de surface. En vert très clair apparaissent des parcelles de forêts, de petites tailles sur cette image. Sur ces parcelles les effets de volume sont maximum.
- En bleu apparaissent les diffusions de type double rebonds, qui sont naturellement induites par les bâtiments ou les objets humains formant des angles droits avec la surface de sol. A l'exception de quelques bâtiments atypiques, la plupart des bâtiments de l'image apparaissent dans cette couleur sur la représentation de Pauli.

Nous faisons remarquer la présence d'une ligne horizontale qui traverse l'ensemble de l'image, située au premier tiers de l'image, dans laquelle la mesure de polarisation est partiellement déformée et qui correspond à un artefact de mesure. L'interprétation doit être faite avec précaution dans cette zone.

Après calcul des différentes décompositions, nous pouvons faire les observations suivantes :

- Sur la figure 2.2, nous comparons les paramètres d'entropie et de dépolarisation. Nous pouvons observer que les paramètres de dépolarisation et d'entropie sont fortement corrélés. Cependant la dynamique est ici en faveur de la dépolarisation qui permet de bien mieux visualiser les différentes zones de l'image. On retrouve ainsi plus facilement la différence que nous observions sur la représentation de Pauli entre les zones dépolarisantes (en vert) et les terrains nus (en rouge). Cette discrimination est moins évidente avec l'entropie.
- Sur la figure 2.3, nous comparons le paramètre d'anisotropie et celui de diatténuation. La corrélation est ici assez faible. L'ensemble des bâtiments de l'image ont une forte anisotropie, qui correspond aux zones qui apparaissaient en bleu sur la représentation de Pauli. Certains bâtiments ont aussi une forte diatténuation (en

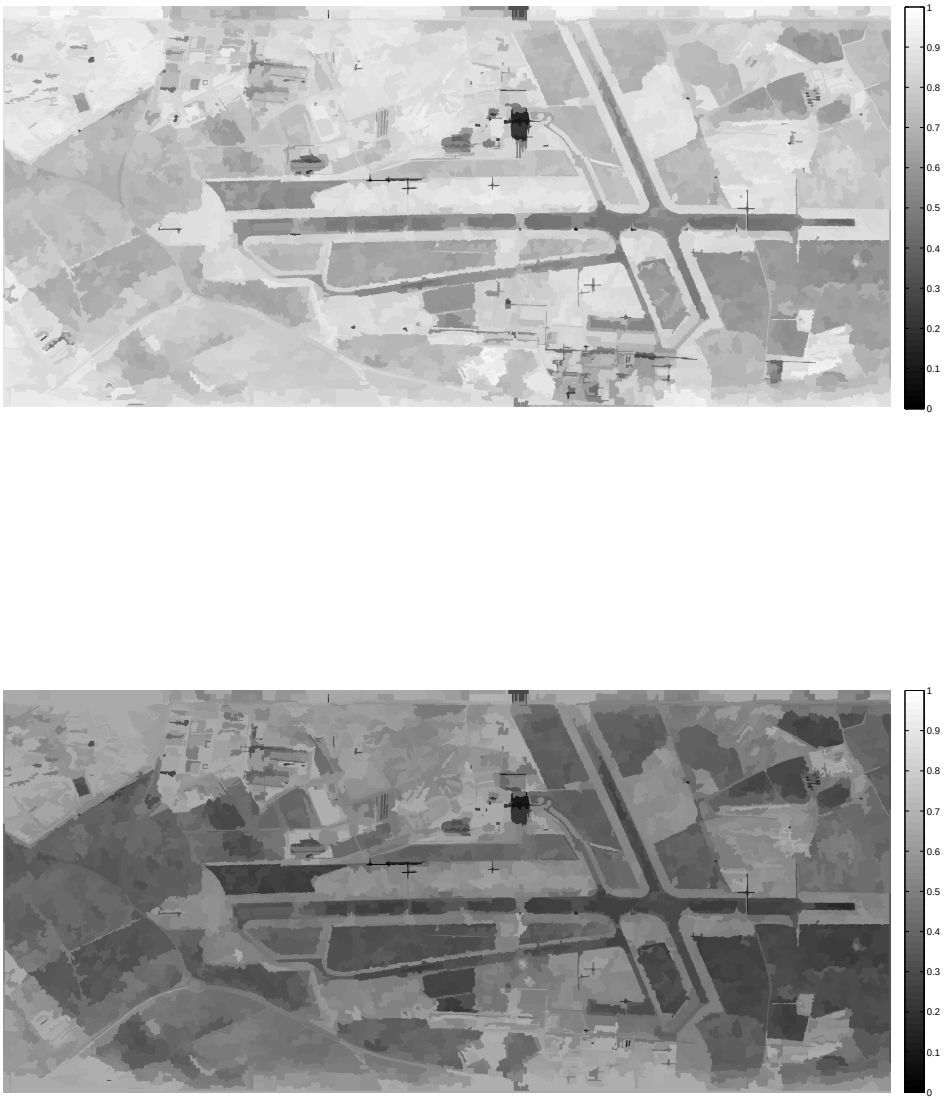


FIGURE 2.2 – Entropie (en haut) et dépolarisation (en bas) sur l'image de Brétigny

bas et quelques bâtiments isolés au milieu). La diatténuation est aussi sensiblement forte sur toutes les routes et les pistes de décollage ce qui n'est pas le cas pour l'anisotropie, qui prend d'ailleurs des valeurs assez aléatoires sur ces zones.
– Sur la figure 2.4, nous comparons les paramètres de retard (normalisé entre 0 et 1) et celui de l'angle α (aussi normalisé entre 0 et 1). Les informations sont encore une fois assez décorrélées. Le retard prend des valeurs très contrastées. Il est nul sur les bâtiments, et très proche de 1 pour le reste de l'image. Il permet ainsi simplement d'identifier rapidement les zones urbaines. L'angle α prend lui un grand nombre de valeurs différentes. Il est maximum pour les bâtiments, et les zones urbaines. Il prend des valeurs moyennes voire fortes pour les champs et forêts, et très faibles pour les terrains nus. L'angle α permet a priori aussi de séparer les zones comprenant des double-rebonds, mais avec un risque de confusion avec les zones dépolarisantes bien plus important que le retard.

Cette comparaison sur l'image de Brétigny nous permet de conclure que les paramètres de Lu et Chipman apportent une vision différente des paramètres classiques en radar, mais pertinente, de l'information polarimétrique. La dépolarisation est semble-t-il plus intéressante que l'entropie, de par sa dynamique plus contrastée. La diatténuation vient compléter l'anisotropie et permet d'observer routes et pistes sans exploiter l'information de puissance. Enfin le retard, de par sa dynamique, semble plus adapté pour discriminer les rebonds pairs des rebonds impairs que l'angle alpha.

2.1.3 Exemple de forêt en monostatique : Nezer, image monostatique de forêt en bande P

Cet exemple est une image de la forêt de Nezer, située en France dans les Landes, acquise en configuration monostatique, en bande P, avec une résolution de 1.4 m. La bande P est une gamme de fréquence choisie traditionnellement pour l'étude des forêts car elle permet à l'onde de pénétrer dans le feuillage et ainsi de mieux 'sonder' leur structure. Cette forêt propose en fait un cas d'école car il s'agit d'une forêt artificielle, constituée de parcelles rectangulaires utilisées pour l'exploitation du bois. Ces données ont donc permis de valider de nombreux algorithmes polarimétriques et interférométriques comme dans [79]. Notons que les arbres plantés ici sont des pins landais, dont le houppier est peu dense et les troncs longs et fins.

Intéressons nous d'abord à la représentation de Pauli sur la figure 2.5. Nous pouvons y apporter les commentaires suivants :
– Les parcelles "très vertes" correspondent à des arbres assez âgés. Le volume de branches est important et il n'y a que très peu d'effets de clairières ou de double rebond. Les parcelles de vert moins franc correspondent à des parcelles d'arbres plus jeunes, et on observe donc une augmentation des effets de clairière (rouge), et des troncs qui deviennent naturellement plus visibles (bleu pour le double rebond).
– Les parcelles qui apparaissent clairement en violet (même si certaines tendent aussi vers le marron) sont plantées d'arbres beaucoup plus jeunes. Les branchages sont très peu denses et ne créent pas de dépolarisation. Les arbres sont toutefois suffisamment grands pour générer des effets de double rebond qui se mélangent alors aux effets de

FIGURE 2.3 – Anisotropie (en haut) et Diattenuation (en bas) sur l'image de Brétigny

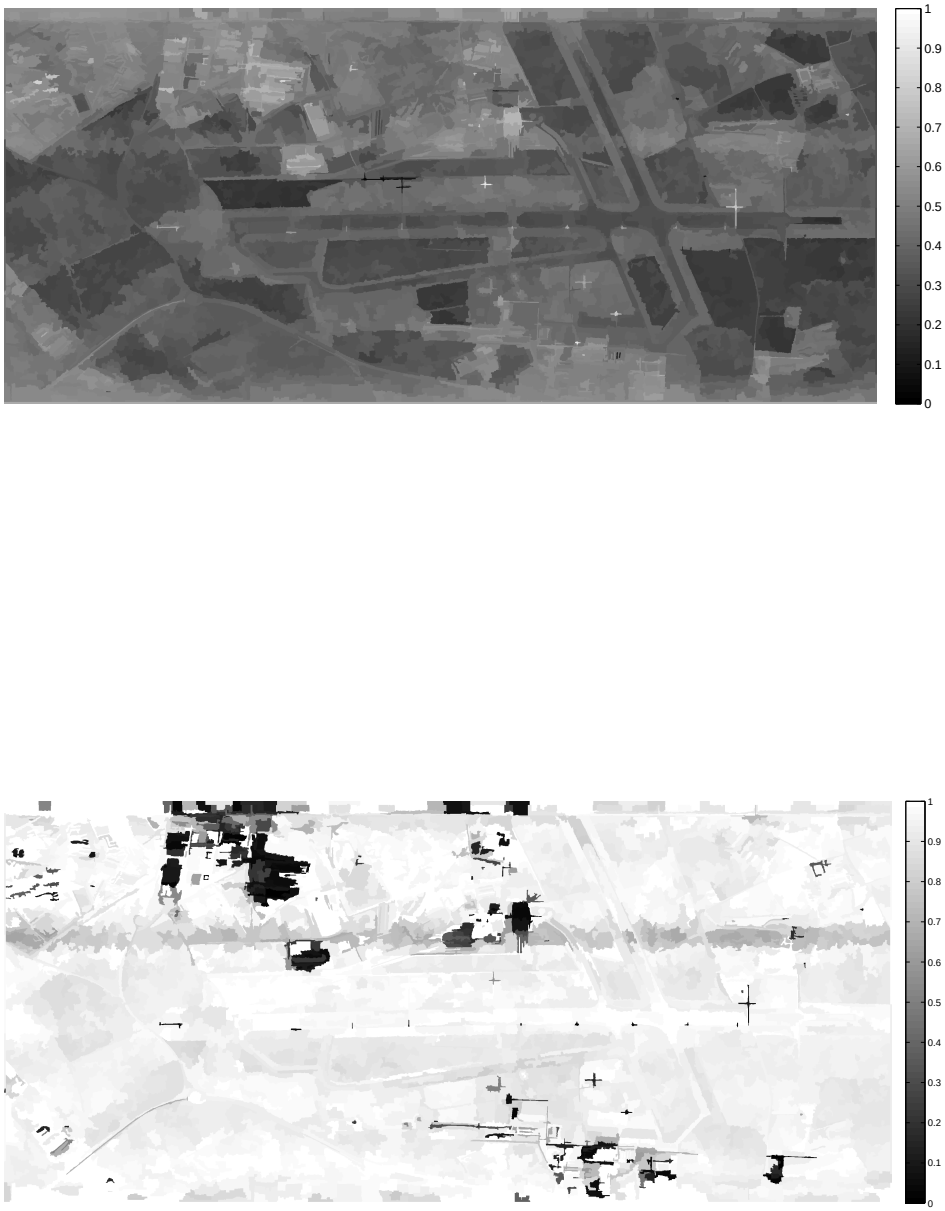


FIGURE 2.4 – Angle alpha (en haut) et Retard (en bas) sur l'image de Brétigny

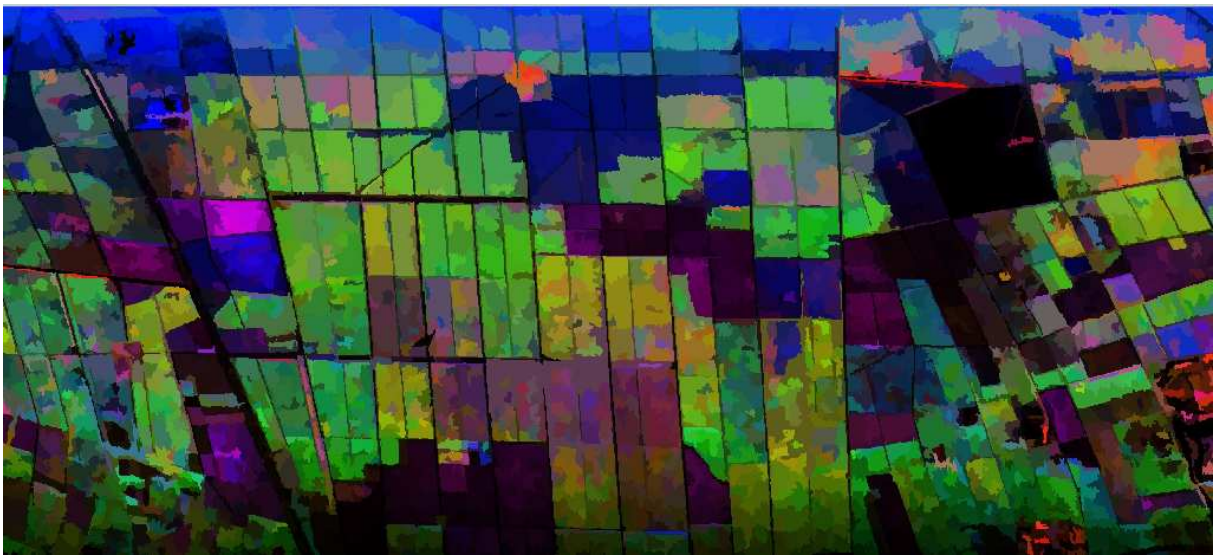


FIGURE 2.5 – Représentation de l'image de la forêt de Nezer dans la base de Pauli (avec normalisation)

simples diffusion du sol. De manière générale, plus la parcelle se rapproche du rouge, plus les arbres sont petits et la contribution du sol est importante. Il faut noter que nous n'observons pas vraiment de zones clairement rouge comme dans l'image de Brétigny. En fait ces zones où les arbres sont extrêmement jeunes apparaissent en noir sur notre image. C'est simplement parce que l'image a été acquise en bande P (longueur d'onde de l'ordre du mètre). La rugosité, à cette échelle, est beaucoup plus faible et donc le signal rétrodiffusée aussi.

