



**HAL**  
open science

# Détection et Estimation en Environnement non Gaussien

Frédéric Pascal

► **To cite this version:**

Frédéric Pascal. Détection et Estimation en Environnement non Gaussien. Traitement du signal et de l'image [eess.SP]. Université de Nanterre - Paris X, 2006. Français. NNT: . tel-00128438

**HAL Id: tel-00128438**

**<https://theses.hal.science/tel-00128438v1>**

Submitted on 1 Feb 2007

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

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

# Détection et Estimation en Environnement Non Gaussien

## THÈSE

présentée et soutenue publiquement le 4 Décembre 2006

pour l'obtention du

Doctorat de l'université de Nanterre  
(spécialité Traitement du Signal)

par

Frédéric PASCAL

### Composition du jury

*Président :* FUCHS Jean-Jacques

*Rapporteurs :* BESSON Olivier  
MARCOS Sylvie

*Examineurs :* GINI Fulvio  
LARZABAL Pascal  
FORSTER Philippe (Directeur de Thèse)  
OVARLEZ Jean-Philippe (Encadrant ONERA)



# Remerciements

Cette longue aventure doctorale de trois années se termine et me voilà face au redoutable mais néanmoins extrêmement agréable exercice de style que sont les remerciements. Comme le disait mon auteur favori MP, "*Soyons reconnaissants aux personnes qui nous donnent du bonheur ; elles sont les charmants jardiniers par qui nos âmes sont fleuries.*"

Ma thèse a été effectuée au sein de l'équipe de Traitement de Signal du Département d'Electro-Magnétisme et de Radar de l'ONERA en collaboration avec le laboratoire du GEA de l'IUT de Ville d'Avray et le laboratoire SATIE de l'ENS-Cachan, j'adresse donc à ces personnes mes salutations les plus sincères ainsi que toute ma sympathie.

Cette thèse est en très grande partie le fruit d'un encadrement parfait, autant sur le plan scientifique qu'affectif. Je commencerai donc par remercier Jean-Philippe Ovarlez pour sa compétence scientifique, sa disponibilité, sa tolérance, sa gentillesse et surtout, son enseignement approfondi des échecs. Merci, Jean-Phi.

Mes deux directeurs de thèse, Philippe Forster et Pascal Larzabal, ont largement contribué à l'évolution positive de ces travaux. Ce tandem de choc a su m'apprendre les rouages subtiles du monde de la recherche ainsi que tellement d'autres choses, je vous en remercie chaleureusement. Ce triptyque d'encadrants est d'une qualité inestimable à mes yeux, merci encore à vous trois !!!

J'adresse également mes sincères remerciements aux membres du jury. Merci à Sylvie Marcos et à Olivier Besson d'avoir accepté la lourde tâche d'être rapporteur de mon manuscrit. Votre réputation de chercheur n'étant plus à faire, ce fut un honneur de vous avoir comme rapporteur. Je tiens également à remercier Jean-Jacques Fuchs, le président du jury, qui a toujours suivi, dans les rues de New York ou à Louvain la Neuve, l'avancée de mes travaux. Enfin, c'est un très grand honneur que Fulvio Gini, maître incontesté dans le domaine du radar, ait accepté de participer au jury de ma thèse. Thank you so much, Fulvio.

Tous les "matins", ce fut un plaisir d'aller bosser à l'ONERA dans une équipe qui au fil des années est devenue un groupe solide d'amis. D'un partage de voiture est née une réelle amitié, merci Yoann et bon voyage en Afrique du Sud. Qu'on se le dise, la recherche avance autour d'un café, ou d'une bière, cela dépend de l'heure et ces moments sont devenus sacrés grâce à Marc, Christèle, Philippe, Gilles et Jean-Phi. Vous avez toute mon affection, surtout ne changez rien, vous êtes parfaits ! La recherche avance aussi et surtout grâce aux thésards, merci à Gio, Seb, David, Elise, Vincent, Micka, Aurore et Antoine, encadrés par une secrétaire exceptionnelle, Elisabeth. Enfin, merci à tous ceux qui ont partagé cette aventure onérianne, j'espère n'oublier personne. Merci à Luc, Dodo, Michel, Cécile et Christian. Je n'oublierai jamais ces trois années passées avec vous, c'était du pur bonheur, de la balle et encore trop chant-mé...

## *Remerciements*

---

Je souhaite également remercier sincèrement Yacine Chitour pour m'avoir aidé dans les démonstrations théoriques grâce à ses compétences reconnues de chercheur en mathématique. Merci Yacine, et à la prochaine démo...

Un chaleureux merci à Eric Chaumette pour notre courte collaboration qui, je l'espère, en fera naître beaucoup d'autres, et pour toutes ces discussions enrichissantes aux quatre coins du monde.

Enfin, j'adresse toute ma sympathie aux personnes que j'ai côtoyées à l'ENS, Alex, nouveau docteur et ami, Fanny, Jean-Pierre, Benoît, Marc... ainsi que celles de l'IUT de Ville d'Avray, Sébastien, Guillaume, Françoise et Franck.

Cette thèse ne se serait certainement pas déroulée aussi bien sans des amis toujours présents... Je commencerai par remercier Thieums qui a toujours été présent quelle que soit la difficulté des moments. Il a en plus de son amitié, utilisé ses compétences pour corriger ce manuscrit ainsi que beaucoup d'autres travaux... "Un immense merci Matthieu!", ton ami de toujours...

J'aimerais aussi remercier Pierre et Hugo, pour leur participation au travers de longues discussions à propos de la recherche toujours menées à bien autour d'une bière et d'une table de billard. Merci à tous les deux.

J'aimerais aussi remercier sincèrement mon frère et mes parents, ainsi que la "petite famille" de Flora pour leurs encouragements, leur présence et leur amour.

Enfin, la meilleure pour la fin, ma petite Flore, qui m'a donné tout ce dont je pouvais rêver... Merci, merci, merci ! Cette thèse t'es dédiée. Emiatej...

Fred

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 État de l'Art</b>	<b>5</b>
1.1 Théorie de la détection radar . . . . .	5
1.1.1 Le détecteur optimal Gaussien ou OGD . . . . .	7
1.1.2 Mise en défaut de l'OGD en environnement non Gaussien . . . . .	8
1.2 Les vecteurs aléatoires sphériquement invariants ou SIRV . . . . .	10
1.2.1 Définition d'un SIRV . . . . .	10
1.2.2 Conditions d'identifiabilité : un problème de normalisation . . . . .	11
1.2.3 Détecteur associé aux SIRV : le BORD ou GLRT-LQ . . . . .	12
1.3 Synthèse et problématique . . . . .	14
<b>2 Estimation de la Matrice de Covariance</b>	<b>17</b>
2.1 Comment choisir les estimateurs ? . . . . .	17
2.1.1 Une normalisation appropriée . . . . .	18
2.1.2 La propriété CFAR-texture . . . . .	18
2.1.3 Les performances statistiques . . . . .	19
2.1.4 La mise en œuvre . . . . .	19
2.2 Deux estimateurs classiques, un estimateur théorique . . . . .	20
2.2.1 La SCM : le "benchmark" idéal . . . . .	20
2.2.2 La NSCM : estimateur adapté au problème . . . . .	21
2.2.3 Le TML : l'estimateur théorique parfait . . . . .	21
2.3 Performances statistiques de $\hat{\mathbf{M}}_{SCM}$ , $\hat{\mathbf{M}}_N$ , $\hat{\mathbf{M}}_{NSCM}$ et $\hat{\mathbf{M}}_{TML}$ . . . . .	23
2.3.1 Biais . . . . .	23
2.3.2 Consistance . . . . .	25
2.3.3 Moment d'ordre 2 et distribution asymptotique . . . . .	25
2.3.4 Synthèse . . . . .	28

<b>2A Preuves du chapitre 2</b>	<b>29</b>
2A.1 Preuve du théorème 2.3.1.1 : "Biais de $\widehat{\mathbf{M}}_N$ "	29
2A.2 Preuve du théorème 2.3.1.2 : "Biais de $\widehat{\mathbf{M}}_{NSCM}$ "	31
2A.3 Preuve du théorème 2.3.2.1 : "Consistance de $\widehat{\mathbf{M}}_N$ "	34
2A.4 Preuve du théorème 2.3.3.2 : "Distribution asymptotique de $\widehat{\mathbf{M}}_N$ "	35
2A.5 Preuve du théorème 2.3.3.3 : "Distribution asymptotique de $\widehat{\mathbf{M}}_{TML}$ "	36
<b>3 Un Estimateur Innovant : la Matrice du Point Fixe</b>	<b>39</b>
3.1 Introduction et justification	39
3.1.1 Cas d'un $\tau$ déterministe mais inconnu	40
3.1.2 Cas du SIRV : $\tau$ aléatoire	40
3.1.3 L'équation clé	42
3.2 Propriétés statistiques	44
3.2.1 Biais	44
3.2.2 Consistance	44
3.2.3 Distribution asymptotique	44
3.3 Propriétés "radaristiques"	46
3.3.1 Mise en œuvre	46
3.3.1.1 Influence des paramètres	48
3.3.1.2 Comportement de l'algorithme en fonction de $\rho$	50
3.3.2 Les propriétés CFAR	53
3.3.2.1 Propriété CFAR-texture	54
3.3.2.2 Propriété CFAR-matrice	54
<b>3A Preuves du chapitre 3</b>	<b>57</b>
3A.1 Preuve du théorème 3.2.1.1 : "Biais de $\widehat{\mathbf{M}}_{FP}$ "	57
3A.2 Preuve du théorème 3.2.2.1 : "Consistance de $\widehat{\mathbf{M}}_{FP}$ "	59
3A.3 Preuve du théorème 3.2.3.1 : "Distribution asymptotique de $\widehat{\mathbf{M}}_{FP}$ "	61
3A.4 Preuve du théorème 3.3.2.2 : "Propriété CFAR-matrice de $\widehat{\Lambda}(\widehat{\mathbf{M}}_{FP})$ "	65
<b>4 Bilan</b>	<b>67</b>
4.1 Les propriétés statistiques	67
4.2 Les propriétés "radaristiques"	68
4.3 Décision : quel estimateur retenir ?	69
<b>5 Application à la Détection Radar</b>	<b>71</b>
5.1 Régulation de la fausse alarme	71
5.1.1 Calcul du seuil de détection $\lambda$ pour une $P_{fa}$ donnée	71

---

5.1.1.1	Cas de d'une matrice de covariance estimée par $\widehat{\mathbf{M}}_N$ . . . . .	72
5.1.1.2	Cas de d'une matrice de covariance estimée par $\widehat{\mathbf{M}}_{FP}$ . . . . .	74
5.1.2	Simulations . . . . .	75
5.1.3	Conjecture de la distribution statistique de $\widehat{\mathbf{M}}_{FP}$ . . . . .	81
5.2	Détection dans un environnement SIRV . . . . .	86
5.2.1	Présentation des données . . . . .	86
5.2.2	Traitement des données . . . . .	86
5.3	Détection dans des zones de transition de fouillis . . . . .	93
	<b>Conclusion</b> . . . . .	<b>97</b>
	<b>Annexe A Généralités</b> . . . . .	<b>101</b>
A.1	Loi Gaussienne ou normale $\mathcal{N}(m, \sigma^2)$ définie sur $\mathbb{R}$ . . . . .	101
A.2	Densité de probabilité du vecteur Gaussien $\mathcal{N}(\mathbf{m}, \mathbf{D})$ . . . . .	101
A.3	K-distribution . . . . .	101
A.4	Loi de Weibull . . . . .	101
A.5	Loi Uniforme . . . . .	102
A.6	Loi du $\chi^2$ centré . . . . .	102
A.7	Loi Beta de première et de seconde espèce . . . . .	102
A.8	Loi de Wishart . . . . .	102
A.9	Loi faible des Grands Nombres . . . . .	103
A.10	Théorème Centrale Limite . . . . .	103
	<b>Annexe B Lemme B.0.0.1</b> . . . . .	<b>105</b>
	<b>Annexe C "First and Second Order Moments of the Normalized Sample Covariance Matrix of Spherically Invariant Random Vectors"</b> . . . . .	<b>107</b>
	<b>Annexe D Preuves des théorèmes 3.1.3.1 et 3.1.3.2 : "Covariance Structure Maximum Likelihood Estimates in Compound Gaussian Noise : Existence and Algorithm Analysis"</b> . . . . .	<b>113</b>
	<b>Publications</b> . . . . .	<b>129</b>
	<b>Bibliographie</b> . . . . .	<b>131</b>



# Acronymes

## Organismes :

MIT	Massachusetts Institute of Technology
ONERA	Office National d'Etudes et de Recherches Aérospatiales
TAD	THALES Air Defence

## Acronymes généraux :

BORD	Bayesian Optimum Radar Detector
FP	Fixed Point
GLRT	Generalized Likelihood Ratio Test
GLRT - LQ	GLRT - Linear Quadratic
LfGN	Loi faible des Grands Nombres
LRT	Likelihood Ratio Test
MV / ML	Maximum de Vraisemblance / Maximum Likelihood
MVA / AML	Maximum de Vraisemblance Approché / Approximate Maximum Likelihood
NSCM	Normalized Sample Covariance Matrix
OGD	Optimum Gaussian Detector
PDF	Probability Density Function
RSB / SNR	Rapport Signal-à-Bruit / Signal to Noise Ratio
RV	Rapport de Vraisemblance
RVG	Rapport de Vraisemblance Généralisé
SAR	Synthetic Aperture Radar
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
SCM	Sample Covariance Matrix
TCL / CLT	Théorème Central Limite / Central Limit Theorem
TFAC / CFAR	Taux de Fausse Alarme Constant / Constant False Alarm Rate
TML	Theoretical Maximum Likelihood



# Symboles et Notations

$\mathbb{N}$	Ensemble des nombres entiers
$\mathbb{R}$	Ensemble des nombres réels
$\mathbb{C}$	Ensemble des nombres complexes
$M_m(\mathbb{C})$	Ensemble des matrices de dimension $m \times m$ et composée d'éléments complexes
$\hat{\theta}$	Valeur estimée de $\theta$
$\Theta$	Ensemble des $\theta$
$\hat{\mathbf{M}}$	Estimateur de la matrice de covariance $\mathbf{M}$
$\hat{\mathbf{M}}_{SCM}$	Estimateur par SCM
$\hat{\mathbf{M}}_N$	Estimateur par SCM normalisée
$\hat{\mathbf{M}}_{NSCM}$	Estimateur par NSCM
$\hat{\mathbf{M}}_{FP}$	Estimateur du point fixe
$\hat{\mathbf{M}}_{TML}$	Estimateur par TML
$B(\hat{\mathbf{M}})$	Biais de l'estimateur $\hat{\mathbf{M}}$
$\Re(z)$	Partie réelle de $z$
$\Im(z)$	Partie imaginaire de $z$
$H$	Opérateur de transposition conjugaison
$T$	Opérateur de transposition
$\mathbf{I}$	Matrice identité de dimension appropriée au contexte
$\text{Tr}(\mathbf{M})$	Trace de la matrice $\mathbf{M}$
$\ \mathbf{M}\ $	Norme matricielle quelconque de $\mathbf{M}$
$\mathbb{P}(A)$	Probabilité de l'événement $A$
$p_X(x)$	Densité de probabilité de la variable aléatoire $X$
$p_{X Y}(x)$	Densité de probabilité de $X$ "sachant" $Y$
$\mathbb{E}(X)$	Espérance mathématique de la variable aléatoire $X$
$\mathcal{L}(X)$	Loi de la variable aléatoire $X$
$\xrightarrow[N \rightarrow +\infty]{dist.}$	Convergence en loi quand tend vers l'infini
$\xrightarrow[N \rightarrow +\infty]{Pr}$	Convergence en probabilité quand tend vers l'infini
$\xrightarrow[N \rightarrow +\infty]$	Convergence classique (presque sûre) quand tend vers l'infini

$H_0$	Hypothèse "bruit seul" du test de détection
$H_1$	Hypothèse "cible + bruit" du test de détection
$\Lambda(\mathbf{y}/H_0)$	Rapport de vraisemblance de l'observation $\mathbf{y}$ sous l'hypothèse $H_0$
$\hat{\Lambda}(\hat{\mathbf{M}})$	Rapport de vraisemblance adaptatif construit avec l'estimateur $\hat{\mathbf{M}}$
$P_{fa}$	Probabilité de fausse alarme
$P_{nd}$	Probabilité de non détection
$P_d$	Probabilité de détection
$\Gamma(x)$	Fonction Gamma de $x$
$\delta(\cdot)$	distribution de Dirac ( $\delta(x - a) = 1$ si $x = a$ , 0 sinon)
$\mathbb{1}_{[a,b]}(x)$	Fonction indicatrice de $x$ sur l'intervalle $[a, b]$
${}_bF_a(\cdot)$	Fonction hypergéométrique $F$ à $a$ et $b$ paramètres en arguments
$\mathcal{N}(m, \sigma^2)$	Loi Normale réelle univariée (moyenne $m$ , variance $\sigma^2$ )
$\mathcal{N}(\mathbf{m}, \mathbf{M})$	Loi Normale réelle multivariée (moyenne $\mathbf{m}$ , covariance $\mathbf{M}$ )
$\mathcal{CN}(\mathbf{m}, \mathbf{M})$	Loi Normale complexe multivariée (moyenne $\mathbf{m}$ , covariance $\mathbf{M}$ )
$\mathcal{U}([a, b])$	Loi Uniforme sur l'intervalle $[a, b]$
$\chi^2(n)$	Loi du $\chi^2$ à $n$ degrés de liberté
$\beta_{a,b}^1$	Loi Beta de première espèce et de paramètres $a$ et $b$
$\beta_{a,b}^2$	Loi Beta de deuxième espèce et de paramètres $a$ et $b$
$\mathcal{W}(N, m; \mathbf{M})$	Loi de Wishart de paramètres $N$ , $m$ et de matrice $\mathbf{M}$
$\text{diag}(\mathbf{x})$	Opérateur qui transforme un vecteur $\mathbf{x}$ en une matrice diagonale
$\text{vec}(\mathbf{M})$	Opérateur qui transforme une matrice en un vecteur
$\otimes$	Produit de Kronecker

# Table des figures

1.1	Mise en défaut de l'OGD dans du fouillis non Gaussien, de même puissance que le bruit thermique - Ajustement du seuil de détection . . . . .	8
1.2	Mise en défaut de l'OGD dans du fouillis non Gaussien, de même puissance que le bruit thermique - Probabilité de détection . . . . .	9
1.3	Propriété CFAR-texture du aBORD . . . . .	14
3.1	Influence de la matrice $\widehat{\mathbf{M}}(0)$ d'initialisation de l'algorithme 3.15, pour $m = 8$ et $N = 20$	47
3.2	Influence du nombre $N$ de données secondaires, pour $m = 8$ et $\widehat{\mathbf{M}}(0) = \mathbf{I}$ . . . . .	49
3.3	Influence de la taille $m$ des observations, pour $N = 200$ et $\widehat{\mathbf{M}}(0) = \mathbf{I}$ . . . . .	51
3.4	Comportement de l'algorithme pour différents coefficients de corrélation $\rho$ . . . . .	52
3.5	Comportement de la convergence de l'algorithme : $C(150)$ en fonction de $\rho$ . . . . .	52
3.6	Comportement de la convergence de l'algorithme : Nombre $k$ d'itérations, nécessaire pour atteindre $C(k) = 10^{-5}$ , en fonction de $\rho$ . . . . .	53
3.7	Propriété CFAR-matrice de $\widehat{\Lambda}(\widehat{\mathbf{M}}_{FP})$ . . . . .	55
5.1	Illustration de la convergence théorique de l'équation " $P_{fa}$ -seuil" (5.3) (i.e. quand $\mathbf{M}$ est estimée par $\widehat{\mathbf{M}}_N$ ), vers l'équation " $P_{fa}$ -seuil" (1.21) (i.e. quand $\mathbf{M}$ est connue), quand $N$ tend vers l'infini . . . . .	76
5.2	Comportement des courbes " $P_{fa}$ -seuil" pour des valeurs extrêmes de $P_{fa}$ et de seuils . . . . .	77
5.3	Validation Monte-Carlo de l'équation (5.3) . . . . .	78
5.4	Illustration de la convergence théorique de l'équation " $P_{fa}$ -seuil" (5.6) (i.e. quand $\mathbf{M}$ est estimée par $\widehat{\mathbf{M}}_{FP}$ ), vers l'équation " $P_{fa}$ -seuil" (1.21) (i.e. quand $\mathbf{M}$ est connue) quand $N$ tend vers l'infini . . . . .	78
5.5	Relation " $P_{fa}$ -seuil" obtenue pour l'estimateur du point fixe par simulations Monte-Carlo, pour $m = 10$ et pour des faibles valeurs de $N$ : $N = 20$ , $N = 50$ et $N = 100$ . . . . .	79
5.6	Zoom de la figure (Fig. 5.5) . . . . .	80
5.7	Distribution du premier élément de la diagonale, i.e. l'élément $\widehat{\mathbf{M}}(1, 1)$ dans le cas d'une matrice de Wishart et de l'estimateur du point fixe non normalisé . . . . .	82
5.8	Distribution de l'élément $\widehat{\mathbf{M}}(1, 2)$ dans le cas d'une matrice de Wishart et de l'estimateur du point fixe non normalisé . . . . .	83
5.9	Comparaison des distributions des 8 plus grandes valeurs propres des matrices $\widehat{\mathbf{M}}_W$ et $\widehat{\mathbf{M}}_{FP\ bis}$ . . . . .	84
5.10	Échos radar de fouillis de sol : 8 impulsions . . . . .	87
5.11	Échos radar de fouillis de sol : Impulsion n°1 . . . . .	88
5.12	Masque CFAR de dimension $5 \times 5$ . . . . .	89
5.13	Comparaison des relations " $P_{fa}$ -seuil" pour le GLRT-LQ construit avec l'estimateur du point fixe, dans les cas théoriques et à partir de données réelles . . . . .	90

5.14	Comparaison des relations " $P_{fa}$ -seuil" pour le GLRT-LQ construit avec l'estimateur du point fixe et construit avec d'autres estimateurs de $\mathbf{M}$ ( $\hat{\mathbf{M}}_N$ et $\hat{\mathbf{M}}_{NSCM}$ ) pour la fréquence Doppler nulle . . . . .	91
5.15	Régulation du taux de fausses alarmes pour l'OGD et le GLRT-LQ dans 5 différentes zones de fouillis, allant du bruit Gaussien jusqu'à un bruit très impulsif . . . . .	93
5.16	Performances de détection de l'OGD et du GLRT-LQ dans 5 différentes zones de fouillis, allant du bruit Gaussien jusqu'à un bruit très impulsif, tout en garantissant une $P_{fa} = 10^{-3}$ . . . . .	94
5.17	Carte du GLRT-LQ construit avec l'estimateur du point fixe . . . . .	95

# Liste des tableaux

4.1	Récapitulatif de l'étude statistique des estimateurs de la matrice de covariance $\mathbf{M}$ . . . . .	67
4.2	Récapitulatif des propriétés du GLRT-LQ construits avec les différents estimateurs de la matrice de covariance $\mathbf{M}$ . . . . .	68



# Introduction

La fonction principale d'un radar est de détecter la présence d'objets d'intérêt noyés dans un bruit ambiant considéré comme processus aléatoire. Ce bruit perturbe la qualité de détection et provient généralement de l'électronique du radar (bruit thermique), modélisé à juste titre comme du bruit Gaussien blanc. Dans certains cas, le radar doit faire face à l'environnement situé tout autour de l'objet à détecter : c'est le cas des radars terrestres de surveillance qui scrutent l'horizon du sol. Ceux-ci sont alors perturbés par les échos indésirables des réflecteurs constitutifs du sol, du sur-sol ou encore de la mer. Ces signaux parasites sont généralement assimilés à du bruit aléatoire se superposant au bruit thermique, et constituent ce que l'on nomme le fouillis. Dans le cas où ce fouillis s'apparente également à du bruit Gaussien, les détecteurs conventionnels, construits sur l'hypothèse Gaussienne, permettent alors une détection optimale.

Dans de nombreuses situations, la statistique du fouillis ne peut plus être modélisée simplement par une loi Gaussienne. Ce phénomène se produit, par exemple, pour des radars à fort pouvoir de résolution distance, ou bien lors de visées à site bas. Les exemples en pratique sont nombreux : on peut citer les problèmes de reconnaissance de cibles dans les scènes SAR (Synthetic Aperture Radar), la détection de cibles mobiles à partir de traitement STAP HR (Space Time Adaptive Processing Haute Résolution), ou encore la détection de cibles sous couvert végétal (FOPEN). Dans ces cas-là, le signal reçu provenant de chaque cellule d'analyse se compose d'un nombre trop faible de composantes pour l'assimiler à une statistique Gaussienne. Ce phénomène se produit également dans des zones où la nature du terrain est inhomogène, ce qui se traduit par une variation aléatoire spatiale de la puissance du fouillis et qui nécessite de considérer le processus de fouillis comme Gaussien composé. Une bonne description de ce phénomène est donnée dans [30] pour des échos de fouillis de mer. Dans chacune de ces situations, l'hypothèse Gaussienne n'est plus validée et le caractère impulsionnel des signaux de fouillis non Gaussien engendre une augmentation du taux de fausses alarmes, c'est à dire de fausses détections. Pour remédier à ce problème, le radar doit alors adapter son seuil de détection pour réduire le niveau de fausses alarmes, mais les performances en détection sont par conséquent fortement dégradées.

De nombreux auteurs se sont intéressés au comportement de ces détecteurs classiques en présence de fouillis non Gaussien. Les méthodes proposées ont principalement pour but de maintenir un Taux de Fausses Alarmes Constant tout au long du traitement (procédure TFAC ou CFAR - Constant False Alarm Rate - en anglais), en adaptant automatiquement le calcul du seuil de détection à la variation locale de la puissance du fouillis. Ces diverses techniques TFAC, issues d'un traitement non-cohérent sur les données radar, sont pour la plupart adaptatives. La stratégie de détection initiale ayant été obtenue sous une hypothèse Gaussienne, ces traitements ne sont pas optimaux.

La recherche de stratégies optimales de détection a été initiée par les travaux de A. Farina et al.

[15], portant, pour la première fois, sur le problème de l'élimination du fouillis non Gaussien par des traitements cohérents. Les échos radar sont alors considérés en tant que processus complexes, au sens mathématique, et toutes les informations contenues dans les composantes en quadrature du radar sont alors conservées. A partir d'une modélisation réaliste de l'environnement par des processus complexes non Gaussiens comme les SIRV (Spherically Invariant Random Vectors), A. Farina, ainsi que d'autres auteurs, ont développé des stratégies optimales de détection cohérente pour des environnements de statistique donnée. Il est ainsi possible de construire les détecteurs optimaux pour du fouillis K-distribué, Weibull, Laplace, Student-t, Cauchy, Rice,  $\chi$ , Rayleigh généralisé, et également de prendre en compte le bruit thermique Gaussien qui s'ajoute à ces perturbations.

L'utilisation des modèles SIRV a été justifiée lors de campagnes de mesures de fouillis (de sol ou de mer), réalisées par des organismes tel le MIT (Massachusetts Institute of Technology), l'ONERA (Office National d'Etudes et Recherches Aérospatiales) ou encore par [50, 28, 52, 53, 7, 8, 41, 24]. Ces modèles se sont révélés être en adéquation avec les mesures de fouillis, ce qui pourrait permettre de construire des détecteurs adaptés à leur statistique. Cependant, les stratégies optimales de détection, obtenues après modélisation de l'environnement par les SIRV, sont basées sur une parfaite connaissance *a priori* de sa statistique, et les détecteurs ne sont alors optimaux que si l'environnement suit le modèle effectivement choisi.

Ce travail a pour objectif de mieux caractériser l'environnement du radar, afin d'élaborer une stratégie de détection adaptée à tous les types de fouillis. Le détecteur n'est alors plus dépendant du choix du modèle d'environnement et peut s'utiliser dans toutes les situations. Pour ce faire, l'approche adoptée tout au long de ce document, consiste à supposer que l'environnement est modélisé par un processus SIRV. Les SIRV sont des processus aléatoires dits "composés", interprétés comme le produit de deux composantes aléatoires appelées *speckle* et *texture*. Le speckle, dont le terme a été introduit par la communauté du SAR, est représenté par un processus Gaussien complexe circulaire multi-dimensionnel, caractéristique des propriétés de "cohérence" du radar. La texture, interprétée comme la variance du processus Gaussien, est une variable aléatoire positive dont la densité caractérise spatialement les variations locales de la puissance du fouillis. Un grand nombre de lois s'obtiennent par cette modélisation, comme par exemple la K-distribution ou la loi de Weibull, représentatives de la statistique de fouillis de mer ou de terre.

La thèse d'E. Jay [31] a largement contribué à l'analyse des SIRV et à la mise en place de stratégies de détection optimales selon la densité de la texture. Cette analyse a cependant des limites puisqu'elle fait l'hypothèse de connaissance *a priori* des paramètres du fouillis, comme la texture et la matrice de covariance des données, ce qui ne peut être le cas dans la pratique. De plus, [31] a introduit également un détecteur invariant à la loi de la texture (BORD, Bayesian Optimum Radar Detector), mais toujours construit sous l'hypothèse de la matrice de covariance connue, extrêmement contraignante pour les applications.

Nous proposons, dans cette thèse, une solution au travers de l'étude de différentes procédures d'estimation de cette matrice de covariance, tout en n'ayant aucun *a priori* sur la distribution de la texture. Ces estimateurs sont ensuite utilisés dans les schémas de détection appropriés à la modélisation SIRV, tels que le GLRT-LQ (Generalized Likelihood Ratio Test - Linear Quadratic). Ces travaux s'articulent autour de deux grandes directions : une étude théorique des différents outils utilisés, puis plusieurs applications sur données réelles.

Ce mémoire est construit en cinq chapitres, au fil desquels chaque point de la démarche employée

---

est détaillé. Après un bref rappel des principes de la théorie de la détection, le chapitre 1 présente un état de l'art sur la détection radar. L'expression du détecteur classique, communément appelé l'OGD (pour Optimum Gaussian Detector), est rappelée. Les diverses méthodes visant à l'amélioration de ce détecteur sont ensuite explicitées, lorsque son optimalité est mise en défaut par la statistique non Gaussienne de l'environnement. Ces méthodes, dites TFAC, sont définies à partir du critère utilisé pour l'estimation de la puissance locale du fouillis, réalisée sur un petit nombre de données de référence. Le seuil de détection s'adapte ainsi localement aux variations du fouillis et le taux de fausses alarmes peut être respecté. Enfin, la stratégie optimale de détection, obtenue pour un environnement donné et modélisé par un SIRV, est décrite ainsi qu'une synthèse des résultats théoriques relatifs aux processus SIRV. Contrairement aux méthodes TFAC, cette approche garantit l'optimalité des détecteurs dans leur environnement désigné, et augmente très nettement le pouvoir de détection d'un radar. Mais ce cas de figure est obtenu uniquement lorsque tous les paramètres sont connus, ce qui n'est pas le cas en pratique : la texture étant aléatoire et la matrice de covariance inconnue.

Dans les chapitres suivants, les théorèmes originaux introduits et démontrés dans cette thèse sont encadrés.

Dans le chapitre 2, une synthèse des résultats théoriques relatifs aux estimateurs de matrice de covariance est décrite. Parmi ces résultats, figure une étude statistique complète des estimateurs classiquement utilisés dans la littérature. Après une brève discussion sur le choix d'un estimateur et ses conditions d'utilisation, ce chapitre présente quatre estimateurs : la très répandue SCM (pour Sample Covariance Matrix) notée  $\hat{\mathbf{M}}_{SCM}$  et distribuée selon une loi de Wishart,  $\hat{\mathbf{M}}_N$  qui est un dérivé de la SCM approprié au problème sous-jacent d'estimation, la NSCM (pour Normalized Sample Covariance Matrix), notée  $\hat{\mathbf{M}}_{NSCM}$ , et enfin l'acolyte théorique de cette dernière, le TML (pour Theoretical Maximum Likelihood). Pour chaque estimateur, un bilan statistique complet est dressé au travers de l'étude de la consistance, du biais et de la distribution asymptotique de chacun.

Afin de ne pas alourdir ce manuscrit, une annexe au chapitre 2, succédant directement l'étude statistique, contient l'ensemble des preuves les plus longues du chapitre 2.

Le chapitre 3 s'inscrit dans la continuité des deux premiers chapitres et constitue un des principaux apports de ce travail de recherche en estimation et détection radar. Un nouvel estimateur de la matrice de covariance du fouillis est introduit,  $\hat{\mathbf{M}}_{FP}$ , l'estimateur du point fixe (ou Fixed Point en anglais). Ce nom provient de la forme particulière de cet estimateur, pour lequel aucune expression analytique n'est disponible. Il est cependant défini comme étant l'unique point fixe d'une fonction  $f$ . La mise en place de cet estimateur occupe la première partie de ce chapitre. En effet, de par sa définition implicite, il est essentiel de prouver, d'une part, que cet estimateur existe et, d'autre part qu'il est unique. Puis, un algorithme itératif permettant la construction pratique de  $\hat{\mathbf{M}}_{FP}$  est proposé et analysé. Ensuite, une nouvelle étude statistique est effectuée : consistance, biais et distribution asymptotique, grâce à des techniques de perturbations. Enfin, les qualités "radaristiques" de cet estimateur pour les applications réelles sont mises en avant au travers des propriétés CFAR-texture et CFAR-matrice du détecteur construit avec le point fixe et des simulations illustrent ces propriétés. Puis, afin de "conclure" ce chapitre, une conjecture concernant la distribution de l'estimateur du point fixe est proposée et légèrement argumentée.

Comme précédemment, une annexe au chapitre 3 contient les preuves des théorèmes relatifs à l'analyse statistique et "radaristique". Quant à la mise en place de l'estimateur (existence, unicité et algorithme associé), ceci a fait l'objet d'une publication à paraître [40] qui est jointe dans l'annexe D. Ce choix est motivé par le fait que les démonstrations dans [40] font appel à de nombreuses notations et outils d'algèbre linéaire, qui pourraient alourdir la lisibilité de ce manuscrit.

Le chapitre 4 est un chapitre récapitulatif qui dresse le bilan complet des deux chapitres qui le précèdent. Au travers de tableaux récapitulatifs, il permet de comparer les propriétés de chaque estimateur afin de sélectionner celui qui sera le plus approprié au problème.