– Pour les parcelles qui apparaissent en bleu, il faut distinguer celles du haut de l'image de celles qui sont situées au milieu. Les parcelles en haut de l'image doivent être ignorées, en effet l'incidence varie le long de l'axe vertical, et sur le haut de l'image, zone la plus proche de l'avion, l'incidence est nettement moins rasante. Ainsi la distance parcourue dans le houppier est beaucoup plus courte. Avec cette incidence les effets de double rebonds dominent alors, ce qui explique la couleur bleu. Pour les autres zones qui apparaissent en bleu, il s'agit donc d'arbres de grandes tailles et induisant de nombreux effets de double rebonds, mais dont les branchages sont encore trop peu denses pour générer de la dépolarisation.

Nous nous intéressons maintenant aux résultats de décompositions.

– La dépolarisation et l'entropie sur la figure 2.6 sont extrêmement corrélées. Toutes les deux permettent de discriminer les terrains nus des parcelles contenant des arbres. Il n'y a cependant que peu de nuances visibles au sein des parcelles boisées.
– La diatténuation apporte ici visuellement plus d'information que l'anisotropie sur la figure 2.7. Les valeurs prises par l'anisotropie ne permettent pas d'identifier un comportement particulier au sein des parcelles plantées. La diatténuation permet clairement d'identifier les parcelles d'arbres jeunes, par contraste avec les parcelles de terrain nus et les parcelles d'arbres âgés. Elle est maximum lorsque les troncs

d'arbres sont le plus visibles. De façon très intéressante elle ne semble pas affectée la variation d'incidence, puisque la diatténuation des parcelles en haut de l'image reste très faible.
– Le retard, comme sur l'image de Brétigny, prend des valeurs très contrastées. Il permet de différentier très clairement les parcelles où les arbres sont âgés du reste de l'image (arbres jeunes et terrains nus). L'angle alpha quant à lui reste peu contrasté (figure 2.8).

L'exemple sur la forêt de Nezer est extrêmement valorisant pour les paramètres issus des décompositions optiques. C'est d'ailleurs un résultat attendu dans la mesure où les modèles multiplicatifs sont physiquement plus adaptés à des milieux stratifiés comme les forêts. Ici la dépolarisation n'apporte pas plus d'information que l'entropie, mais la diatténuation et le retard apportent des informations pertinentes par rapport à l'anisotropie et l'angle alpha. La diatténuation permet ainsi de repérer toutes les parcelles d'arbres jeunes, et ses nuances pourraient peut-être distinguer deux parcelles d'âges différents. Le retard permet quant à lui de bien séparer les parcelles nues (simple diffusion) des parcelles boisées.

Sur un extrait de cette image de forêt, nous voulons également illustrer l'apport de la polarimétrie dans notre algorithme de segmentation. Sur la figure 2.9, nous représentons l'image dans la base de Pauli, sans utiliser de segmentation. Nous voyons à quel point les différences de contraste en intensité ou en couleurs (polarimétrie) sont vraiment faibles.

Sur la figure 2.10, nous utilisons l'image de texture segmentée sur l'image entière et montrée précédemment, réduite à une portion d'images de quelques parcelles de forêts. Ces parcelles montrent parfois un léger contraste polarimétrique (indécelable à l'oeil nu sans une normalisation d'histogramme sur image segmentée). Nous superposons en blanc les frontières des régions trouvées, avec ou sans polarisation. Pour le cas sans polarisation, nous avons utilisé une image acquise en HH pour plusieurs passages de l'avion. Ainsi, statistiquement, le nombre de signaux exploité reste le même. Le gain apporté par la polarimétrie est bien visible sur le bas de la figure 2.10, où les parcelles sont mieux segmentées. L'apport de la polarimétrie n'est pas lié ici seulement au nombre de réalisations disponibles, mais bien à une information physique.

Enfin, sur la figure 2.11, nous voyons à quel point l'algorithme de segmentation échoue dans le cas où l'on désactive le critère de forme fondé sur la détection de contour. Bien sûr ce cas est un cas particulièrement difficile, mais il illustre l'importance de chacune des informations prises en compte.

2.2 Tissus cancéreux

La polarimétrie peut jouer un rôle important dans le dépistage du cancer en imagerie médicale. Le Laboratoire de Physique des Interfaces et Couches Minces a notamment travaillé sur le dépistage du cancer du colon et du cancer du col de l'utérus. L'idée générale est que la polarimétrie permet, lorsque l'onde pénètre dans les tissus, de mesurer des paramètres liés à la structure cellulaire ainsi qu'à leurs formes. La structure cellulaire étant très différente entre un tissu sain et cancéreux, l'information polarimétrique permet alors de mieux déterminer la présence et l'ampleur des tumeurs. Dans d'autres cas, il est

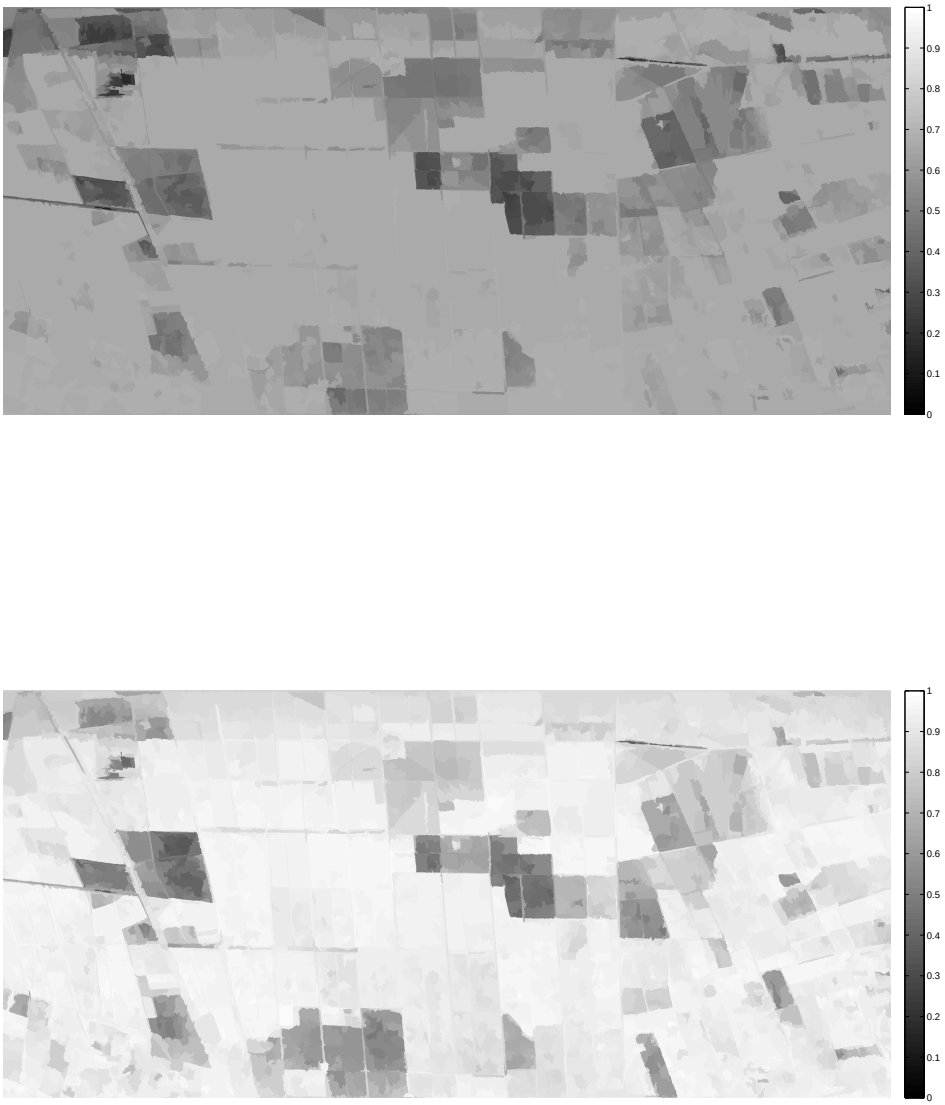


FIGURE 2.6 – Dépolarisation (en haut) et Entropie (en bas) sur l'image de Nezer

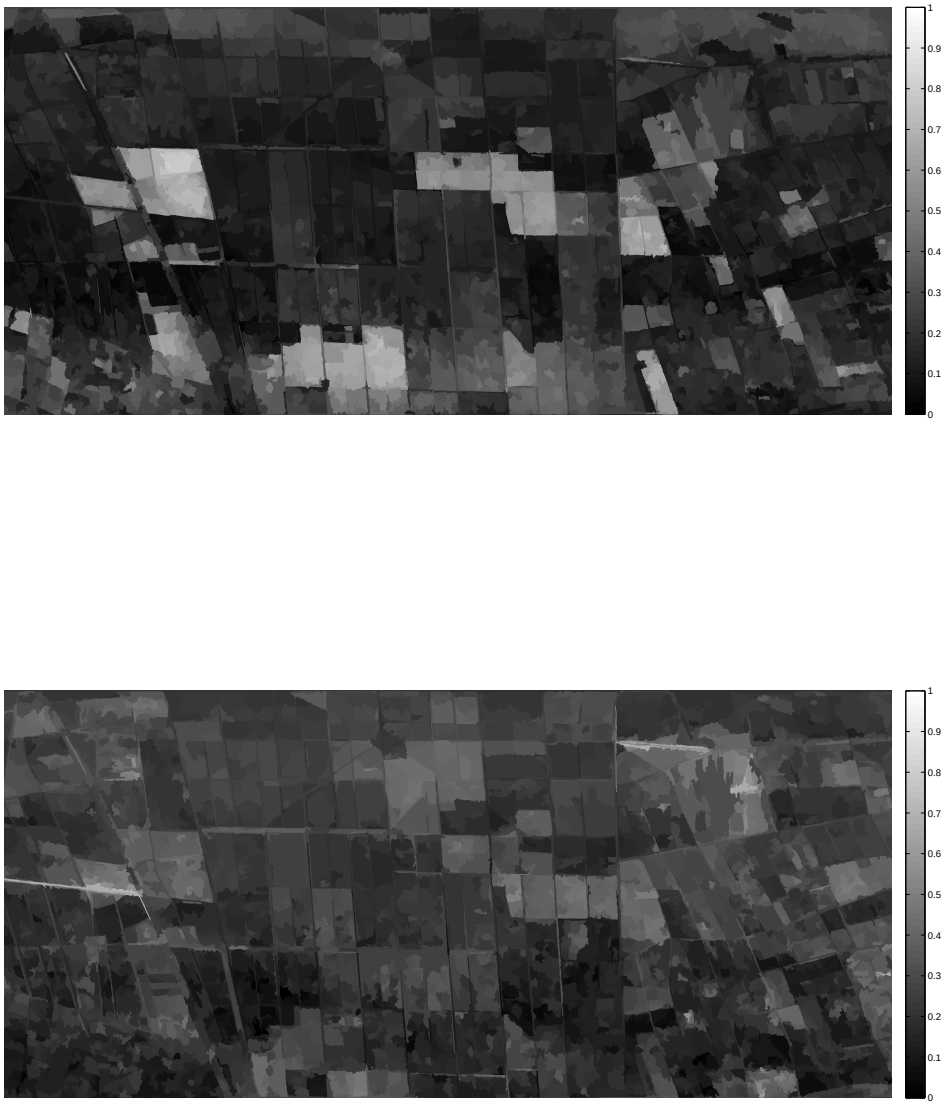


FIGURE 2.7 – Diattenuation (en haut) et Anisotropie (en bas) sur l'image de Nezer