Le chapitre 5 présente un aspect plus appliqué de ce travail en testant les résultats théoriques sur des données tout d'abord simulées, puis sur des données réelles de fouillis de terre. Tout d'abord, une méthode théorique de régulation du taux de fausses alarmes est proposée en fonction de l'estimateur de la matrice de covariance utilisé. Ceci permet de calculer le seuil de détection afin de garantir une probabilité de fausse alarme fixée par l'opérateur. Cette méthode est ensuite testée sur des données SIRV simulées, puis sur des données réelles de fouillis de sol fournies par Thales Air Defense (TAD). Enfin, la dernière partie de ce chapitre met en évidence la robustesse du détecteur associé au problème dans des zones de transition de fouillis en comparant ses performances de détection au détecteur classiquement utilisé, l'OGD.

Enfin, nous concluons ce travail en essayant de dégager les points importants de cette recherche et les perspectives à envisager.

Les deux publications majeures résultant de ce travail de thèse ont été jointes dans l'annexe principale de ce manuscrit tandis que la liste complète des publications est détaillée dans une section précédant la bibliographie.

# Chapitre 1

## État de l'Art

*Ce chapitre présente la problématique étudiée dans cette thèse. Les principaux résultats concernant la détection radar sont d'abord rappelés dans un cadre général. Ensuite, la modélisation utilisée pour le fouillis additif non Gaussien est introduite, principalement au travers de la définition et des principales propriétés des SIRV. Enfin, la présentation des différents détecteurs utilisés dans cette thèse occupe la dernière partie de ce chapitre.*

### 1.1 Théorie de la détection radar

En détection radar, un des problèmes fondamentaux consiste à détecter, pour une case d'analyse fixée, dans un vecteur d'observation  $\mathbf{y}$  de dimension  $m$ , un signal complexe  $\mathbf{s}$  connu, caractérisant une cible, corrompu par un bruit de fouillis  $\mathbf{c}$  additif. Ce problème se formalise généralement par un test d'hypothèses binaires :

$$\begin{cases} H_0 : \mathbf{y} = \mathbf{c} & \mathbf{y}_i = \mathbf{c}_i \quad i = 1, \dots, N \\ H_1 : \mathbf{y} = \mathbf{s} + \mathbf{c} & \mathbf{y}_i = \mathbf{c}_i \quad i = 1, \dots, N \end{cases} \quad (1.1)$$

où les  $\mathbf{y}_i$  sont des vecteurs d'observations supposés indépendants, au sens probabiliste du terme, de  $\mathbf{y}$  et sont appelés "données secondaires" par les radaristes. Ils sont supposés contenir uniquement le bruit de fouillis additif et permettent d'estimer les paramètres inconnus du fouillis.

Sous l'hypothèse  $H_0$ , le signal complexe  $\mathbf{y}$  reçu par le radar est supposé ne contenir que les échos (bruit du fouillis) indésirables et néfastes à une bonne détection, provenant des réflexions sur les différents éléments de l'environnement, ainsi que le bruit thermique. Ces parasites sont supposés admettre une certaine densité de probabilité ou PDF (Probability Density Function), notée  $p_{\mathbf{y}}(\mathbf{y}/H_0) = p_{\mathbf{c}}(\mathbf{y}/H_0)$ .

Sous l'hypothèse  $H_1$ , le signal reçu  $\mathbf{y}$  est supposé contenir le signal  $\mathbf{s}$  provenant des échos de la cible mais noyé parmi les mêmes échos parasites que sous  $H_0$ . Sa densité de probabilité est alors notée  $p_{\mathbf{y}}(\mathbf{y}/H_1)$ .

L'objectif de la détection est de déterminer laquelle des deux hypothèses est la plus vraisemblable, tout en minimisant les deux erreurs suivantes :

- Décider  $H_0$  alors que  $H_1$  est vraie. Dans ce cas, on parle de *non-détection*, avec la probabilité  $P_{nd} = \mathbb{P}(H_0/H_1) = 1 - \mathbb{P}(H_1/H_1) = 1 - P_d$ , où  $P_d$  représente la probabilité de détection ;

- Décider  $H_1$  alors que  $H_0$  est vraie. Dans ce cas, on parle de *fausse alarme*, avec la probabilité  $P_{fa} = \mathbb{P}(H_1/H_0)$ .

Dans la pratique, il est très difficile de s'affranchir totalement de ces erreurs, à moins de connaître parfaitement la statistique de l'environnement du radar ainsi que la nature de la cible à détecter.

Un bon compromis est donné par le *critère de Neymann-Pearson* qui vise à maximiser la probabilité de détection  $P_d$  pour un taux de fausses alarmes  $P_{fa} = \alpha$  fixé. Dans la pratique, ce taux est donné par l'utilisateur et, en général, est très faible.

Le test du rapport de vraisemblance (RV), donné par

$$\Lambda(\mathbf{y}) = \frac{p(\mathbf{y}/H_1)}{p(\mathbf{y}/H_0)} \underset{H_0}{\underset{H_1}{\gtrless}} \eta, \quad (1.2)$$

est alors le test optimal (voir par exemple [33]) pour le critère de Neymann-Pearson.

Dans ce test,  $\eta$ , le *seuil de détection*, est déterminé pour avoir une probabilité de fausse alarme  $P_{fa} = \alpha$  fixée, et se calcule en résolvant l'une des deux équations suivantes :

$$P_{fa} = \mathbb{P}(\Lambda(\mathbf{y}; H_0) > \eta) = \alpha, \quad (1.3)$$

$$P_{fa} = \int_{\mathcal{D}_1} p(\mathbf{y}/H_0) d\mathbf{y}, \quad (1.4)$$

où  $\mathcal{D}_1$  est l'ensemble des  $\mathbf{y}$  contenus dans le domaine de décision de l'hypothèse  $H_1$ . Il est souvent difficile d'obtenir une expression analytique de ces équations, et le calcul de  $\eta$  s'effectue soit en résolvant numériquement (1.3) ou (1.4), soit par Monte-Carlo. Mais parfois, des miracles s'opèrent...

### Remarque 1.1.0.1

La relation entre la  $P_{fa}$  et le seuil de détection  $\eta$  n'est autre que la fonction de répartition complémentaire du détecteur  $\Lambda(\mathbf{y}; H_0)$ , elle caractérise donc complètement la distribution statistique de ce détecteur sous l'hypothèse  $H_0$ .

Une fois le seuil de détection déterminé, le calcul de  $P_d$  peut s'effectuer de deux manières :

$$P_d = \mathbb{P}(\Lambda(\mathbf{y}; H_1) > \eta), \quad (1.5)$$

$$P_d = \int_{\mathcal{D}_1} p(\mathbf{y}/H_1) d\mathbf{y}, \quad (1.6)$$

où  $\mathcal{D}_1$  est l'ensemble des  $\mathbf{y}$  contenus dans le domaine de décision de l'hypothèse  $H_1$ . Comme précédemment, il est très rare d'obtenir une expression analytique de  $P_d$ .

Ainsi, le processus de détection va se décomposer en deux étapes distinctes : la première consiste à réguler les fausses alarmes en déterminant, de manière analytique ou par Monte-Carlo, le seuil de détection pour une probabilité de fausse alarme  $P_{fa}$  donnée, tandis que la seconde étape, est le calcul de la probabilité de détection, pour la même  $P_{fa}$  et pour différents Rapports Signal-à-Bruit (RSB ou Signal-to-Noise Ratio, SNR en anglais).

L'expression du RV donnée par (1.2) suppose que le signal "cible"  $\mathbf{s}$  est déterministe et connu, ce qui n'a lieu d'être que dans une minorité de cas pratiques. Lorsque le signal "cible", supposé ici déterministe, est inconnu, un rapport de vraisemblance généralisé (RVG) est alors mis en place. Sa qualification

de *généralisé* provient du fait que les paramètres d'intérêt du signal "cible" ont été estimés au sens du maximum de vraisemblance (MV) et que ces estimateurs ont remplacé les vrais paramètres dans l'expression (1.2) du RV.

L'ensemble  $\Theta$  des paramètres d'intérêt du signal "cible" est constitué, en général, de l'amplitude complexe  $A$  du signal, de la fréquence Doppler  $f_D$  de la cible (indicatrice de la vitesse de la cible), du retard  $\tau$  du signal (indicateur de la distance de la cible) et de la direction de visée du radar. Le RV devient alors une fonction de  $\theta$  et se note  $\Lambda(\mathbf{y}, \theta)$ . De même, en écrivant le signal  $\mathbf{s}$  comme une fonction de ses paramètres,  $\mathbf{s} = f(\theta)$ , l'estimation de  $\theta$ , au sens du MV s'écrit :

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \Lambda(\mathbf{y}, \theta) = \underset{\theta}{\operatorname{argmax}} p(\mathbf{y} - f(\theta)/H_0), \quad (1.7)$$

qui se résout difficilement dans son intégralité [2].

Dans la suite, le signal "cible"  $\mathbf{s}$  est noté  $\mathbf{s} = A \mathbf{p}$ , où  $\mathbf{p}$  est le "steering vector" spécifié généralement par l'application, comprenant les valeurs estimées des paramètres autres que l'amplitude  $A$  (Doppler, direction d'arrivée, ...). Ainsi, le RVG devient

$$\Lambda(\mathbf{y}, \hat{A}) = \frac{p(\mathbf{y}, \hat{A}/H_1)}{p(\mathbf{y}/H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (1.8)$$

Il n'existe pas d'expression générale pour les estimations au sens du MV. Elles dépendent de la forme de la vraisemblance des données, dictée par la loi du vecteur  $\mathbf{y}$  sous l'hypothèse  $H_1$ .

### 1.1.1 Le détecteur optimal Gaussien ou OGD

Sous l'hypothèse classique de bruit additif Gaussien, i.e.  $\mathbf{c} \sim \mathcal{CN}(\mathbf{0}, \mathbf{M})$ ,  $\mathbf{c}$  est un vecteur Gaussien complexe de taille  $m$ , centré et de matrice de covariance  $\mathbf{M}$ , l'expression de l'estimateur du MV de  $A$  est donnée par :

$$\hat{A} = \frac{\mathbf{p}^H \mathbf{M}^{-1} \mathbf{y}}{\mathbf{p}^H \mathbf{M}^{-1} \mathbf{p}}, \quad (1.9)$$

où  $^H$  désigne l'opérateur transposé conjugué.

Il suffit maintenant de reporter l'expression de  $\hat{A}$  dans le RVG associé (1.8) pour obtenir le classique Optimum Gaussian Detector (OGD) [33] défini par :

$$\Lambda(\mathbf{y}) = \frac{|\mathbf{p}^H \mathbf{M}^{-1} \mathbf{y}|^2}{\mathbf{p}^H \mathbf{M}^{-1} \mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (1.10)$$

pour un seuil de détection  $\eta$  dont l'expression analytique est

$$\eta = \sqrt{-\ln P_{fa}}, \quad (1.11)$$

pour une probabilité de fausse alarme  $P_{fa}$  fixée.

#### Remarque 1.1.1.1 (Détecteur ou Test de RV)

Les termes de "détecteur" et "test de RV" sont indifféremment employés dans ce travail puisque qu'ils désignent le même objet mathématique mais dans des communautés différentes : "détecteur" est principalement utilisé par les radaristes tandis que "test de RV" est l'expression utilisée par les statisticiens.

La communauté de Traitement du Signal étant la jonction naturelle entre ces deux disciplines, il est plus simple d'employer ces deux termes selon le contexte, "détecteur" pour les situations plus appliquées et "test de RV" pour les démonstrations théoriques.

### 1.1.2 Mise en défaut de l'OGD en environnement non Gaussien

D'après le critère de Neymann-Pearson, l'OGD est le détecteur optimal pour une hypothèse Gaussienne. Cependant, dans un fouillis impulsionnel, ses performances de détection sont dégradées.

En effet, comme l'illustre la figure (Fig. 1.1), sur laquelle sont représentées des vraisemblances, pour un bruit thermique Gaussien (Fig. 1.1.a), puis pour un fouillis impulsionnel (Fig. 1.1.b), ici une K-distribution de paramètre  $\nu = 0.1$ , le réglage du seuil de détection  $\eta$ , défini par (1.11), courbe noire, optimal pour du bruit Gaussien, génère des fausses alarmes dans un bruit non Gaussien de même puissance. Ce dernier doit donc être "rehaussé" (souvent manuellement par l'utilisateur) pour garantir la probabilité de fausse alarme attendue, mais cette augmentation du seuil dégrade les performances de détection. Ce nouveau seuil optimal  $\eta_{opt}$  est représenté par la courbe rouge et a été calculé par une méthode Monte-Carlo. Dans le cas du bruit thermique,  $\eta_{opt}$  correspondait parfaitement au seuil théorique  $\eta$ .

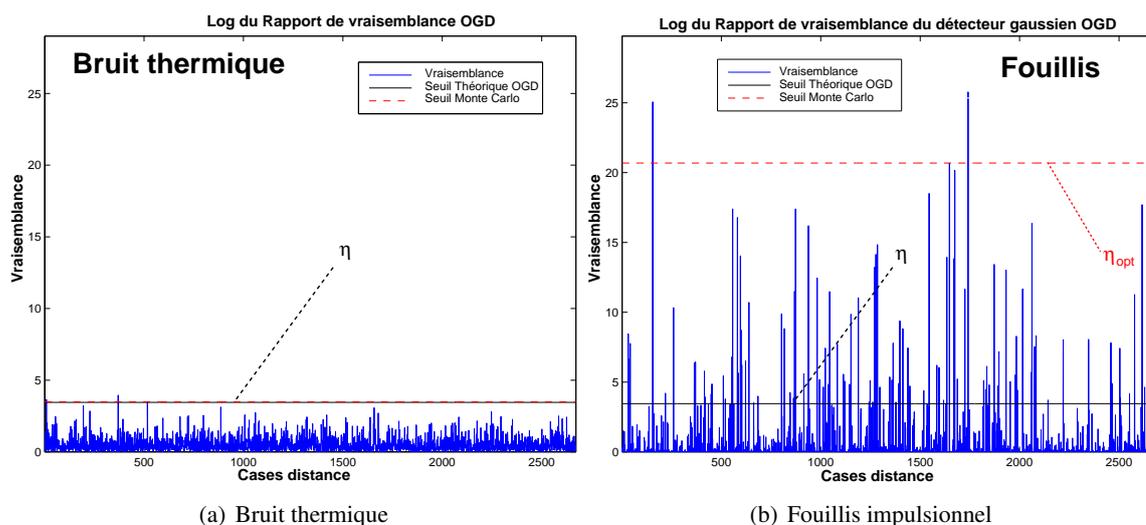


FIG. 1.1 – Mise en défaut de l'OGD dans du fouillis non Gaussien, de même puissance que le bruit thermique - Ajustement du seuil de détection

Le seuil optimal étant réglé sous ces deux hypothèses de bruit additif, la figure (Fig. 1.2) met en valeur la dégradation des performances de l'OGD. La courbe bleue représente la probabilité de détection de l'OGD dans le bruit thermique tandis que la courbe rouge représente cette même probabilité dans le fouillis non Gaussien. Le constat est sans appel : pour un même RSB, 0 db par exemple, la  $P_d$  dans le bruit thermique est proche de 0.7 mais ne vaut que  $10^{-2}$  dans le fouillis K-distribué.

Les dégradations des performances de détection sont conséquentes sous des hypothèses d'environnement non valides.

Il existe cependant des techniques d'optimisation du détecteur classique OGD en environnement non

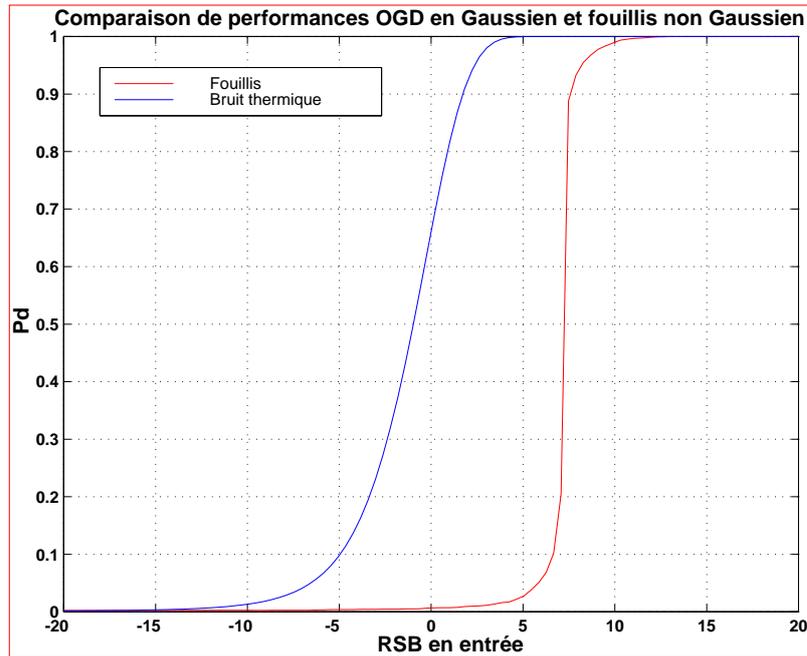


FIG. 1.2 – Mise en défaut de l’OGD dans du fouillis non Gaussien, de même puissance que le bruit thermique - Probabilité de détection

Gaussien rappelées dans [31] mais, bien qu’elles permettent de conserver la structure simple du détecteur classique en réduisant partiellement le taux de fausses alarmes, l’optimalité du détecteur n’est jamais atteinte, dans ces environnements de fouillis pour lesquels il n’est pas adapté.

Il est donc indispensable de bien connaître la statistique de cet environnement, ou dans une certaine mesure, d’utiliser une modélisation de fouillis proche de ce que révèlent les mesures réelles afin de développer des détecteurs optimisés relativement à la statistique des bruits rencontrés.

## 1.2 Les vecteurs aléatoires sphériquement invariants ou SIRV

Dans cette perspective de modélisation, des mesures expérimentales de clutter (terme anglais pour désigné le fouillis) réalisées par des organismes tels que le MIT [6] ont montré que ces données pouvaient être correctement décrites par certains modèles non Gaussiens. Une des modélisations les plus élégantes et générales, et, de surcroît, plutôt "maniable" analytiquement, provient des fameux *processus aléatoires sphériquement invariants* ou SIRP (*Spherically Invariant Random Process*). Afin de respecter l'aspect vectoriel du modèle (1.1) utilisé dans ce document, il sera en fait question de *vecteurs aléatoires sphériquement invariants* ou SIRV (*Spherically Invariant Random Vectors*) [54, 28, 7, 44, 4].

Ces modèles se caractérisent par la modulation aléatoire au cours du temps de la puissance instantanée d'un bruit Gaussien. Ils décrivent parfaitement certaines situations expérimentales (fouillis de mer, de terre), et de plus, leur loi de probabilité est connue dans de nombreux cas.

A partir de cette modélisation, le RV peut être obtenu quelle que soit la statistique de l'environnement. Son expression dépend de la loi de la variance, appelée *densité de probabilité caractéristique* du SIRV, ou plus simplement la *loi de la texture*. A chaque SIRV correspond une loi de texture, connue analytiquement ou non. Par exemple, une K-distribution correspond à un SIRV possédant une texture de loi Gamma. En revanche, une loi de Weibull correspond à un SIRV ne pouvant être caractérisé par la loi de sa texture puisque l'expression de cette dernière n'existe pas sous forme analytique. De plus, un SIRV peut être Gaussien, dans le cas spécifique d'une loi de texture distribuée selon une loi de Dirac.

À partir de cette modélisation, de nombreuses études ont été menées sur les SIRV les plus utilisés en pratique, la K-distribution et la loi de Weibull, produisant ainsi des stratégies de détection optimales pour chaque cas particulier de SIRV, et par conséquent, une littérature abondante [16, 8, 17, 18, 19].

Les SIRV ont été largement étudiés dans la littérature, et une excellente synthèse de leurs propriétés et des détecteurs construits pour les différentes lois de probabilité, est disponible dans [31], chapitre 2.

Dans la suite du document, la modélisation par des SIRV est utilisée mais en supposant la loi de la texture inconnue, ce qui permet de conserver un cadre pleinement général et réaliste dans la description du fouillis.

### 1.2.1 Définition d'un SIRV

Un processus aléatoire sphériquement invariant est un processus Gaussien non homogène de puissance aléatoire. Plus précisément, un SIRV  $\mathbf{c}$  [54] est le produit d'une variable aléatoire  $\tau$  scalaire et positive, appelée la texture, et d'un vecteur aléatoire complexe Gaussien  $\mathbf{x}$ , appelé *speckle* de dimension  $m$ , centré et ayant pour matrice de covariance  $\mathbf{M} = \mathbb{E}[\mathbf{xx}^H]$  où  $\mathbb{E}[\cdot]$  désigne l'espérance mathématique ou moyenne statistique d'une variable aléatoire :

$$\mathbf{c} = \sqrt{\tau} \mathbf{x}. \quad (1.12)$$

Il est important de remarquer que ce problème n'est pas identifiable (au sens statistique). Une condition de normalisation sur la matrice de covariance  $\mathbf{M}$  doit être imposée ; ceci fait l'objet de la partie suivante.

D'autre part, la PDF  $p_{\mathbf{c}}(\cdot)$  de  $\mathbf{c}$  est connue et son expression est la suivante :

$$p_{\mathbf{c}}(\mathbf{c}) = \int_0^{+\infty} g_m(\mathbf{c}, \tau) p_{\tau}(\tau) d\tau, \quad (1.13)$$

où  $p_{\tau}(\cdot)$  désigne la PDF (inconnue) de la texture  $\tau$  et la fonction  $g_m(\cdot, \cdot)$ , dépendant de la taille  $m$  des vecteurs est définie par :

$$g_m(\mathbf{c}, \tau) = \frac{1}{(\pi \tau)^m |\mathbf{M}|} \exp\left(-\frac{\mathbf{c}^H \mathbf{M}^{-1} \mathbf{c}}{\tau}\right). \quad (1.14)$$

Afin de définir précisément la problématique de cette thèse et d'introduire les résultats originaux de ce travail, il est nécessaire de s'attarder sur l'expression de la PDF du SIRV  $\mathbf{c}$  définie par l'équation (1.13).

En effet, cette PDF est à l'origine de la construction du test de RV ou en anglais, Likelihood Ratio Test (LRT), qui sera en fait un Rapport de Vraisemblance Généralisé (RVG ou GLRT en anglais, pour Generalized Likelihood Ratio) puisque plusieurs paramètres dans (1.13) sont inconnus : la texture  $\tau$  et la matrice de covariance  $\mathbf{M}$  du vecteur Gaussien  $\mathbf{x}$ . C'est cette hypothèse de "non connaissance" de ces deux paramètres qui permet de conserver une problématique générale et réaliste.

Tout d'abord, concernant  $\mathbf{M}$ , il faut définir la condition de normalisation adaptée. Ensuite, le détecteur associé aux SIRV et utilisé dans ces travaux, sera établi à partir de deux procédures d'estimation différentes de la texture.

## 1.2.2 Conditions d'identifiabilité : un problème de normalisation

"Identifiable", voilà un mot employé pour qualifier beaucoup d'objets, dans de nombreux domaines et dont la définition se situe entre le flou et l'obscur... Dans ce document, "identifiable" qualifie la modélisation statistique des SIRV et la définition théorique abstraite de ce terme est la suivante :

### Définition 1.2.2.1

Un modèle  $(\mathcal{X}, \mathcal{A}, \mathbb{P}_{\theta}, \theta \in \Theta)$  est dit **identifiable** si l'application de  $\Theta$  dans l'espace des probabilités sur  $(\mathcal{X}, \mathcal{A})$  qui à  $\theta$  associe  $\mathbb{P}_{\theta}$  est *injective*, avec

- $(\mathcal{X}, \mathcal{A})$ , l'espace (et la tribu associée) des réalisations de la variable aléatoire  $X$  définie sur  $(\Omega, \mathcal{F})$ ,
- $\Theta$ , l'ensemble des paramètres,
- $\mathbb{P}_{\theta}$ , une loi de probabilité sur  $(\Omega, \mathcal{F})$ .

Dans le problème étudié, l'identifiabilité du modèle SIRV porte sur le paramètre  $\mathbf{M}$ . Ainsi, pour deux matrices de covariance différentes  $\mathbf{M}_1$  et  $\mathbf{M}_2$ , les SIRV caractérisés par ces deux paramètres peuvent être les mêmes ; par exemple, pour  $\mathbf{M}_1 = 5\mathbf{M}_2$ , le SIRV qui en résulte est unique en prenant  $\tau_1 = \frac{1}{\sqrt{5}} \tau_2$ . Ainsi, les deux couples de paramètres  $(\tau, \mathbf{M})$  et  $(1/\sqrt{5} \tau, 5 \mathbf{M})$  caractérisent le même SIRV. Par conséquent, il faut définir une condition d'identifiabilité, en fait une normalisation pour la matrice de covariance  $\mathbf{M}$ , puisque que  $\tau$  est une variable aléatoire, de densité inconnue.

Dans la grande majorité des études de modèles SIRV [10, 21], la matrice de covariance  $\mathbf{M}$  est normalisée selon  $\text{Tr}(\mathbf{M}) = m$ , appelée par la suite la  $\mathbf{M}$ -normalisation, où  $\text{Tr}(\cdot)$  représente la trace d'une matrice.

### 1.2.3 Détecteur associé aux SIRV : le BORD ou GLRT-LQ

Pour construire le détecteur correspondant à la modélisation SIRV, il faut revenir à l'expression (1.8) du rapport de vraisemblance et calculer la vraisemblance sous  $H_0$  et celle sous  $H_1$  grâce à la PDF (1.13) du SIRV.

Une première approche, développée dans [32], propose une estimation bayésienne de la PDF de la texture  $\tau$  à partir d'un prior non informatif ou prior de Jeffrey (voir par exemple [45]) défini par :

$$g(\tau) = \frac{1}{\tau}. \quad (1.15)$$

En effet, pour un environnement radar inconnu, la seule connaissance apportée par la modélisation SIRV est la positivité de la variable de texture. Dans pareil cas, le prior de Jeffrey non informatif, proportionnel à la racine carrée de l'information de Fisher, est choisi afin d'éviter toute restriction sur le domaine réel de validité de la variable. Malgré le caractère *impropre* de la densité du prior de Jeffrey, définie sur  $\mathbb{R}_+$  par (1.15) [45, 46, 47, 49], son expression garanti la positivité de la variable tout en respectant le comportement en  $1/\tau$  de la variance dans l'expression de la loi Gaussienne. Ainsi défini le prior de Jeffrey est valable pour les paramètres d'échelle en général.

Ainsi, l'estimateur de  $p_\tau(\tau)$  est obtenu à partir des  $N$  données de référence  $\mathbf{c}_i$  ou données secondaires définies par le modèle (1.1) :

$$\hat{p}_N(\tau) = \frac{\tau^{-m-1}}{2^m \Gamma(m) N} \sum_{i=1}^N (\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i)^m \exp\left(-\frac{\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i}{2\tau}\right), \quad (1.16)$$

où  $\Gamma(\cdot)$  désigne la fonction Gamma donnée  $\forall z > 0$  par :

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} \exp(-t) dt, \quad (1.17)$$

avec ce cas particulier bien connu,  $\forall m \in \mathbb{N}$ ,  $\Gamma(m) = (m-1)! = (m-1).(m-2) \dots 2.1$ , le ! désignant l'opérateur factoriel.

Ensuite, en réinjectant l'estimateur de la PDF de la texture (1.16) dans le RVG (1.8), le Bayesian Optimum Radar Detector (BORD) est obtenu, mais, n'étant pas très maniable, il est préférable d'étudier sa version asymptotique (aBORD) en terme de  $N$  qui a la forme suivante :

$$\Lambda(\mathbf{M}) = \frac{|\mathbf{p}^H \mathbf{M}^{-1} \mathbf{y}|^2}{(\mathbf{p}^H \mathbf{M}^{-1} \mathbf{p})(\mathbf{y}^H \mathbf{M}^{-1} \mathbf{y})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \quad (1.18)$$

où le seuil de détection  $\lambda$  est défini par [31]

$$\lambda = 1 - \eta^{-\frac{1}{m}}. \quad (1.19)$$

L'expression du BORD Asymptotique ou aBORD coïncide avec l'expression de détecteurs obtenus par d'autres auteurs. Par exemple, K. J. Sangston et al. dans [48] l'obtiennent en considérant la texture du SIRV comme un paramètre déterministe et en la remplaçant par son estimée au sens du MV sous chacune des hypothèses.

Ce détecteur, communément appelé le GLRT-LQ pour Generalized Likelihood Ratio Test - Linear Quadratic, peut s'exprimer comme un filtre adapté classique (ou OGD) dont la puissance moyenne de

fouillis est estimée sous l'hypothèse  $H_0$ . Cela est montré dans [17], où F. Gini en donne une expression similaire.

Cette expression a également été obtenue par E.Conte et al. dans [10] après un développement asymptotique en présence de K-distribution. Antérieurement, V.A. Korado dans [35], et B. Picinbono et G. Vezzosi dans [42] avaient déjà obtenu cette expression mais sous des hypothèses de bruit Gaussien.

### 1.3 Synthèse et problématique

Le GLRT-LQ a été largement étudié dans la littérature [10, 11, 24, 32, 36] et de nombreuses "bonnes" propriétés ont été établies.

Ainsi, il apparaît rapidement que ce détecteur ne dépend pas de la loi de la texture  $\tau$  sous l'hypothèse  $H_0$ . En effet,  $\Lambda(\mathbf{M})$  définie par (1.18) peut être réécrit uniquement à partir de la variable Gaussienne  $\mathbf{x}$  du SIRV :

$$\Lambda(\mathbf{M}) = \frac{|\mathbf{p}^H \mathbf{M}^{-1} \mathbf{x}|^2}{(\mathbf{p}^H \mathbf{M}^{-1} \mathbf{p})(\mathbf{x}^H \mathbf{M}^{-1} \mathbf{x})} \underset{H_0}{\overset{H_1}{\geq}} \lambda. \quad (1.20)$$

Cette propriété aussi appelée CFAR-texture ou TFAC-texture est mise en évidence par la figure (Fig. 1.3). Celle-ci représente, en échelle logarithmique, des simulations Monte-Carlo des courbes de seuils de détection en fonction de la probabilité de fausse alarme, pour 4 différents SIRV : un Gaussien, une K-distribution de paramètre  $\nu = 0.1$ , une texture distribuée selon une loi de Weibull et un Student-t. La relation théorique (1.21) entre le seuil de détection et la  $P_{fa}$  a aussi été tracée sur ce graphique. Les 5 courbes étant pratiquement confondues, cela confirme la non dépendance de  $\Lambda(\mathbf{M})$  et de la texture.

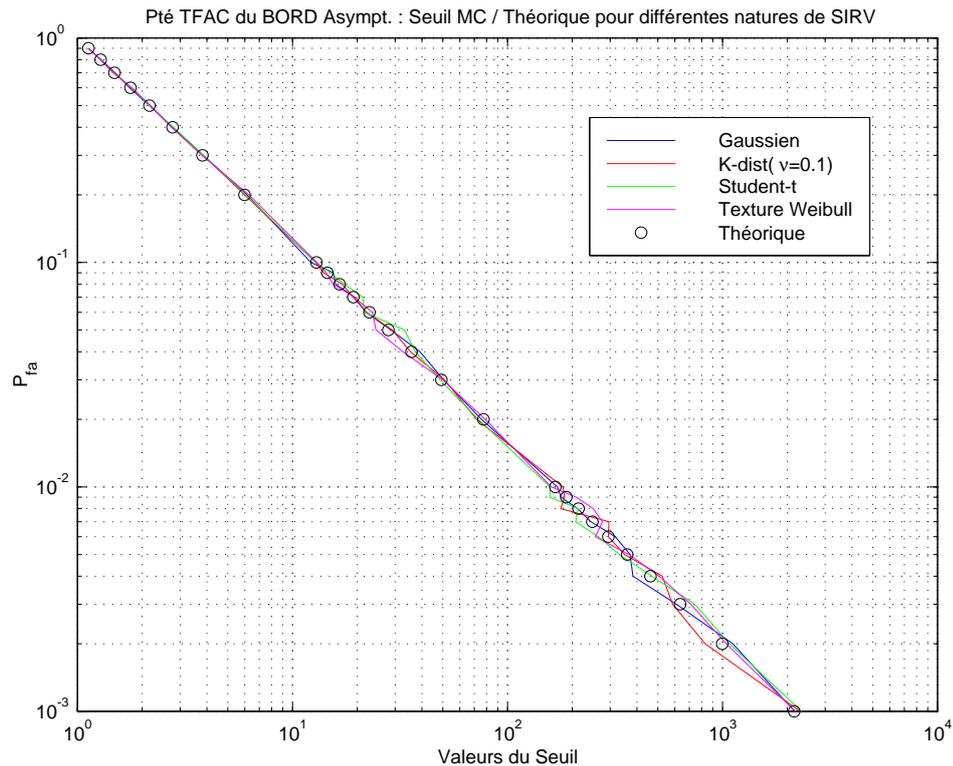


FIG. 1.3 – Propriété CFAR-texture du aBORD

De plus, la PDF de  $\Lambda(\mathbf{M})$  (1.20) a été obtenue, sous l'hypothèse de bruit seul, dans la thèse [31], ce RVG suit une loi  $\beta$  de première espèce. Ainsi, une relation analytique entre le seuil de détection et la probabilité de fausse alarme fixée a pu être établie :

$$\eta = P_{fa}^{\frac{m}{1-m}}, \quad (1.21)$$

ou de manière équivalente

$$\lambda = 1 - P_{fa}^{\frac{-1}{1-m}}. \quad (1.22)$$

Cependant, ces résultats ne sont guère applicables sur des données réelles car la matrice de covariance du fouillis n'est pas connue en pratique : il faut par conséquent l'estimer. Cet estimateur, noté  $\hat{\mathbf{M}}$ , sera ensuite injecté dans le GLRT-LQ pour donner une version adaptative du GLRT-LQ et définie par :

$$\hat{\Lambda}(\hat{\mathbf{M}}) = \frac{|\mathbf{p}^H \hat{\mathbf{M}}^{-1} \mathbf{x}|^2}{(\mathbf{p}^H \hat{\mathbf{M}}^{-1} \mathbf{p})(\mathbf{x}^H \hat{\mathbf{M}}^{-1} \mathbf{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \quad (1.23)$$

Les deux chapitres suivants sont consacrés à l'étude de différents estimateurs de matrice de covariance, puis à la mise en place d'une procédure originale d'estimation : l'estimateur du point fixe (FP). Ces estimateurs de  $\mathbf{M}$  sont construits à partir des  $N$  données secondaires modélisées par des SIRV,  $\mathbf{c}_i = \sqrt{\tau_i} \mathbf{x}_i$ , pour  $i = 1, \dots, N$ .

Ensuite, ces estimateurs sont utilisés dans la version adaptative du GLRT-LQ pour aboutir à l'application sur différentes données réelles, ce qui permet de valider leur comportement théorique.



## Chapitre 2

# Estimation de la Matrice de Covariance

*Dans ce chapitre, les estimateurs de matrice de covariance adaptés au problème sont introduits, en justifiant leur choix. Ensuite, une étude théorique complète de leurs performances est réalisée établissant ainsi les premiers résultats nouveaux de cette thèse. Enfin, leurs inconvénients respectifs sont mis en lumière dans la recherche d'un estimateur idéal.*

*Afin de ne pas alourdir le corps de ce chapitre, les principales démonstrations des théorèmes sont développées dans l'annexe de ce chapitre.*

### 2.1 Comment choisir les estimateurs ?

Le choix des estimateurs de la matrice de covariance du fouillis peut être décisif dans les performances du détecteur qui en résulte. Il est donc important que ce choix puisse être justifié en vue de résultats attrayants. L'estimateur parfait étant impossible à construire dès lors que le bruit n'est plus Gaussien, il faut faire certains compromis...

Que peut-on attendre d'un estimateur d'une matrice de covariance d'un détecteur ? Ou, plus précisément, comment peut-on le construire ? Plusieurs points sont à prendre en compte :

- Le paramètre à estimer étant normalisé pour des raisons d'identifiabilité du problème, il faut définir une condition de normalisation pour l'estimateur  $\hat{\mathbf{M}}$  : la même que pour  $\mathbf{M}$  ou une autre ?
- Le GLRT-LQ possède la propriété CFAR-texture ; si  $\mathbf{M}$  est remplacé par  $\hat{\mathbf{M}}$  dans le détecteur, il serait intéressant que ce dernier conserve cette propriété, puisque cela permet de s'affranchir d'une variable aléatoire inconnue et donc de sa PDF.
- En termes statistiques, l'estimateur doit converger vers le paramètre qu'il estime : il doit être consistant. De plus, s'il est sans biais, les performances de détection n'en seront qu'améliorées. Enfin, selon la vitesse de convergence de son moment d'ordre 2 vers 0, un estimateur peut être meilleur que les autres.
- Enfin, la procédure d'estimation ne doit pas être trop lourde, en termes de temps de calcul ou de charge de calcul, car elle s'inscrit dans une procédure de détection et par conséquent, le ou les estimateurs retenus doivent être "maniabiles".

### 2.1.1 Une normalisation appropriée

Afin de satisfaire des conditions d'identifiabilité du modèle SIRV, la matrice de covariance  $\mathbf{M}$  du fouillis est normalisée selon  $\text{Tr}(\mathbf{M}) = m$  comme détaillé dans la section 1.2. Il paraît donc naturel de normaliser les estimateurs  $\hat{\mathbf{M}}$  de  $\mathbf{M}$  selon la  $\mathbf{M}$ -normalisation :  $\text{Tr}(\hat{\mathbf{M}}) = m$ , d'autant que cette normalisation est extrêmement simple à appliquer en pratique.

Cependant, comme nous le verrons plus tard, la normalisation appropriée utilisée dans cette thèse sera  $\text{Tr}(\mathbf{M}^{-1}\hat{\mathbf{M}}) = m$ , appelée la  $\hat{\mathbf{M}}$ -normalisation pour chaque estimateur  $\hat{\mathbf{M}}$  de  $\mathbf{M}$ . Il est essentiel de remarquer que cette normalisation est judicieuse pour les estimateurs consistants mais est beaucoup plus contraignante que la  $\mathbf{M}$ -normalisation, voire impossible à mettre en pratique car l'estimateur  $\hat{\mathbf{M}}$  de  $\mathbf{M}$  ne peut-être normalisé par une fonction de  $\mathbf{M}$ ...

#### Définition 2.1.1.1

Un estimateur  $\hat{\mathbf{M}}$  de  $\mathbf{M}$  est dit consistant s'il converge en probabilité vers  $\mathbf{M}$  quand  $N$  tend vers l'infini :

$$\forall \varepsilon > 0, \mathbb{P}(\|\hat{\mathbf{M}} - \mathbf{M}\| \geq \varepsilon) \xrightarrow{N \rightarrow +\infty} 0, \quad (2.1)$$

où  $N$  est le nombre de données secondaires  $\mathbf{c}_i$  utilisées pour estimer  $\mathbf{M}$  et  $\|\cdot\|$  désigne une norme matricielle quelconque.

Ainsi, pour un estimateur  $\hat{\mathbf{M}}$  consistant,  $\text{Tr}(\mathbf{M}^{-1}\hat{\mathbf{M}})$  tend vers  $\text{Tr}(\mathbf{I})$  quand  $N$  tend vers l'infini et  $\mathbf{I}$ , désignant la matrice identité, est de trace  $m$ .

Pourtant, il est impossible d'utiliser la  $\hat{\mathbf{M}}$ -normalisation en pratique dès lors que  $\mathbf{M}$  est inconnue. Ceci ne pose aucun problème pour vérifier les résultats à l'aide de simulations mais s'avère problématique dans le cas de données réelles où il deviendra alors nécessaire d'utiliser la  $\mathbf{M}$ -normalisation.

De plus, dans de nombreux problèmes, notamment celui de la détection radar, le facteur de normalisation de l'estimateur  $\hat{\mathbf{M}}$  n'a aucune influence sur le résultat final. En effet, le détecteur étudié, le GLRT-LQ adaptatif, est un rapport de vraisemblance homogène en terme de  $\hat{\mathbf{M}}$ . Ainsi, quel que soit le facteur de normalisation utilisé, ce dernier se simplifie dans le rapport. Ceci fait l'objet de la proposition suivante :

#### Proposition 2.1.1.1

$$\forall \alpha \in \mathbb{C}, \hat{\Lambda}(\alpha \hat{\mathbf{M}}) = \hat{\Lambda}(\hat{\mathbf{M}}).$$

La proposition 2.1.1.1 donne, dans ce contexte de détection radar en environnement non Gaussien, une certaine flexibilité dans le choix de la normalisation de l'estimateur. Ainsi, pour des raisons théoriques, la normalisation utilisée est la  $\hat{\mathbf{M}}$ -normalisation définie par :

$$\text{Tr}(\mathbf{M}^{-1}\hat{\mathbf{M}}) = m. \quad (2.2)$$

### 2.1.2 La propriété CFAR-texture

Le détecteur GLRT-LQ construit avec la vraie matrice de covariance  $\mathbf{M}$  possède la propriété CFAR-texture définie par :

### Définition 2.1.2.1 (Propriété CFAR-texture)

Un élément mathématique construit à partir des données (détecteur, estimateur, ...) possède la propriété CFAR-texture si sa distribution ne dépend pas de celle de la texture.

Cette propriété est d'une importance capitale puisque dans la modélisation SIRV retenue, la texture est distribuée selon une loi inconnue. Elle permet donc de s'affranchir d'un paramètre aléatoire inconnu et par conséquent, il tout est naturel de vouloir choisir des estimateurs de  $\mathbf{M}$  qui, réinjectés dans le détecteur étudié (la version adaptative du GLRT-LQ), permettent à ce dernier d'être CFAR-texture.

Pour respecter cette condition, une solution peut être de construire des estimateurs  $\hat{\mathbf{M}}$  ne dépendant que des  $\mathbf{x}_i$  dans les données secondaires.

### 2.1.3 Les performances statistiques

Afin de sélectionner un estimateur, les performances statistiques de chaque estimateur proposé sont étudiées :

- la consistance, condition indispensable et définie par 2.1.1.1,
- le biais,
- le moment d'ordre 2,
- le comportement asymptotique, plus précisément, la distribution asymptotique.

### 2.1.4 La mise en œuvre

Un élément essentiel est encore à mentionner dans le choix de l'estimateur.  $\hat{\mathbf{M}}$  doit être "maniable", i.e. son expression analytique doit être suffisamment simple pour permettre une analyse théorique de ses propriétés statistiques et pour interpréter physiquement cet estimateur. De plus, étant donné que  $\hat{\mathbf{M}}$  s'intègre à un processus de détection, le temps et la charge de calcul nécessaires à sa construction doivent être les plus faibles possibles, ceci afin de ne pas pénaliser l'étape de détection.

En effet, l'estimateur choisi, bien qu'indispensable, n'est qu'un outil et est ensuite utilisé sur données réelles. Il ne sera donc retenu, par les radaristes, que s'il ne pénalise pas la détection...

## 2.2 Deux estimateurs classiques, un estimateur théorique

Dans cette section, les deux estimateurs les plus en vue dans la littérature ainsi qu'un nouvel estimateur à vocation purement théorique, sont présentés puis analysés.

### 2.2.1 La SCM : le "benchmark" idéal

Tout d'abord, si les  $N$  réalisations indépendantes des données Gaussiennes  $\mathbf{x}_k$ , présentes dans les SIRV  $\mathbf{c}_k = \sqrt{T_k} \mathbf{x}_k$ , étaient observables, la théorie de l'estimation par maximum de vraisemblance fournirait comme estimateur la Sample Covariance Matrix (SCM) bien connue, qui suit une loi de Wishart et qui est définie par :

$$\widehat{\mathbf{M}}_{SCM} = \frac{1}{N} \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^H. \quad (2.3)$$

En pratique, ce sont seulement les  $N$  réalisations indépendantes des SIRV,  $\mathbf{c}_1, \dots, \mathbf{c}_N$  qui sont observables et il est impossible d'isoler le processus Gaussien  $\mathbf{x}$ . Cet estimateur, dont l'utilisation est largement répandue, est étudié et utilisé comme référence ou benchmark, auquel les autres estimateurs sont comparés. De plus, l'analyse des performances de  $\widehat{\mathbf{M}}_{SCM}$  autorise une interprétation fort intéressante des résultats théoriques obtenus pour les autres estimateurs.

Afin de respecter la  $\widehat{\mathbf{M}}$ -normalisation, l'estimateur  $\widehat{\mathbf{M}}_N$  défini de la façon suivante ,

$$\widehat{\mathbf{M}}_N = \frac{m}{\text{Tr}(\mathbf{M}^{-1} \widehat{\mathbf{M}}_{SCM})} \widehat{\mathbf{M}}_{SCM}, \quad (2.4)$$

sera étudié. Il peut se réécrire comme

$$\widehat{\mathbf{M}}_N = \frac{m}{N} \frac{\sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^H}{\sum_{k=1}^N \mathbf{x}_k^H \mathbf{M}^{-1} \mathbf{x}_k}. \quad (2.5)$$

En effet, en utilisant la propriété  $\text{Tr}(\mathbf{A}\mathbf{B}) = \text{Tr}(\mathbf{B}\mathbf{A})$  pour toutes matrices  $\mathbf{A}, \mathbf{B}$ , on a

$$\begin{aligned} \text{Tr}(\mathbf{M}^{-1} \widehat{\mathbf{M}}_{SCM}) &= \text{Tr}(\mathbf{M}^{-1/2} \widehat{\mathbf{M}}_{SCM} \mathbf{M}^{-1/2}) \\ &= \sum_{k=1}^N \text{Tr}(\mathbf{M}^{-1/2} \mathbf{x}_k \mathbf{x}_k^H \mathbf{M}^{-1/2}), \text{ par linéarité de la trace,} \\ &= \sum_{k=1}^N \text{Tr}(\mathbf{x}_k^H \mathbf{M}^{-1/2} \mathbf{M}^{-1/2} \mathbf{x}_k) \\ &= \sum_{k=1}^N \mathbf{x}_k^H \mathbf{M}^{-1} \mathbf{x}_k, \text{ car } \mathbf{x}_k^H \mathbf{M}^{-1} \mathbf{x}_k \text{ est un scalaire.} \end{aligned} \quad (2.6)$$

Les deux estimateurs précédents (2.3) et (2.4) ont uniquement un intérêt théorique puisque que les  $\mathbf{x}_k$  ne sont pas disponibles. Les estimateurs utilisables sont des fonctions des  $\mathbf{c}_k$  et, parmi ces derniers, les

”bons” estimateurs ne devraient pas dépendre des  $\tau_k$ .

$\widehat{\mathbf{M}}_{SCM}$  a été très largement étudié dans la littérature [29, 51]. Ces propriétés statistiques sont rappelées par la suite tandis que l’annexe A.8 est consacrée à un rappel succinct de la théorie des matrices aléatoires et, en particulier, de la distribution de Wishart. Le caractère de benchmark attribué à  $\widehat{\mathbf{M}}_{SCM}$  provient de toutes ses bonnes propriétés : consistance, absence de biais et distribution asymptotique Gaussienne.

### 2.2.2 La NSCM : estimateur adapté au problème

Le second candidat à l’estimation de  $\mathbf{M}$  est la Normalized Sample Covariance Matrix (NSCM) très largement répandue et introduite par [9] dans le contexte des modèles SIRV.

$\widehat{\mathbf{M}}_{NSCM}$  est définie par :

$$\widehat{\mathbf{M}}_{NSCM} = \frac{m}{N} \sum_{k=1}^N \frac{\mathbf{c}_k \mathbf{c}_k^H}{\mathbf{c}_k^H \mathbf{c}_k}, \quad (2.7)$$

qui peut être réécrite uniquement en termes de  $\mathbf{x}_k$

$$\widehat{\mathbf{M}}_{NSCM} = \frac{m}{N} \sum_{k=1}^N \frac{\mathbf{x}_k \mathbf{x}_k^H}{\mathbf{x}_k^H \mathbf{x}_k}. \quad (2.8)$$

Plusieurs remarques s’imposent en considérant l’expression de  $\widehat{\mathbf{M}}_{NSCM}$  : tout d’abord, cet estimateur ne dépendant pas des  $\tau_i$ , le détecteur associé  $\widehat{\Lambda}(\widehat{\mathbf{M}}_{NSCM})$  possède la propriété CFAR-texture. D’autre part, l’équation (2.8) montre que  $\widehat{\mathbf{M}}_{NSCM}$  est, par construction, normalisé selon  $\text{Tr}(\widehat{\mathbf{M}}_{NSCM}) = m$ . Afin de ne pas alourdir la définition de  $\widehat{\mathbf{M}}_{NSCM}$ , la  $\mathbf{M}$ -normalisation est conservée pour  $\widehat{\mathbf{M}}_{NSCM}$ . Enfin, cet estimateur étant biaisé et, par conséquent, non consistant, il sera rapidement écarté de l’étude des estimateurs potentiels de  $\mathbf{M}$ .

### 2.2.3 Le TML : l’estimateur théorique parfait

Le Theoretical Maximum Likelihood,  $\widehat{\mathbf{M}}_{TML}$  est défini de la façon suivante :

$$\widehat{\mathbf{M}}_{TML} = \frac{m}{N} \sum_{k=1}^N \frac{\mathbf{c}_k \mathbf{c}_k^H}{\mathbf{c}_k^H \mathbf{M}^{-1} \mathbf{c}_k},$$

qui peut aussi se réécrire uniquement à partir des  $\mathbf{x}_k$

$$\widehat{\mathbf{M}}_{TML} = \frac{m}{N} \sum_{k=1}^N \frac{\mathbf{c}_k \mathbf{c}_k^H}{\mathbf{c}_k^H \mathbf{M}^{-1} \mathbf{c}_k}, \quad (2.9)$$

où les  $\mathbf{x}_k$  sont les données Gaussiennes du SIRV  $\mathbf{c}_k$ . Cet estimateur n’a aucun sens en pratique, puisqu’il dépend du paramètre  $\mathbf{M}$  qu’il doit estimer. Cependant, son analyse a révélé de très bonnes propriétés théoriques, il est donc présenté comme la NSCM ”idéale”. En effet, le TML peut s’exprimer à l’aide de la NSCM, il suffit pour cela de blanchir les  $\mathbf{x}_k$ ’s selon  $\mathbf{y}_k = \mathbf{M}^{-1/2} \mathbf{x}_k$  :

$$\mathbf{M}^{-1/2} \widehat{\mathbf{M}}_{TML} \mathbf{M}^{-1/2} = \frac{m}{N} \sum_{k=1}^N \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k} = \widehat{\mathbf{M}}_{NSCM, \mathbf{I}}, \quad (2.10)$$

où  $\widehat{\mathbf{M}}_{NSCM, \mathbf{I}}$  est l'estimateur NSCM de la matrice identité  $\mathbf{I}$ .

De plus,  $\widehat{\mathbf{M}}_{TML}$  est, par définition, normalisé selon la  $\widehat{\mathbf{M}}$ -normalisation et conduit à la propriété CFAR-texture du GLRT-LQ.

## 2.3 Performances statistiques de $\widehat{\mathbf{M}}_{SCM}$ , $\widehat{\mathbf{M}}_N$ , $\widehat{\mathbf{M}}_{NSCM}$ et $\widehat{\mathbf{M}}_{TML}$

Cette section a pour l'objectif l'analyse des performances des 4 estimateurs  $\widehat{\mathbf{M}}_{SCM}$ ,  $\widehat{\mathbf{M}}_N$ ,  $\widehat{\mathbf{M}}_{NSCM}$  et  $\widehat{\mathbf{M}}_{TML}$  : biais, consistance, moment d'ordre 2 ainsi que distribution asymptotique.

La notion de consistance est rappelée par la définition 2.1.1.1 tandis que celle du biais est la suivante :

### Définition 2.3.0.1

Le biais  $B(\widehat{\mathbf{M}})$  d'un estimateur  $\widehat{\mathbf{M}}$  de  $\mathbf{M}$  est défini par :

$$B(\widehat{\mathbf{M}}) = \mathbb{E}[\widehat{\mathbf{M}}] - \mathbf{M}. \quad (2.11)$$

De plus, lorsque  $B(\widehat{\mathbf{M}}) = 0$ , l'estimateur  $\widehat{\mathbf{M}}$  est appelé estimateur sans biais ou non biaisé de  $\mathbf{M}$ .

### 2.3.1 Biais

La Sample Covariance Matrix  $\widehat{\mathbf{M}}_{SCM}$  est évidemment un estimateur sans biais de  $\mathbf{M}$ . Maintenant, il faut analyser le biais de  $\widehat{\mathbf{M}}_N$ , ceci fait l'objet du théorème suivant :

**Théorème 2.3.1.1 (Biais de  $\widehat{\mathbf{M}}_N$ )**  
 $\widehat{\mathbf{M}}_N$  est un estimateur sans biais de  $\mathbf{M}$ .

#### Preuve 2.3.1.1

Voir l'annexe 2A.1 du chapitre.

Le biais de  $\widehat{\mathbf{M}}_{NSCM}$  est étudié pour deux cas distincts, selon les hypothèses faites sur  $\mathbf{M}$ . Dans le cas le plus général, le théorème 2.3.1.2 donne une expression analytique de  $B(\widehat{\mathbf{M}}_{NSCM})$  tandis que le théorème 2.3.1.3 prouve que  $\widehat{\mathbf{M}}_{NSCM}$  est un estimateur sans biais dans le cas particulier où  $\mathbf{M} = \mathbf{I}$ , ce qui signifie que  $\widehat{\mathbf{M}}_{TML}$  est un estimateur sans biais.

**Théorème 2.3.1.2 (Biais de  $\widehat{\mathbf{M}}_{NSCM}$  quand  $\mathbf{M}$  a des valeurs propres distinctes)**

Sous l'hypothèse que  $\mathbf{M}$  a des valeurs propres distinctes, le biais de  $\widehat{\mathbf{M}}_{NSCM}$  est donné par

$$B(\widehat{\mathbf{M}}_{NSCM}) = \mathbf{U}(\text{diag}(\mathbf{E}\mathbf{d}))\mathbf{U}^H - \mathbf{M}, \quad (2.12)$$

où

- l'opérateur  $\text{diag}$  transforme un vecteur de dimension  $m$   $\mathbf{v} = (v_j)_{1 \leq j \leq m}$  en une matrice diagonale de dimension  $m \times m$   $\mathbf{V} = (V_{ij})_{1 \leq i, j \leq m}$  ayant comme éléments  $V_{ii} = v_i$ ,
- $\mathbf{U}$  désigne la matrice orthogonale contenant les  $m$  vecteurs propres de  $\mathbf{M}$ ,
- $\mathbf{E} = (E_{ij})_{1 \leq i, j \leq m}$  avec  $E_{ij} = \frac{\log(\lambda_j/\lambda_i)}{\lambda_j/\lambda_i - 1} - \frac{\lambda_i}{\lambda_j}$  si  $i \neq j$  et  $E_{ij} = 0$  si  $i = j$ , où  $\lambda_i$  est le  $i^{\text{ème}}$  valeur propre de  $\mathbf{M}$ ,
- $\mathbf{d} = (d_j)_{1 \leq j \leq m}$  avec  $d_j = m \prod_{k \neq j} \frac{1}{1 - \lambda_k/\lambda_j}$ .

La preuve de ce théorème a fait l'objet d'une publication [5], jointe dans l'annexe C, puisque cet article contient aussi l'étude du moment d'ordre 2 de  $\widehat{\mathbf{M}}_{NSCM}$  quand  $\mathbf{M}$  a des valeurs propres distinctes, étude qui ne sera pas développée dans ce manuscrit pour des raisons expliquées dans la remarque 2.3.2.1.

**Preuve 2.3.1.2**

Voir l'annexe 2A.2 du chapitre.

**Théorème 2.3.1.3 (Biais de  $\widehat{\mathbf{M}}_{TML}$  et biais de  $\widehat{\mathbf{M}}_{NSCM}$  quand  $\mathbf{M} = \mathbf{I}$ )**

$\widehat{\mathbf{M}}_{TML}$  est un estimateur sans biais de  $\mathbf{M}$ , quelle que soit  $\mathbf{M}$ .

$\widehat{\mathbf{M}}_{NSCM}$  est un estimateur sans biais de  $\mathbf{M}$ , quand  $\mathbf{M} = \mathbf{I}$ .

**Preuve 2.3.1.3**

Avec le même raisonnement que dans la preuve du théorème 2.3.1.2 et la relation 2.10 entre le TML et la NSCM, il résulte que  $\mathbb{E} \left[ \mathbf{M}^{-1/2} \widehat{\mathbf{M}}_{TML} \mathbf{M}^{-1/2} \right]$  est une matrice diagonale dont les éléments sont

$$A_{kk} = m \mathbb{E} \left[ \frac{\chi_k^2}{\sum_{j=1}^m \chi_j^2} \right],$$

où  $\chi_j^2 / \sum_{j=1}^m \chi_j^2$  est une variable aléatoire distribué selon une loi Beta de première espèce de paramètres

$1$  et  $m$ , détaillée à l'annexe A.7. De plus, la moyenne statistique d'une  $\beta(1, m)$  étant  $\frac{1}{m}$ , on obtient  $A_{kk} = 1$ , ce qui termine cette preuve.

### 2.3.2 Consistance

La consistance d'un estimateur  $\widehat{\mathbf{M}}$  de  $\mathbf{M}$  est définie par la définition 2.1.1.1.

#### Remarque 2.3.2.1

Quand les valeurs propres de  $\mathbf{M}$  sont toutes différentes, le théorème 2.3.1.2 assure que  $\widehat{\mathbf{M}}_{NSCM}$  est un estimateur biaisé de  $\mathbf{M}$ , analytiquement non "débiaisable". De plus, ce biais ne dépend pas du nombre  $N$  de données secondaires  $\mathbf{x}_i$ . Ainsi, il est évident que  $\widehat{\mathbf{M}}_{NSCM}$  n'est pas consistant. Par la suite, cet estimateur est mis à l'écart, dû à ces précédents inconvénients (biaisé et non consistant).

D'autre part, l'analyse des performances statistiques dans le cas très particulier où  $\mathbf{M}$  est la matrice identité, est étudié au travers de l'estimateur "plus général",  $\widehat{\mathbf{M}}_{TML}$ .

Sous hypothèses Gaussiennes, la SCM est un estimateur consistant de  $\mathbf{M}$ . Ce résultat est établi, par exemple, dans [3] pp. 80-81.

**Théorème 2.3.2.1 (Consistance de  $\widehat{\mathbf{M}}_N$ )**  
 $\widehat{\mathbf{M}}_N$  est un estimateur consistant de  $\mathbf{M}$ .

#### Preuve 2.3.2.1

Voir l'annexe 2A.3 du chapitre.

**Théorème 2.3.2.2 (Consistance de  $\widehat{\mathbf{M}}_{TML}$ )**  
 $\widehat{\mathbf{M}}_{TML}$  est un estimateur consistant de  $\mathbf{M}$ .

#### Preuve 2.3.2.2

Le théorème 2.3.1.3 et la Loi faible des Grands Nombres (LfgN) impliquent que  $\widehat{\mathbf{M}}_{TML}$  est un estimateur consistant de  $\mathbf{M}$ .

### 2.3.3 Moment d'ordre 2 et distribution asymptotique

Dans cette section, une analyse de perturbations est mise en œuvre pour obtenir la distribution asymptotique des estimateurs, tous désignés par  $\widehat{\mathbf{M}}$  dans un souci de clarté. Dans cette perspective,  $\widehat{\mathbf{M}}$  est réécrit de la façon suivante :

$$\widehat{\mathbf{M}} = \mathbf{M} + \delta\mathbf{M}. \quad (2.13)$$

Les quantités suivantes, essentielles dans notre analyse, sont définies par :

- $\Delta = \mathbf{M}^{-1/2} (\delta \mathbf{M}) \mathbf{M}^{-1/2} = \mathbf{M}^{-1/2} \widehat{\mathbf{M}} \mathbf{M}^{-1/2} - \mathbf{I}$ ,
- $\delta = \text{vec}(\Delta)$  où  $\delta$  est le vecteur contenant tous les éléments de  $\Delta$  et  $\text{vec}$  désigne l'opérateur qui transforme les matrices de dimension  $m \times n$  en un vecteur colonne de dimension  $mn$ .

En effet, la distribution asymptotique de  $\widehat{\mathbf{M}}$  s'obtient directement de celle de  $\delta$  grâce à la proposition suivante :

**Proposition 2.3.3.1**

$$\text{vec}(\widehat{\mathbf{M}}) = \left( \mathbf{M}^{1/2} \otimes \mathbf{M}^{\top/2} \right) \delta + \text{vec}(\mathbf{M}), \quad (2.14)$$

où  $\otimes$  représente le produit de Kronecker et  $^{\top}$  désigne l'opérateur transposé.

**Preuve 2.3.3.1**

$\widehat{\mathbf{M}} = \mathbf{M}^{1/2} \Delta \mathbf{M}^{1/2} + \mathbf{M}$ , et en utilisant la propriété,  $\text{vec}(\mathbf{EFG}) = (\mathbf{E} \otimes \mathbf{G}^{\top}) \text{vec}(\mathbf{F})$  quelles que soient les matrices  $\mathbf{E}, \mathbf{F}, \mathbf{G}$ , (voir par exemple [29] p.9), ceci mène à la conclusion de la proposition 2.3.3.1.

Le but de cette section est d'établir la distribution asymptotique de  $\delta$ , i.e. en d'autres termes, la distribution de  $\begin{pmatrix} \text{Re}(\delta) \\ \text{Im}(\delta) \end{pmatrix}$ , où  $\text{Re}(\delta)$  désigne la partie réelle du vecteur complexe  $\delta$  et  $\text{Im}(\delta)$  sa partie imaginaire.

Comme nous le verrons plus tard, cette distribution est Gaussienne et, donc, est complètement caractérisée par sa matrice de covariance asymptotique  $\mathbf{C}$ . Cette matrice peut évidemment être calculée à partir des deux quantités  $\mathbb{E}[\delta \delta^{\top}]$  et  $\mathbb{E}[\delta \delta^H]$ . Cependant, dans ce cas de figure  $\delta$  est le vec d'une matrice Hermitienne et, par conséquent,  $\mathbb{E}[\delta \delta^{\top}]$  peut facilement être obtenu à partir de  $\mathbb{E}[\delta \delta^H]$ .

Les résultats suivants font appel aux notations  $\mathbf{C}_1$  et  $\mathbf{C}_2$ , définies par :

$$\mathbf{C}_1 = \frac{m}{m+1} \left( \mathbf{P} - \frac{1}{m} \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^{\top} \right), \quad (2.15)$$

$$\mathbf{C}_2 = \frac{m}{m+1} \left( \mathbf{I} - \frac{1}{m} \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^{\top} \right), \quad (2.16)$$

où  $\mathbf{P}$  est défini par, pour  $1 \leq p, p' \leq m$ ,

- $P_{kl} = 1$ , pour  $k = l = p + m(p-1)$ ,
- $P_{kl} = 1$ , pour  $k = p + m(p'-1)$  et  $l = p' + m(p-1)$ ,
- $P_{kl} = 0$ , sinon.

La NSCM a été rejetée à cause de ses inconvénients constatés par l'analyse du moment d'ordre 1. Cependant, cet estimateur étant largement répandu dans la littérature et utilisé dans d'autres contextes que la détection radar, l'analyse de ses performances est essentiel : ceci a donné lieu à une publication disponible dans l'annexe C.

Concernant le benchmark, le moment d'ordre 2 ainsi que la distribution asymptotique de la SCM sont analysés dans [3]. Le théorème 2.3.3.1 rappelle ce résultat à l'aide d'une écriture vectorielle :

**Théorème 2.3.3.1 (Distribution asymptotique de  $\text{vec}(\widehat{\mathbf{M}}_{SCM})$ )**

1.  $\sqrt{N} \begin{pmatrix} \text{Re}(\boldsymbol{\delta}_{SCM}) \\ \text{Im}(\boldsymbol{\delta}_{SCM}) \end{pmatrix} \xrightarrow[N \rightarrow +\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_{SCM})$ , où  $\xrightarrow{dist.}$  représente la convergence en loi ;
2.  $N \mathbb{E} [\boldsymbol{\delta}_{SCM} \boldsymbol{\delta}_{SCM}^\top] \xrightarrow[N \rightarrow +\infty]{} \mathbf{P}$  où  $\mathbf{P}$  est défini ci-dessus ;
3.  $N \mathbb{E} [\boldsymbol{\delta}_{SCM} \boldsymbol{\delta}_{SCM}^H] \xrightarrow[N \rightarrow +\infty]{} \mathbf{I}$ .

Comme les précédentes propriétés statistiques (consistance et sans biais) de  $\widehat{\mathbf{M}}_{SCM}$  sont aussi valables pour  $\widehat{\mathbf{M}}_N$ , il est possible d'étudier sa distribution asymptotique.

**Théorème 2.3.3.2 (Distribution asymptotique de  $\text{vec}(\widehat{\mathbf{M}}_N)$ )**

1.  $\sqrt{N} \begin{pmatrix} \text{Re}(\boldsymbol{\delta}_N) \\ \text{Im}(\boldsymbol{\delta}_N) \end{pmatrix} \xrightarrow[N \rightarrow +\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_N)$  ;
2.  $N \mathbb{E} [\boldsymbol{\delta}_N \boldsymbol{\delta}_N^\top] \xrightarrow[N \rightarrow +\infty]{} \left( \frac{m+1}{m} \right) \mathbf{C}_1$  ;
3.  $N \mathbb{E} [\boldsymbol{\delta}_N \boldsymbol{\delta}_N^H] \xrightarrow[N \rightarrow +\infty]{} \left( \frac{m+1}{m} \right) \mathbf{C}_2$ .

Les deux matrices  $\mathbf{C}_1$  et  $\mathbf{C}_2$  sont définies par les équations (2.15) et (2.16).

**Preuve 2.3.3.2**

Voir l'annexe 2A.4 du chapitre.

**Théorème 2.3.3.3 (Distribution asymptotique de  $\text{vec}(\widehat{\mathbf{M}}_{TML})$ )**

1.  $\sqrt{N} \begin{pmatrix} \text{Re}(\boldsymbol{\delta}_{TML}) \\ \text{Im}(\boldsymbol{\delta}_{TML}) \end{pmatrix} \xrightarrow[N \rightarrow +\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_{TML})$  ;
2.  $N \mathbb{E} [\boldsymbol{\delta}_{TML} \boldsymbol{\delta}_{TML}^\top] \xrightarrow[N \rightarrow +\infty]{} \mathbf{C}_1$  ;
3.  $N \mathbb{E} [\boldsymbol{\delta}_{TML} \boldsymbol{\delta}_{TML}^H] \xrightarrow[N \rightarrow +\infty]{} \mathbf{C}_2$ .

**Preuve 2.3.3.3**

Voir l'annexe 2A.5 du chapitre.

### 2.3.4 Synthèse

Ce chapitre, dédié à l'étude d'estimateurs de la matrice de covariance  $\mathbf{M}$ , propose un bilan mitigé : en effet,  $\widehat{\mathbf{M}}_{SCM}$ ,  $\widehat{\mathbf{M}}_N$  et  $\widehat{\mathbf{M}}_{TML}$  peuvent être retenus comme estimateurs de  $\mathbf{M}$  pour leurs bonnes propriétés, mais ne sont pas applicables à des cas pratiques puisqu'ils reposent sur les vecteurs Gaussiens des données secondaires qui ne sont pas accessibles ou bien parce qu'ils utilisent la vraie matrice de covariance  $\mathbf{M}$  dans leur définition.

La NSCM, estimateur utilisé jusqu'à présent dans la littérature, convient au problème par certains côtés : une mise en œuvre assez simple et la conservation de la propriété CFAR-texture du détecteur associé. Cependant, ses mauvaises propriétés statistiques (estimateur non consistant et biaisé) ne permettent pas d'en faire le candidat idéal à l'estimation de  $\mathbf{M}$ .

Pour ces raisons, un nouvel estimateur de covariance est introduit et analysé dans le chapitre suivant : l'estimateur du Point Fixe ou FP.