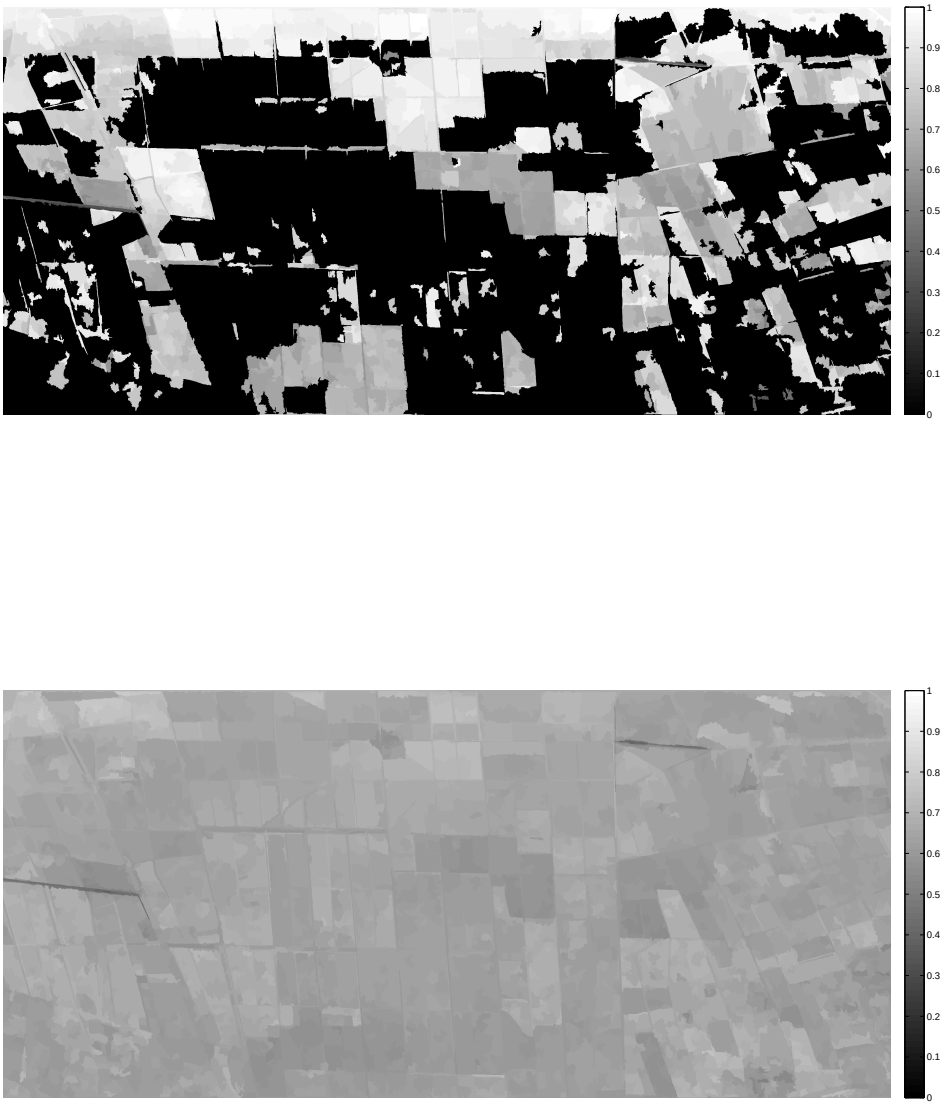


FIGURE 2.8 – Retard (en haut) et angle alpha sur (en bas) l'image de Nezer, les deux angles sont ramenés entre 0 et 1.

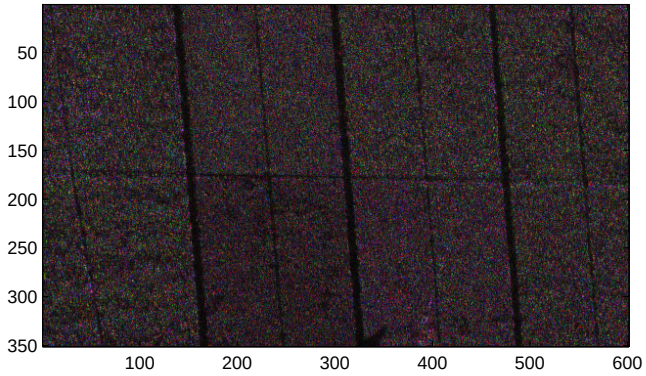


FIGURE 2.9 – Représentation colorée d'un extrait d'image, sans normalisation et sans segmentation

aussi possible de mesurer, au fur et a mesure de la progression de la tumeur, l'exposition progressive des différentes couches de tissus, ayant chacune des comportements polarimétriques très différents. Dans le cas du cancer du colon, cette information permet de juger si la tumeur est encore contrainte localement, ou si la disparition de certaines couches de tissus n'empêche plus la propagation de la tumeur.

Naturellement, tous ces phénomènes sont très complexes à modéliser, et nécessitent une très bonne compréhension à la fois physique de la diffusion par les cellules, mais aussi médicale ce qui implique une collaboration étroite avec les médecins. Comme pour l'étude des forêts, des modélisations fines sont nécessaires. Les décompositions polarimétriques permettent ici encore de proposer une représentation rapide et une compréhension simple des comportements polarimétriques.

L'exemple que nous présentons ici est un tissu de col de l'utérus. Celui-ci comprend une zone saine, une zone d'électropion ainsi qu'une zone de dysplasie. La description du tissu imagé ainsi que la "vérité" associée est montrée sur la figure 2.12.

L'ectropion est une malformation du col de l'utérus qui provoque des sécrétions de glaires. En principe, une frontière se trouve au niveau du col de l'utérus entre deux muqueuses. Celle que l'on voit de l'extérieure, l'exocol, est rose, épaisse, faite de plusieurs rangées de cellules, elle ressemble à la muqueuse de la joue. En termes plus techniques, il s'agit de l'épithélium malpighien exocervical. L'autre est plus rouge, beaucoup plus fine. Elle n'est composée que d'une seule couche de cellules. Elle est fripée et elle sécrète la glaire. Elle est, en principe, à peine visible au niveau de l'orifice. Le tissu qui recouvre la partie interne constitue l'épithélium glandulaire endocervical (ou épithélium cylindrique) Lorsque cette partie interne de la muqueuse du col dépasse par l'orifice par-dessus la muqueuse de l'exocol, n'étant pas habituée à l'acidité de la flore vaginale, elle se met à sécréter du mucus en excès. L'ectropion est précisément la surface occupée par cette

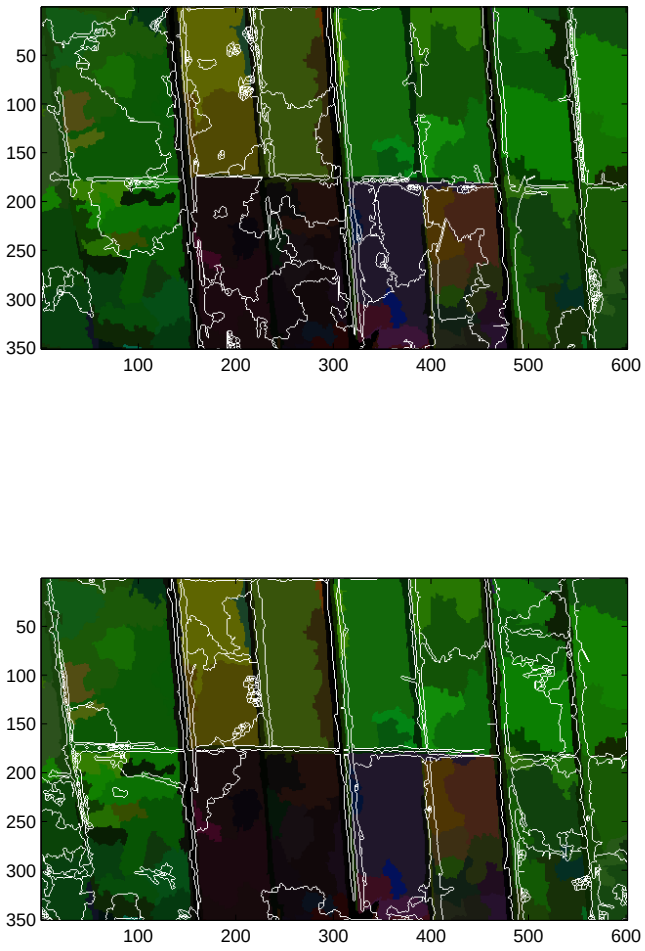


FIGURE 2.10 – Résultats de segmentation sans puis avec l'information polarimétrique sur quelques parcelles de forêts.

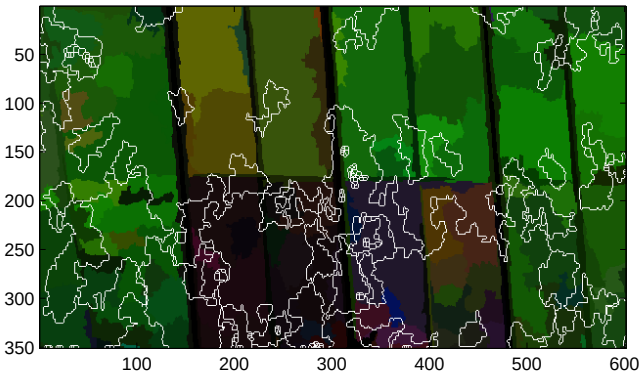


FIGURE 2.11 – Résultats de segmentation sans critères de forme sur les contours

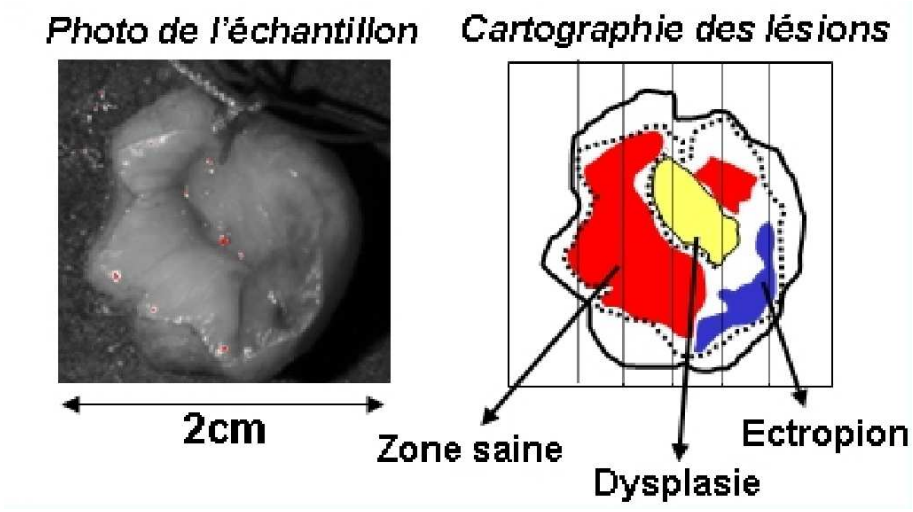
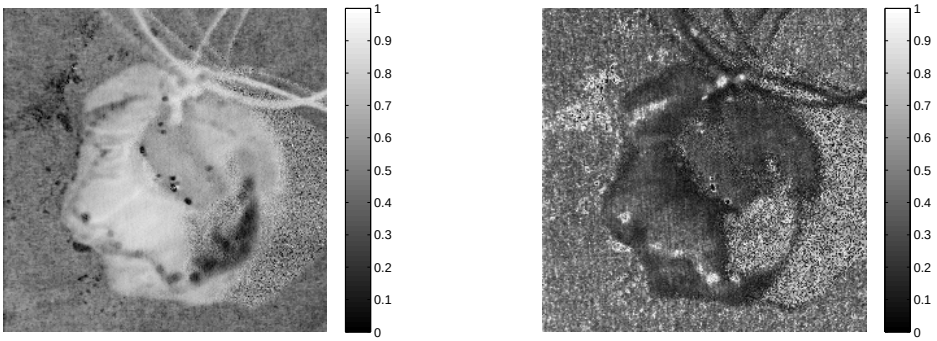


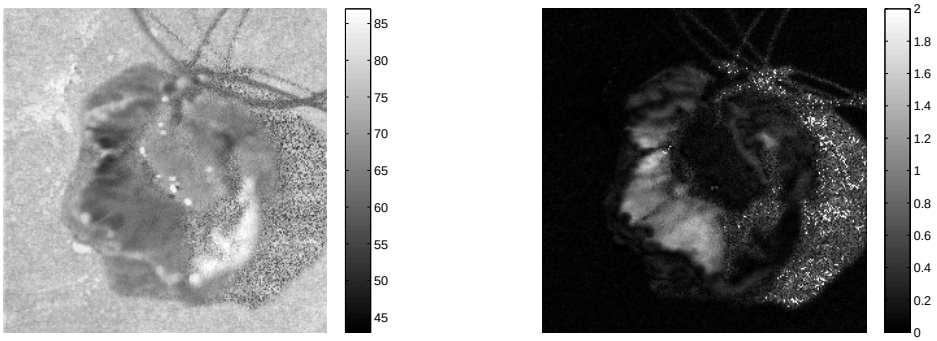
FIGURE 2.12 – Vérité terrain sur le tissu de col d'utérus

muqueuse interne qui deborde vers l'extérieur. Les cellules constituant le col de l'utérus peuvent également se transformer et constituer une dysplasie. C'est le terme médical utilisé pour désigner la présence de cellules anormales sur le col de l'utérus dues à l'HPV (papillomavirus humain).

Sur la figure 2.13 nous présentons certains paramètres polarimétriques qui offre des contrastes très intéressant sur cette image, alors que ces contrastes sont absents de la photographie du tissu utérin.



Dépolarisation et Anisotropie



Alpha (°) et Retardance (radian)

FIGURE 2.13 – Quatre paramètres polarimétriques pertinents. En haut à gauche, la dépolarisation issue de la décomposition de Lu et Chipman. En haut à droite l'image d'anisotropie issue de la décomposition de Cloude-Pottier. En bas à gauche le paramètre alpha issu de cette même décomposition, et enfin en bas à droite le paramètre de retardance.

Comme pour les images radar, les paramètres de dépolarisation et d'entropie sont très comparables ; le paramètre de dépolarisation restant plus contrasté avec une meilleur dynamique. Nous n'avons représenté ici que cette dépolarisation.

La surprise vient du paramètre de biréfringence. Il s'avère que la valeur de la biréfringence est particulièrement élevée sur la zone saine, ce qui pour l'heure n'est pas encore bien expliqué. Quant au paramètre alpha, il offre ici un contraste très intéressant entre les trois zones (éctropion, tissu sain et dysplasie), alors que le contraste encore ectropion

et tissu sain n'est pas franc sur le retard.

Enfin, si le paramètre de diatténuation offre très peu de contraste, la diatténuation restant faible sur l'ensemble du tissu, il s'avère que le paramètre radar d'anisotropie offre lui aussi des informations pertinentes.

A la suite de nos observations, nous avons proposé sur la figure 2.14 une représentation colorée à l'aide des trois paramètres suivants : dépolarisation, alpha et anisotropie. Cette représentation fait apparaître qualitativement des différences de couleurs entre les trois zones, alors que celles -ci sont difficilement discernables sur la photographie de l'échantillon.

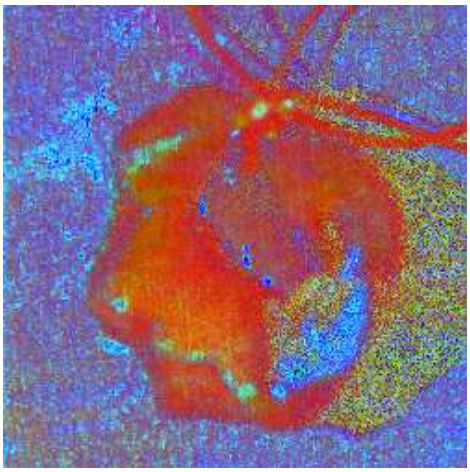


FIGURE 2.14 – Représentation polarimétrique colorée à l'aide de différents paramètres polarimétriques : en rouge, la dépolarisation ; en vert, le paramètre alpha, en bleu l'anisotropie.

2.3 Apport de la configuration bistatique et perspectives

L'étude des forêts se fait encore aujourd'hui exclusivement en configuration monostatique. Dans cette configuration, les troncs et leur interaction avec le sol, comme nous l'avons décrit plus haut, créent des effets de double rebonds qui, par l'amplitude de leur réponse, occultent parfois partiellement les autres comportements polarimétriques. Parmi les intérêts envisagés des mesures radar bistatiques, l'étude des forêts est une des applications potentielles majeures. Dans une configuration bistatique, les effets de double rebonds engendrés par les troncs peuvent être très fortement réduits, ce qui d'après les simulations en cours devrait permettre d'améliorer l'exploitation de l'information polarimétrique. Actuellement tous les travaux concernant les forêts en configuration bistatique

sont uniquement réalisées à l'aide de simulations et/ou par des mesures en chambre anéchoïque.

Dans cette section nous allons dans un premier temps illustrer la décomposition symétrique d'Ossikovski sur un exemple de mesure en chambre anéchoïque ainsi qu'avec certaines simulations simples. Nous chercherons à exhiber l'intérêt de cette décomposition pour l'étude des forêts en configuration bistatique. Dans un second temps nous présenterons les résultats de mesures réalisées sur une forêt de nanotubes. Notre idée est d'exploiter les capacités actuelles de l'optique à réaliser des mesures bistatiques, afin d'étudier dès maintenant, mais à des échelles différentes, des scénarios futurs.

2.3.1 Une forêt de clous

Nous utilisons ici des résultats de simulation d'un nuage de cylindres avec l'approximation du cylindre infini. Cette simulation correspond à la mesure effectuée en chambre anéchoïque sur la centaine de clous métalliques, déjà présentée en partie 2. Les clous ont donc des répartitions spatiales aléatoires, avec une orientation proche de la verticale. Les résultats de la décomposition d'Ossikovski sont comparés aux résultats obtenus à l'aide de la décomposition de Lu et Chipman sur la figure 2.15.

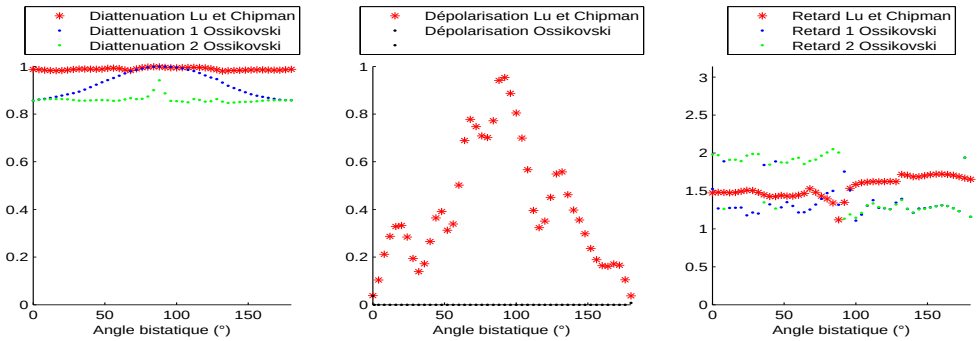


FIGURE 2.15 – Effet de la décomposition d'Ossikovski sur la modélisation simple d'un nuage de cylindres (sans prise en compte des interactions).

Sur cette figure, nous pouvons voir au sujet de la décomposition d'Ossikovski, que

- L'un des paramètres de diatténuation est affecté par la diattenuation induite, et l'autre l'est beaucoup moins.
- La dépolarisation d'Ossikovski est quasiment nulle, alors que la dépolarisation classique ne l'est pas. Cet effet est directement du au fait que la décomposition d'Ossikovski mène à un élémént de dépolarisation sans polarisance, au contraire de la décomposition classique.
- L'un des retards semble subir une discontinuité autour d'un angle de bistatisme de $\frac{\pi}{2}$, tandis que l'autre retard ne semble pas en subir.

Ce genre de constatation a été généralisé à d'autres simulations.

2.3.2 Une forêt de nanotubes

Le problème principal de l'étude des forêts en polarimétrie radar bistatique est l'absence aujourd'hui de campagnes de mesures bistatiques en polarimétrie complète. Très récemment, un couple de satellites TANDEM-X a été mis en orbite et a fonctionné pendant un temps dans une configuration bistatique en polarimétrie partielle, mais avec des angles de bistatisme très faibles (au maximum quatre degrés en azimut) qui ne permettront probablement pas d'observer une différence importante par rapport au monostatisme. Des campagnes bistatiques ont aussi été effectuées lors d'une collaboration impliquant l'ONERA et le FOI (Agence de la Recherche pour la défense suédoise) mais elles n'ont été effectuées qu'en polarimétrie partielle et ne sont pas étalonnées polarimétriquement à l'heure actuelle.

Notre idée ici est d'exploiter la polarimétrie optique comme moyen de mesure d'anticipation. Nous effectuons des mesures en polarimétrie de Mueller, en configuration bistatique sur des cibles ayant les mêmes structures que des forêts à l'échelle optique. Ainsi, à l'instar des chambres anéchoïques qui sont utilisées pour faire des mesures à échelle réduite, nous avons proposé d'effectuer des mesures qualitatives sur des forêts de nanotubes.

Ce principe de mesure à l'échelle, et des similitudes entre forêt réelle et forêt de nanotubes est illustré sur la figure 2.16.

FIGURE 2.16 – De l'étude de la forêt de Nezer en radar, aux forêts de Nanotubes en optique.

Nous présentons ici une démonstration de principe plus qu'une étude complète et renseignée des comportements observés. Le LPICM ayant justement une activité de recherche centrée sur les nanotubes, il n'a pas été difficile de nous procurer un échantillon contenant une forte densité de nanotubes. Cependant nous n'avions que des informations approximatives sur cette densité ainsi que sur les permittivités diélectriques. C'est pourquoi ces mesures n'ont pas été confrontées à des simulations. Pour ce qui est du rapport d'échelle, nous avons calculé que par rapport à une mesure réalisée avec une fréquence radar comprise entre (600-800MHz) et des arbres de $20m$ de hauteurs, des nanotubes de $27\mu m$, correspondent à la gamme de longueur d'onde de (500-650nm) qu'il nous était possible de mesurer. Les bandes P et L (400 et 1500 MHz) sont traditionnellement utilisées en radar pour l'étude des forêts. Nous pouvons donc considérer que nous sommes bien dans les mêmes domaines de rapport d'échelle.

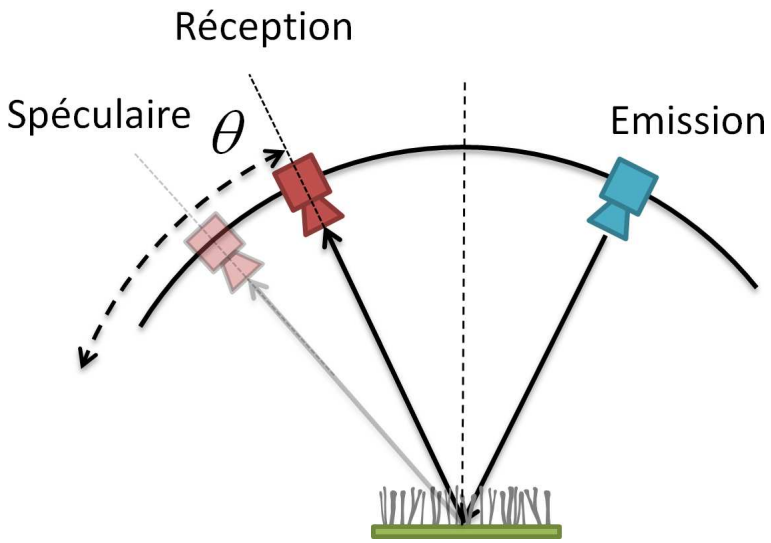


FIGURE 2.17 – Dispositif expérimental pour la mesure polarimétrique d'un échantillon de forêt de nanotube à l'aide d'un polarimètre de Mueller

L'échantillon est disposé sur un plan horizontal. L'émission est fixée au dessus de l'échantillon et la réception peut être positionnée de manière variable comme illustré sur la figure 2.17. Dans ce dispositif experimental il n'est toutefois pas possible d'effectuer des mesures pour un pas angulaire régulier, ni de contrôler avec précision la position de la réception. Ainsi un nombre réduit de positions de réception ont été mesurées, en s'écartant progressivement de la position dite "spéculaire" où la normale à l'échantillon forme la bissectrice de l'angle formé par l'émission et la réception.

Nous représentons les positions mesurées sur la figure 2.18. La position *gradient* est légèrement éloignée de la position *spéculaire*, alors que la position dite *rasante* l'est beaucoup plus nettement. Ces positions ne sont malheureusement pour l'instant décrite que qualitativement. Pour une étude quantitative, l'élaboration d'une configuration expérimentale dédiée serait nécessaire.