## Chapitre 2A

# Preuves du chapitre 2

### 2A.1 Preuve du théorème 2.3.1.1 : "Biais de $\widehat{\mathbf{M}}_N$ "

Pour démontrer le théorème 2.3.1.1, l'expression suivante, issue de la définition de  $\widehat{\mathbf{M}}_N$  (2.5), doit être calculée :

$$\mathbb{E} \left[ \widehat{\mathbf{M}}_N \right] = m \sum_{i=1}^N \mathbb{E} \left[ \frac{\mathbf{x}_i \mathbf{x}_i^H}{\sum_{k=1}^N \mathbf{x}_k^H \mathbf{M}^{-1} \mathbf{x}_k} \right]. \quad (2A.1)$$

Il faut montrer que  $\mathbb{E} \left[ \widehat{\mathbf{M}}_N \right] = \mathbf{M}$ . Tout d'abord, blanchissons les  $\mathbf{x}_k$  dans l'équation (2A.1), selon  $\mathbf{y}_k = \mathbf{M}^{-1/2} \mathbf{x}_k$

$$\mathbb{E} \left[ \widehat{\mathbf{M}}_N \right] = m \mathbf{M}^{1/2} \left( \sum_{i=1}^N \mathbb{E} \left[ \frac{\mathbf{y}_i \mathbf{y}_i^H}{\sum_{k=1}^N \mathbf{y}_k^H \mathbf{y}_k} \right] \right) \mathbf{M}^{1/2}. \quad (2A.2)$$

Pour  $i = 1, \dots, N$ , posons

$$\mathbf{A}_i = \frac{\mathbf{y}_i \mathbf{y}_i^H}{\sum_{k=1}^N \mathbf{y}_k^H \mathbf{y}_k}. \quad (2A.3)$$

Puisque  $\mathbf{y}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$  quel que soit  $k$ , des considérations basiques de probabilité permettent de réécrire les  $y_k^{(j)}$ , où  $y_k^{(j)}$  est la notation utilisée pour le  $j^{\text{ème}}$  élément de  $\mathbf{y}_k$ , comme

$$y_k^{(j)} = \sqrt{(1/2) \chi_k^{2(j)}} \exp(i\theta_k^{(j)}),$$

où pour  $1 \leq k \leq N$  and  $1 \leq j \leq m$ ,  $\chi_k^{2(j)} \sim \chi^2(2)$  et  $\theta_k^{(j)} \sim \mathcal{U}([0, 2\pi])$ , où  $\chi^2(2)$  représente la loi du Chi 2 avec 2 degrés de liberté et  $\mathcal{U}([0, 2\pi])$ , la loi uniforme sur l'intervalle  $[0, 2\pi]$ , et  $\chi_k^{2(j)}$  et  $\theta_k^{(j)}$  sont

indépendants.

Donc, en remplaçant les  $\mathbf{y}_k$  dans l'équation (2A.3), l'élément  $pq$  de la matrice  $\mathbf{A}_i$  est

$$A_i^{(pq)} = \frac{\sqrt{\chi_i^{2(p)} \chi_i^{2(q)}}}{\sum_{k=1}^N \sum_{r=1}^m \chi_k^{2(r)}} \exp(i(\theta_k^{(p)} - \theta_k^{(q)})).$$

Maintenant, comme  $\mathbb{E} \left[ \exp(i\theta_k^{(p)}) \right] = 0$ , quels que soient  $k$  et  $p$ , il vient directement que les éléments hors de la diagonale de  $\mathbb{E} [\mathbf{A}_i]$  sont nuls. Ensuite, l'élément diagonal  $pp$  est

$$\mathbb{E} \left[ A_i^{(pp)} \right] = \mathbb{E} \left[ \frac{\chi_i^{2(p)}}{\sum_{k=1}^N \sum_{r=1}^m \chi_k^{2(r)}} \right],$$

où  $\chi_i^{2(p)} / \sum_{k=1}^N \sum_{r=1}^m \chi_k^{2(r)}$  est une variable aléatoire de loi Beta de première espèce, ayant comme para-

mètres 1 et  $Nm$ . Comme l'espérance d'une  $\beta(1, Nm)$  est  $\frac{1}{Nm}$ , alors,  $\mathbb{E} \left[ A_i^{(pp)} \right] = \frac{1}{Nm}$  pour  $i = 1, \dots, N$  et pour  $p = 1, \dots, m$  et ainsi,  $\mathbb{E} [\mathbf{A}_i] = (1/Nm) \mathbf{I}$ . En remplaçant maintenant dans l'équation (2A.2), on obtient

$$\mathbb{E} \left[ \widehat{\mathbf{M}}_N \right] = m \mathbf{M}^{1/2} \left( \sum_{i=1}^N (1/Nm) \mathbf{I} \right) \mathbf{M}^{1/2},$$

ou de manière équivalente,

$$\mathbb{E} \left[ \widehat{\mathbf{M}}_N \right] = \mathbf{M},$$

ce qui conclut la preuve.

## 2A.2 Preuve du théorème 2.3.1.2 : "Biais de $\widehat{\mathbf{M}}_{NSCM}$ "

Dans cette preuve, afin de ne pas alourdir les notations,  $\widehat{\mathbf{M}}$  désigne  $\widehat{\mathbf{M}}_{NSCM}$ .

$\mathbf{M}$  étant une matrice Hermitienne définie positive, elle peut être réécrite comme  $\mathbf{M} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$  où  $\mathbf{\Lambda}$  est la matrice diagonale des valeurs propres de  $\mathbf{M}$ ,  $\lambda_1, \dots, \lambda_m$  et  $\mathbf{U}$  est la matrice unitaire constituée par les vecteurs propres de  $\mathbf{M}$ .

Puisque les  $\mathbf{x}_k$  sont identiquement distribués,  $\mathbb{E}[\widehat{\mathbf{M}}]$  peut s'écrire :

$$\mathbb{E}[\widehat{\mathbf{M}}] = \frac{m}{N} \sum_{k=1}^N \mathbb{E} \left[ \begin{array}{c} \mathbf{x}_k \mathbf{x}_k^H \\ \mathbf{x}_k^H \mathbf{x}_k \end{array} \right] = m \mathbb{E} \left[ \begin{array}{c} \mathbf{x} \mathbf{x}^H \\ \mathbf{x}^H \mathbf{x} \end{array} \right], \quad (2A.4)$$

où  $\mathbf{x}$  suit la même loi que les  $\mathbf{x}_k$ .

Maintenant, en blanchissant  $\mathbf{x}$  selon

$$\mathbf{y} = \left( \mathbf{\Lambda}^{-1/2} \mathbf{U} \right)^H \mathbf{x}, \quad (2A.5)$$

où  $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ , i.e.  $\mathbf{y}$  est un vecteur Gaussien complexe centré, de matrice de covariance  $\mathbf{I}$ , puis en remplaçant  $\mathbf{x}$  par (2A.5), (2A.4) devient

$$\mathbb{E}[\widehat{\mathbf{M}}] = m \mathbb{E} \left[ \mathbf{U} \frac{\mathbf{\Lambda}^{-1/2} \mathbf{y} \mathbf{y}^H \mathbf{\Lambda}^{-1/2}}{\mathbf{y}^H \mathbf{\Lambda}^{-1} \mathbf{y}} \mathbf{U}^H \right] = m \mathbf{U} \mathbb{E} \left[ \frac{\mathbf{\Lambda}^{-1/2} \mathbf{y} \mathbf{y}^H \mathbf{\Lambda}^{-1/2}}{\mathbf{y}^H \mathbf{\Lambda}^{-1} \mathbf{y}} \right] \mathbf{U}^H = m \mathbf{U} \mathbb{E}[\mathbf{A}] \mathbf{U}^H. \quad (2A.6)$$

Afin d'établir le biais  $B(\widehat{\mathbf{M}})$ , il faut se concentrer sur la distribution de la matrice aléatoire suivante

$$\mathbf{A} = \frac{\mathbf{\Lambda}^{-1/2} \mathbf{y} \mathbf{y}^H \mathbf{\Lambda}^{-1/2}}{\sum_{j=1}^m \lambda_j |y_j|^2}, \quad (2A.7)$$

où les  $y_j$  sont les  $m$  composantes du vecteur  $\mathbf{y}$ .

Grâce à (2A.7),  $B(\widehat{\mathbf{M}})$  peut être réécrit comme

$$B(\widehat{\mathbf{M}}) = \mathbf{U} (m \mathbb{E}[\mathbf{A}] - \mathbf{\Lambda}) \mathbf{U}^H. \quad (2A.8)$$

Comme  $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ , les mêmes considérations probabilistes basiques que dans la preuve précédente permettent de réécrire les  $y_j$  comme

$$y_j = \sqrt{(1/2) \chi_j^2} \exp(i\theta_j), \quad (2A.9)$$

où pour  $1 \leq j \leq m$ ,  $\chi_j^2 \sim \chi^2(2)$  et  $\theta_j \sim \mathcal{U}([0, 2\pi])$ , où  $\chi^2(2)$  désigne la distribution du Chi-deux avec 2 degrés de liberté et  $\mathcal{U}([0, 2\pi])$ , la distribution uniforme sur l'intervalle  $[0, 2\pi]$  and  $i = \sqrt{-1}$ . Dans (2A.9), toutes les variables aléatoires sont indépendantes.

Tout d'abord, concernant les éléments, hors de la diagonale, de  $\mathbf{A} = (A_{kl})_{1 \leq k, l \leq m}$ , on a

$$\mathbb{E}[A_{kl}] = \sqrt{\lambda_k \lambda_l} \mathbb{E} \left[ \frac{\sqrt{\chi_k^2 \chi_l^2}}{\sum_{j=1}^m \lambda_j \chi_j^2} \right] \mathbb{E}[\exp(i\theta_k)] \mathbb{E}[\exp(-i\theta_l)] \quad \text{pour } k \neq l. \quad (2A.10)$$

Comme  $\mathbb{E}[\exp(i\theta_k)] = 0$  pour  $1 \leq k \leq m$ ,  $\mathbb{E}[\mathbf{A}]$  est une matrice diagonale.

Les éléments diagonaux de  $\mathbb{E}[\mathbf{A}]$  sont donnés par

$$\mathbb{E}[A_{kk}] = \mathbb{E} \left[ \frac{\lambda_k \chi_k^2}{\lambda_1 \chi_1^2 + \dots + \lambda_m \chi_m^2} \right]. \quad (2A.11)$$

(2A.11) peut être réécrit

$$\mathbb{E}[A_{kk}] = \mathbb{E} \left[ \frac{1}{1 + G} \right], \quad (2A.12)$$

où  $G = \frac{1}{\lambda_k \chi_k^2} \sum_{j \neq k} \lambda_j \chi_j^2$ .

$G$  est le rapport de 2 variables aléatoires indépendantes  $G_1 = \sum_{j \neq k} \lambda_j \chi_j^2$  et  $G_2 = \lambda_k \chi_k^2$ . Pour obtenir la distribution de  $G$ , la distribution de  $G_1$  doit être calculée. Sa fonction caractéristique est donnée par

$$\phi_{G_1}(u) = \prod_{j \neq k} \frac{1}{1 - 2i \lambda_j u}. \quad (2A.13)$$

Puisque toutes les valeurs propres de  $\mathbf{M}$  sont supposées différentes, le développement en fractions de  $\phi_{G_1}$  donne

$$\phi_{G_1}(u) = \sum_{j \neq k} \frac{\beta_j}{1 - 2i \lambda_j u}, \quad (2A.14)$$

où

$$\beta_j = \prod_{\substack{l \neq j \\ l \neq k}} \frac{1}{1 - \lambda_l / \lambda_j}. \quad (2A.15)$$

La PDF de  $G_1$  découle en prenant la transformée inverse de Fourier de (2A.14)

$$f_{G_1}(x) = \frac{1}{2} \sum_{j \neq k} \left( \prod_{\substack{l \neq j \\ l \neq k}} \frac{1}{1 - \lambda_l / \lambda_j} \right) \frac{1}{\lambda_j} \exp \left( -\frac{x}{2 \lambda_j} \right). \quad (2A.16)$$

Ceci permet d'obtenir la distribution  $f_G(x)$  de  $G$

$$f_G(x) = \sum_{j \neq k} \left( \prod_{\substack{l \neq j \\ l \neq k}} \frac{1}{1 - \lambda_l / \lambda_j} \right) \frac{1}{\lambda_j} \frac{1}{\lambda_k} \left( \frac{1}{(x/\lambda_j + 1/\lambda_k)^2} \right). \quad (2A.17)$$

Maintenant,  $\mathbb{E}[A_{kk}]$  peut être calculé à partir de (2A.17) par

$$\mathbb{E}[A_{kk}] = \int_0^\infty \frac{1}{1+u} f_G(u) du, \quad (2A.18)$$

ce qui équivaut à

$$\mathbb{E}[A_{kk}] = \sum_{j \neq k} \left( \prod_{\substack{l \neq j \\ l \neq k}} \frac{1}{1 - \lambda_l / \lambda_j} \right) \lambda_k \left( \frac{\log(\lambda_j) - \log(\lambda_k)}{\lambda_j - \lambda_k} - \frac{1}{\lambda_j} \right). \quad (2A.19)$$

Avec les notations du théorème (2.3.1.2), cette expression conduit à  $\mathbb{E}[\mathbf{A}] = \text{diag}(\mathbf{E} \mathbf{c})$ , ce qui termine la preuve.

### 2A.3 Preuve du théorème 2.3.2.1 : "Consistance de $\widehat{\mathbf{M}}_N$ "

En blanchissant les  $\mathbf{x}_i$  dans l'équation 2.5, on a :

$$\widehat{\mathbf{M}}_N = m \mathbf{M}^{1/2} \left( \frac{\sum_{i=1}^N \mathbf{y}_i \mathbf{y}_i^H}{N} \right) \mathbf{M}^{1/2},$$

avec  $\mathbf{y}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ .

Ensuite, la LfGN assure

$$\frac{1}{N} \sum_{i=1}^N \mathbf{y}_i \mathbf{y}_i^H \xrightarrow[N \rightarrow +\infty]{Pr} \mathbf{I},$$

$$\frac{1}{N} \sum_{k=1}^N \mathbf{y}_k^H \mathbf{y}_k \xrightarrow[N \rightarrow +\infty]{Pr} \mathbb{E} [\mathbf{y}^H \mathbf{y}] = m.$$

où la notation  $\xrightarrow[N \rightarrow +\infty]{Pr}$  désigne la convergence en probabilité, cette notation a exactement la même signification que la notation utilisée dans la définition 2.1.1.1 de la consistance. Les deux notations sont donc indifféremment utilisées.

Finalement, les résultats précédents ainsi que des propriétés classiques sur la convergence des rapports de variables aléatoires entraînent que :

$$\widehat{\mathbf{M}}_N \xrightarrow[N \rightarrow +\infty]{Pr} \mathbf{M},$$

ce qui signifie que  $\widehat{\mathbf{M}}_N$  est un estimateur consistant de  $\mathbf{M}$ .

## 2A.4 Preuve du théorème 2.3.3.2 : "Distribution asymptotique de $\widehat{\mathbf{M}}_N$ "

Dans cette démonstration, les résultats obtenus pour  $\widehat{\mathbf{M}}_{SCM}$  sont exploités afin d'en déduire ceux concernant  $\widehat{\mathbf{M}}_N$ . En effet, on a :

$$\Delta_N = \mathbf{M}^{-1/2} \widehat{\mathbf{M}}_N \mathbf{M}^{-1/2} - \mathbf{I}.$$

En utilisant la définition de  $\widehat{\mathbf{M}}_N$ , équation (2.4), on obtient :

$$\Delta_N = \frac{m}{\text{Tr}(\Delta_{SCM} + \mathbf{I})} (\Delta_{SCM} + \mathbf{I}) - \mathbf{I} = \frac{1}{1 + \frac{\text{Tr}(\Delta_{SCM})}{m}} (\Delta_{SCM} + \mathbf{I}) - \mathbf{I}.$$

De plus,  $N$  est supposé suffisamment grand pour assurer la validité des expressions d'ordre 1, en terme de  $\Delta_{SCM}$ , et ainsi

$$\Delta_N \simeq \left( 1 - \frac{\text{Tr}(\Delta_{SCM})}{m} \right) (\Delta_{SCM} + \mathbf{I}) - \mathbf{I}.$$

Puis en négligeant le terme d'ordre 2, i.e.  $\Delta_{SCM} \frac{\text{Tr}(\Delta_{SCM})}{m}$ , l'équation précédente devient

$$\Delta_N \simeq \Delta_{SCM} - \mathbf{I} \frac{\text{Tr}(\Delta_{SCM})}{m}.$$

Maintenant, avec les notations présentées au début de la section 2.3.3, on a :

$$\delta_N \simeq \delta_{SCM} - \frac{\text{Tr}(\Delta_{SCM})}{m} \text{vec}(\mathbf{I}).$$

Puis, à partir de l'expression de  $\mathbf{C}_2$  dans l'équation (2.16), il vient que

$$\delta_N \simeq \frac{m+1}{m} \mathbf{C}_2 \delta_{SCM}. \quad (2A.20)$$

Ensuite, l'équation (2A.20) assure que

$$\sqrt{N} \begin{pmatrix} \text{Re}(\delta_N) \\ \text{Im}(\delta_N) \end{pmatrix} \xrightarrow[N \rightarrow +\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_N),$$

et il reste juste à calculer les deux quantités  $\mathbb{E}[\delta_N \delta_N^\top]$  et  $\mathbb{E}[\delta_N \delta_N^H]$ . Mais, puisque l'équation (2A.20) établit une relation entre  $\delta_N$  et  $\delta_{SCM}$ , le résultat découle rapidement :

$$\begin{aligned} N \mathbb{E}[\delta_N \delta_N^\top] &\xrightarrow[N \rightarrow +\infty]{} \left( \frac{m+1}{m} \right) \mathbf{C}_1, \\ N \mathbb{E}[\delta_N \delta_N^H] &\xrightarrow[N \rightarrow +\infty]{} \left( \frac{m+1}{m} \right) \mathbf{C}_2, \end{aligned}$$

où les deux matrices  $\mathbf{C}_1$  et  $\mathbf{C}_2$  sont définies par (2.15) et (2.16).

## 2A.5 Preuve du théorème 2.3.3.3 : "Distribution asymptotique de $\widehat{\mathbf{M}}_{TML}$ "

Après le blanchiment des  $\mathbf{x}_i$ ,  $\widehat{\mathbf{M}}_{TML}$  défini par (2.9) devient

$$\widehat{\mathbf{M}}_{TML} = \mathbf{M}^{1/2} \left( \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} \right) \mathbf{M}^{1/2}. \quad (2A.21)$$

où  $\mathbf{y}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ .

Ainsi, l'expression de  $\delta$  dans le cas du TML devient :

$$\delta_{TML} = \text{vec} \left( \mathbf{M}^{-1/2} \widehat{\mathbf{M}}_{TML} \mathbf{M}^{-1/2} - \mathbf{I} \right) = \text{vec} \left( \frac{m}{N} \sum_{i=1}^N \left( \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} \right) - \mathbf{I} \right).$$

Le Théorème Central Limite (TCL) assure le premier point du théorème 2.3.3.3 :

$$\sqrt{N} \begin{bmatrix} \text{Re}(\delta_{TML}) \\ \text{Im}(\delta_{TML}) \end{bmatrix} \xrightarrow[N \rightarrow +\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_{TML}),$$

où  $\mathbf{C}_{TML}$  est la matrice de covariance de  $\begin{bmatrix} \text{Re}(\delta_{TML}) \\ \text{Im}(\delta_{TML}) \end{bmatrix}$ .

Il ne reste maintenant plus qu'à calculer les deux quantités  $\mathbb{E} [\delta_{TML} \delta_{TML}^\top]$  et  $\mathbb{E} [\delta_{TML} \delta_{TML}^H]$ .

Tout d'abord, calculons  $E [\delta_{TML} \delta_{TML}^\top]$ . Pour tout  $N$ , on a :

$$\mathbb{E} [\delta_{TML} \delta_{TML}^\top] = \frac{m^2}{N} \mathbb{E} \left[ \text{vec} \left( \frac{\mathbf{y} \mathbf{y}^H}{\mathbf{y}^H \mathbf{y}} \right) \text{vec} \left( \frac{\mathbf{y} \mathbf{y}^H}{\mathbf{y}^H \mathbf{y}} \right)^\top \right] - \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^\top, \quad (2A.22)$$

avec  $\mathbf{y} = (y_1, \dots, y_m)^\top \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ .

Ensuite, il faut se concentrer sur la variable suivante :

$$\mathbf{B} = \mathbb{E} \left[ \text{vec} \left( \frac{\mathbf{y} \mathbf{y}^H}{\mathbf{y}^H \mathbf{y}} \right) \text{vec} \left( \frac{\mathbf{y} \mathbf{y}^H}{\mathbf{y}^H \mathbf{y}} \right)^\top \right].$$

En réécrivant à nouveau les  $y_j$  selon  $y_j = \sqrt{\chi_j^2/2} \exp(i\theta_j)$  où pour  $j = 1, \dots, m$ ,  $\chi_j^2$  et  $\theta_j$  sont des variables indépendantes, avec  $\chi_j^2 \sim \chi^2(2)$  et  $\theta_j \sim \mathcal{U}([0, 2\pi])$ , chaque élément de la matrice  $\mathbf{B}$  devient

$$B_{kl} = \mathbb{E} \left[ \frac{\sqrt{\chi_p^2 \chi_q^2 \chi_{p'}^2 \chi_{q'}^2}}{\left( \sum_{j=1}^m \chi_j^2 \right)^2} \right] \mathbb{E} [\exp(i(\theta_p - \theta_q + \theta_{q'} - \theta_{p'}))].$$

On remarque rapidement que  $B_{kl} = 0$ , sauf pour

1.  $k = l = p + m(p - 1)$ ,
2.  $k = p + m(p - 1), l = p' + m(p' - 1)$  and  $p \neq p'$ ,
3.  $k = p + m(p' - 1), l = p' + m(p - 1)$  and  $p \neq p'$ ,

et, pour ces indices, on a

1.  $B_{p+m(p-1), p+m(p-1)} = \frac{2}{m(m+1)}$ ,
2.  $B_{p+m(p-1), p'+m(p'-1)} = \frac{1}{m(m+1)}$ ,
3.  $B_{p+m(p'-1), p'+m(p-1)} = \frac{1}{m(m+1)}$ .

En réinjectant ces résultats dans l'équation (2A.22), on obtient

$$N \mathbb{E} \left[ \boldsymbol{\delta}_{TML} \boldsymbol{\delta}_{TML}^\top \right] \xrightarrow{N \rightarrow +\infty} \mathbf{C}_1 = \frac{m}{m+1} \left( \mathbf{P} - \frac{1}{m} \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^\top \right),$$

où  $\mathbf{P}$  est défini, pour  $1 \leq p, p' \leq m$ , par

- $P_{kl} = 1$ , pour  $k = l = p + m(p - 1)$ ,
- $P_{kl} = 1$ , pour  $k = p + m(p' - 1)$  et  $l = p' + m(p - 1)$ ,
- $P_{kl} = 0$ , sinon.

En répétant le même raisonnement, on obtient

$$N \mathbb{E} \left[ \boldsymbol{\delta}_{TML} \boldsymbol{\delta}_{TML}^H \right] \xrightarrow{N \rightarrow +\infty} \mathbf{C}_2 = \frac{m}{m+1} \left( \mathbf{I} - \frac{1}{m} \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^\top \right).$$

Ceci achève la preuve du théorème 2.3.3.3.



## Chapitre 3

# Un Estimateur Innovant : la Matrice du Point Fixe

*Ce chapitre est consacré à un des principaux apports de ce travail de recherche. Tout d'abord, la matrice du Point Fixe est introduite dans le contexte des SIRV. La forme particulière de cet estimateur est étudiée : comme il est défini uniquement implicitement comme solution d'une équation particulière, son existence et son unicité sont tout d'abord démontrées, et un algorithme de construction est développé. Ensuite, une analyse complète de ses performances statistiques (biais, consistance et distribution asymptotique) est réalisée, puis, ses avantages sont mis en évidence pour son utilisation dans la version adaptative du GLRT-LQ, notamment au travers d'une nouvelle propriété du détecteur : la propriété CFAR-matrice. Ce chapitre se conclut par une conjecture sur la loi de cet estimateur. Les démonstrations concernant les propriétés statistiques du Point Fixe sont reportées dans l'annexe consacrée à ce chapitre.*

### 3.1 Introduction et justification

Depuis le choix du modèle SIRV, il n'a jamais été question d'estimateur de  $\mathbf{M}$  par la théorie du Maximum de Vraisemblance, bien que ce dernier s'avère toujours être un estimateur présentant beaucoup d'avantages.

Cette absence s'explique par le fait que lors de la construction de la vraisemblance à partir de la PDF du SIRV  $\mathbf{c}$  définie par

$$p_{\mathbf{c}}(\mathbf{c}) = \int_0^{+\infty} g_m(\mathbf{c}, \tau) p_{\tau}(\tau) d\tau ,$$

la PDF de la texture  $p_{\tau}(\tau)$  intervenant sous l'intégrale est inconnue.

Cette approche d'estimation par MV va être présentée dans deux cas généraux, tout d'abord, en supposant la texture déterministe mais inconnue, puis en la supposant aléatoire et inconnue, ce qui est le cas dans la modélisation SIRV.

### 3.1.1 Cas d'un $\tau$ déterministe mais inconnu

Cette approche a été développée dans [14] où les  $\tau_i$ 's sont supposés être des paramètres inconnus. La fonction de vraisemblance correspondante à maximiser en  $\mathbf{M}$  et en  $\tau_i$  est donnée par :

$$p_C(\mathbf{c}_1, \dots, \mathbf{c}_N; \mathbf{M}, \tau_1, \dots, \tau_N) = \frac{1}{(\pi)^{mN} |\mathbf{M}|^N} \prod_{i=1}^N \frac{1}{\tau_i^m} \exp\left(-\frac{\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i}{\tau_i}\right). \quad (3.1)$$

Pour un  $\mathbf{M}$  fixé (mais inconnu), la maximisation en  $\tau_i$  conduit à

$$\hat{\tau}_i = \frac{\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i}{m}. \quad (3.2)$$

Ensuite, en remplaçant les  $\tau_i$ 's dans (3.1) leur estimateur du MV  $\hat{\tau}_i$ 's défini par (3.2), la fonction de vraisemblance réduite est donnée par :

$$\hat{p}_C(\mathbf{c}_1, \dots, \mathbf{c}_N; \mathbf{M}) = \frac{1}{(\pi)^{mN} |\mathbf{M}|^N} \prod_{i=1}^N \frac{m^m \exp(-m)}{(\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i)^m}.$$

Enfin, maximiser  $\hat{p}_C(\mathbf{c}_1, \dots, \mathbf{c}_N; \mathbf{M})$  en  $\mathbf{M}$  revient à maximiser la fonction suivante  $F$  :

$$F(\mathbf{M}) = \frac{1}{|\mathbf{M}|^N} \prod_{i=1}^N \frac{1}{(\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i)^m}.$$

L'estimateur du MV de  $\mathbf{M}$  est le  $\hat{\mathbf{M}}$  qui annule le gradient de  $F$ , ce qui revient à l'équation en  $\hat{\mathbf{M}}$  suivante :

$$\hat{\mathbf{M}} = f(\hat{\mathbf{M}}), \quad (3.3)$$

où  $f$  est définie comme

$$f(\hat{\mathbf{M}}) = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{c}_i \mathbf{c}_i^H}{\mathbf{c}_i^H \hat{\mathbf{M}}^{-1} \mathbf{c}_i}. \quad (3.4)$$

Notons que si cette équation possède une solution unique, cette dernière sera l'estimateur du MV associé au problème : "bruit additif Gaussien modulé à chaque cas par une puissance  $\tau_i$  inconnue".

### 3.1.2 Cas du SIRV : $\tau$ aléatoire

Dans ce cas, la fonction de vraisemblance à maximiser pour obtenir le MV est :

$$p_C(\mathbf{c}_1, \dots, \mathbf{c}_N; \mathbf{M}) = \prod_{i=1}^N \int_0^{+\infty} \frac{1}{(\pi \tau_i)^m |\mathbf{M}|} \exp\left(-\frac{\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i}{\tau_i}\right) p_\tau(\tau_i) d\tau_i. \quad (3.5)$$

Dans [27], Gini et al. ont calculé, en annulant le gradient de (3.5), le MV obtenu comme la solution de l'équation suivante :

$$\hat{\mathbf{M}}_{MV} = \frac{1}{N} \sum_{i=1}^N \frac{h_{m+1}(\mathbf{c}_i^H \hat{\mathbf{M}}_{MV}^{-1} \mathbf{c}_i)}{h_m(\mathbf{c}_i^H \hat{\mathbf{M}}_{MV}^{-1} \mathbf{c}_i)} \mathbf{c}_i \mathbf{c}_i^H, \quad (3.6)$$

où

$$h_m(q) = \int_0^{+\infty} \tau^{-m} \exp(-q/\tau) p_\tau(\tau) d\tau. \quad (3.7)$$

Après plusieurs approximations détaillées dans [27], Gini a obtenu un MV Approché  $\widehat{\mathbf{M}}$  (MVA ou, en anglais, AML pour Approximate Maximum Likelihood) qui se ramène à nouveau à l'équation (3.3) :

$$\widehat{\mathbf{M}} = f(\widehat{\mathbf{M}}), \quad (3.8)$$

avec  $f$  une fois encore définie par (3.4).

Remarquons que  $f$  peut être réécrite uniquement en termes de  $\mathbf{x}_i$ , en simplifiant les  $\tau_i$  dans l'équation (3.4) :

$$f(\widehat{\mathbf{M}}) = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{M}}^{-1} \mathbf{x}_i}. \quad (3.9)$$

Une autre approche, peut-être plus en accord avec ce qui a été établi dans la section 1.2 pour la détection en environnement non Gaussien, est d'utiliser dans (3.6) le prior de Jeffrey (1.15), quelque peu modifié afin de respecter les conditions d'intégrabilité,

$$p_\tau(\tau) = \frac{1}{\tau} \left( \frac{1}{\ln(A) - \ln(\varepsilon)} \right) \mathbb{1}_{[\varepsilon, A]}(\tau),$$

où  $\mathbb{1}_{[a, b]}(x)$  est la fonction indicatrice sur l'intervalle  $[a, b]$  définie par :

$$\mathbb{1}_{[a, b]}(x) = \begin{cases} 1 & \text{si } x \in [a, b], \\ 0 & \text{sinon.} \end{cases}$$

Ensuite, grâce au changement de variable  $z = \frac{q}{\tau}$ , on obtient

$$h_m(q) = \left( \frac{1}{\ln(A) - \ln(\varepsilon)} \right) q^{-m} \int_{\frac{q}{A}}^{\frac{q}{\varepsilon}} z^{m-1} \exp(-z) dz.$$

Puis,

$$\frac{h_{m+1}(q)}{h_m(q)} = q^{-1} \frac{\int_{\frac{q}{A}}^{\frac{q}{\varepsilon}} z^m \exp(-z) dz}{\int_{\frac{q}{A}}^{\frac{q}{\varepsilon}} z^{m-1} \exp(-z) dz}. \quad (3.10)$$

De plus, d'après (1.17),

$$\int_0^{+\infty} z^{m-1} \exp(-z) dz = (m-1)!.$$

Maintenant, en faisant tendre  $\varepsilon$  vers 0 et  $A$  vers l'infini, (3.10) devient

$$\frac{h_{m+1}(q)}{h_m(q)} = \frac{m}{q},$$

et l'on retrouve (3.8) et (3.3)

$$\widehat{\mathbf{M}} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{c}_i \mathbf{c}_i^H}{\mathbf{c}_i^H \widehat{\mathbf{M}}^{-1} \mathbf{c}_i},$$

qui a été précédemment réécrit uniquement en termes de  $\mathbf{x}_i$

$$\widehat{\mathbf{M}} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{M}}^{-1} \mathbf{x}_i}. \quad (3.11)$$

### 3.1.3 L'équation clé

Les deux hypothèses sur  $\tau$  conduisent à la même équation (3.8), dont la solution est, pour un  $\tau$  déterministe, l'estimateur du MV, et, pour un  $\tau$  aléatoire, un estimateur approché du MV mais parfaitement en accord, dans la construction, avec le détecteur GLRT-LQ. Le  $\widehat{\mathbf{M}}$  solution de (3.8) est appelée l'estimateur du point fixe  $\widehat{\mathbf{M}}_{FP}$ . Le nom de cet estimateur provient du fait que la solution de (3.8) est aussi le point fixe de fonction  $f_{N,\mathbf{M}}$ , définie par (3.4), et donc de manière plus détaillée par :

$$f_{N,\mathbf{M}} : \begin{cases} \mathcal{D} & \longrightarrow \mathcal{D} \\ \mathbf{A} & \longrightarrow \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \mathbf{A}^{-1} \mathbf{x}_i} \end{cases} \quad (3.12)$$

où  $\mathcal{D} = \{\mathbf{A} \in M_m(\mathbb{C}) \mid \mathbf{A}^H = \mathbf{A}, \mathbf{A} \text{ définie positive}\}$  avec  $M_m(\mathbb{C}) = \{\text{matrices } m \times m \text{ dont les éléments appartiennent à } \mathbb{C}\}$ .

La notation  $f_{N,\mathbf{M}}$  souligne la dépendance en  $N$  et  $\mathbf{M}$  impliquée par les  $\mathbf{x}_i$ .

On peut aussi remarquer que des fonctions similaires ont été étudiées dans le cadre de la théorie des  $\mathbf{M}$ -estimateurs [37, 34], dans lequel peut s'inscrire l'estimateur du Point Fixe.

Le premier résultat concernant  $\widehat{\mathbf{M}}_{FP}$  est présenté par le théorème suivant et concerne l'existence et l'unicité de cet estimateur :

**Théorème 3.1.3.1 (Existence et Unicité du Point Fixe)**

$\widehat{\mathbf{M}}_{FP}$ , défini comme la solution de l'équation,

$$\widehat{\mathbf{M}} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{M}}^{-1} \mathbf{x}_i},$$

existe et est unique à un facteur près, pour  $N > m$ .