Dans le cadre de l'étude des forêts, il a notamment été montré qu'une configuration bistatique s'éloignant d'une dizaine de degrés de la configuration monostatique permet

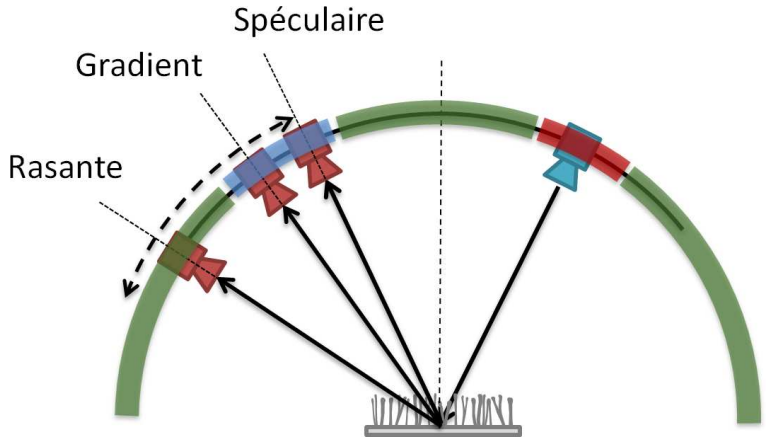


FIGURE 2.18 – Positions de la réception pour l'étude polarimétrique de la forêt de nanotube

de réduire fortement l'influence de la double diffusion sur les troncs [67]. Cet effet, en monostatique, a tendance à eclipser la diffusion volumique liée aux branches et qui permet d'extraire une information relative à la densité de biomasse. Cet effet est illustré sur la partie gauche de la figure 2.19. La configuration strictement bistatique est supposée permettre ainsi d'améliorer la mesure de biomasse.

Dans notre configuration d'acquisition la mesure polarimétrique en rétrodiffusion n'est pas réalisable, il nous est impossible ainsi d'observer des effets du type "double rebond" et leur éventuelle disparition pour des configurations bistatique voisines.

La configuration spéculaire est toutefois intéressante car elle va permettre d'exhiber un phénomène similaire. En effet dans cette configuration, c'est la diffusion spéculaire par le sol qui va eclipser la diffusion volumique, origine de la dépolarisation. Une configuration bistatique qui s'éloigne de la configuration spéculaire permet alors d'observer une importance relative de la diffusion volumique (qui a tendance à être moins directive) et donc une dépolarisation croissante. Cet effet est illustré sur la partie droite de la figure 2.19.

Nous nous intéressons ici à la dépolarisation, obtenue par une simple décomposition de Lu et Chipman. Nous représentons sur la figure 2.20 la dépolarisation pour les trois positions de la réception pour notre échantillon de forêt de Nanotubes. Nous observons bien une évolution très nette des effets de dépolarisation en fonction de la configuration géométrique considérée. Pour la configuration spéculaire, la dépolarisation est nulle, comme attendu : la diffusion par le substrat en silicium (qui joue ici le rôle du sol) est majoritaire, la diffusion de volume et donc la dépolarisation n'est pas observable. Quand on s'écarte de cette configuration, la diffusion spéculaire diminue, et on observe une part plus importante de la diffusion volumique. Ainsi on voit apparaître de la dépolarisation même si elle reste assez faible (autour de 0.2). Quand on s'écarte très nettement de la configuration spéculaire, la diffusion volumique devient majoritaire, ainsi on observe un dépolarisation très importante (autour de 0.9).

De cette série de mesure préliminaire nous pouvons donc déduire les conclusions sui-

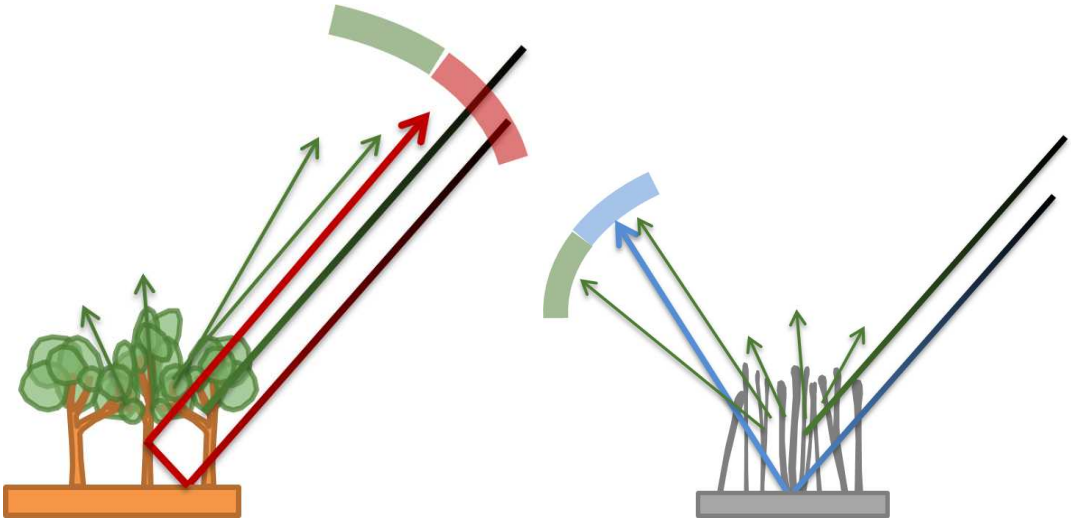


FIGURE 2.19 – En configuration spéculaire ou de rétrodiffusion, la diffusion par le sol, ou par double rebond sur le sol et les troncs, "eclipse" la diffusion volumique synonyme de dépolarisation.

vantes :

- L'impact des configurations géométrique pour l'observation de la diffusion de volume dans le cadre de l'étude des forêt a pu être illustrer expérimentalement. La diffusion spéculaire par le sol, joue ici un rôle analogue à la double diffusion par le sol et les troncs en rétrodiffusion. Elle peut être réduite, voir annulée, par un choix adéquat de la position de réception. L'analyse plus fine du signal polarimétrique afin d'en extraire des paramètres liés par exemple à la biomasse sera donc bien plus précise.
- L'exploitation d'outils de mesure polarimétriques optiques peut permettre de compléter des outils de mesures comme les chambres anéchoïques pour l'anticipation des futures campagnes polarimétriques bistatiques.

L'exploitation plus poussée de cette idée nécessiterait une modélisation plus précise de la diffusion par une forêt de nanotubes (en mesurant très précisément la densité, les dimensions et la permitivitté des nanotubes). En outre il serait souhaitable de disposer d'un outil de mesure capable de mesure précisément l'ensemble des positions angulaires, comprenant idéalement le cas de la rétrodiffusion.

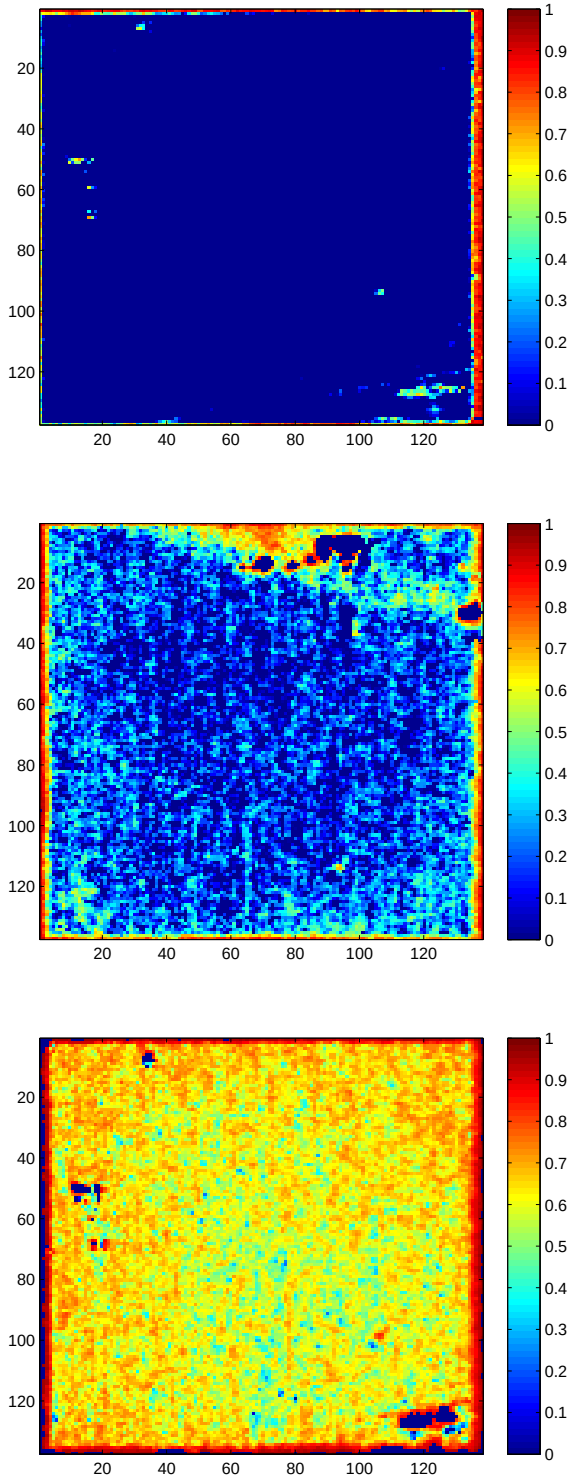


FIGURE 2.20 – Dépolarisation issue de la décomposition de Lu et Chipman pour les trois positions : *Spéculaire* en haut, *Gradient* au milieux, et *Rasante* en bas. Mesure effectuée à 650nm.

Chapitre 3

Exploitation du contraste entre milieux déterministes et milieux naturels

Dans ce chapitre nous nous intéressons à la séparation des milieux déterministes et des milieux dépolarisants. Sur des images radar par exemple, on cherchera à différencier les milieux naturels des milieux artificiels comme les bâtiments, ou véhicules. En imagerie hyperspectrale, certains travaux ont aussi été réalisés sur la détection de mines dans un environnement végétal.

La polarimétrie permet ici d'exploiter la structure souvent très ordonnée des objets manufacturés qui contraste alors du milieu végétal dont la structure est parfois désordonnée, chaotique voir fractale. cette structure entraîne une forte dépolarisation et une forte entropie. Au contraire les objets humains sont très peu dépolarisants. Ils présentent également souvent une anisotropie et une diatténuation importante. La phase ou la retardance est aussi affectée par les effets de double ou de triple rebonds. Cependant elle aura plutôt tendance à prendre des valeurs discrètes (0 ou π) qui permettront uniquement de discriminer la parité des rebonds.

Nous nous sommes rendus compte que la dépolarisation pouvait parfois être trompeuse en milieu urbain. La présence d'objets aux réponses très différentes peut induire une dépolarisation n'ayant pas une origine physique, mais uniquement liée au mélange statistique réalisé lors de l'estimation des matrices de cohérence. La segmentation est alors très utile pour minimiser l'impact de ces effets. Par ailleurs, à certaines longueurs d'onde, les surfaces d'objets manufacturés peuvent être vues comme rugueuses et dépolarisantes.

Compte tenu des données à notre disposition dans cette section nous ne présenterons que des applications radars, sur une image en bande X de Toulouse.

3.1 Milieux urbains

Comme dans le contexte agricole et forestier, les images polarimétriques urbaines peuvent donner lieu à des applications en classification. Les décompositions polarimétriques sont souvent assez efficaces pour que leurs paramètres soient utilisés dans des clas-

sifieurs classiques (kppv, réseaux de neurones etc.). Nous présenterons ainsi sur l'image de Toulouse les paramètres issus de plusieurs décompositions.

Les décompositions polarimétriques permettent d'effectuer rapidement des classifications visuelles des images et de retrouver de façon franche la majorité des bâtiments qui contraste fortement avec le reste de l'image.

Les paramètres de Cloude-Pottier représentés sur la figure 3.1 montrent que l'entropie est très élevée sur toute l'image. Une entropie faible est bien préservée sur les échos de double-rebond qui ont été bien segmentés ; par contre les segments de toits sont des surfaces souvent rugueuses (toits recouverts de gravier) dont l'entropie reste donc élevée. Ces mêmes surfaces présentent des effets de diatténuation importants, ce qui pourrait être exploité pour améliorer la discrimination. Des efforts de classification sont en cours afin de séparer au mieux les bâtiments du reste de l'image.

3.2 PolinSAR

3.2.1 Combinaison d'une segmentation polarimétrique avec une méthode d'optimisation de cohérence

Une des autres applications de la polarimétrie dans les milieux urbains se fait en complément des données interférométriques. La combinaison de la polarimétrie et de l'interférométrie est appelée PolInSAR. Sans rentrer dans les détails il faut comprendre que la synthèse d'une image SAR se fait, dans une dimension, en fonction de la distance entre l'avion ou du satellite et le point de la scène. Dans une scène en trois dimensions, il est tout à fait possible d'avoir deux objets distincts de la scène à la même distance de l'avion. Dans ce cas les deux objets vont se superposer de façon cohérente sur l'image (La cellule de résolution contiendra la somme des vecteurs de diffusion). C'est ce qu'on appelle l'effet de renversement des pentes ou *layover*. L'interférométrie exploite la phase absolue du vecteur de diffusion, ou plutôt la phase relative de deux acquisitions du vecteur de diffusion polarimétrique. Cette phase permet de retrouver la hauteur du point de l'image. L'interférométrie est donc très sensible aux effets de *layover* parce qu'alors la phase mesurée est un mélange des phases des deux points situés à la même distance de l'avion.

En milieu urbain cet effet de layover concerne souvent les bâtiments et il n'est pas rare d'observer une proportion importante de pixels où se superposent les vecteurs de diffusion du sol, et du toit du bâtiment. Le principe du layover en milieu urbain est présenté sur la figure 3.2. La polarimétrie permet, lorsque la signature polarimétrique du sol et du toit sont différentes, de séparer la contribution en phase des deux points de la scène, et de retrouver leurs hauteurs propres. C'est le principe des algorithmes d'optimisation de cohérence développés dans [80], [81]. Nous présenterons dans ce chapitre uniquement des résultats d'application croisée avec l'outil de segmentation présenté dans la troisième partie.

La thèse [79] a étudié le concept d'une optimisation de cohérence dans le cadre des calculs de hauteurs pour des cibles déterministes, comme des bâtiments. Ces cibles peuvent être considérées comme une superposition d'au plus trois mécanismes élémentaires différents, et la méthode d'optimisation de cohérence permet au mieux de calculer les hauteurs

FIGURE 3.1 – Décomposition de Cloude Pottier - entropie, alpha et anisotropie

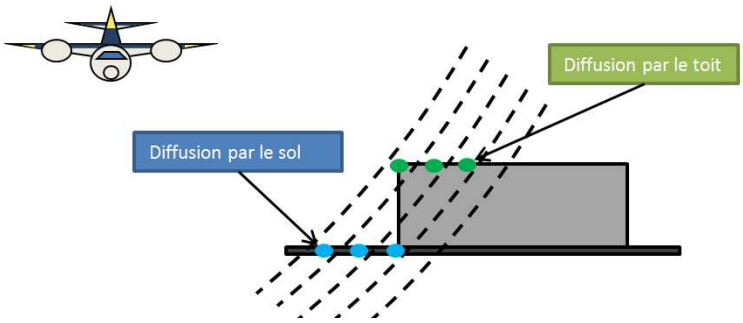


FIGURE 3.2 – Principe du layover en milieux urbains, exemple sur un bâtiment. Les lignes en pointillé représentent les lignes équidistantes à l'avion.

de ces différentes contributions. Cette méthode a montré de bonnes perspectives pour la reconstruction 3D en milieux urbains. Mais comme déjà mentionné lors de la thèse, l'un des enjeux majeurs dans l'application de cette méthode est le choix des échantillons statistiques utilisés pour calculer la cohérence de la cible. La méthode communément employée est toujours la problématique méthode des fenêtres glissantes. La phase étant un paramètre dont la variance est extrêmement importante il est très difficile d'obtenir un compromis entre la qualité de l'estimation et la préservation de la résolution. Nous pouvons observer par exemple sur la figure 3.4 qu'une fenêtre (5x5) est bien loin d'être suffisante.

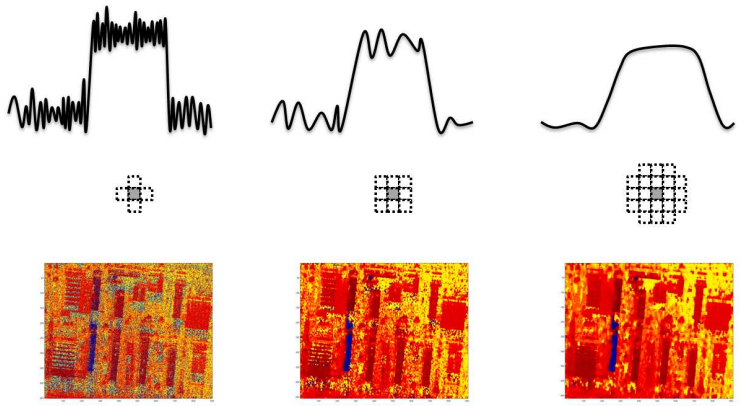


FIGURE 3.3 – Phase estimée par une fenêtre glissante sur une sous image de Toulouse. Pour différente taille de fenêtre et les effets sur la hauteur des bâtiments.

Comme pour l'estimation des matrices de cohérence polarimétrique, l'interférométrie polarimétrique est un domaine qui va très largement bénéficier de l'utilisation d'outils de segmentation pour l'estimation de ses matrices de cohérence. C'est particulièrement vrai dans le cadre de l'imagerie SAR urbaine, où l'information de hauteur prend des valeurs très contrastées et très bruitées. Nous avons donc naturellement utilisé nos résultats de segmentation sur l'image de Toulouse pour estimer les hauteurs des points de l'image par

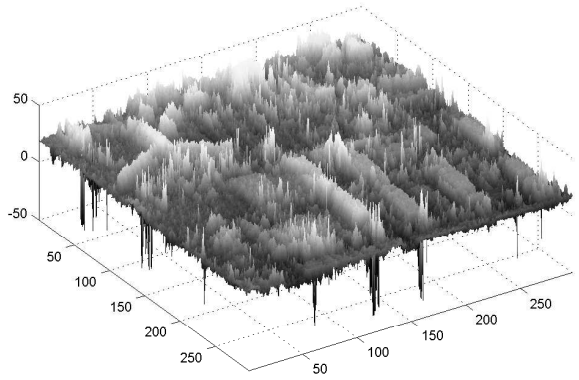


FIGURE 3.4 – Hauteurs estimées par une fenêtre glissante sur une sous image de Toulouse.

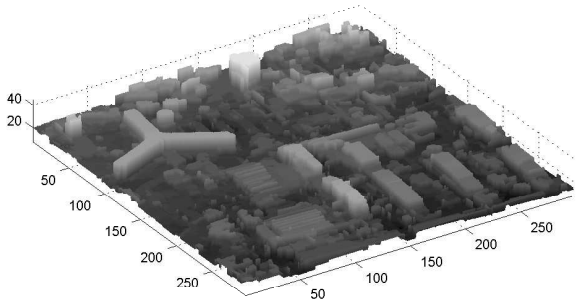


FIGURE 3.5 – Hauteur estimée sur chacune des régions issue de la segmentation polarimétrique de l'image.

région et non plus par pixels. L'algorithme d'optimisation de cohérence est alors appliqué sur les matrices de cohérence polarimétrique-interférométrique (de dimension 6), et permet de,

- Réduire la variance et améliorer l'estimation lorsque la région ne correspond qu'à une seule contribution (sol seul, bâtiment seul).
- Séparer les contributions et estimer les deux hauteurs lorsque la région contient un layover partiel ou total.

Nous pouvons ainsi observer les résultats de la méthode proposée, à comparer aux résultats précédents sur la figure 3.8.

Nous avons pu obtenir une partie de la vérité terrain d'une douzaine de bâtiments de l'image de Toulouse. Nous confrontons sur le tableau 3.1 cette vérité aux valeurs obtenues par notre estimation. La vérité terrain ainsi qu'un rendu en 3D de la scene sont par ailleurs présentés sur la figure 3.6.

FIGURE 3.6 – Vérité terrain publique obtenue grace à Google Earth, et représentée sur Google SketchUp, à comparer à une carte en 3D obtenue à l'aide du radar.

Nous pouvons constater sur ces exemples que la précision obtenue est très pertinente et que le gain est donc à la fois quantitatif et qualitatif.

3.2.2 De la compression des données en segmentation vers la restitution 3D

Dans notre processus d'estimation de la carte de hauteur, après avoir calculé la hauteur de chacune des régions de l'image, nous effectuons une seconde étape de segmentation. Celle-ci est basée uniquement sur la hauteur des régions et permet de simplifier à nouveau

Verité terrain	hauteur estimée
22	23.7
19	16.3
16	13.9
23	18.4
24	24.8
24	24.2
24	25.3
10	13.9
23	24.6
21	19.5

TABLE 3.1 – Estimation des hauteurs par rapport à la vérité terrain

la carte des hauteurs. En effet les critères utilisés pour la première segmentation garantissent une différence polarimétrique entre les différentes régions, mais nous rencontrons fréquemment des zones très différentes polarimétriquement, mais ayant des hauteurs très proches. Cette étape est très importante à la fois pour réduire le nombre de facettes nécessaires pour décrire la carte en trois dimensions mais aussi pour améliorer le rendu visuel en supprimant des effets type "marches d'escalier" ou "puzzle", sur des surfaces planes.

Enfin lors d'une dernière étape nous appliquons un algorithme classique de réduction de facette. Celui-ci à été adapté par nos soins afin de contraindre la position des sommets du maillage 3D (ce sont les *vertices*, points dans l'espace qui définissent chacun des faces triangulaires) sur un pavage régulier sur les trois dimensions. On évite ainsi des effets d'inclinaisons qui pourraient apparaître sur les bâtiments et assurant donc bien un rendu ou chacune des régions possède une hauteur constante sur sa surface. L'effet de cette réduction de facette est illustré sur la figure 3.7.

La compression 3D est une étape indispensable à la restitution de grandes images comme les nôtres. Pour donner un ordre de grandeur, après la réduction initiale liée à la première segmentation de l'ordre de 95%, une deuxième segmentation sur la hauteur de l'ordre de 50% et une réduction de facette de l'ordre de 90%, nous atteignons encore un nombre de facette de l'ordre de 5.10^5 pour l'ensemble de l'image de Toulouse (3960 * 19800 pixels, donc théoriquement 156 millions de polygones !, c'est environ le nombre moyen de polygones affichés dans le film Avatar de James Cameron). Après projection de l'image sur la carte 3D, un ordinateur moderne équipé de cartes d'accélération graphiques peine encore à produire des rendus en temps réel de l'image (Nous obtenons ainsi une poignée d'images par secondes).

A titre d'illustrations finales de cette thèse, nous proposons d'abord un extrait de la restitution 3D produite en noir et blanc (Figure 3.8), en couleur (Figure 3.9), puis une restitution de notre image sous forme d'anaglyphe sur la figure 3.10. Il s'agit d'une image créée pour être vue en relief, à l'aide de deux filtres de couleurs différentes (lunettes 3D rouge et cyan) disposés devant chacun des yeux de l'observateur. Ce principe est fondé sur la notion de stéréoscopie qui permet à notre cerveau d'utiliser le décalage entre nos

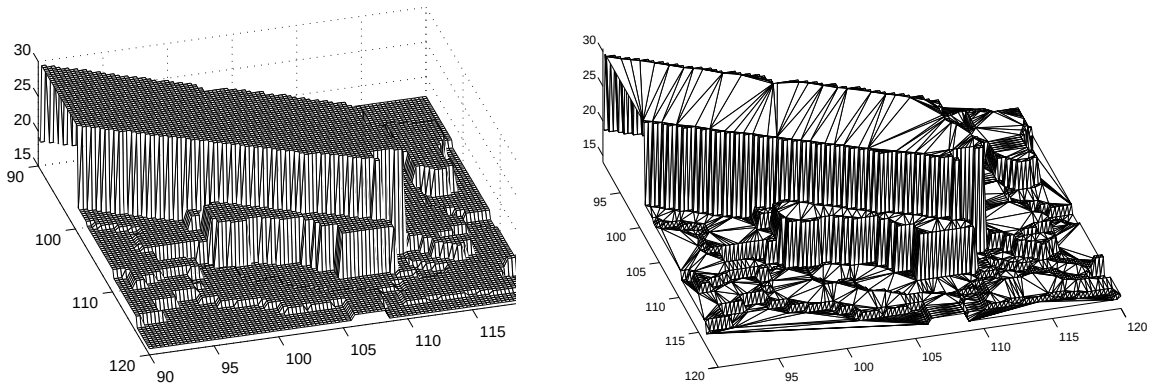


FIGURE 3.7 – Triangulation d'un bâtiment en utilisant d'abord tous les pixels, puis seulement les polygones issus de la segmentation.