Une petite remarque sur l'unicité du point fixe s'impose. En effet, si  $\widehat{\mathbf{M}}_{FP}$  est solution de (3.11), alors, pour tout  $\alpha$  complexe,  $\alpha \widehat{\mathbf{M}}_{FP}$  est aussi solution de (3.11). L'unicité concerne donc la structure de la matrice de covariance sans imposer de condition de normalisation. Cependant, dans la suite,  $\widehat{\mathbf{M}}_{FP}$  est normalisé selon la  $\widehat{\mathbf{M}}$ -normalisation,  $\text{Tr}(\mathbf{M}^{-1} \widehat{\mathbf{M}}_{FP}) = m$ . Mais, une fois encore, si on considère l'application spécifique au détecteur GLRT-LQ, on constate que cette normalisation n'a que peu d'importance au vu de la proposition 2.1.1.1 :

$$\forall \alpha \in \mathbb{C}, \widehat{\Lambda}(\alpha \widehat{\mathbf{M}}) = \widehat{\Lambda}(\widehat{\mathbf{M}}).$$

**Preuve 3.1.3.1**

Les démonstrations d'existence et d'unicité étant assez longues et faisant appel à de nombreuses considérations d'algèbre matricielle, elles sont reportées en annexe [D](#) par l'intermédiaire de [\[40\]](#).

D'autre part,  $\widehat{\mathbf{M}}_{FP}$  n'est pas analytiquement calculable, mais un algorithme itératif convergent est proposé dans [\[40\]](#), ceci fait l'objet du théorème suivant.

**Théorème 3.1.3.2 (Algorithme de construction de  $\widehat{\mathbf{M}}_{FP}$ )**

Soit, pour  $k \in \mathbb{N}$ , l'algorithme défini par :

$$\widehat{\mathbf{M}}(k+1) = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H (\widehat{\mathbf{M}}^{(k)})^{-1} \mathbf{x}_i}. \quad (3.13)$$

Alors cet algorithme converge, quelque soit le point de départ  $\widehat{\mathbf{M}}(0)$ , vers  $\widehat{\mathbf{M}}_{FP}$  quand  $k$  tend vers l'infini.

**Preuve 3.1.3.2**

La preuve de ce théorème fait suite à la preuve du théorème [3.1.3.1](#) dans l'annexe [D](#).

## 3.2 Propriétés statistiques

Les propriétés statistiques sont à nouveau étudiées par l'intermédiaire du biais, de la consistance et de la distribution asymptotique.

Cette analyse fait l'objet des trois théorèmes suivants.

### 3.2.1 Biais

**Théorème 3.2.1.1 (Biais de  $\widehat{\mathbf{M}}_{FP}$ )**  
 $\widehat{\mathbf{M}}_{FP}$  est un estimateur sans biais de  $\mathbf{M}$ .

#### Preuve 3.2.1.1

Voir l'annexe [3A.1](#) du chapitre.

### 3.2.2 Consistance

**Théorème 3.2.2.1 (Consistance de  $\widehat{\mathbf{M}}_{FP}$ )**  
 $\widehat{\mathbf{M}}_{FP}$  est un estimateur consistant de  $\mathbf{M}$ .

#### Preuve 3.2.2.1

Voir l'annexe [3A.2](#) du chapitre.

### 3.2.3 Distribution asymptotique

Les notations utilisées dans cette section pour la distribution asymptotique sont les mêmes que celles utilisées pour la distribution asymptotique de  $\widehat{\mathbf{M}}_N$  et  $\widehat{\mathbf{M}}_{SCM}$ , à la section [2.3.3](#).

Dans [\[38\]](#), les résultats suivants concernant  $\widehat{\mathbf{M}}_{FP}$  ont été établis.

**Théorème 3.2.3.1 (Distribution asymptotique de  $\text{vec}(\widehat{\mathbf{M}}_{FP})$ )**

$$1. \sqrt{N} \begin{pmatrix} \text{Re}(\boldsymbol{\delta}_{FP}) \\ \text{Im}(\boldsymbol{\delta}_{FP}) \end{pmatrix} \xrightarrow[N \rightarrow +\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_{FP}) ;$$

$$2. N \mathbb{E} [\boldsymbol{\delta}_{FP} \boldsymbol{\delta}_{FP}^{\top}] \xrightarrow[N \rightarrow +\infty]{} \left( \frac{m+1}{m} \right)^2 \mathbf{C}_1 ;$$

$$3. N \mathbb{E} [\boldsymbol{\delta}_{FP} \boldsymbol{\delta}_{FP}^H] \xrightarrow[N \rightarrow +\infty]{} \left( \frac{m+1}{m} \right)^2 \mathbf{C}_2 ;$$

où  $\mathbf{C}_1$  (respectivement  $\mathbf{C}_2$ ) est défini par l'équation (2.15) (respectivement (2.16)) et où  $\mathbf{C}_{FP}$  est complètement caractérisé par ces deux quantités.

**Preuve 3.2.3.1**

Voir l'annexe 3A.3 du chapitre.

### 3.3 Propriétés "radaristiques"

L'étude des performances statistiques du Point Fixe a montré que cet estimateur a beaucoup d'avantages à être retenu comme l'estimateur de la matrice de covariance  $\mathbf{M}$  du modèle SIRV, puisqu'il est consistant, sans biais et que son comportement asymptotique est extrêmement proche de celui de  $\widehat{\mathbf{M}}_N$ , qui est une des matrices aléatoires les plus étudiées actuellement au travers de la distribution de Wishart.

Il ne faut cependant pas négliger les autres caractéristiques de cet estimateur. Tout d'abord, la mise en œuvre, à partir d'un algorithme itératif, peut être un handicap, elle est donc analysée en termes de temps et de charge de calcul. Ensuite, certaines propriétés spécifiques à l'utilisation du point fixe dans le GLRT-LQ vont être établies : la propriété CFAR-texture déjà vue ainsi que la propriété CFAR-matrice.

#### 3.3.1 Mise en œuvre

$\widehat{\mathbf{M}}_{FP}$  n'étant pas défini de manière explicite mais uniquement comme étant le point fixe d'une fonction  $f_{N,\mathbf{M}}$  (3.12), il est indispensable de construire cet estimateur à l'aide d'un algorithme itératif, présenté par le théorème 3.1.3.2. Afin d'apprécier le comportement de cet algorithme, des simulations sont présentées selon le point d'initialisation de l'algorithme, la matrice de covariance  $\mathbf{M}$  à estimer, le nombre  $N$  de données secondaires et la taille  $m$  des observations.

Tout d'abord, le choix des vraies matrices de covariance  $\mathbf{M}$  s'est arrêté sur les matrices de Toeplitz car elles sont fréquemment rencontrées dans les cas de processus stationnaires. Nous avons choisi dans notre étude des matrices de la forme :

$$M_{ij} = \rho^{|i-j|},$$

pour  $1 \leq i, j \leq m$  et pour  $0 < \rho < 1$  et avec, par convention, pour  $\rho = 0$ ,  $\mathbf{M} = \mathbf{I}$ . On peut aussi noter que des matrices de ce type modélisent bien le clutter.

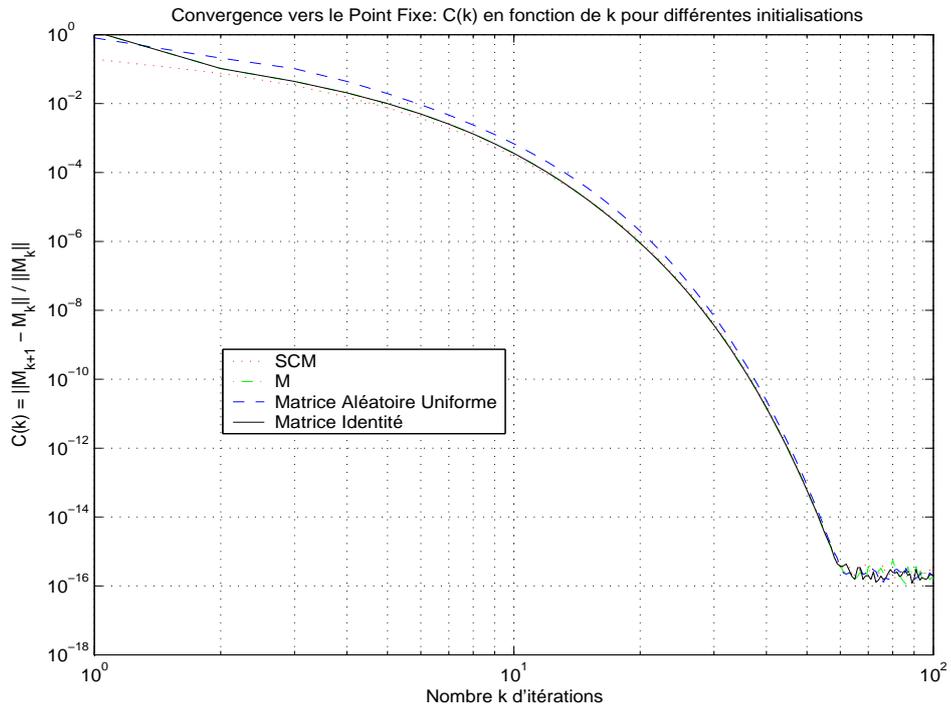
La matrice  $\mathbf{M}$  est donc complètement définie par le paramètre  $\rho$ , qui caractérise la corrélation de chaque vecteur de données : pour  $\rho = 0$ , chaque vecteur est constitué de composantes indépendantes tandis que plus  $\rho$  augmente, plus ces composantes sont corrélées. La convergence de l'algorithme est analysée en évaluant le critère  $C$ , appelé aussi erreur relative, comme une fonction de l'itération  $k$  :

$$C(k) = \frac{\|\widehat{\mathbf{M}}(k+1) - \widehat{\mathbf{M}}(k)\|}{\|\widehat{\mathbf{M}}(k)\|}. \quad (3.14)$$

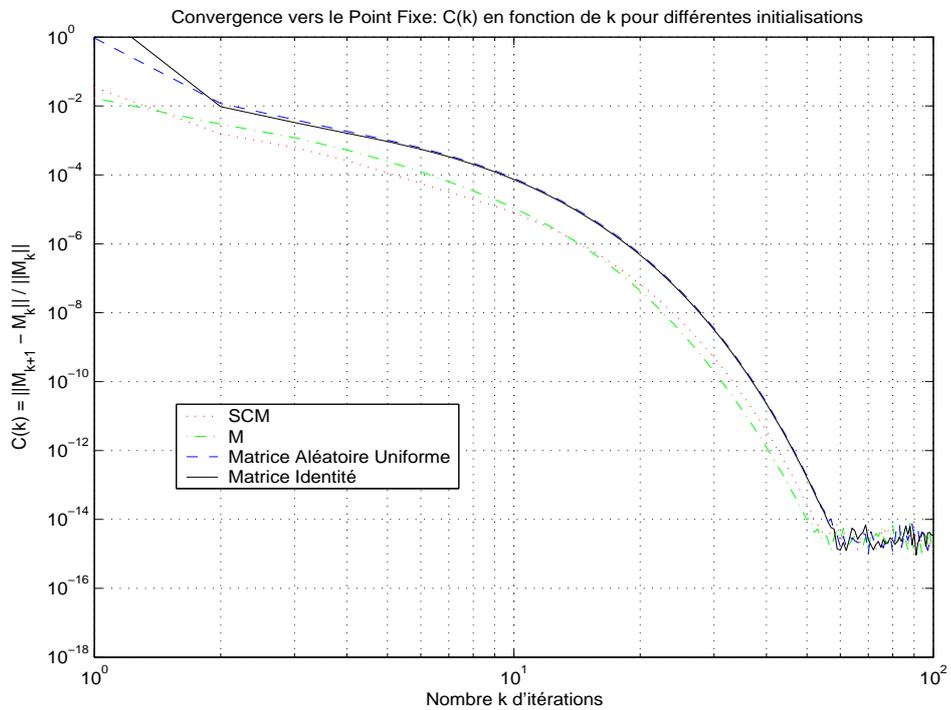
#### Remarque 3.3.1.1

Étant donné que l'erreur relative s'obtient en comparant deux itérations consécutives  $\widehat{\mathbf{M}}(k+1)$  et  $\widehat{\mathbf{M}}(k)$ , il est essentiel que ces deux matrices possèdent la même normalisation : la  $\mathbf{M}$ -normalisation, car ces simulations mettent en valeur la mise en œuvre pratique de l'estimateur du point fixe. Le nouvel algorithme utilisé est alors, pour  $k \in \mathbb{N}$ , :

$$\widehat{\mathbf{M}}(k+1) = \frac{m}{\text{Tr}(f_{N,\mathbf{M}}(\widehat{\mathbf{M}}(k)))} f_{N,\mathbf{M}}(\widehat{\mathbf{M}}(k)). \quad (3.15)$$



(a)  $\rho = 0.001$



(b)  $\rho = 0.999$

FIG. 3.1 – Influence de la matrice  $\hat{M}(0)$  d'initialisation de l'algorithme 3.15, pour  $m = 8$  et  $N = 20$

### 3.3.1.1 Influence des paramètres

Cet algorithme est utilisé pour mettre en valeur l'influence des différents paramètres utilisés. Les 3 figures suivantes étudient l'influence de la matrice  $\widehat{\mathbf{M}}(0)$  d'initialisation de l'algorithme (Fig. 3.1), le rôle du nombre  $N$  de données secondaires (Fig. 3.2) et l'influence de la taille  $m$  des observations (Fig. 3.3), pour deux valeurs de  $\rho$  :

- $\rho = 0.001$ , et ainsi la matrice de covariance  $\mathbf{M}$  est très proche de la matrice identité ;
- $\rho = 0.999$ , ce qui produit des données très corrélées et une matrice  $\mathbf{M}$  plus difficilement inversible.

Analyse de la figure (Fig. 3.1) :

Pour étudier, l'influence du point d'initialisation de l'algorithme 3.15, 4 matrices ont été utilisées :  $\widehat{\mathbf{M}}_{SCM}$ ,  $\mathbf{M}$ , une matrice aléatoire dont les éléments sont distribués uniformément sur  $[0, 1]$  et  $\mathbf{I}$ . Le critère  $C(k)$  a été tracé, en échelle logarithmique, en fonction de  $k = 0, \dots, 100$  pour  $m = 8$  et  $N = 20$ . Le choix des paramètres  $m$  et  $N$  est proche de ce qui peut arriver dans la pratique tout en respectant la contrainte théorique  $N \geq 2m$ .

Pour les 2 valeurs de  $\rho$ , le critère  $C(k)$  converge vers 0 (ici,  $10^{-16}$  représente la précision "machine" limite) pour  $k \simeq 50$  itérations, ce qui signifie que le point fixe est atteint pour ce nombre  $k$  d'itérations. En effet,

$$C(k) \simeq 0 \iff \widehat{\mathbf{M}}(k+1) \simeq \widehat{\mathbf{M}}(k) \simeq \widehat{\mathbf{M}}_{FP}.$$

Notons ici que le sens donné au mot "précision" est l'écart relatif entre  $\widehat{\mathbf{M}}(k)$  et  $\widehat{\mathbf{M}}(k+1)$ .

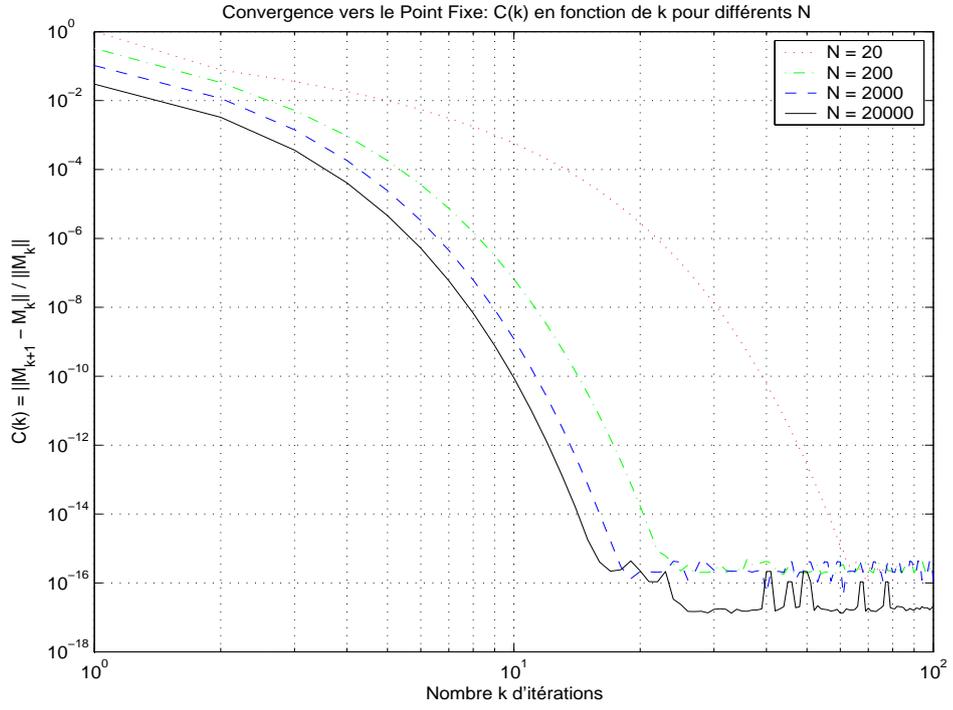
Comme l'affirmait la théorie, l'algorithme converge vers le point fixe quelle que soit la matrice  $\widehat{\mathbf{M}}_0$  d'initialisation. Sur la figure (Fig. 3.1), les différentes courbes se rapprochent très rapidement malgré des valeurs initiales parfois très éloignées. Ce phénomène est encore plus visible quand la matrice  $\mathbf{M}$  à estimer est proche de l'identité. Ceci peut s'expliquer par le fait que cette matrice est très simplement inversible, ce qui n'est pas le cas quand  $\rho = 0.999$ , et comme l'inversion de matrice intervient dans le dénominateur de l'algorithme, ceci peut dégrader légèrement sa précision.

L'initialisation de l'algorithme n'ayant aucune influence sur son comportement, dans la suite des simulations, dans un souci de simplicité, la matrice  $\widehat{\mathbf{M}}(0)$  d'initialisation choisie sera dorénavant la matrice identité.

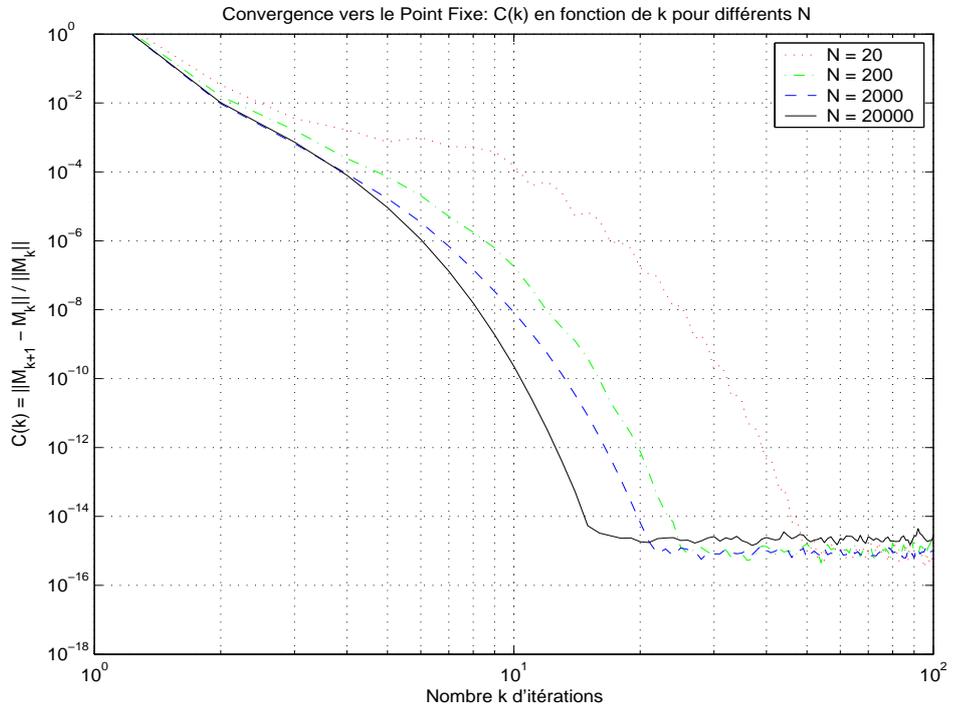
Analyse de la figure (Fig. 3.2) :

Ce graphique étudie l'influence du nombre  $N$  de données secondaires sur le comportement de l'algorithme de construction du point fixe. Les valeurs de  $N$  sont 20, 200, 2000 et 20000, quant à  $m$ , sa valeur est fixée à 8 afin de toujours respecter  $N \geq 2m$ .

Le premier constat illustré par ce graphique est le suivant : pour les deux valeurs de  $\rho$ , plus  $N$  est grand et plus l'algorithme converge rapidement vers le point fixe. En effet, quand  $N = 20000$ , il faut seulement une dizaine d'itérations pour atteindre la limite de l'algorithme, i.e. le point fixe, alors qu'il faut près de 50 itérations quand  $N = 20$ . De plus, quelle que soit la valeur de  $N$ , la précision de  $C(k)$  est la même, environ  $10^{-16}$  quand  $\rho = 0.001$  et  $10^{-15}$  quand  $\rho = 0.999$ . L'influence de  $\rho$  est étudiée plus en détail dans la suite de cette section. Cependant, sur la figure (Fig. 3.2.a), pour  $\rho = 0.001$ , on peut remarquer que la stabilité du critère quand l'algorithme a convergé, est perturbée quand  $N$  est grand.



(a)  $\rho = 0.001$



(b)  $\rho = 0.999$

FIG. 3.2 – Influence du nombre  $N$  de données secondaires, pour  $m = 8$  et  $\hat{M}(0) = \mathbf{I}$

Ainsi, quand  $N = 20000$ , la courbe noire, après la 10<sup>ème</sup> itération, oscille entre  $10^{-17}$  et  $10^{-16}$ , ce qui n'a aucune importance au vu de ces valeurs extrêmement faibles.

Analyse de la figure (Fig. 3.3) :

Sur cette figure, pour deux valeurs de  $\rho$ , les critères  $C(k)$  ont été tracés en fonction du nombre  $k$  d'itérations pour plusieurs valeurs du paramètre  $m$  : 3, 5, 8, 20, 50 et 100,  $N$  étant fixé au préalable à  $N = 200$ .

Dans ce cas, la valeur de  $\rho$  est importante puisque l'influence de  $m$  sur le comportement de l'algorithme est différente selon que  $\rho = 0.001$  ou que  $\rho = 0.999$ . Dans le premier cas (Fig. 3.3.a), la convergence est atteinte plus rapidement, entre 20 et 30 itérations, quand  $m$  est faible. En revanche, il faut près de 40 itérations, quand  $m = 100$ . La précision de l'algorithme est toujours sensiblement la même,  $C(k) \simeq 10^{-16}$  pour  $\rho = 0.001$ . En revanche, pour  $\rho = 0.999$  (Fig. 3.3.b), le paramètre n'a aucune influence sur le comportement de l'algorithme : les courbes sont pratiquement toutes confondues, avec une précision oscillant entre  $10^{-14}$  et  $10^{-15}$  selon les valeurs de  $m$  : meilleure précision pour des valeurs de  $m$  faibles.

En fait, une conclusion peut être extraite des deux précédentes analyses :  $m$  a peu d'influence sur le comportement de l'algorithme 3.15 mais, en revanche, l'écart entre  $m$  et  $N$  joue un rôle non négligeable : plus  $N/m$  est proche de 2, plus le nombre d'itérations augmente. Plusieurs simulations complémentaires ont montré que  $N$  était le paramètre le plus influent sur le comportement de l'algorithme, ce qui est "intuitivement" en accord avec le rôle de  $\hat{\mathbf{M}}_{FP}$  : estimer  $\mathbf{M}$ , avec une meilleure estimation pour des valeurs élevées de  $N$ .

### 3.3.1.2 Comportement de l'algorithme en fonction de $\rho$

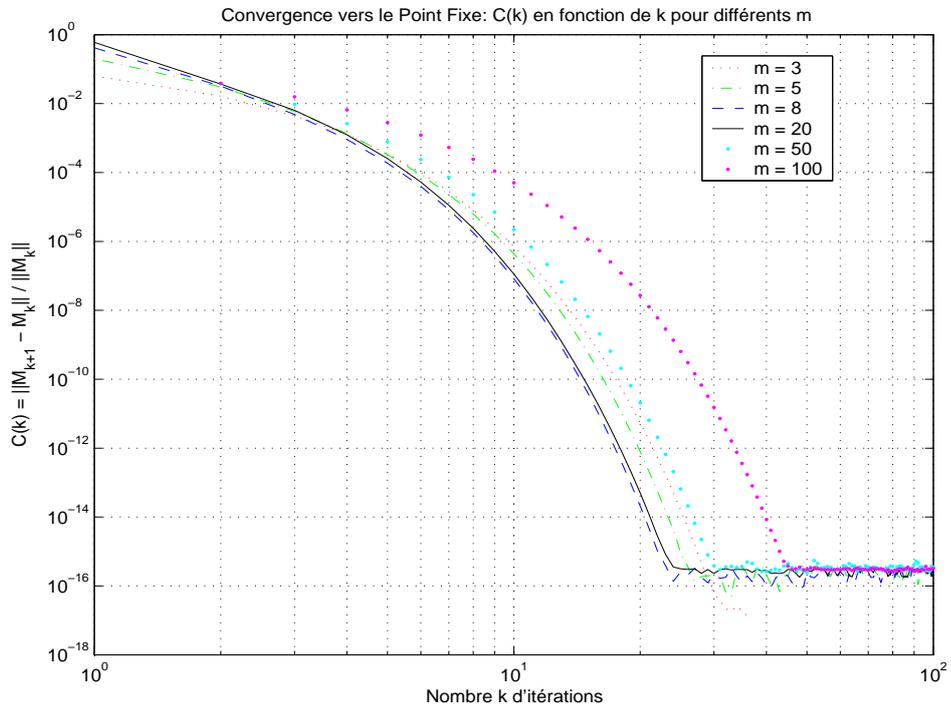
Sur la figure (Fig. 3.4), le critère  $C(k)$  est tracé en fonction du nombre  $k$  d'itérations pour l'ensemble des paramètres suivants :  $m = 8$ ,  $N = 20$  et l'initialisation  $\hat{\mathbf{M}}(0) = \mathbf{I}$ . Cette simulation va mettre en évidence le rôle de la corrélation des données observées sur le comportement de l'algorithme.

Quand  $\rho$  est faible ( $10^{-5}$ ), la matrice de covariance  $\mathbf{M}$  est proche de l'identité et les données sont très peu corrélées. Dans ce cas, l'algorithme converge plus lentement (courbe bleue), environ 60 itérations, mais avec une meilleure précision :  $C(k)$  est proche de  $10^{-16}$ .

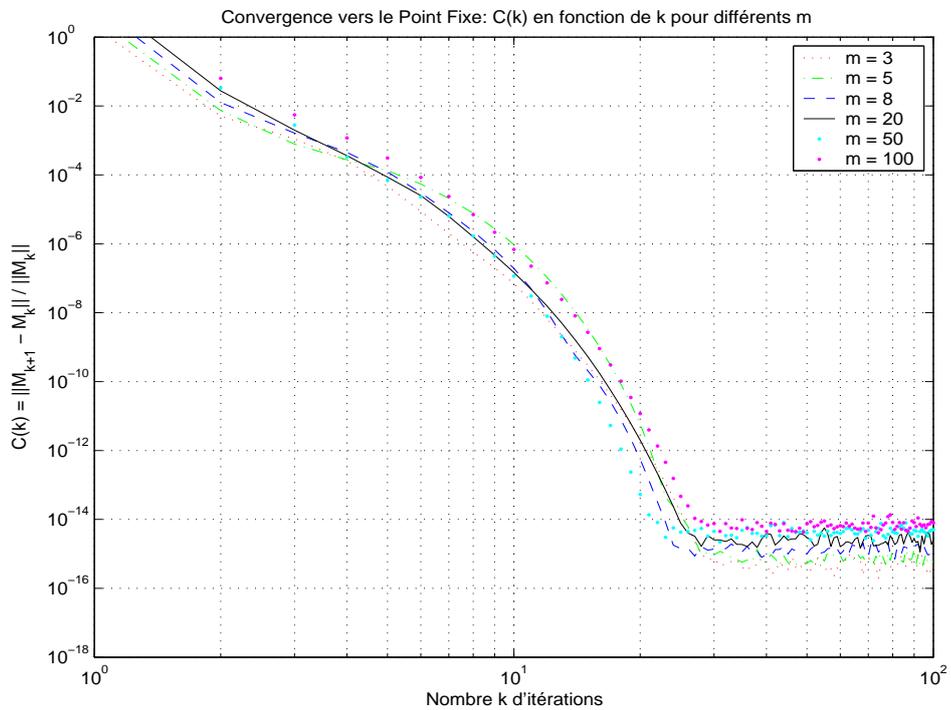
En revanche, l'augmentation de  $\rho$  traduit une plus forte corrélation entre les données, ce qui peut être plus proche de la réalité. Dans ce contexte, la précision de l'algorithme diminue quelque peu (jusqu'à  $10^{-14}$ ) mais gagne en vitesse de convergence. La précision de  $10^{-14}$  est atteinte dès 40 itérations.

On peut donc conclure à la robustesse de l'algorithme par rapport à la corrélation des données.

Les 2 figures (Fig. 3.5 et Fig. 3.6) permettent de corroborer plus en détail les conclusions de la figure Fig.(3.4). En effet, la figure (Fig. 3.5) présente le critère  $C(150)$  en fonction de  $\rho$ , i.e. la valeur de  $C$  alors que la convergence a eu lieu, ceci met à nouveau en évidence, plus en détail, l'influence de  $\rho$  sur la précision de l'algorithme : on constate que la meilleure précision possible est celle de la machine, i.e.  $10^{-16}$ , quand  $\rho$  est proche de 0, puis la précision diminue quand  $\rho$  augmente. Elle diminue jusqu'à  $10^{-4}$  pour des valeurs de  $\rho$  proche de  $1 - 10^{-7}$ , ce qui n'est pas très réaliste puisque la matrice composée



(a)  $\rho = 0.001$



(b)  $\rho = 0.999$

FIG. 3.3 – Influence de la taille  $m$  des observations, pour  $N = 200$  et  $\hat{M}(0) = \mathbf{I}$

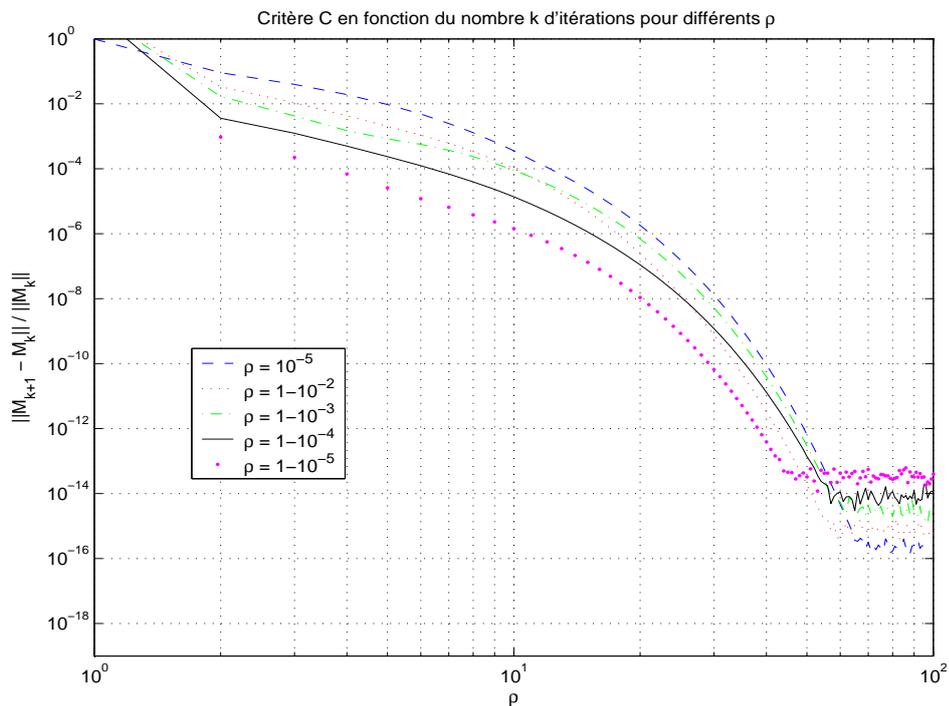


FIG. 3.4 – Comportement de l'algorithme pour différents coefficients de corrélation  $\rho$

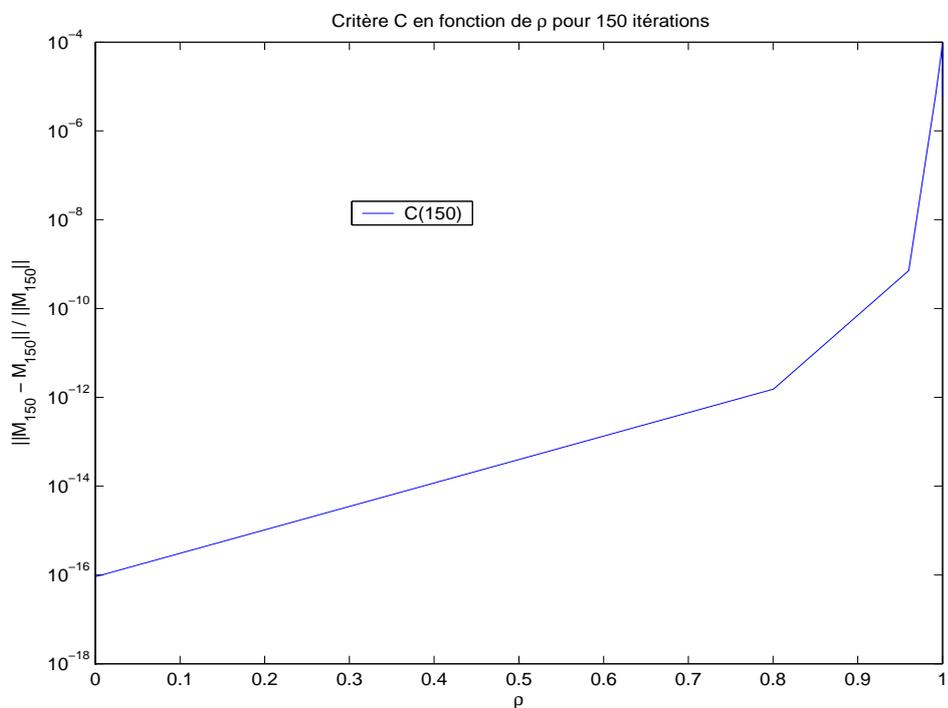


FIG. 3.5 – Comportement de la convergence de l'algorithme :  $C(150)$  en fonction de  $\rho$

uniquement de 1 n'est pas inversible...