FIGURE 3.8 – Illustation de la restitution 3D avec l'image d'intensité comme texture

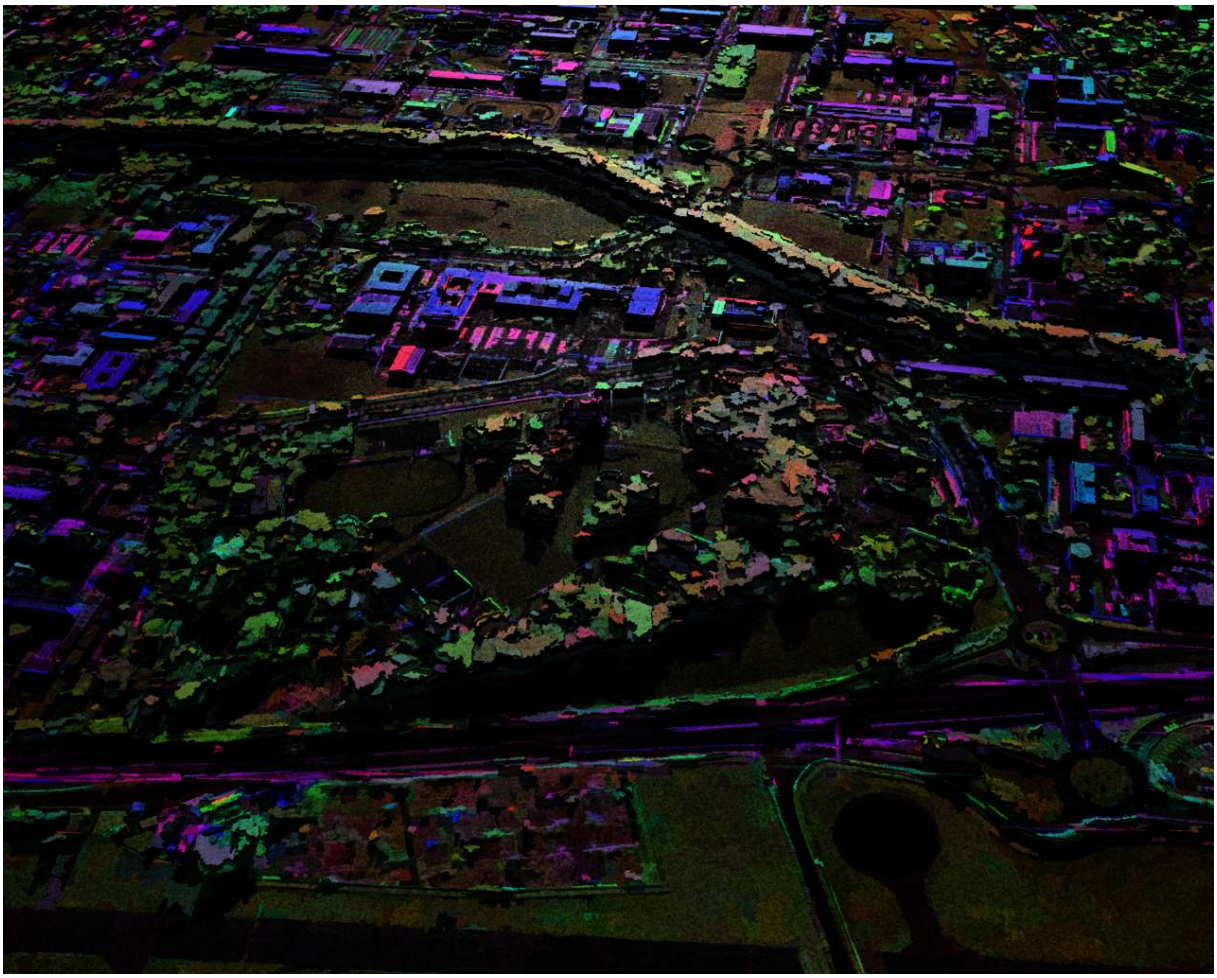


FIGURE 3.9 – Illustation de la restitution 3D avec une représentation de Pauli comme texture

deux yeux pour percevoir le relief.

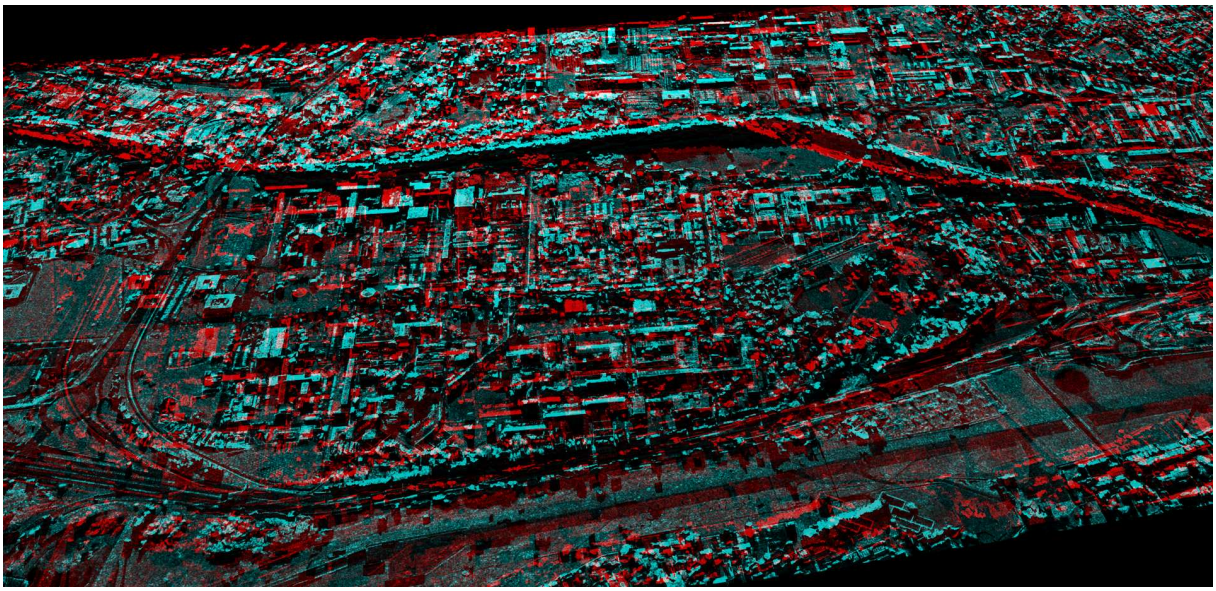


FIGURE 3.10 – Figure anaglyphe de l'image de Toulouse

Conclusion

Dans cette partie, nous avons vu que les champs d'investigation sur les applications de la polarimétrie sont multiples. Nous n'avons présenté que quelques uns d'entre elles, sur lesquels nous avons illustré l'intérêt de nos algorithmes de prétraitements comme la segmentation. Il a été également montré que l'application de décompositions traditionnellement employées en optique n'est pas absurde en radar, au contraire. Des perspectives intéressantes semblent s'offrir sur la décomposition symétrique d'Ossikovki appliquée à des images bistatiques radar. Par ailleurs, la compréhension des phénomènes de dépolarisation reste à creuser à l'aide des données optiques. Il serait intéressant de poursuivre les mesures faites sur les nanotubes avec des configurations plus proches des configurations traditionnelles radar, ou du moins mieux maîtrisées.

Conclusion Générale

Ce travail de thèse a été l'occasion d'aborder un grand nombre de problématiques liées à la polarimétrie optique et à la polarimétrie radar. En suivant l'ordre logique d'un protocole de mesure, et toutes les étapes préalables et nécessaires à l'extraction de l'information recherchée, nous avons pu confronter les outils et méthodes des deux mondes. Pour la majorité des étapes, nous avons contribué soit par l'adaptation ou l'application croisée de méthodes inédites, soit par des idées nouvelles contribuant à la valorisation de la polarimétrie.

Dans un premier temps nous avons abordé l'impact de la géométrie bistatique sur la mesure polarimétrique. Nous avons démontré l'intérêt du plan de diffusion, déjà couramment utilisé en optique, pour les études radar en configuration bistatique. Cette approche diffère naturellement des méthodes classiques utilisées en radar pour s'affranchir de l'influence de la géométrie. Nous pouvons alors, dans des configurations bistatiques, distinguer l'influence de l'orientation de cible, de celles des antennes, chose impossible avec la méthode convenue en radar.

Cette étude s'appuie sur un nombre conséquent de mesures réalisées en chambre anéchoïque, comme des volumes de billes, de clous et plusieurs sphères. Ces mesures nous ont permis alors, une fois placé dans le plan de diffusion, d'observer les effets de l'angle de bistatisme sur les paramètres polarimétriques classiques. Nous avons ainsi pu exhiber des effets induits par des angles de bistatisme importants (entre 90 et 120 degrés) qui affectent notamment des paramètres comme la diatténuation ou le retard qui dépendent des valeurs singulières des matrices de diffusion. Ces résultats nous permettent d'anticiper des problèmes qui pourraient survenir dans des futures mesures réelles en polarimétrie bistatique. Elles permettent aussi, dans le cadre de mesure bistatique en configuration mono-polar bistatiques, de déterminer la polarisation optimale à émettre en fonction des positions relatives des antennes d'émission, de réception et de la cible.

La seconde étape du travail concerne avant tout une problématique exclusive au radar qui est l'estimation des matrices de cohérence polarimétrique. Ce thème est abordé dans le contexte des images polarimétriques urbaines en haute résolution. Nous envisageons ainsi l'utilisation de modèles non gaussiens, et proposons une étude très détaillée des mesures de distances polarimétriques à la fois à l'aide de modèles gaussien et de modèles non gaussien.

Cette étude se synthétise par l'élaboration d'un outil de segmentation par croissance de régions contrainte, utilisé par la suite comme méthode d'estimation des matrices de cohérence polarimétrique sur nos images radar. Cet outil est innovant notamment par l'apport de règles et de contraintes de formes qui ont été adaptées précisément au contexte de l'imagerie urbaine.

L'étude des modèles non gaussiens est complétée par une réflexion poussée sur l'application de ces modèles dans le cadre de la détection polarimétrique de cibles dans les images SAR. De ce travail émerge des solutions innovantes et performantes dans le cadre

de la polarimétrie, mais aussi de manière plus générale dans le cadre détection multivariée en environnement non gaussien. Ces nouveaux outils de détection, appelés détecteurs hybrides sont alors testés sur des images synthétiques. Notre algorithme de segmentation est enfin mis à contribution, l'estimation des matrices de cohérence étant nécessaire le calcul des tests de détection.

La dernière partie du travail s'est concentrée sur les applications de la polarimétrie en optique et en radar. Au travers d'une étude sur les décompositions polarimétriques optique et radar nous avons pu encore démontrer l'apport des techniques optique sur des images SAR, mais aussi l'application de méthodes radar sur des données optiques.

Dans le cadre des etudes sur la polarimétrie radar en configuration bistatique nous avons pu montrer que les instruments de mesure optique sont une alternative viable aux chambres anéchoïque.

Enfin l'algorithme de segmentation développé dans la deuxième partie a été mis en oeuvre dans le cadre plus vaste de la polarimétrie-interférométrie radar, et appliqué pour la restitution de scènes urbaines en trois dimensions. De tels résultats de reconstruction en trois dimensions à cette échelle n'avaient jamais été produits.

Bibliographie

[1] N. Trouve and E. Colin-Koeniguer, "Polarimetric study of an anisotropic cloud of cylinders in a bistatic configuration," in *International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry, Italy*, 2009.

[2] N. Trouve, E. Colin-Koeniguer, P. Fargette, and A. De Martino, "Influence of geometrical configurations and polarization basis definitions on the analysis of bistatic polarimetric measurements," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 6, pp. 2238–2250, 2011.

[3] F. T. Ulaby and C. Elachi, *Radar polarimetry for geoscience applications*, F. T. Ulaby, Ed. Artech House, 1990.

[4] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. IEEE Computer Society Press, 1978.

[5] S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 498–517, 1996.

[6] S.-Y. Lu and R. A. Chipman, "Homogeneous and inhomogeneous jones matrices," *J. Opt. Soc. Am. A*, vol. 11, no. 2, pp. 766–773, 1994. [Online]. Available : http://josaa.osa.org/abstract.cfm?URI=josaa-11-2-766

[7] G. T. Ruck, *Radar cross section handbook*, Ruck G.T, Ed. Plenum, 1970.

[8] G. Mie, "Beiträge zur optik trüber medien, speziell kolloidaler metallösungen," in *Ann. Phys*, vol. 25, 1908, pp. 377–445.

[9] H. V. D. Hulst, *Light scattering by small particles*. Dover, 1981.

[10] G. E. Heath, "Bistatic scattering reflection asymmetry, polarization reversal asymmetry, and polarization reversal reflection symmetry," *IEEE Transactions on Antennas and Propagation*, vol. 29, pp. 429–434, May 1981.

[11] S.-Y. Lu and R. A. Chipman, "Interpretation of mueller matrices based on polar decomposition," *J. Opt. Soc. Am. A*, vol. 13, no. 5, p. 1106, 1996. [Online]. Available : http://josaa.osa.org/abstract.cfm?URI=josaa-13-5-1106

[12] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric sar," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, no. 1, pp. 68–78, Jan 1997.

[13] J. Lee and E. Pottier, *Polarimetric radar imaging : from basics to applications*. CRC, 2009, vol. 142.

[14] G. Vasile, J.-P. Ovarlez, F. Pascal, and C. Tison, "Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 1809–1826, 2010.

[15] P. Formont, F. Pascal, G. Vasile, J. Ovarlez, and L. Ferro-Famil, "A statistical approach to polarimetric sar data classifiaction," in *2nd SONDRA Workshop, Cargèse*, 2010.

[16] O. Harant, L. Bombrun, G. Vasile, L. Ferro-Famil, and M. Gay, "Maximum likelihood texture tracking in highly heterogeneous polsar clutter," in *Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2010, pp. 4031–4034.

[17] N. Goodman, "Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction)," *Annals of Mathematical Statistics*, pp. 152–177, 1963.

[18] K. Yao, "A representation theorem and its applications to spherically-invariant random processes," *IEEE Transactions on Information Theory*, vol. 19, no. 5, pp. 600–608, 1973.

[19] S. Yueh, J. Kong, J. Jao, R. Shin, and L. Novak, "K-distribution and polarimetric terrain radar clutter," *Journal of Electromagnetic Waves and Applications*, vol. 3, no. 8, pp. 747–768, 1989.

[20] L. Bombrun and J. Beaulieu, "Fisher distribution for texture modeling of polarimetric SAR data," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 3, 2008.

[21] J. Billingsley, A. Farina, F. Gini, M. Greco, and L. Verrazzani, "Statistical analyses of measured radar ground clutter data," *IEEE Trans. Aerospace and Electronic Systems*, vol. 35, no. 2, pp. 579–593, 1999.

[22] A. Farina, F. Gini, M. Greco, and L. Verrazzani, "High resolution sea clutter data : statistical analysis of recorded live data," in *Radar, Sonar and Navigation, IEE Proceedings-*, vol. 144, no. 3. IET, 1997, pp. 121–130.

[23] E. Stacy, "A generalization of the gamma distribution," *The Annals of Mathematical Statistics*, pp. 1187–1192, 1962.

[24] Y. Nagahara, "The pdf and cf of pearson type iv distributions and the ml estimation of the parameters," *Statistics & probability letters*, vol. 43, no. 3, pp. 251–264, 1999.

[25] G. Vasile, J. Ovarlez, F. Pascal, and M. Gay, "Estimation of the normalized coherency matrix through the SIRV model. Application to high resolution POLSAR data," *POLinSAR 2009, European Space Agency, Frascati, Italy*, vol. 668, p. 8, 2009.

[26] F. Pascal, P. Forster, J. Ovarlez, and P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2206–2217, 2008.

[27] J. Nicolas, "Introduction aux statistiques de deuxième espèce : applications des logs-moments et des logs-cumulants à l'analyse des lois d'images radar," *TS. Traitement du signal*, vol. 19, no. 3, pp. 139–167, 2002.

[28] S. Anfinsen and T. Eltoft, "Application of the matrix-variate mellin transform to analysis of polarimetric radar images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 6, pp. 2281–2295, 2011.

[29] S. Anfinsen, R. Jenssen, and T. Eltoft, "Spectral clustering of polarimetric SAR data with Wishart-derived distance measures," in *Proc. POLINSAR*, 2007.

[30] J. Neyman and E. S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Royal Society of London Philosophical Transactions Series A*, vol. 231, pp. 289–337, 1933.

[31] S. M. Kay, *Fundamentals of Statistical Signal Processing : Estimation Theory*. Upper Saddle River, New Jersey, 07458 : Prentice-Hall, 1993, vol. I.

[32] K. Conradsen, A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 1, pp. 4–19, 2003.

[33] M. Bartlett, "The statistical analysis of spatial pattern," *Advances in Applied Probability*, pp. 336–358, 1974.

[34] P. Kersten, J. Lee, and T. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, p. 519, 2005.

[35] F. Calderero and F. Marques, "General region merging approaches based on information theory statistical measures," in *15th IEEE International Conference on Image Processing, 2008. ICIP 2008*, 2008, pp. 3016–3019.

[36] F. Calderero Patino and F. Marqués Acosta, "Region merging techniques using information theory statistical measures," *IEEE transactions on Image Processing"*, vol. 19, no. 6, pp. 1567–1586, 2010.

[37] J. Morio, P. Réfrégier, F. Goudail, P. Dubois-Fernandez, and X. Dupuis, "Information theory-based approach for contrast analysis in polarimetric and/or interferometric sar images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 8, pp. 2185–2196, 2008.

[38] J. Morio, "Analyse d'images polinsar à l'aide de techniques statistiques et issues de la théorie de l'information," Ph.D. dissertation, Université Aix Marseille, 2007.

[39] F. Goudail and P. Refregier, "Contrast definition for optical coherent polarimetric images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 7, pp. 947–951, 2004.

[40] S. Kullback and R. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[41] S. Kullback, *Information theory and statistics*. Dover Pubns, 1997.

[42] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc*, vol. 35, no. 99-109, p. 4, 1943.

[43] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 493–507, 1952.

[44] J. Tilton, "Image segmentation by region growing and spectral clustering with a natural convergence criterion," in *IGARSS'98- 1998 IEEE International Geoscience and Remote Sensing Symposium, Seattle, WA*, 1998, pp. 1766–1768.

[45] L. Bombrun, G. Vasile, M. Gay, and F. Totir, "Hierarchical segmentation of polarimetric sar images using heterogeneous clutter models," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 2, pp. 726–737, 2011.

[46] J. Beaulieu and R. Touzi, "Segmentation of textured polarimetric SAR scenes by likelihood approximation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, p. 2063, 2004.

[47] J. Beaulieu, "Pseudo-convex contour criterion for hierarchical segmentation of sar images," in *Computer and Robot Vision, 2006. The 3rd Canadian Conference on*. IEEE, 2006, pp. 29–29.

[48] B. Bouchon-Meunier and L. Zadeh, *La logique floue et ses applications*. Addison-Wesley France, 1995.

[49] O. Rabaste and N. Trouvé, "Gaussian and Non Gaussian Hybrid Detectors based on Complementarity of Acceptance Regions," *submitted to IEEE Trans. Signal Processing*, under review.

[50] ——, "A new family of robust non gaussian detectors based on a geometric heuristic," in *EUSIPCO 2011 Conference*, August-September 2011.

[51] S. Kay, *Fundamentals of Statistical Signal Processing. Detection theory*. Prentice Hall, 1998.

[52] E. Jay, J.-P. Ovarlez, D. Declercq, and P. Duvaut, "BORD : Bayesian optimum radar detector," *Signal Processing*, vol. 83, pp. 1151–1162, 2003.

[53] H. Poor, *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 1988.

[54] E. Conte, M. Lops, and G. Ricci, "Asymptotically Optimum Radar Detection in Compound-Gaussian Clutter," *IEEE Trans. Aerospace and Electronic Systems*, vol. 31, no. 2, pp. 617–625, 1995.

[55] L. Scharf and B. Friedlander, "Matched Subspace Detectors," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 2146–2157, 1994.

[56] F. Bandiera, O. Besson, D. Orlando, G. Ricci, and L. Scharf, "GLRT-Based Direction Detectors in Homogeneous Noise and Subspace Interference," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2386–2394, 2007.

[57] A. De Maio, "Robust Adaptive Radar Detection in the Presence of Steering Vector Mismatches," *IEEE Trans. Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1322–1337, 2005.

[58] F. Bandiera, A. De Maio, and G. Ricci, "Adaptive CFAR Radar Detection With Conic Rejection," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2533–2541, 2007.

[59] E. Krogager, "Aspects of Polarimetric Radar Imaging," Ph.D. dissertation, Univ. Denmark, Copenhagen, 1993.

[60] A. De Maio, S. De Nicola, and A. Farina, "GLRT Versus MFLRT for Adaptive CFAR Radar Detector With Conic Uncertainty," *IEEE Signal Processing Letters*, vol. 16, no. 8, pp. 707–710, 2009.

[61] A. De Maio, S. De Nicola, Y. Huang, S. Zhang, and A. Farina, "Adaptive Detection and Estimation in the Presence of Useful Signal and Interference Mismatches," *IEEE Trans. on signal Processing*, vol. 57, no. 2, pp. 436–450, 2009.

[62] M. Desai and R. Mangoubi, "Robust Gaussian and Non-Gaussian Matched Subspace Detection," *IEEE Transactions on Signal Processing*, vol. 51, no. 12, pp. 3115–3127, 2003.

[63] E. Kelly, "An adaptive detection algorithm," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-22, no. 2, pp. 115 –127, march 1986.

[64] C. Richmond, "Performance of the Adaptive Sidelobe Blanker Detection Algorithm in Homogeneous Environments," *IEEE Trans. Signal Processing*, vol. 48, no. 5, pp. 1235–1247, 2000.

[65] C. Cochin, P. Pouliguen, B. Delahaye, D. Le Hellard, P. Gosselin, and F. Aubineau, "MOCEM - An 'all in one' tool to simulate SAR image," in *EUSAR 2008 Conference*, June 2008.

[66] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric sar," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, no. 1, pp. 68–78, Jan. 1997.

[67] L. Thirion-Lefevre, E. Colin-Koeniguer, and C. Dahon, "Bistatic scattering from forest components. part i : coherent polarimetric modelling and analysis of simulated results," *Waves in Random and Complex Media*, vol. 20, pp. 36–61, 2010.

[68] E. Colin-Koeniguer and L. Thirion-Lefevre, "Bistatic scattering from forest components. part ii : first validation of a bistatic polarimetric forest model in the vhf-uhf band [225-475 mhz] using indoor measurements," *Waves in Random and Complex Media*, vol. 20, pp. 62–85, 2010.

[69] J. Praks, E. C. Koeniguer, and M. T. Hallikainen, "Alternatives to Target Entropy and Alpha Angle in SAR Polarimetry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 2262–2274, Jul. 2009.

[70] R. Touzi, "Target scattering decomposition in terms of roll-invariant target parameters," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 1, pp. 73 –84, jan. 2007.

[71] A. Freeman and S. Durden, "A three-component scattering model for polarimetric sar data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 36, no. 3, pp. 963–973, 1998.

[72] J. Lee, M. Grunes, T. Ainsworth, L. Du, D. Schuler, and S. Cloude, "Unsupervised classification using polarimetric decomposition and the complex wishart classifier," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 37, no. 5, pp. 2249–2258, 1999.

[73] J. Souyris and C. Tison, "Polarimetric analysis of bistatic sar images from polar decomposition : A quaternion approach," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 9, pp. 2701–2714, 2007.

[74] J. Morio and F. Goudail, "Influence of the order of diattenuator, retarder, and polarizer in polar decomposition of mueller matrices," *Optics letters*, vol. 29, no. 19, pp. 2234–2236, 2004.

[75] R. Ossikovski, A. De Martino, and S. Guyot, "Forward and reverse product decompositions of depolarizing mueller matrices," *Optics letters*, vol. 32, no. 6, pp. 689–691, 2007.

[76] R. Ossikovski, "Analysis of depolarizing Mueller matrices through a symmetric decomposition," *JOSA A*, vol. 26, no. 5, pp. 1109–1118, 2009.

[77] L. Thirion, "Modélisation de l'interaction cohérente des ondes électromagnétiques avec des couverts forestiers," Ph.D. dissertation, Université Paul Sabatier, Toulouse, 2003.

[78] K. P. Papathanassiou, "Polarimetric sar interferometry," Ph.D. dissertation, Institut fur Hochfrequenztechnik Oberpfaffenhofen, 1999.

[79] E. Koeniguer, "Apport de la polarimétrie à l'interférometrie radar pour l'estimation de hauteurs de cibles et de paramètres de forêt," Ph.D. dissertation, Université Pierre et Marie Curie, Paris 6, 2005.

[80] E. Colin, C. Titin-Schnaider, and W. Tabbara, "An interferometric coherence optimization method in radar polarimetry for high-resolution imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 1, p. 167, 2006.

[81] E. Colin and F. Garestier, "Polarimetric interferometry applied to high resolution imagery and urban areas," in *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*, july 2006, pp. 4028 –4031.

Résumé

L'information polarimétrique des ondes électromagnétiques est fondamentalement semblable pour les micros ondes (radar), l'infrarouge et le visible (optique). Cependant, les échelles et les moyens de mesures impliqués varient très fortement ce qui explique que les divers outils développés (étalonnage, configuration géométrique, filtrage du bruit) divergent entre les différentes communautés. Diffèrent notamment : la géométrie d'acquisition, classiquement bistatique en optique et majoritairement monostatique en radar, la nature de la mesure, cohérente pour le radar (intensité et phase) et en énergie pour l'optique ; les traitements. Cette thèse exploite l'utilisation transverse des outils appartenant à ces deux mondes afin d'améliorer notre compréhension globale de la polarimétrie et son exploitation. Nous avons d'abord adapté une méthode de calibration inspirée de l'optique au cas du radar bistatique. Ensuite nous avons analysé l'impact des géométries bistatiques et des choix de référence, pour proposer le choix du plan de diffusion dans l'analyse des images radar. Puis nous nous sommes penchés sur l'estimation de la matrice de covariance polarimétrique dans le cadre du radar. Ceci a permis de proposer une méthode originale de segmentation des images SAR ainsi que de nouveaux types de détecteurs. Enfin, nous présentons les différents "traitements polarimétriques" à proprement parler, et diverses applications auxquelles ces traitements peuvent servir : caractérisation de milieux dépolarisants en optique (tissus) ou radar (forêt), aide à la restitution 3D.

Mots-clés: polarimétrie, optique, radar, SAR

Abstract

The polarimetric information is fundamentally similar for the micro-wave (radar), infrared and visible (optics) domains. However, measurement instruments vary considerably and therefore all the developed tools (calibration, geometric conventions, noise filtering) differ between the different communities. These differences are illustrated in particular by the following : the acquisition geometry, typically bistatic in optics and mainly monostatic in radar ; the nature of the measure, coherent for the radar (intensity and phase) and incoherent for optics ; the processing. This thesis attempts to compare the use of these tools in these two worlds to improve our understanding and use of polarimetry. A calibration method inspired by optics was first adapted to the case of bistatic radar. Then the impact of bistatic geometries and choice of reference polarimetric bases were analyzed, first for simple targets and then for depolarizing targets. That leads us to choose the diffusion plane as reference plane for conducting polarimetric analysis on bistatic radar images. Then we focus on the estimation of the covariance matrix in radar and on the statistical pre-processing. We propose here an original polarimetric segmentation scheme and a new family of polarimetric detectors. Finally, we discuss the different polarimetric decompositions, and various applications : characterization of depolarizing optical media (biological tissues) or radar (forest), help for 3D rendering.

Keywords: polarimetry,interferometry, SAR imagery, Mueller optics