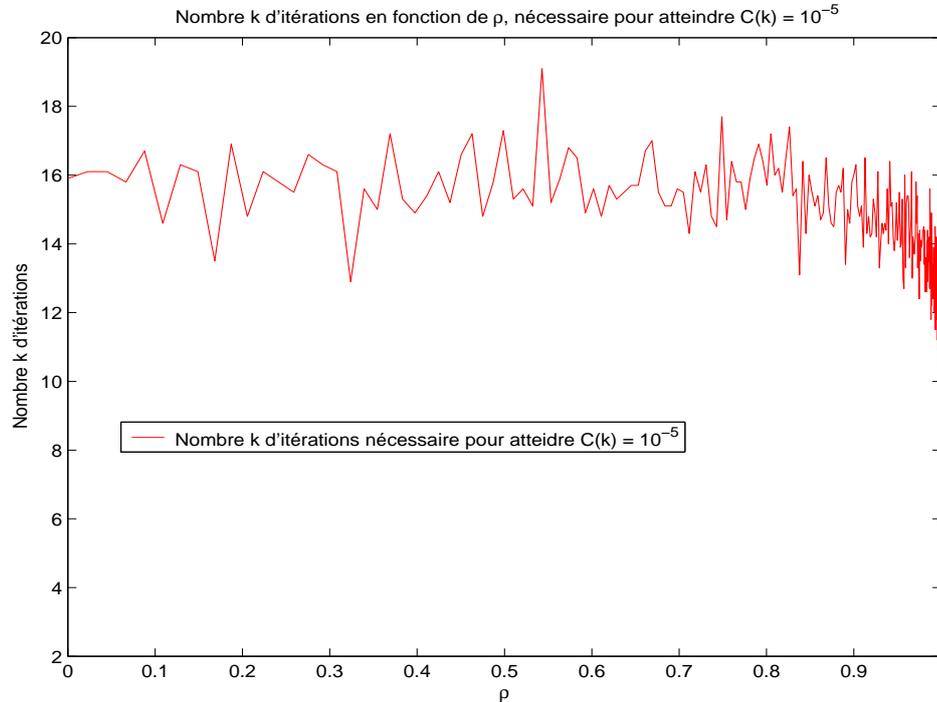


FIG. 3.6 – Comportement de la convergence de l'algorithme : Nombre  $k$  d'itérations, nécessaire pour atteindre  $C(k) = 10^{-5}$ , en fonction de  $\rho$

Sur la figure (Fig. 3.6), cette fois, c'est le nombre  $k$  d'itérations, nécessaire pour avoir  $C(k) = 10^{-5}$ , qui est tracé en fonction de  $\rho$ . Cette figure confirme, elle aussi, les précédentes affirmations : la courbe rouge a une tendance décroissante, ce qui signifie qu'il faut moins d'itérations quand  $\rho$  est proche de 1, pour obtenir  $C(k) = 10^{-5}$  ; en d'autres termes, l'algorithme converge avec des données extrêmement corrélées,  $\rho = 1 - 10^{-7}$ . Ce point est assez intéressant pour les cas pratiques où la matrice de covariance se rapproche fortement d'une matrice non inversible : l'algorithme, bien que sa précision soit légèrement dégradée, converge tout de même.

La conclusion de ce paragraphe est, que dans les conditions normales (en pratique) d'utilisation de cet algorithme, ce dernier converge assez rapidement, entre 30 et 40 itérations, avec une excellente précision de  $10^{-16}$ .

Il serait donc intéressant de faire une analyse numérique approfondie (temps de calcul, coût de calcul,...) de cet algorithme.

### 3.3.2 Les propriétés CFAR

Pour être utilisé en pratique, l'estimateur du point fixe doit être réinjecté dans le détecteur GLRT-LQ, qui devient alors :

$$\hat{\Lambda}(\hat{\mathbf{M}}_{FP}) = \frac{|\mathbf{p}^H \hat{\mathbf{M}}_{FP}^{-1} \mathbf{y}|^2}{(\mathbf{p}^H \hat{\mathbf{M}}_{FP}^{-1} \mathbf{p})(\mathbf{y}^H \hat{\mathbf{M}}_{FP}^{-1} \mathbf{y})} \underset{H_0}{\underset{H_1}{\gtrless}} \lambda, \quad (3.16)$$

où le seuil de détection  $\lambda$  est défini par

$$\lambda = 1 - \eta^{-\frac{1}{m}}.$$

et  $\hat{\mathbf{M}}_{FP}$  est défini par l'équation

$$\hat{\mathbf{M}}_{FP} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \hat{\mathbf{M}}_{FP}^{-1} \mathbf{x}_i} = f_{N, \mathbf{M}}(\hat{\mathbf{M}}_{FP}), \quad (3.17)$$

où, pour tout  $i$ ,  $\mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{M})$ .

### 3.3.2.1 Propriété CFAR-texture

$\hat{\mathbf{M}}_{FP}$  ne dépendant que des  $\mathbf{x}_i$ , le théorème suivant découle :

**Théorème 3.3.2.1 (Propriété CFAR-texture de  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$ )**  
 $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$  possède la propriété CFAR-texture sous  $H_0$  : sa loi de probabilité ne dépend pas de la texture sous  $H_0$ .

#### Preuve 3.3.2.1

*Évidente.*

### 3.3.2.2 Propriété CFAR-matrice

L'objet de ce paragraphe est de mettre en place la propriété CFAR-matrice de  $\hat{\Lambda}(\hat{\mathbf{M}})$  :

*"la loi du détecteur, construit avec l'estimateur  $\hat{\mathbf{M}}_{FP}$  de  $\mathbf{M}$ , ne dépend pas de  $\mathbf{M}$ ."*

Cette proposition peut paraître absurde mais elle est d'un intérêt inestimable dans la pratique...

#### Définition 3.3.2.1 (Propriété CFAR-matrice)

Un détecteur adaptatif  $\hat{\Lambda}(\hat{\mathbf{M}})$  vérifie la propriété "CFAR-matrice" si sa distribution statistique est indépendante de la matrice  $\mathbf{M}$  de covariance estimée par  $\hat{\mathbf{M}}$ .

Cette propriété fait l'objet du théorème suivant :

**Théorème 3.3.2.2 (Propriété CFAR-matrice de  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$ )**

Soient  $\mathbf{M}_1$  et  $\mathbf{M}_2$  deux matrices de covariance distinctes ( $\mathbf{M}_1 \neq \mathbf{M}_2$ ), soit  $\hat{\mathbf{M}}_{FP,1}$  (respectivement  $\hat{\mathbf{M}}_{FP,2}$ ) l'estimateur du point fixe de  $\mathbf{M}_1$  (respectivement de  $\mathbf{M}_2$ ),

alors, sous l'hypothèse  $H_0$ ,

$$\mathcal{L}(\hat{\Lambda}(\hat{\mathbf{M}}_{FP,1})) = \mathcal{L}(\hat{\Lambda}(\hat{\mathbf{M}}_{FP,2})) \quad (3.18)$$

où  $\mathcal{L}(X)$  représente la loi de la variable aléatoire  $X$ .

Ainsi, le théorème 3.3.2.2 établit la propriété CFAR-matrice du détecteur GLRT-LQ construit avec l'estimateur du point fixe  $\hat{\mathbf{M}}_{FP}$ .

**Preuve 3.3.2.2**

Voir l'annexe 3A.4 du chapitre.

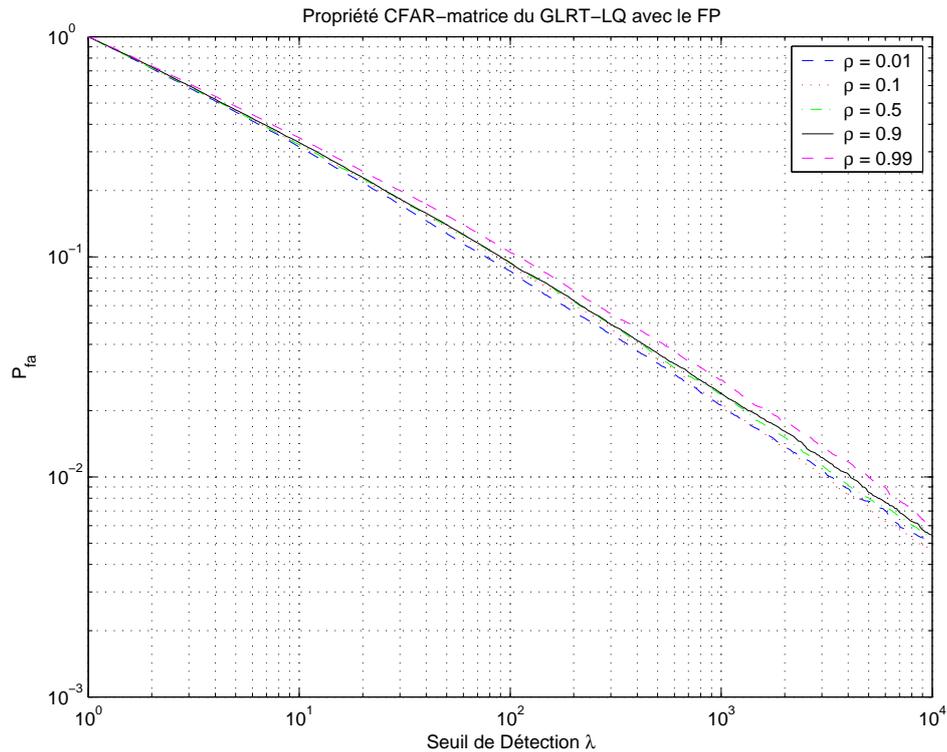


FIG. 3.7 – Propriété CFAR-matrice de  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$

Afin d'illustrer le résultat de ce théorème, la figure (Fig. 3.7) présente des courbes empiriques de probabilités de fausse alarme en fonction du seuil de détection  $\lambda$ , pour le détecteur GLRT-LQ construit avec l'estimateur du point fixe  $\hat{\mathbf{M}}_{FP}$ . Les paramètres utilisés sont :  $m = 4$ ,  $N = 10$ , l'algorithme de

construction du point fixe est initialisé par la matrice identité et 80 itérations sont effectuées pour obtenir  $\hat{\mathbf{M}}_{FP}$ .

Les courbes " $P_{fa}$ -seuil" ont été tracées pour différentes matrices de covariance  $\mathbf{M}$ . Comme dans les simulations précédentes,  $\mathbf{M}$  est une matrice de Toeplitz, dont les éléments sont de la forme :

$$M_{ij} = \rho^{|i-j|},$$

pour  $1 \leq i, j \leq m$  et pour  $0 < \rho < 1$  et avec, par convention, pour  $\rho = 0$ ,  $\mathbf{M} = \mathbf{I}$ .

Sur la figure (Fig. 3.7), 5 valeurs de  $\rho$  ont été utilisées :  $\rho = 0.01$ ,  $\rho = 0.1$ ,  $\rho = 0.5$ ,  $\rho = 0.9$  et  $\rho = 0.99$ . Pour toutes ces valeurs de  $\rho$ , les courbes " $P_{fa}$ -seuil" sont pratiquement confondues, ce qui met en évidence l'indépendance de la distribution du détecteur GLRT-LQ, construit avec l'estimateur du point fixe, avec la vraie matrice de covariance  $\mathbf{M}$ . La remarque 1.1.0.1 permet en effet d'affirmer que la distribution d'un détecteur est complètement caractérisée par la relation " $P_{fa}$ -seuil", qui représente la fonction de répartition complémentaire du détecteur.

#### **Remarque 3.3.2.1**

*Cette propriété du détecteur d'être CFAR-matrice est très importante en pratique : en effet, sur une zone de fouillis hétérogène en terme de matrice de covariance (localement, chaque matrice de covariance est différente), ce qui caractérise bien la réalité, la distribution du détecteur reste invariante. Ainsi, la régulation de la fausse alarme est la même quelle que soit la taille de la carte de fouillis étudiée. De plus, puisque cette propriété d'invariance concerne aussi la texture, ceci ouvre des perspectives intéressantes pour la détection de cibles dans des zones de transition de fouillis, comme par exemple, les zones terremer (côtes, îles, ...) ou les zones de lisières de forêt.*

# Chapitre 3A

## Preuves du chapitre 3

### 3A.1 Preuve du théorème 3.2.1.1 : "Biais de $\widehat{\mathbf{M}}_{FP}$ "

Pour la clarté de cette preuve,  $\widehat{\mathbf{M}}_{FP}$  est simplement noté  $\widehat{\mathbf{M}}$ , et défini par l'intermédiaire de l'équation (3.11) :

$$\widehat{\mathbf{M}} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{M}}^{-1} \mathbf{x}_i}. \quad (3A.1)$$

Tout d'abord,

$$\mathbb{E} [\widehat{\mathbf{M}}] = \mathbb{E} \left[ \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{M}}^{-1} \mathbf{x}_i} \right] = m \mathbb{E} \left[ \frac{\mathbf{x} \mathbf{x}^H}{\mathbf{x}^H \widehat{\mathbf{M}}^{-1} \mathbf{x}} \right], \quad (3A.2)$$

par linéarité de l'espérance mathématique et car les  $\mathbf{x}_i$  sont i.i.d (indépendants et identiquement distribués) et où  $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{M})$ .

En blanchissant les  $\mathbf{x}_i$  selon  $\mathbf{y} = \mathbf{M}^{-1/2} \mathbf{x}$ , (3A.1) devient

$$\widehat{\mathbf{M}} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{M}^{1/2} \mathbf{y}_i \mathbf{y}_i^H \mathbf{M}^{1/2}}{\mathbf{y}_i^H \widehat{\mathbf{T}}^{-1} \mathbf{y}_i},$$

avec

$$\widehat{\mathbf{T}} = \mathbf{M}^{-1/2} \widehat{\mathbf{M}} \mathbf{M}^{-1/2}.$$

Et ainsi,

$$\widehat{\mathbf{T}} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \widehat{\mathbf{T}}^{-1} \mathbf{y}_i}. \quad (3A.3)$$

$\widehat{\mathbf{T}}$  est donc l'unique point fixe (à un facteur près) de la matrice identité. Sa loi est clairement indépendante de  $\mathbf{M}$  puisque les  $\mathbf{y}_i$  sont distribués selon  $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ .

De plus, pour toute matrice unitaire  $\mathbf{U}$ , on a

$$\mathbf{U} \widehat{\mathbf{T}} \mathbf{U}^H = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{z}_i \mathbf{z}_i^H}{\mathbf{z}_i^H (\mathbf{U} \widehat{\mathbf{T}} \mathbf{U}^H)^{-1} \mathbf{z}_i}, \quad (3A.4)$$

où  $\mathbf{z}_i = \mathbf{U} \mathbf{y}_i$  est aussi  $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ . Par conséquent,  $\mathbf{U} \widehat{\mathbf{T}} \mathbf{U}^H$  possède la même distribution que  $\widehat{\mathbf{T}}$  et donc,

$$\mathbb{E} [\widehat{\mathbf{T}}] = \mathbf{U} \mathbb{E} [\widehat{\mathbf{T}}] \mathbf{U}^H, \text{ pour toute matrice unitaire } \mathbf{U}. \quad (3A.5)$$

Puisque  $\mathbb{E} [\widehat{\mathbf{T}}]$  est évidemment différent de  $\mathbf{0}$ , le lemme B.0.0.1, détaillé dans l'annexe B, assure que  $\mathbb{E} [\widehat{\mathbf{T}}] = \alpha \mathbf{I}$ . Ensuite, comme  $\widehat{\mathbf{T}} = \mathbf{M}^{-1/2} \widehat{\mathbf{M}} \mathbf{M}^{-1/2}$ , on a  $\mathbb{E} [\widehat{\mathbf{M}}] = \alpha \mathbf{M}$  et la  $\widehat{\mathbf{M}}$ -normalisation,  $\text{Tr}(\mathbf{M}^{-1} \widehat{\mathbf{M}}) = m$ , implique que  $\alpha = 1$  et la conclusion en découle :  $\widehat{\mathbf{M}}$  est un estimateur sans biais de  $\mathbf{M}$ , pour tout entier  $N$ .

### 3A.2 Preuve du théorème 3.2.2.1 : "Consistance de $\widehat{\mathbf{M}}_{FP}$ "

Afin de montrer la consistance de  $\widehat{\mathbf{M}}_{FP}$ , noté dans cette preuve  $\widehat{\mathbf{M}}(N)$  pour rendre compte de la dépendance entre  $\widehat{\mathbf{M}}_{FP}$  et le nombre  $N$  de  $\mathbf{x}_i$ , il est nécessaire d'étudier les propriétés de la fonction  $f_{N,\mathbf{M}}$  définie par (3.12). Tout d'abord, une nouvelle fonction  $g_N$  est introduite :

$$g_N : \begin{cases} \mathcal{D} & \longrightarrow \mathcal{D} \\ \mathbf{A} & \longrightarrow g_N = A - f_{N,\mathbf{M}}(A) \end{cases} \quad (3A.6)$$

où  $\mathcal{D} = \{\mathbf{A} \in M_m(\mathbb{C}) \mid \mathbf{A}^H = \mathbf{A}, \mathbf{A} \text{ défini positive}\}$  avec  $M_m(\mathbb{C}) = \{\text{matrices } m \times m \text{ dont les éléments appartiennent à } \mathbb{C}\}$ .

Si  $\widehat{\mathbf{M}}(N)$  est un point fixe de la fonction  $f_{N,\mathbf{M}}$ , alors il est l'unique zéro, à un facteur près, de la fonction aléatoire  $g_N$ .

Afin de démontrer la consistance de  $\widehat{\mathbf{M}}(N)$ , nous allons utiliser le théorème 5.9 pp. 46 de [51]. Il faut donc vérifier les hypothèses de ce théorème.

Tout d'abord, grâce à la LfGN, on a :

$$\forall \mathbf{A} \in \mathcal{D}, g_N(\mathbf{A}) \xrightarrow[N \rightarrow +\infty]{Pr} g(\mathbf{A}), \quad (3A.7)$$

où  $\xrightarrow[N \rightarrow +\infty]{Pr}$  désigne la convergence en probabilité et

$$\forall \mathbf{A} \in \mathcal{D}, g(\mathbf{A}) = \mathbf{A} - m \mathbb{E} \left[ \frac{\mathbf{x}\mathbf{x}^H}{\mathbf{x}^H \mathbf{A}^{-1} \mathbf{x}} \right], \quad (3A.8)$$

pour  $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{M})$ .

Ensuite,  $g$ , définie par (3A.8), est réécrite avec un changement de variable approprié appliqué à  $\mathbf{x}$ . Ainsi, en posant  $\mathbf{y} = \mathbf{A}^{-1/2} \mathbf{x}$ ,  $g$  devient :

$$\forall \mathbf{A} \in \mathcal{D}, g(\mathbf{A}) = \mathbf{A}^{1/2} \left( \mathbf{I} - m \mathbb{E} \left[ \frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H \mathbf{y}} \right] \right) \mathbf{A}^{1/2},$$

où  $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{A}^{-1/2} \mathbf{M} \mathbf{A}^{-1/2})$ .

Ensuite, en appliquant le même changement de variable à  $g_N$ , définie par (3A.6), on obtient :

$$\forall \mathbf{A} \in \mathcal{D}, g_N(\mathbf{A}) = \mathbf{A}^{1/2} \left( \mathbf{I} - \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} \right) \mathbf{A}^{1/2},$$

où, pour  $i = 1, \dots, N$ ,  $\mathbf{y}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{A}^{-1/2} \mathbf{M} \mathbf{A}^{-1/2})$ .

Maintenant, il faut vérifier les deux hypothèses (H1) et (H2) du théorème, i.e. pour tout  $\varepsilon > 0$ ,

$$(H1) : \sup_{\mathbf{A} \in \mathcal{D}} \{\|g_N(\mathbf{A}) - g(\mathbf{A})\|\} \xrightarrow[N \rightarrow +\infty]{Pr} 0,$$

$$(H2) : \inf_{\mathbf{A} : \|\mathbf{A} - \mathbf{M}\| \geq \varepsilon} \{\|g(\mathbf{A})\|\} > 0 = g(\mathbf{M}).$$

Grâce au précédent changement de variable appliqué à  $g$  et  $g_N$ , on a pour tout  $\mathbf{A} \in \mathcal{D}$  :

$$\|g_N(\mathbf{A}) - g(\mathbf{A})\| = m \|\mathbf{A}\| \left\| \frac{1}{N} \sum_{i=1}^N \left( \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} - \mathbb{E} \left[ \frac{\mathbf{y} \mathbf{y}^H}{\mathbf{y}^H \mathbf{y}} \right] \right) \right\|,$$

et donc, la Loi Forte des Grands Nombres, appliquée aux  $N$  variables i.i.d  $\mathbf{Y}_i = \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i}$ , avec le même moment d'ordre 1, implique (H1).

Maintenant, pour montrer (H2), il suffit d'utiliser le biais de  $\widehat{\mathbf{M}}_{NSCM}$  énoncé par le théorème 2.3.1.2. En effet, pour tout  $\mathbf{A} \in \mathcal{D}$ , avec  $\mathbf{A} \neq \mathbf{M}$ ,

$$\|g(\mathbf{A})\| > 0.$$

L'équation précédente est issue de

$$\|g(\mathbf{A})\| = \|B(\widehat{\mathbf{M}}_{NSCM, \mathbf{A}^{-1/2} \mathbf{M} \mathbf{A}^{-1/2}})\|,$$

où  $\widehat{\mathbf{M}}_{NSCM, \mathbf{A}^{-1/2} \mathbf{M} \mathbf{A}^{-1/2}}$  est l'estimateur par NSCM de  $\mathbf{A}^{-1/2} \mathbf{M} \mathbf{A}^{-1/2}$  et  $B(\widehat{\mathbf{M}}_{NSCM, \mathbf{A}^{-1/2} \mathbf{M} \mathbf{A}^{-1/2}})$  le biais de  $\widehat{\mathbf{M}}_{NSCM, \mathbf{A}^{-1/2} \mathbf{M} \mathbf{A}^{-1/2}}$  défini par le théorème 2.3.1.2.

Finalement, le théorème 5.9 pp. 46 de [51] conclut cette preuve et  $\widehat{\mathbf{M}}_{FP} \xrightarrow[N \rightarrow +\infty]{Pr} \mathbf{M}$ , ce qui est la définition de la consistance de  $\widehat{\mathbf{M}}_{FP}$ .

### 3A.3 Preuve du théorème 3.2.3.1 : "Distribution asymptotique de $\widehat{\mathbf{M}}_{FP}$ "

Tout d'abord, écrivons  $\widehat{\mathbf{M}} = \mathbf{M} + \delta\mathbf{M}$  où  $\widehat{\mathbf{M}} = \widehat{\mathbf{M}}_{FP}$ . Pour un  $N$  suffisamment grand,  $\delta\mathbf{M} \simeq \mathbf{0}$  car  $\widehat{\mathbf{M}}$  est un estimateur consistant de  $\mathbf{M}$  (Théorème 3.2.2.1).

$N$  est donc supposé suffisamment grand afin d'assurer la validité des expressions d'ordre 1.

On peut ainsi écrire

$$\begin{aligned}\widehat{\mathbf{M}}^{-1} &= (\mathbf{M} (\mathbf{I} + \mathbf{M}^{-1}(\delta\mathbf{M})))^{-1} \\ &= (\mathbf{I} + \mathbf{M}^{-1}(\delta\mathbf{M}))^{-1} \mathbf{M}^{-1} \\ &\simeq (\mathbf{I} - \mathbf{M}^{-1}(\delta\mathbf{M})) \mathbf{M}^{-1} \\ &\simeq (\mathbf{M}^{-1} - \mathbf{M}^{-1}(\delta\mathbf{M})\mathbf{M}^{-1}) .\end{aligned}$$

Pour  $N$  suffisamment grand, ceci implique que

$$\widehat{\mathbf{M}} \simeq \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H (\mathbf{M}^{-1} - \mathbf{M}^{-1}(\delta\mathbf{M})\mathbf{M}^{-1}) \mathbf{x}_i},$$

et donc,

$$\delta\mathbf{M} \simeq \frac{m}{N} \sum_{i=1}^N \left( \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H (\mathbf{M}^{-1} - \mathbf{M}^{-1}(\delta\mathbf{M})\mathbf{M}^{-1}) \mathbf{x}_i} \right) - \mathbf{M}.$$

Soit  $\mathbf{y}_i = \mathbf{M}^{-1/2} \mathbf{x}_i$  où les  $\mathbf{y}_i$  sont des vecteurs Gaussiens centrés i.i.d et de matrice de covariance l'identité. Ensuite,

$$\mathbf{M}^{-1/2}(\delta\mathbf{M})\mathbf{M}^{-1/2} \simeq \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H (\mathbf{I} - \mathbf{M}^{-1/2}(\delta\mathbf{M})\mathbf{M}^{-1/2}) \mathbf{y}_i} - \mathbf{I},$$

ou, de manière équivalente en utilisant l'expression  $\Delta = \mathbf{M}^{-1/2}(\delta\mathbf{M})\mathbf{M}^{-1/2}$ ,

$$\Delta \simeq \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i \left( 1 - \frac{\mathbf{y}_i^H \Delta \mathbf{y}_i}{\mathbf{y}_i^H \mathbf{y}_i} \right)} - \mathbf{I}.$$

Ainsi, pour  $N$  suffisamment grand, on obtient à l'ordre 1

$$\Delta \simeq \frac{m}{N} \sum_{i=1}^N \left( \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} \left( 1 + \frac{\mathbf{y}_i^H \Delta \mathbf{y}_i}{\mathbf{y}_i^H \mathbf{y}_i} \right) \right) - \mathbf{I}.$$

Pour trouver une expression explicite de  $\Delta$  en termes de données, l'expression précédente doit être réorganisée comme suit :

$$\Delta - \frac{m}{N} \sum_{i=1}^N \left( \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} \frac{\mathbf{y}_i^H \Delta \mathbf{y}_i}{\mathbf{y}_i^H \mathbf{y}_i} \right) \simeq \frac{m}{N} \sum_{i=1}^N \left( \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} \right) - \mathbf{I}.$$

Pour résoudre ce système de taille  $m^2$ , l'équation ci-dessus est réécrite selon :

$$\mathbf{B} \boldsymbol{\delta} \simeq \text{vec} \left( \frac{m}{N} \sum_{i=1}^N \left( \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} \right) - \mathbf{I} \right), \quad (3A.9)$$

où

- $\boldsymbol{\delta} = \text{vec}(\boldsymbol{\Delta})$ ,
- $\mathbf{B} = \mathbf{I} - \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{D}_i}{(\mathbf{y}_i \mathbf{y}_i^H)^2}$ ,
- $\mathbf{D}_i$  est la matrice  $m^2 \times m^2$  définie par  $\mathbf{D}_i = (d_{kl})_{1 \leq k, l \leq m^2}^{(i)}$  avec

$$d_{kl} = y_p \bar{y}_q \bar{y}_{p'} y_{q'}$$

$$\text{et } \begin{cases} k = p + m(q-1) & \text{avec } 1 \leq p, q \leq m \\ l = p' + m(q'-1) & \text{avec } 1 \leq p', q' \leq m \end{cases} .$$

Posons dans l'équation (3A.9)

$$\mathbf{A} = \text{vec} \left( \frac{m}{N} \sum_{i=1}^N \left( \frac{\mathbf{y}_i \mathbf{y}_i^H}{\mathbf{y}_i^H \mathbf{y}_i} \right) - \mathbf{I} \right) .$$

Ainsi, l'équation (3A.9) devient

$$\mathbf{B} \boldsymbol{\delta} \simeq \mathbf{A} . \quad (3A.10)$$

En appliquant le Théorème Central Limite (TCL), le membre de droite de (3A.10) vérifie

$$\sqrt{N} \begin{bmatrix} \text{Re}(\mathbf{A}) \\ \text{Im}(\mathbf{A}) \end{bmatrix} \xrightarrow[N \rightarrow +\infty]{\text{dist.}} \mathcal{N}(\mathbf{0}, \mathbf{G}) , \quad (3A.11)$$

où  $\mathbf{G}$  est la matrice de covariance de  $\begin{bmatrix} \text{Re}(\mathbf{A}) \\ \text{Im}(\mathbf{A}) \end{bmatrix}$ .

De plus,  $\mathbf{B}$ , dans le membre de gauche de (3A.10), par la LfGN, possède la propriété suivante :

$$\mathbf{B} \xrightarrow[N \rightarrow +\infty]{Pr} \mathbf{C}_2 = \mathbf{I} - m \mathbb{E} \left[ \frac{\mathbf{D}}{(\mathbf{y}^H \mathbf{y})^2} \right] , \quad (3A.12)$$

où  $\mathbf{y} = (y_1, \dots, y_m)^\top \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$  et où  $\mathbf{D}$  est la matrice  $m^2 \times m^2$  définie par

$$\mathbf{D} = (d_{kl})_{1 \leq k, l \leq m^2} , \text{ avec } d_{kl} = y_p \bar{y}_q \bar{y}_{p'} y_{q'} .$$

Donc, par des considérations classiques sur la convergence en probabilité, le premier point du théorème 3.2.3.1 est démontré.

Il reste maintenant à calculer l'expression analytique de  $\mathbf{C}_2$  défini par (3A.12) :

$$\mathbf{C}_2 = \mathbf{I} - m \mathbf{E} , \text{ où } \mathbf{E} = \mathbb{E} \left[ \frac{\mathbf{D}}{(\mathbf{y}^H \mathbf{y})^2} \right] .$$

En réécrivant les  $y_j$  comme  $y_j = \sqrt{\chi_j^2/2} \exp(i\theta_j)$ , où pour  $j = 1, \dots, m$ ,  $\chi_j^2$  et  $\theta_j$  sont des variables indépendantes, avec  $\chi_j^2 \sim \chi^2(2)$  et  $\theta_j \sim \mathcal{U}([0, 2\pi])$ , chaque élément de la matrice  $\mathbf{E}$  devient

$$E_{kl} = \mathbb{E} \left[ \frac{\sqrt{\chi_p^2 \chi_q^2 \chi_{p'}^2 \chi_{q'}^2}}{\left(\sum_{j=1}^m \chi_j^2\right)^2} \right] \mathbb{E} [\exp(i(\theta_p - \theta_q + \theta_{q'} - \theta_{p'}))].$$

Il suffit maintenant de remarquer que  $\mathbb{E} [\exp(i(\theta_p - \theta_q + \theta_{q'} - \theta_{p'}))] \neq 0$  si et seulement si

1.  $k = l = p + m(p - 1)$ ,
2.  $k = p + m(p - 1), l = p' + m(p' - 1)$  et  $p \neq p'$ , i.e.  $k \neq l$
3.  $k = p + m(q - 1), l = p + m(q - 1)$  et  $p \neq q$ .

Ainsi, les éléments non nuls de la matrice  $\mathbf{E}$  sont :

1.  $E_{p+m(p-1), p+m(p-1)} = \mathbb{E} \left[ \frac{(\chi_p^2)^2}{\left(\sum_{j=1}^m \chi_j^2\right)^2} \right] = \frac{2}{m(m+1)}$ ,
2.  $E_{p+m(p-1), p'+m(p'-1)} = \mathbb{E} \left[ \frac{\chi_p^2 \chi_{p'}^2}{\left(\sum_{j=1}^m \chi_j^2\right)^2} \right] = \frac{1}{m(m+1)}$ ,
3.  $E_{p+m(q-1), p+m(q-1)} = \mathbb{E} \left[ \frac{\chi_p^2 \chi_q^2}{\left(\sum_{j=1}^m \chi_j^2\right)^2} \right] = \frac{1}{m(m+1)}$ ,

et donc,  $\mathbf{C}_2$  est donné par (2.16) :

$$\mathbf{C}_2 = \frac{m}{m+1} \left( \mathbf{I} - \frac{1}{m} \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^\top \right). \quad (3A.13)$$

Pour achever la preuve du théorème 3.2.3.1, il faut calculer les deux quantités  $\mathbb{E} [\boldsymbol{\delta} \boldsymbol{\delta}^\top]$  et  $\mathbb{E} [\boldsymbol{\delta} \boldsymbol{\delta}^H]$ . Dans cette perspective, à partir de l'équation (3A.10), on a, pour  $N$  grand,

$$\begin{cases} \mathbf{C}_2 \mathbb{E} [\boldsymbol{\delta} \boldsymbol{\delta}^\top] \mathbf{C}_2 = \mathbb{E} [\mathbf{A} \mathbf{A}^\top] \\ \mathbf{C}_2 \mathbb{E} [\boldsymbol{\delta} \boldsymbol{\delta}^H] \mathbf{C}_2 = \mathbb{E} [\mathbf{A} \mathbf{A}^H] \end{cases}. \quad (3A.14)$$

Malheureusement,  $\mathbf{C}_2$ , défini par (3A.13) n'est pas une matrice inversible. Cependant, le lemme suivant apporte une solution adéquate.

#### Lemme 3A.3.0.1

$$\mathbf{C}_2 \boldsymbol{\delta} = \mathbf{A} \quad \text{est équivalent à} \quad \boldsymbol{\delta} = \frac{m+1}{m} \mathbf{A} \quad (3A.15)$$

#### Preuve 3A.3.0.3

En effet, de l'expression de  $\mathbf{C}_2$  (3A.13), il vient

$$\mathbf{C}_2 \boldsymbol{\delta} = \frac{m}{m+1} \left( \boldsymbol{\delta} - \frac{\text{Tr}(\boldsymbol{\Delta})}{m} \text{vec}(\mathbf{I}) \right),$$

où  $\delta = \text{vec}(\Delta)$  et  $\Delta = \mathbf{M}^{-1/2} \widehat{\mathbf{M}} \mathbf{M}^{-1/2} - \mathbf{I}$ .

Alors, la  $\widehat{\mathbf{M}}$ -normalisation ( $\text{Tr}(\mathbf{M}^{-1} \widehat{\mathbf{M}}) = m$ ) assure que  $\text{Tr}(\Delta) = 0$  et

$$\mathbf{C}_2 \delta = \frac{m}{m+1} \delta.$$

Ceci conclut la preuve du lemme.

En appliquant le lemme 3A.3.0.1, l'équation (3A.14) devient

$$\begin{cases} \mathbb{E} [\delta \delta^\top] &= \left( \frac{m+1}{m} \right)^2 \mathbb{E} [\mathbf{A} \mathbf{A}^\top] \\ \mathbb{E} [\delta \delta^H] &= \left( \frac{m+1}{m} \right)^2 \mathbb{E} [\mathbf{A} \mathbf{A}^H] \end{cases}.$$

Il reste donc à calculer  $\mathbb{E} [\mathbf{A} \mathbf{A}^\top]$  et  $\mathbb{E} [\mathbf{A} \mathbf{A}^H]$ . Ceci est fait dans la preuve du théorème 2.3.3.3 car  $\mathbf{A}$  est exactement  $\delta_{TML}$  dans le cas de  $\widehat{\mathbf{M}} = \widehat{\mathbf{M}}_{TML}$  et donc, on a

$$\begin{cases} N \mathbb{E} [\delta \delta^\top] &\xrightarrow{N \rightarrow +\infty} \left( \frac{m+1}{m} \right)^2 \mathbf{C}_1 \\ N \mathbb{E} [\delta \delta^H] &\xrightarrow{N \rightarrow +\infty} \left( \frac{m+1}{m} \right)^2 \mathbf{C}_2 \end{cases}.$$

avec

$$\mathbf{C}_1 = \frac{m}{m+1} \left( \mathbf{P} - \frac{1}{m} \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^\top \right),$$

où  $\mathbf{P}$  est la matrice  $m^2 \times m^2$  nilpotente définie par :

$$P_{ij} = 0 \text{ et pour } 1 \leq p, p' \leq m :$$

- $P_{p+m(p-1), p+m(p-1)} = 1$ ,
- $P_{p+m(p'-1), p'+m(p-1)} = 1$  ;

et

$$\mathbf{C}_2 = \frac{m}{m+1} \left( \mathbf{I} - \frac{1}{m} \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^\top \right).$$

### 3A.4 Preuve du théorème 3.3.2.2 : "Propriété CFAR-matrice de $\widehat{\Lambda}(\widehat{\mathbf{M}}_{FP})$ "

Il faut montrer que la distribution de  $\widehat{\Lambda}(\widehat{\mathbf{M}}_{FP,1})$  ne dépend pas de la matrice de covariance  $\mathbf{M}_1$ .

Considérons tout d'abord  $N$  données secondaires ayant pour matrice de covariance  $\mathbf{M}_1$ , i.e. pour  $i = 1, \dots, N$ ,

$$\mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{M}_1).$$

L'estimateur du point fixe de  $\mathbf{M}_1$  est défini par

$$\widehat{\mathbf{M}}_{FP,1} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{M}}_{FP,1}^{-1} \mathbf{x}_i},$$

et le détecteur est défini par

$$\widehat{\Lambda}(\widehat{\mathbf{M}}_{FP,1}) = \frac{|\mathbf{p}^H \widehat{\mathbf{M}}_{FP,1}^{-1} \mathbf{z}|^2}{(\mathbf{p}^H \widehat{\mathbf{M}}_{FP,1}^{-1} \mathbf{p})(\mathbf{z}^H \widehat{\mathbf{M}}_{FP,1}^{-1} \mathbf{z})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda,$$

où  $\mathbf{z}$  représente le vecteur d'observation, sous l'hypothèse  $H_0$ , et donc  $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \mathbf{M}_1)$ .

La première étape consiste à blanchir les données secondaires par le changement de variables  $\mathbf{y}_i = \mathbf{M}_1^{-1/2} \mathbf{x}_i$ . Alors,  $\mathbf{y}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ . Ce changement de variable est ensuite appliqué à l'estimateur du point fixe

$$\widehat{\mathbf{M}}_{FP,1} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{M}_1^{1/2} \mathbf{y}_i \mathbf{y}_i^H \mathbf{M}_1^{1/2}}{\mathbf{y}_i^H (\mathbf{M}_1^{-1/2} \widehat{\mathbf{M}}_{FP,1} \mathbf{M}_1^{-1/2})^{-1} \mathbf{y}_i}.$$

En posant maintenant

$$\widehat{\mathbf{T}} = \mathbf{M}_1^{-1/2} \widehat{\mathbf{M}}_{FP,1} \mathbf{M}_1^{-1/2},$$

on constate que  $\widehat{\mathbf{T}}$  est l'unique estimateur du point fixe (à un facteur près) dans le cas de données  $\mathbf{y}_i$  ayant comme matrice de covariance, l'identité  $\mathbf{I}$ .

On s'intéresse maintenant au détecteur, auquel on applique aussi le changement de variable,  $\mathbf{y} = \mathbf{M}_1^{-1/2} \mathbf{z}$ , ce qui donne

$$\widehat{\Lambda}(\widehat{\mathbf{M}}_{FP,1}) = \frac{|\mathbf{p}_1^H \widehat{\mathbf{T}}^{-1} \mathbf{y}|^2}{(\mathbf{p}_1^H \widehat{\mathbf{T}}^{-1} \mathbf{p}_1)(\mathbf{y}^H \widehat{\mathbf{T}}^{-1} \mathbf{y})} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda,$$

où  $\mathbf{p}_1 = \mathbf{M}_1^{-1/2} \mathbf{p}$ .

La distribution de  $\widehat{\Lambda}(\widehat{\mathbf{M}})$  ne dépend pas de  $\mathbf{p}$ , et ainsi, la loi de  $\widehat{\Lambda}(\widehat{\mathbf{M}}_{FP,1})$  est la même que la loi de  $\widehat{\Lambda}(\widehat{\mathbf{T}})$  où  $\widehat{\mathbf{T}}$  est l'estimateur du point fixe de l'identité. Ceci conclut la démonstration du théorème 3.3.2.2.



# Chapitre 4

## Bilan

*Ce chapitre dresse un bilan des résultats obtenus dans les chapitres précédents, concernant les différents estimateurs de matrice de covariance disponibles. Les résultats sont rappelés en deux parties : les propriétés statistiques et les propriétés "radaristiques".*

### 4.1 Les propriétés statistiques

Les résultats concernant l'étude statistique des estimateurs présentés dans ces travaux sont énoncés dans le tableau suivant :

	$\hat{\mathbf{M}}_{FP}$	$\hat{\mathbf{M}}_{TML}$	$\hat{\mathbf{M}}_N$	$\hat{\mathbf{M}}_{SCM}$	$\hat{\mathbf{M}}_{NSCM}$
Biais de $\hat{\mathbf{M}}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\neq \mathbf{0}$ (Thm 2.3.1.2)
$\hat{\mathbf{M}}$ est-il consistant ?	oui	oui	oui	oui	non
$E[\delta\delta^\top]$	$\left(\frac{m+1}{m}\right)^2 \mathbf{C}_1$	$\mathbf{C}_1$	$\frac{m+1}{m} \mathbf{C}_1$	$\mathbf{P}$	Voir [5]
$E[\delta\delta^H]$	$\left(\frac{m+1}{m}\right)^2 \mathbf{C}_2$	$\mathbf{C}_2$	$\frac{m+1}{m} \mathbf{C}_2$	$\mathbf{I}$	Voir [5]

TAB. 4.1 – Récapitulatif de l'étude statistique des estimateurs de la matrice de covariance  $\mathbf{M}$

où :

$$- \mathbf{C}_1 = \frac{m}{m+1} \left( \mathbf{P} - \frac{1}{m} \text{vec}(\mathbf{I})\text{vec}(\mathbf{I})^\top \right) ;$$

$$- \mathbf{C}_2 = \frac{m}{m+1} \left( \mathbf{I} - \frac{1}{m} \text{vec}(\mathbf{I})\text{vec}(\mathbf{I})^\top \right) ;$$

-  $\mathbf{P}$  est la matrice nilpotente de dimension  $m^2 \times m^2$  définie par :

$$P_{ij} = 0 \text{ et pour } 1 \leq p, p' \leq m :$$

- $P_{p+m(p-1), p+m(p-1)} = 1$ ,
- $P_{p+m(p'-1), p'+m(p-1)} = 1$ .

Quant à l'article [5], il est disponible à l'annexe C et contient l'ensemble des démonstrations du moment d'ordre 2 de  $\widehat{\mathbf{M}}_{NSCM}$ .

## 4.2 Les propriétés "radaristiques"

	$\widehat{\mathbf{M}}_{FP}$	$\widehat{\mathbf{M}}_{TML}$	$\widehat{\mathbf{M}}_N$	$\widehat{\mathbf{M}}_{SCM}$	$\widehat{\mathbf{M}}_{NSCM}$
Propriété CFAR-texture de $\widehat{\Lambda}(\widehat{\mathbf{M}})$	oui	oui	oui / non	oui / non	oui
Propriété CFAR-matrice de $\widehat{\Lambda}(\widehat{\mathbf{M}})$	oui	oui	oui	oui	non

TAB. 4.2 – Récapitulatif des propriétés du GLRT-LQ construits avec les différents estimateurs de la matrice de covariance  $\mathbf{M}$

### Remarque 4.2.0.2

Notons que  $\widehat{\Lambda}(\widehat{\mathbf{M}}_{SCM})$  et  $\widehat{\Lambda}(\widehat{\mathbf{M}}_{NSCM})$  possèdent la propriété CFAR-texture par construction, puisque leur expression ne dépend que des données Gaussiennes contenues dans les données secondaires. Ces données Gaussiennes ne peuvent cependant pas être observées en pratique.

En fait, si  $\widehat{\mathbf{M}}_{SCM}$  était calculée en pratique, ce qui est actuellement le cas dans la communauté radar, cet estimateur prendrait la forme suivante :

$$\widehat{\mathbf{M}}_{SCM} = \frac{1}{N} \sum_{i=1}^N \mathbf{c}_i \mathbf{c}_i^H = \frac{1}{N} \sum_{i=1}^N \tau_i \mathbf{x}_i \mathbf{x}_i^H.$$

Il résulte de cette équation que  $\widehat{\Lambda}(\widehat{\mathbf{M}}_{SCM})$  dépend évidemment de la texture  $\tau_i$ , et dans ces conditions, ne vérifie plus la propriété CFAR-texture. Ainsi, dans l'exploitation des données réelles faite au chapitre suivant,  $\widehat{\Lambda}(\widehat{\mathbf{M}}_{SCM})$  ne sera plus CFAR-texture.

Ce constat s'applique évidemment de la même manière à  $\widehat{\Lambda}(\widehat{\mathbf{M}}_N)$ .

Le GLRT-LQ, construit avec n'importe quel estimateur (parmi ceux étudiés) possèdent les 2 propriétés CFAR (texture et matrice), on dira dans ce cas que  $\widehat{\Lambda}(\widehat{\mathbf{M}})$  est CFAR-SIRV, car il est indépendant des deux paramètres caractérisant le SIRV.

### 4.3 Décision : quel estimateur retenir ?

Parmi les cinq estimateurs étudiés, trois sont utilisables en pratique :  $\hat{\mathbf{M}}_N$ ,  $\hat{\mathbf{M}}_{NSCM}$  et  $\hat{\mathbf{M}}_{FP}$ . En prenant en compte les propriétés de chacun, il vient tout naturellement que le Point Fixe est l'estimateur le plus judicieux pour approcher la matrice de covariance  $\mathbf{M}$ , puisque  $\hat{\Lambda}(\hat{\mathbf{M}}_{NSCM})$  n'est pas CFAR-matrice et  $\hat{\Lambda}(\hat{\mathbf{M}}_N)$  ne peut être CFAR-texture en pratique.

Le chapitre suivant est consacré aux applications faites dans le cadre de la détection radar. Cependant, l'estimateur du point fixe pourrait être appliqués dans d'autres domaines...



## Chapitre 5

# Application à la Détection Radar

*Ce chapitre traite de l'application des résultats obtenus selon les différents estimateurs de matrice de covariance injectés dans le détecteur GLRT-LQ. Les données utilisées sont de deux natures différentes : tout d'abord, des jeux de données simulées sont analysés et enfin, l'exploitation de données réelles fournies par THALES Air Defence<sup>1</sup> permet la validation des résultats théoriques énoncés dans les chapitres précédents.*

*Le chapitre est organisé de la façon suivante : tout d'abord, l'analyse de l'étape 1 de détection, i.e. la régulation des fausses alarmes, puis l'étude de la robustesse du détecteur dans des zones de transition de fouillis.*

### 5.1 Régulation de la fausse alarme

Comme il a été rappelé dans le premier chapitre de ce manuscrit, la détection radar de cibles se décompose en deux étapes distinctes : tout d'abord, il faut déterminer un seuil de détection pour une probabilité de fausse alarme fixée par l'utilisateur, autrement dit, il faut calculer  $\lambda$  tel que

$$P_{fa} = \mathbb{P}(\hat{\Lambda}(\hat{\mathbf{M}}; H_0) > \lambda) = \alpha, \quad (5.1)$$

où  $\hat{\Lambda}(\hat{\mathbf{M}}; H_0)$  désigne la version adaptative du détecteur GLRT-LQ, construit avec un estimateur  $\hat{\mathbf{M}}$  de la matrice de covariance, et sous l'hypothèse  $H_0$ , i.e. l'observation ne contient que le fouillis.

Une fois cette régulation des fausses alarmes effectuée, la détection consiste à comparer, pour chaque observation  $\mathbf{y}$ , le détecteur au seuil de détection  $\lambda$  calculé précédemment.

#### 5.1.1 Calcul du seuil de détection $\lambda$ pour une $P_{fa}$ donnée

Dans la section 1.1, l'équation (1.11) fournit une relation analytique entre le seuil de détection et la  $P_{fa}$  dans le cas du détecteur optimal Gaussien et pour une matrice de covariance  $\mathbf{M}$  supposée connue.

Ensuite, la section 1.3 introduit, par l'intermédiaire de l'équation (1.22), une autre relation " $P_{fa}$ -seuil" analytique, adaptée cette fois au détecteur GLRT-LQ mais encore sous l'hypothèse de matrice de covariance  $\mathbf{M}$  connue.

---

<sup>1</sup>Les auteurs sont reconnaissants envers THALES Air Defence pour l'exploitation de leurs données

Mais qu'en est-il de ce détecteur quand  $\mathbf{M}$  est inconnue et estimée par un  $\widehat{\mathbf{M}}$  donné ? Ceci fait l'objet du prochain paragraphe.

### 5.1.1.1 Cas de d'une matrice de covariance estimée par $\widehat{\mathbf{M}}_N$

Une première approche consiste à utiliser l'estimateur  $\widehat{\mathbf{M}}_N$  défini par (2.5) :

$$\widehat{\mathbf{M}}_N = \frac{m}{N} \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^H.$$

où les  $\mathbf{x}_k$  sont des vecteurs i.i.d centrés, Gaussien et de matrice de covariance  $\mathbf{M}$ .

Comme  $\frac{m}{N} \sum_{k=1}^N \mathbf{x}_k^H \mathbf{M}^{-1} \mathbf{x}_k$  est un scalaire, le détecteur construit avec  $\widehat{\mathbf{M}}_N$  peut s'écrire de la façon suivante :

$$\widehat{\Lambda}(\widehat{\mathbf{M}}_N) = \frac{|\mathbf{p}^H \widehat{\mathbf{M}}_N^{-1} \mathbf{y}|^2}{(\mathbf{p}^H \widehat{\mathbf{M}}_N^{-1} \mathbf{p})(\mathbf{y}^H \widehat{\mathbf{M}}_N^{-1} \mathbf{y})} = \frac{|\mathbf{p}^H \widehat{\mathbf{M}}_W^{-1} \mathbf{y}|^2}{(\mathbf{p}^H \widehat{\mathbf{M}}_W^{-1} \mathbf{p})(\mathbf{y}^H \widehat{\mathbf{M}}_W^{-1} \mathbf{y})} = \widehat{\Lambda}(\widehat{\mathbf{N}}),$$

avec  $\widehat{\mathbf{M}}_W = \sum_{k=1}^N \mathbf{x}_k \mathbf{x}_k^H$ , et donc,  $\widehat{\mathbf{M}}_W \sim \mathcal{CW}(N, m; \mathbf{M})$ , i.e.  $\widehat{\mathbf{M}}_W$  est distribuée selon la loi de Wishart complexe avec les paramètres  $N$  (nombre de  $\mathbf{x}_k$ ),  $m$  (taille de chaque vecteur) et  $\mathbf{M}$  (matrice de covariance des  $\mathbf{x}_k$ ).

#### **Théorème 5.1.1.1 (Loi de $\widehat{\Lambda}(\widehat{\mathbf{M}}_N)$ quand $\mathbf{M}$ est estimée par $\widehat{\mathbf{M}}_N$ )**

La version adaptative du GLRT-LQ,  $\widehat{\Lambda}(\widehat{\mathbf{M}}_N)$ , est distribuée selon la densité de probabilité  $g_{N,m}$  définie par :

$$g_{N,m}(x) = \frac{(N-m+1)(m-1)}{(N-1)} \frac{{}_2F_1(a, a; b; x)}{(1-x)^{N-m}} \mathbb{1}_{[0,1]}(x) \quad (5.2)$$

avec  $a = N - m + 2$ ,  $b = N + 2$  et  ${}_2F_1$  la fonction hypergéométrique [1] définie par :

$${}_2F_1(a, b; c; x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b+k)}{\Gamma(c+k)} \frac{x^k}{k!}.$$

#### **Preuve 5.1.1.1**

La densité de  $\widehat{\mathbf{M}}_W$  est alors donnée par :

$$f(\mathbf{A}) = \frac{\det(\mathbf{A})^{N-m}}{J(\mathbf{M})} \exp(-\text{Tr}(\mathbf{M}^{-1}\mathbf{A}))$$

où  $J(\mathbf{M})$  s'écrit de la manière suivante :

$$J(\mathbf{M}) = \pi^{\frac{1}{2} m(m-1)} \Gamma(N) \Gamma(N-1) \dots \Gamma(N-m+1) \det(\mathbf{M})^N.$$

L'objectif est maintenant d'obtenir la loi de  $\widehat{\Lambda}(\widehat{\mathbf{M}}_N)$ , et par conséquent, la relation " $P_{fa}$ -seuil" qui n'est rien d'autre que la fonction de répartition de  $\widehat{\Lambda}(\widehat{\mathbf{M}}_N)$ .

Le calcul de cette loi est issu de [36] et est basé sur la théorie de Wishart. Kraut, dans [36], réécrit  $\widehat{\Lambda}(\widehat{\mathbf{M}}_N)$  à l'aide de la variable aléatoire  $\widehat{F}$  :

$$\widehat{\Lambda}(\widehat{\mathbf{M}}_N) = \frac{\widehat{F}}{\widehat{F} + 1},$$

où  $\widehat{F}$  dépend d'une autre variable  $B$ . Conditionnellement à  $B$ ,  $\widehat{F}$  est distribuée selon une loi Beta de 2<sup>ème</sup> espèce, de paramètres 1 et  $N - m + 1$ , i.e.  $\widehat{F}|B \sim \beta_{1, N-m+1}^2$  et  $B$  suit une loi Beta de première espèce, de paramètres  $N - m + 2$  et  $m - 1$ , i.e.  $B \sim \beta_{N-m+2, m-1}^1$ .

Les densités de variables de loi  $\beta_{a,b}^1$  et  $\beta_{a,b}^2$  sont définies dans [1] par :

$$\beta_{a,b}^1(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \mathbb{1}_{[0,1]}(x),$$

et

$$\beta_{a,b}^2(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{x^{a-1}}{(1+x)^{a+b}}, \quad x > 0$$

où  $\mathbb{1}_{[0,1]}(x)$  est la fonction indicatrice sur l'intervalle  $[0, 1]$ , définie par (3.10).

Puis, après quelques manipulations basiques de probabilités, on obtient le théorème 5.1.1.1.

#### **Théorème 5.1.1.2 (Relation " $P_{fa}$ -seuil" quand $\mathbf{M}$ est estimée par $\widehat{\mathbf{M}}_N$ )**

La relation entre la  $P_{fa}$  et le seuil de détection  $\eta$  (ou  $\lambda$ , lié à  $\eta$  par l'équation (1.19)) est donnée par :

$$P_{fa} = \eta^{-\frac{a-1}{m}} {}_2F_1\left(a, a-1; b-1; 1 - \eta^{-\frac{1}{m}}\right), \quad (5.3)$$

$$= (1 - \lambda)^{a-1} {}_2F_1(a, a-1; b-1; \lambda). \quad (5.4)$$

$$(5.5)$$

avec  $a = N - m + 2$  et  $b = N + 2$ .

#### **Preuve 5.1.1.2**

Pour obtenir cette relation, il suffit d'intégrer la densité de  $\widehat{\Lambda}(\widehat{\mathbf{M}}_N)$  donnée par l'équation (5.2) :

$$P_{fa} = \int_{\eta}^{+\infty} g_{N,m}(x) dx.$$

### 5.1.1.2 Cas de d'une matrice de covariance estimée par $\widehat{\mathbf{M}}_{FP}$

Ce paragraphe présente un résultat essentiel de cette thèse. En effet, les chapitres précédents présentent l'étude de différents estimateurs de la matrice de covariance  $\mathbf{M}$  des données secondaires, et celui qui répond le mieux au problème d'un point de vue théorique, comme pratique, est l'estimateur développé dans cette thèse : le point fixe  $\widehat{\mathbf{M}}_{FP}$ .

Il était donc indispensable de pouvoir obtenir une relation analytique entre la  $P_{fa}$  et le seuil de détection quand  $\mathbf{M}$  était estimé par  $\widehat{\mathbf{M}}_{FP}$ . Cependant, il n'existe aucune forme explicite pour  $\widehat{\mathbf{M}}_{FP}$ . Il faut donc utiliser ses propriétés statistiques.

Le chapitre 4 présente une synthèse des résultats obtenus pour tous les estimateurs et il est intéressant de remarquer que  $\widehat{\mathbf{M}}_N$  et  $\widehat{\mathbf{M}}_{FP}$  ont la même distribution asymptotique Gaussienne, mais qui diffèrent uniquement par leur moment d'ordre 2 :

$$\begin{aligned} - N \mathbb{E} [\boldsymbol{\delta}_N \boldsymbol{\delta}_N^\top] &\xrightarrow{N \rightarrow +\infty} \frac{m+1}{m} \mathbf{C}_1 \text{ tandis que } N \mathbb{E} [\boldsymbol{\delta}_{FP} \boldsymbol{\delta}_{FP}^\top] \xrightarrow{N \rightarrow +\infty} \left(\frac{m+1}{m}\right)^2 \mathbf{C}_1, \\ - N \mathbb{E} [\boldsymbol{\delta}_N \boldsymbol{\delta}_N^H] &\xrightarrow{N \rightarrow +\infty} \frac{m+1}{m} \mathbf{C}_2 \text{ tandis que } N \mathbb{E} [\boldsymbol{\delta}_N \boldsymbol{\delta}_N^H] \xrightarrow{N \rightarrow +\infty} \left(\frac{m+1}{m}\right)^2 \mathbf{C}_2. \end{aligned}$$

On peut donc en déduire le théorème suivant, valide pour  $N$  suffisamment grand :

#### **Théorème 5.1.1.3 (Relation " $P_{fa}$ -seuil" quand $\mathbf{M}$ est estimée par $\widehat{\mathbf{M}}_{FP}$ )**

Pour  $N$  suffisamment grand, la relation entre la  $P_{fa}$  et le seuil de détection  $\eta$  (ou  $\lambda$ , lié à  $\eta$  par l'équation (1.19)) est donnée par :

$$P_{fa} = \eta^{-\frac{a-1}{m}} {}_2F_1\left(a, a-1; b-1; 1 - \eta^{-\frac{1}{m}}\right), \quad (5.6)$$

$$= (1 - \lambda)^{a-1} {}_2F_1(a, a-1; b-1; \lambda). \quad (5.7)$$

$$(5.8)$$

$$\text{avec } a = \frac{m}{m+1}N - m + 2 \text{ et } b = \frac{m}{m+1}N + 2.$$

#### **Preuve 5.1.1.3**

Pour  $N$  suffisamment grand,  $\widehat{\mathbf{M}}_{FP}$  se comporte comme une matrice de Wishart mais avec un nombre de degrés de liberté de  $\frac{m}{m+1}N$ .

#### **Remarque 5.1.1.1**

– En fait, cette relation est la même que dans le cas d'une matrice de Wishart mais avec moins de données secondaires,  $\frac{m}{m+1}N$  au lieu de  $N$ . Autrement dit,  $\widehat{\mathbf{M}}_{FP}$  se comporte asymptotiquement comme

$$\frac{m \sum_{k=1}^{\frac{m}{m+1}N} \mathbf{x}_k \mathbf{x}_k^H}{\sum_{k=1}^{\frac{m}{m+1}N} \mathbf{x}_k^H \mathbf{M}^{-1} \mathbf{x}_k}.$$

Ce résultat est assez théorique puisque la matrice précédente ne peut être construite, car elle utilise  $\frac{m}{m+1}N$  données secondaires, ce qui n'est pas forcément un nombre entier...

- Il est évident que la loi de  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$  peut être calculée de la même manière à partir du théorème 5.1.1.1.
- Cette relation est encore valide pour  $\hat{\mathbf{M}}_{TML}$  mais pour un nombre de données secondaires plus grand  $\frac{m+1}{m}N$  (voir tableau 4.1).
- Enfin, il est intéressant de noter que ces relations " $P_{fa}$ -seuil" ne dépendent que du nombre  $N$  de données secondaires et de la taille  $m$  des vecteurs  $\mathbf{x}_k$ . Ceci corrobore les propriétés CFAR de ce détecteur construit avec ces trois estimateurs...

## 5.1.2 Simulations

Afin d'illustrer les précédents résultats, plusieurs simulations sont présentées. Sur chaque figure, la  $P_{fa}$  (axe des ordonnées) est tracée en fonction du seuil de détection (axe des abscisses), en échelle logarithmique, selon différents jeux de paramètres :  $N$ ,  $m$ ,  $\mathbf{M}$  et pour différents estimateurs  $\hat{\mathbf{M}}$ .

Analyse de la figure (Fig. 5.1) :

Le cas d'un estimateur  $\hat{\mathbf{M}}_N$  de la matrice de covariance est traité. La figure (Fig. 5.1) présente des courbes " $P_{fa}$ -seuil" théoriques, i.e. la relation analytique établie par le théorème 5.1.1.2, pour différentes valeurs de  $N$  :  $N = 20$ ,  $N = 50$ ,  $N = 100$  et  $N = 5000$ . Dans la formule donnée par l'équation (5.3), le seul autre paramètre intervenant est  $m$ , ici il est fixé à  $m = 10$ .

La figure (Fig. 5.1) illustre la convergence (attendue) théorique de la relation " $P_{fa}$ -seuil" quand  $\mathbf{M}$  est estimée par  $\hat{\mathbf{M}}_N$ , vers la relation " $P_{fa}$ -seuil" quand  $\mathbf{M}$  est connue, pour  $N$  tendant vers l'infini. Ceci se formalise de la façon suivante :

$$\eta^{-\frac{a-1}{m}} {}_2F_1\left(a, a-1; b-1; 1 - \eta^{-\frac{1}{m}}\right) \xrightarrow{N \rightarrow +\infty} \eta^{\frac{1-m}{m}},$$

$$\text{avec } a = \frac{m}{m+1}N - m + 2 \text{ et } b = \frac{m}{m+1}N + 2.$$

En effet, la courbe noire (o) représente le tracé de la relation " $P_{fa}$ -seuil" quand  $\mathbf{M}$  est connue :

$$P_{fa} = \eta^{\frac{1-m}{m}},$$

tandis que les courbes de couleur représente la relation " $P_{fa}$ -seuil" quand  $\mathbf{M}$  est estimée par  $\hat{\mathbf{M}}_N$ , pour différentes valeurs de  $N$  :

$$P_{fa} = \eta^{-\frac{a-1}{m}} {}_2F_1\left(a, a-1; b-1; 1 - \eta^{-\frac{1}{m}}\right),$$

avec  $a = N - m + 2$ ,  $b = N + 2$  et

- $N = 20$ , courbe rouge,
- $N = 50$ , courbe bleue,
- $N = 100$ , courbe verte,
- $N = 5000$ , courbe de rose.

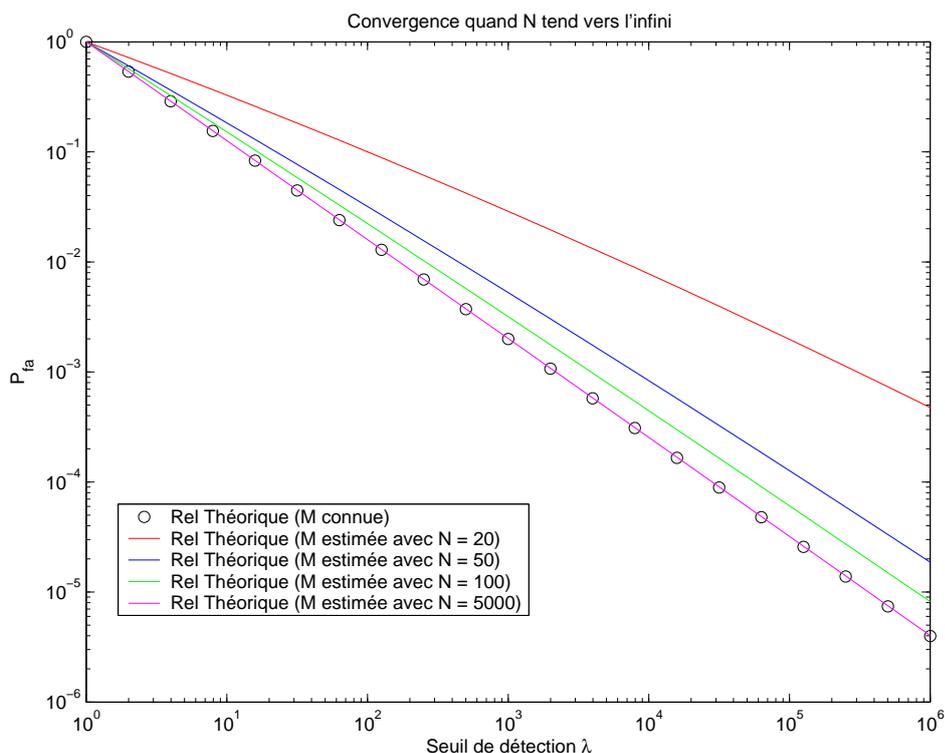


FIG. 5.1 – Illustration de la convergence théorique de l'équation " $P_{fa}$ -seuil" (5.3) (i.e. quand  $\mathbf{M}$  est estimée par  $\hat{\mathbf{M}}_N$ ), vers l'équation " $P_{fa}$ -seuil" (1.21) (i.e. quand  $\mathbf{M}$  est connue), quand  $N$  tend vers l'infini

La conclusion est très nette : plus  $N$  augmente, plus les courbes de couleur se rapprochent de la courbe noire jusqu'à être parfaitement confondues pour  $N = 5000$ .

Ce résultat est assez intuitif car  $N$  représente le nombre de données utilisées pour l'estimation de  $\mathbf{M}$ , et quand  $N$  tend vers l'infini, l'estimateur  $\hat{\mathbf{M}}_N$  tend vers la vraie matrice de covariance  $\mathbf{M}$ . Il est, par conséquent, naturel que ce comportement s'applique aussi aux relations " $P_{fa}$ -seuil" correspondantes.

Analyse de la figure (Fig. 5.2) :

Sur la figure (Fig. 5.1), les valeurs extrêmes du seuil de détection sont aux alentours de  $10^6$  pour assurer une  $P_{fa}$  d'environ  $10^{-4}$ . Ces valeurs correspondent bien à la réalité du radar, les utilisateurs ne réglant pratiquement jamais des  $P_{fa}$  inférieures à  $10^{-6}$ .

Mais, sur le plan purement théorique, qu'en est-il des valeurs extrêmement petites de la  $P_{fa}$  et donc des valeurs extrêmement grandes du seuil de détection  $\eta$ ? Ceci fait l'objet de la figure (Fig. 5.2), sur

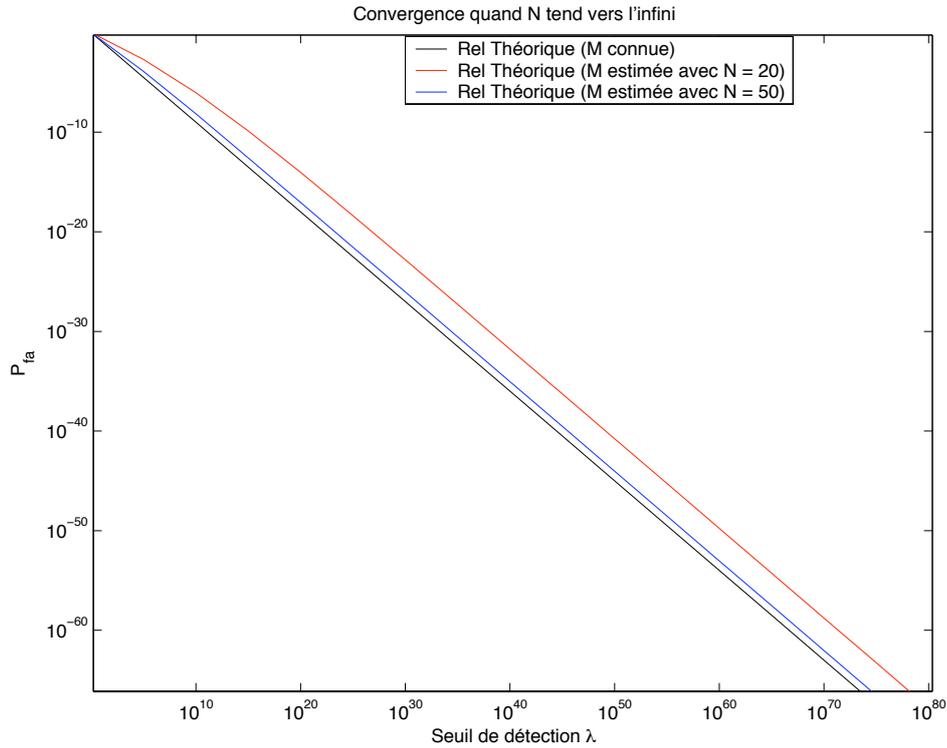


FIG. 5.2 – Comportement des courbes " $P_{fa}$ -seuil" pour des valeurs extrêmes de  $P_{fa}$  et de seuils

laquelle ont été tracées uniquement les courbes correspondant à  $N = 20$  et  $N = 50$  afin de se dégager des problèmes de calculs de machine.

Le constat est relativement surprenant, les courbes deviennent linéaires (en échelle logarithmique) pour des valeurs très grandes du seuil  $\eta$  de détection, proche de  $10^{80}$  alors qu'une courbure nette de ses courbes intervient dans la zone des valeurs qui intéressent les radaristes, i.e. pour des  $P_{fa}$  appartenant à  $[1, 10^{-6}]$ . Une étude plus approfondie permettrait peut-être de comprendre ce comportement "asymptotique"...

Analyse de la figure (Fig. 5.3) :

La figure (Fig. 5.3) compare les relations " $P_{fa}$ -seuil" théoriques (pour  $N = 20$  et  $N = 100$ ) avec des simulations Monte-Carlo. Les données simulées sont des vecteurs Gaussiens centrés, de dimension  $m = 10$  et de matrice de covariance  $\mathbf{M} = \mathbf{I}$ . Le choix de la matrice identité se justifie à l'aide des résultats synthétisés par le chapitre 4 : la distribution du GLRT-LQ construit avec  $\hat{\mathbf{M}}_N$  ne dépendant pas de  $\mathbf{M}$ , il est donc tout naturel de choisir comme matrice de covariance, la matrice la plus simple qui soit, i.e. l'identité, donc  $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ .

Sur la figure (Fig. 5.3), les relations théoriques sont les courbes bleue ( $N = 20$ ) et verte ( $N = 100$ ) tandis que les deux courbes rouges correspondent aux simulations Monte-Carlo. Ces derniers valident parfaitement à la théorie. En effet, les courbes rouges sont confondues respectivement à la courbe bleue, et à la courbe verte selon le nombre  $N$  de  $\mathbf{x}_k$ 's utilisés pour calculer  $\hat{\mathbf{M}}_N$ .

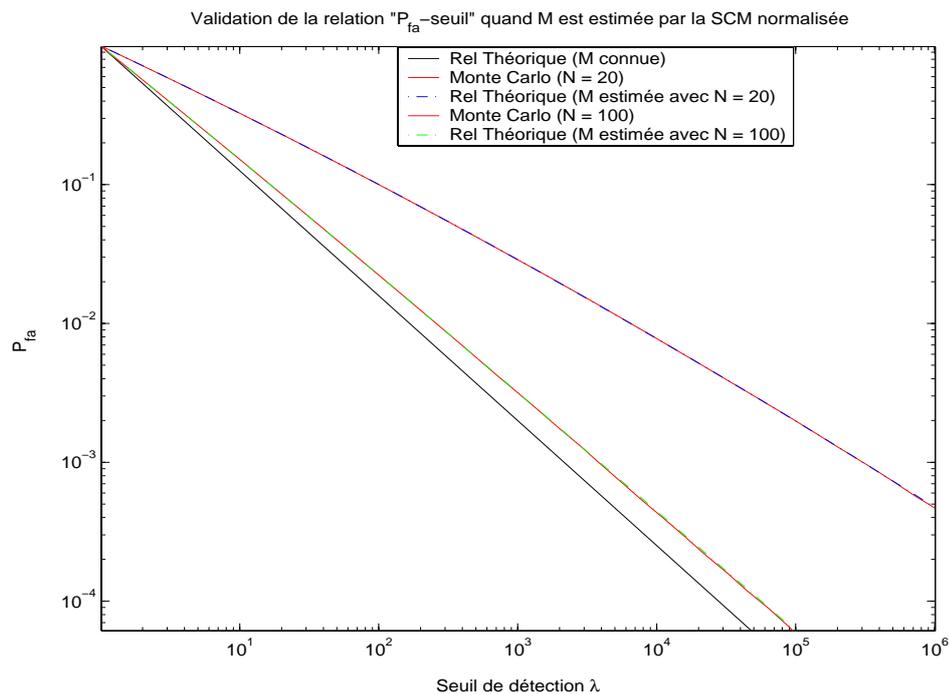


FIG. 5.3 – Validation Monte-Carlo de l'équation (5.3)

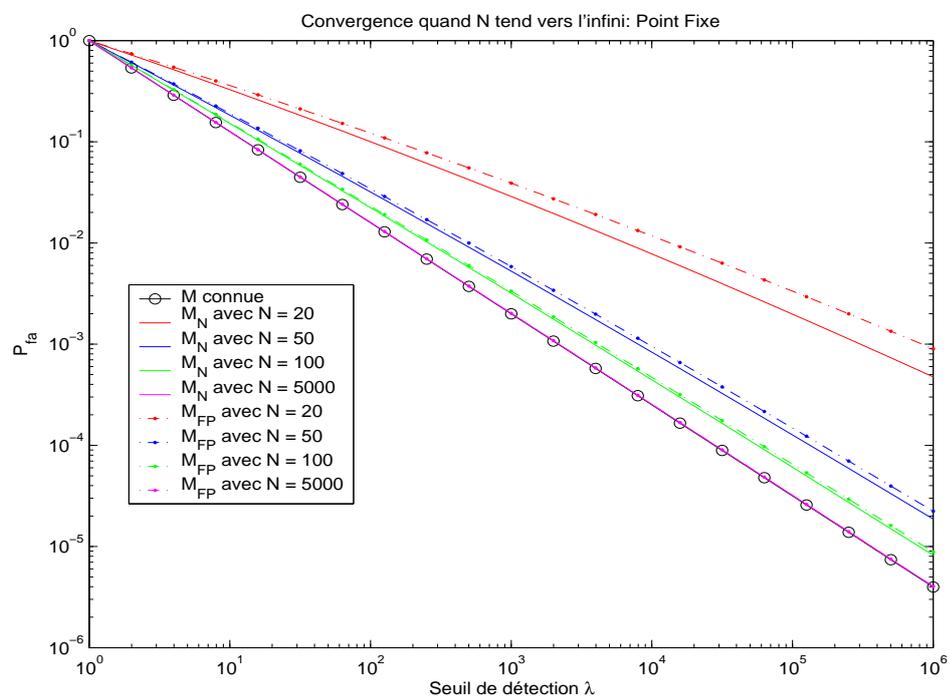


FIG. 5.4 – Illustration de la convergence théorique de l'équation " $P_{fa}$ -seuil" (5.6) (i.e. quand  $M$  est estimée par  $\hat{M}_{FP}$ ), vers l'équation " $P_{fa}$ -seuil" (1.21) (i.e. quand  $M$  est connue) quand  $N$  tend vers l'infini

Analyse de la figure (Fig. 5.4) :

Considérons maintenant le cas où la matrice de covariance est estimée par l'estimateur du point fixe  $\widehat{\mathbf{M}}_{FP}$ , traité par le théorème 5.1.1.3. Attention, ce théorème n'est valable que pour  $N$  suffisamment grand...

Tout d'abord, comme précédemment, la figure (Fig. 5.4), théorique, illustre la convergence de la relation " $P_{fa}$ -seuil" quand  $\mathbf{M}$  est estimée par  $\widehat{\mathbf{M}}_{FP}$  vers la relation " $P_{fa}$ -seuil" obtenue pour  $\mathbf{M}$  connue.

Cette figure reprend la figure (Fig. 5.1), à laquelle ont été rajoutées les courbes " $P_{fa}$ -seuil" obtenues pour  $\widehat{\mathbf{M}}_{FP}$  (même couleur que pour  $\widehat{\mathbf{M}}_N$  mais en pointillé) et valables pour  $N$  suffisamment grand. La première remarque qui s'impose est la convergence de ces courbes vers la courbe " $P_{fa}$ -seuil" pour  $\mathbf{M}$  connue (courbe noire). De plus, on constate que quand  $N$  augmente, l'écart entre les courbes associées à  $\widehat{\mathbf{M}}_{FP}$  et celles associées à  $\widehat{\mathbf{M}}_N$  se ressert, ce qui confirme la validité du théorème 5.1.1.3 quand  $N$  est grand. Cependant, il est intéressant de se demander quel est le comportement du GLRT-LQ construit avec le  $\widehat{\mathbf{M}}_{FP}$  quand  $N = 20$  ou  $50$ , i.e. des valeurs faibles.

Analyse de la figure (Fig. 5.5) :

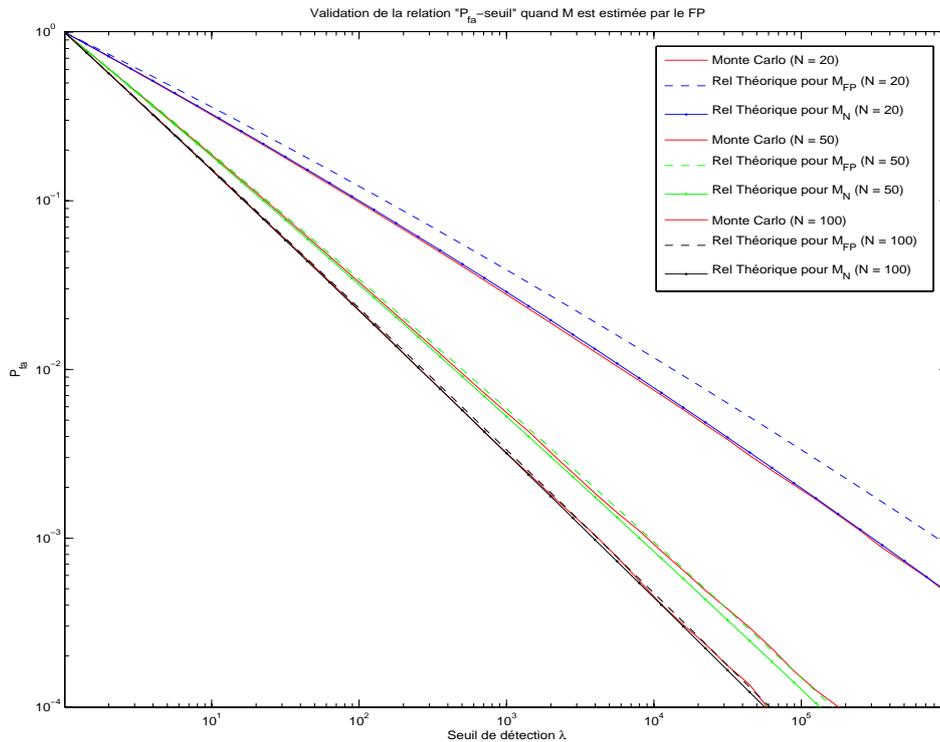


FIG. 5.5 – Relation " $P_{fa}$ -seuil" obtenue pour l'estimateur du point fixe par simulations Monte-Carlo, pour  $m = 10$  et pour des faibles valeurs de  $N$  :  $N = 20$ ,  $N = 50$  et  $N = 100$

- Sur la figure (Fig. 5.5), trois courbes sont tracées pour différentes valeurs de  $N$ , 20, 50 et 100 :
- la relation " $P_{fa}$ -seuil" théorique obtenue quand la matrice de covariance est estimée par  $\hat{\mathbf{M}}_N$ , et donnée par l'équation (5.3), courbe pleine de couleur bleue, verte ou magenta,
  - la relation " $P_{fa}$ -seuil" théorique obtenue quand la matrice de covariance est estimée par  $\hat{\mathbf{M}}_{FP}$ , et donnée par l'équation (5.6), valide seulement pour  $N$  suffisamment grand, courbe en pointillés de couleur bleue, verte ou magenta,
  - les simulations Monte-Carlo quand  $\mathbf{M}$  est estimée par  $\hat{\mathbf{M}}_{FP}$ .

Plusieurs remarques peuvent être faites à partir de cette figure :

- Pour la plus petite valeur de  $N$ , la simulation Monte-Carlo réalisée pour l'estimateur du point fixe, correspond parfaitement avec la courbe pleine, i.e. la relation théorique calculée pour  $\hat{\mathbf{M}}_N$  et non pour  $\hat{\mathbf{M}}_{FP}$ . Ceci n'est ni en accord ni en contradiction avec les précédents résultats puisqu'aucune information sur la loi de  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$  n'est disponible quand  $N$  est petit. Cependant, une conjecture peut être proposée : " $\hat{\mathbf{M}}_{FP}$  a une distribution statistique très proche de celle de  $\hat{\mathbf{M}}_N$ , et donc, puisque les facteurs de normalisation peuvent être négligés, très proche de la distribution de Wishart..."

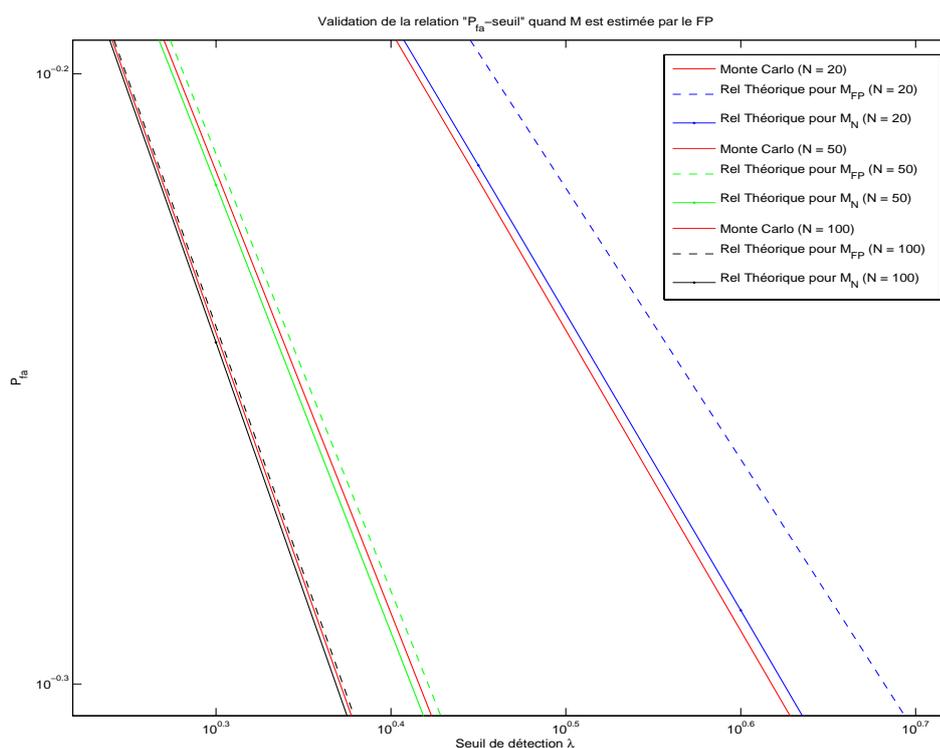


FIG. 5.6 – Zoom de la figure (Fig. 5.5)

Cette conjecture, si elle s'avérait exacte, aurait une importance capitale car la distribution statistique de  $\hat{\mathbf{M}}_{FP}$  pour tout  $N$  permettrait de caractériser complètement cet estimateur. Il ne faut pourtant pas perdre de vue que les relations " $P_{fa}$ -seuil" ne dépendent pas uniquement de l'estimateur de  $\mathbf{M}$  mais aussi du détecteur, puisque ce sont les fonctions de répartition de  $\hat{\Lambda}_{\hat{\mathbf{M}}}$ . Il est donc

tout à fait possible que les distributions statistiques de  $\widehat{\mathbf{M}}_N$  et  $\widehat{\mathbf{M}}_{FP}$  soient différentes et que les relations " $P_{fa}$ -seuil" soient les mêmes.

A partir de cette argumentation, la section suivante a été rajoutée à ce chapitre et contient un début d'analyse de cette conjecture...

- On constate que plus  $N$  augmente, plus la courbe Monte-Carlo se rapproche de la courbe théorique du point fixe. Ceci est en parfait accord avec le théorème 5.1.1.3. De plus, les courbes Monte-Carlo se situent toujours à l'intérieur du cône formé par les deux courbes théoriques, ceci est illustré par la figure (Fig. 5.6), qui est un zoom de la figure précédente pour des  $P_{fa}$  comprises entre  $10^{-0.2}$  et  $10^{-0.3}$ . Ce résultat est important puisqu'il complète le théorème 5.1.1.3 pour des valeurs plus faibles de  $N$  et permet d'ajuster de façon assez précise le seuil de détection pour un taux de fausses alarmes donné.

### 5.1.3 Conjecture de la distribution statistique de $\widehat{\mathbf{M}}_{FP}$

Le fait que la distribution statistique de  $\widehat{\mathbf{M}}_{FP}$  soit très proche de celle de  $\widehat{\mathbf{M}}_N$  a été suggéré par les simulations précédentes. Afin d'explorer cette éventualité, ce paragraphe présente une analyse par simulations des comportements statistiques des éléments des deux matrices pour une faible valeur de  $N$ ,  $N = 20$ , et pour des tailles de vecteurs  $m = 10$ .

Les matrices comparées  $\widehat{\mathbf{M}}_W$  et  $\widehat{\mathbf{M}}_{FP bis}$  sont définies de la manière suivante :

$$\widehat{\mathbf{M}}_W = \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^H, \quad (5.9)$$

et

$$\widehat{\mathbf{M}}_{FP bis} = m \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{M}}_{FP bis}^{-1} \mathbf{x}_i}. \quad (5.10)$$

avec  $\mathbf{x}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{M})$ .

Pourquoi ces matrices sont-elles légèrement différentes des précédentes ? En fait,  $\widehat{\mathbf{M}}_W$  reprend la définition exacte d'une matrice de Wishart présentée au début de ce chapitre,  $\widehat{\mathbf{M}}_W \sim \mathcal{W}(N, m; \mathbf{M})$ , et on conjecture que l'estimateur du point fixe est distribué selon une loi de Wishart, il faut évidemment le comparer à  $\widehat{\mathbf{M}}_W$ , sans aucun terme de normalisation.

Une première approche (assez naïve) est d'étudier les différents éléments caractérisant les deux matrices. Ils se décomposent en trois grands groupes :

- les valeurs propres de  $\widehat{\mathbf{M}}_{FP bis}$  et de  $\widehat{\mathbf{M}}_W$ ,
- les éléments diagonaux de chaque matrice,
- les éléments triangulaires supérieurs (car les matrices sont symétriques) sans ceux de la diagonale de chaque matrice.

Les éléments étudiés, notés  $e_W$  et  $e_{FP}$  selon la matrice qu'il représentent, ont été centrés et réduits de manière à comparer leur distribution sans prendre en compte les moments d'ordre 1 et 2.

Analyse des figures (Fig. 5.7 et Fig. 5.8) :

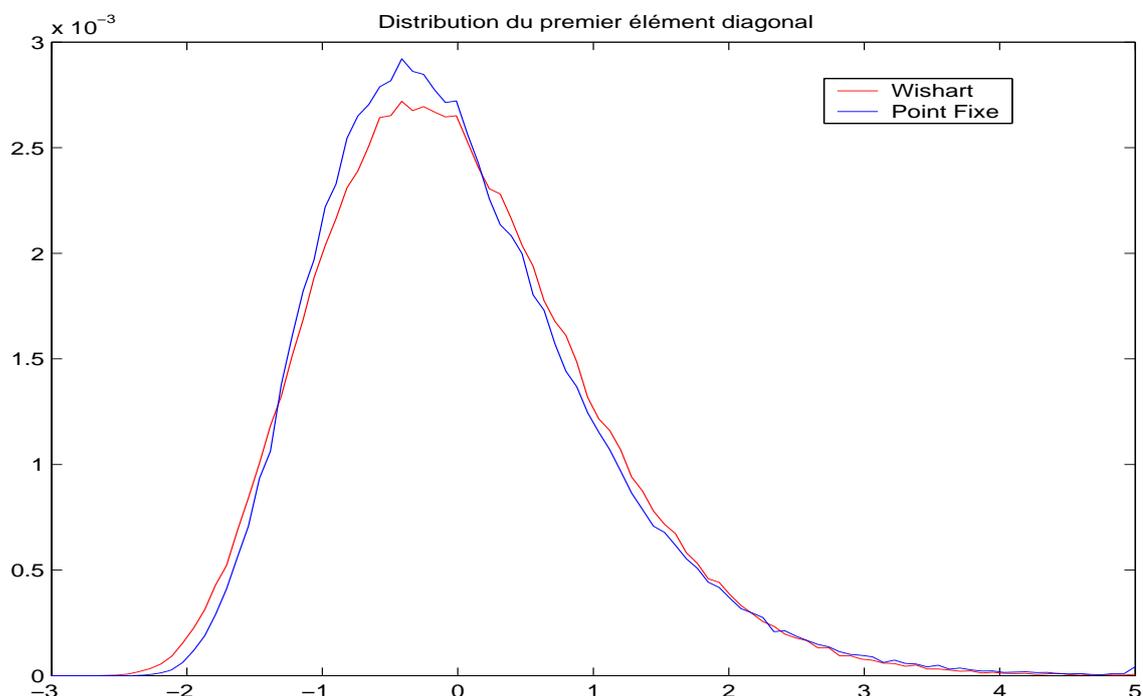


FIG. 5.7 – Distribution du premier élément de la diagonale, i.e. l'élément  $\hat{\mathbf{M}}(1, 1)$  dans le cas d'une matrice de Wishart et de l'estimateur du point fixe non normalisé

Ces premières figures (Fig. 5.7) et (Fig. 5.8) montrent que les distributions de  $\hat{\mathbf{M}}_W(1, 1)$  et de  $\hat{\mathbf{M}}_{FPbis}(1, 1)$  sont légèrement différentes. Quant aux termes  $\hat{\mathbf{M}}_W(1, 1)$  et de  $\hat{\mathbf{M}}_{FPbis}(1, 1)$ , on ne peut rien dire... Cette première démarche n'est pas du tout concluante.

Une autre approche consiste à comparer les distributions des valeurs propres des deux matrices. Ainsi, les distributions des 8 plus grandes valeurs ont été tracées sur la figure (5.9).

Sur ce graphique, pour chaque valeur propre, les courbes rouges et bleues sont pratiquement confondues. On peut donc supposer que les lois marginales des valeurs propres de  $\hat{\mathbf{M}}_{FPbis}$  sont les mêmes que celles de  $\hat{\mathbf{M}}_W$ , ce qui ne signifie pas que les deux vecteurs contenant les valeurs propres soient identiquement distribués. Pour arriver à cette conclusion, il faudrait connaître la corrélation de chaque vecteur et pour ce problème une étude des fonctions copules serait intéressante...

En se plongeant un peu plus dans la théorie des matrices aléatoires [29], on remarque que  $\hat{\mathbf{M}}_{FPbis}$  et  $\hat{\mathbf{M}}_W$  appartiennent toutes deux à une même famille de matrices aléatoires :

**Théorème 5.1.3.1**

$\hat{\mathbf{M}}_{FPbis}$  et  $\hat{\mathbf{M}}_W$  sont des matrices aléatoires ayant une distribution sphérique.

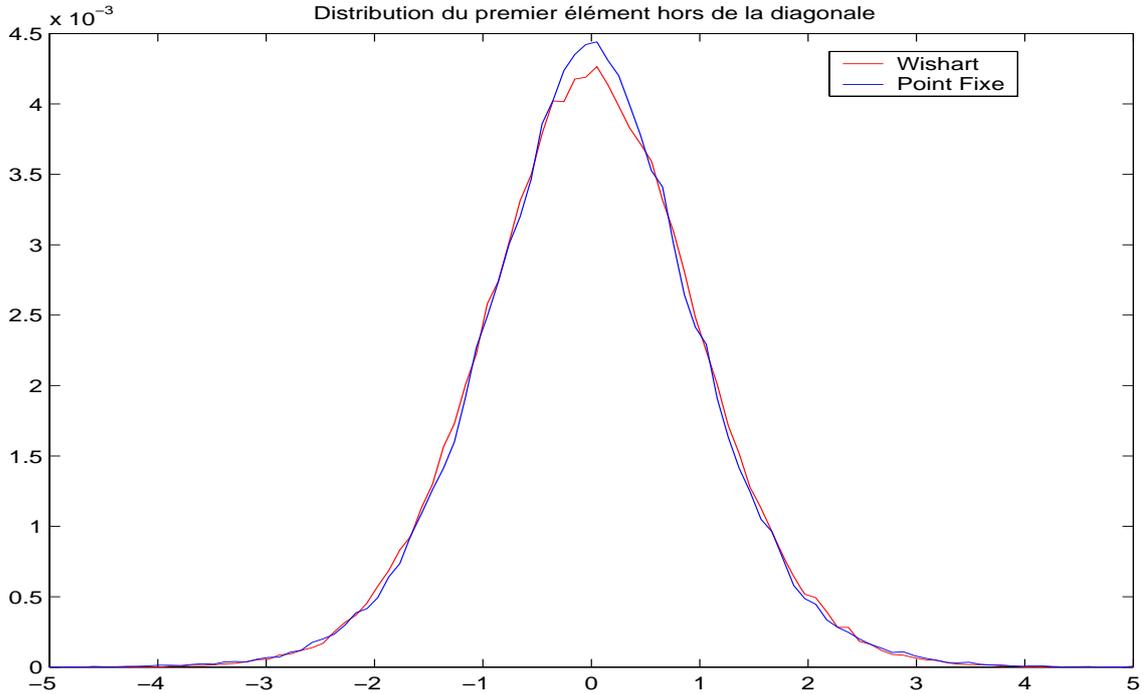


FIG. 5.8 – Distribution de l'élément  $\hat{\mathbf{M}}(1, 2)$  dans le cas d'une matrice de Wishart et de l'estimateur du point fixe non normalisé

#### Définition 5.1.3.1

Une matrice aléatoire  $\mathbf{A}$  de dimension  $n \times p$  possède une distribution sphérique si

$$\forall \mathbf{P} \in \mathcal{O}(n) \text{ et } \forall \mathbf{Q} \in \mathcal{O}(p)$$

$$\mathcal{L}(\mathbf{A}) = \mathcal{L}(\mathbf{PAQ})$$

où  $\mathcal{O}(n)$  désigne l'ensemble des matrices de dimension  $n \times n$  orthogonales et  $\mathcal{L}(X)$  désigne la loi de la variable aléatoire  $X$ .

Grâce à la définition précédente, la preuve du théorème 5.1.3.1 devient une évidence. Cette propriété des matrices  $\mathbf{a}$ , en effet, déjà été utilisée dans la preuve de la propriété CFAR-matrice du détecteur construit avec l'estimateur du point fixe.

Puis toujours dans [29], page 319, un théorème concernant la distribution de ces matrices est proposé :

#### Théorème 5.1.3.2

Si  $\mathbf{A}$  est une matrice sphérique de dimension  $n \times p$ , alors

$$\mathcal{L}(\mathbf{A}) = \mathcal{L}(\mathbf{U}\mathbf{\Lambda}\mathbf{V})$$

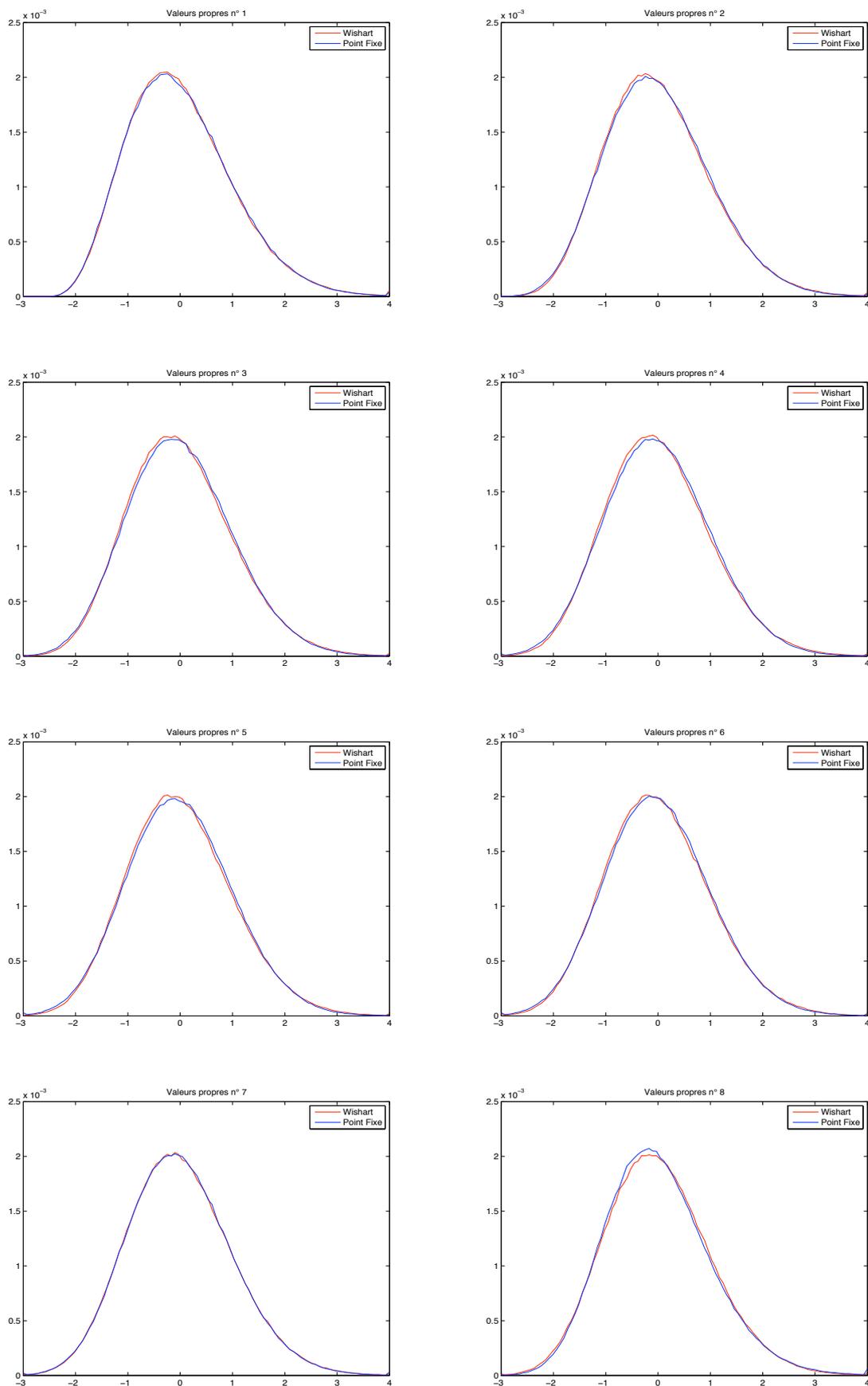


FIG. 5.9 – Comparaison des distributions des 8 plus grandes valeurs propres des matrices  $\hat{\mathbf{M}}_W$  et  $\hat{\mathbf{M}}_{FPbis}$

où  $\mathbf{U} \sim \mathcal{U}_{n,n}$ ,  $\mathbf{V} \sim \mathcal{U}_{p,p}$ , et  $\Lambda$  est la matrice des valeurs propres, ordonnées par ordre décroissant, de  $(\mathbf{A}\mathbf{A}^H)^{1/2}$ , avec  $\mathbf{U}$ ,  $\mathbf{V}$  et  $\Lambda$  sont des matrices aléatoires mutuellement indépendantes et  $\mathcal{U}_{n,n}$  est la notation utilisée pour la loi uniforme dans l'espace des matrices aléatoires.

Ces résultats évoquent la nécessité d'une analyse plus approfondie de cette conjecture, reformulée de la façon suivante :

$$\hat{\mathbf{M}}_{FPbis} \sim \mathcal{W}(N', m; \mathbf{M}) \quad (5.11)$$

où  $N'$  est le degré de liberté à déterminer en fonction de  $N$  et de  $m$ .

## 5.2 Détection dans un environnement SIRV

Dans cette partie, les précédents résultats concernant la régulation de fausse alarme vont être appliqués à des données réelles provenant de THALES Air Defence<sup>2</sup> et représentant du fouillis de sol.

### 5.2.1 Présentation des données

La figure (Fig. 5.10) représente une carte "distance-azimut" de données de fouillis de sol collectées par un radar de THALES Air Defence. Le radar a été positionné à 13 mètres au-dessus du niveau de sol et il éclaire la zone avec un angle faible rasant. Les échos complexes du clutter de sol ont été collectés dans  $N = 868$  cases distances pour 70 angles d'azimut différents et pour  $m = 8$  impulsions, ce qui signifie que la taille des vecteurs est  $m = 8$ . Chaque image correspond donc au retour d'une impulsion émise.

Près du radar, les échos caractérisent du clutter de sol hétérogène et non Gaussien (zone de gauche variant du jaune au rouge sur chaque image) tandis qu'après l'horizon radio-électrique du radar (environ 15 kilomètres), uniquement le bruit thermique Gaussien est présent (la partie de droite en bleue sur chaque image).

Afin de mettre plus en valeur les zones de bruit impulsif, la figure (Fig. 5.11) représente la même carte "distance-azimut" de la figure (Fig. 5.10), pour l'impulsion n°1, en 2 dimensions (Fig. 5.11.a), puis en 3 dimensions (Fig. 5.11.b) où la troisième dimension (verticale) désigne la puissance du bruit. Il apparaît clairement que la zone de gauche est une zone de fouillis hétérogène et impulsif.

### 5.2.2 Traitement des données

Puisqu'aucune cible n'est présente dans cette zone (hypothèse faite d'après les informations reçues de l'opérateur), les résultats obtenus pour la régulation de fausse alarme sont testés sur cette zone de fouillis.

Il y a plusieurs paramètres à régler : la taille de chaque vecteur est fixée,  $m = 8$  d'après le nombre d'échos de la scène de fouillis et le nombre de données secondaires choisi dans un premier temps est  $N = 24$ , ce qui s'explique de la façon suivante : traditionnellement, l'ajustement expérimental du seuil de détection est réalisé par comptage, en déplaçant un masque CFAR rectangulaire, de dimension  $5 \times 5$ . Pour chaque case centrale du masque (i.e. la case testée), la case noire sur la figure (Fig. 5.12), correspondant à l'observation  $\mathbf{y}$  (vecteur de dimension 8), une valeur de  $\hat{\Lambda}(\hat{\mathbf{M}})$  est calculée.

La matrice de covariance  $\hat{\mathbf{M}}$  est estimée grâce à un ensemble de  $N = 24$  vecteurs de dimension 8, considérés comme les données secondaires,  $\mathbf{c}_1, \dots, \mathbf{c}_{24}$ , et situés autour de la case testée. Ces données de références sont les cases colorées en bleu ciel sur la figure (Fig. 5.12).

Ce processus est réitéré pour chaque pixel (i.e. d'observations) sur la carte de fouillis, ici  $868 * 70 = 60760$ .

#### Remarque 5.2.2.1

*Il est intéressant de remarquer ici la non optimalité des traitements CFAR. En effet, nous avons choisi un type particulier de masque CFAR, mais d'autres auraient pu être utilisés, selon leurs formes (rectangu-*

---

<sup>2</sup>Les auteurs sont reconnaissants envers THALES Air Defence pour l'exploitation de leurs données

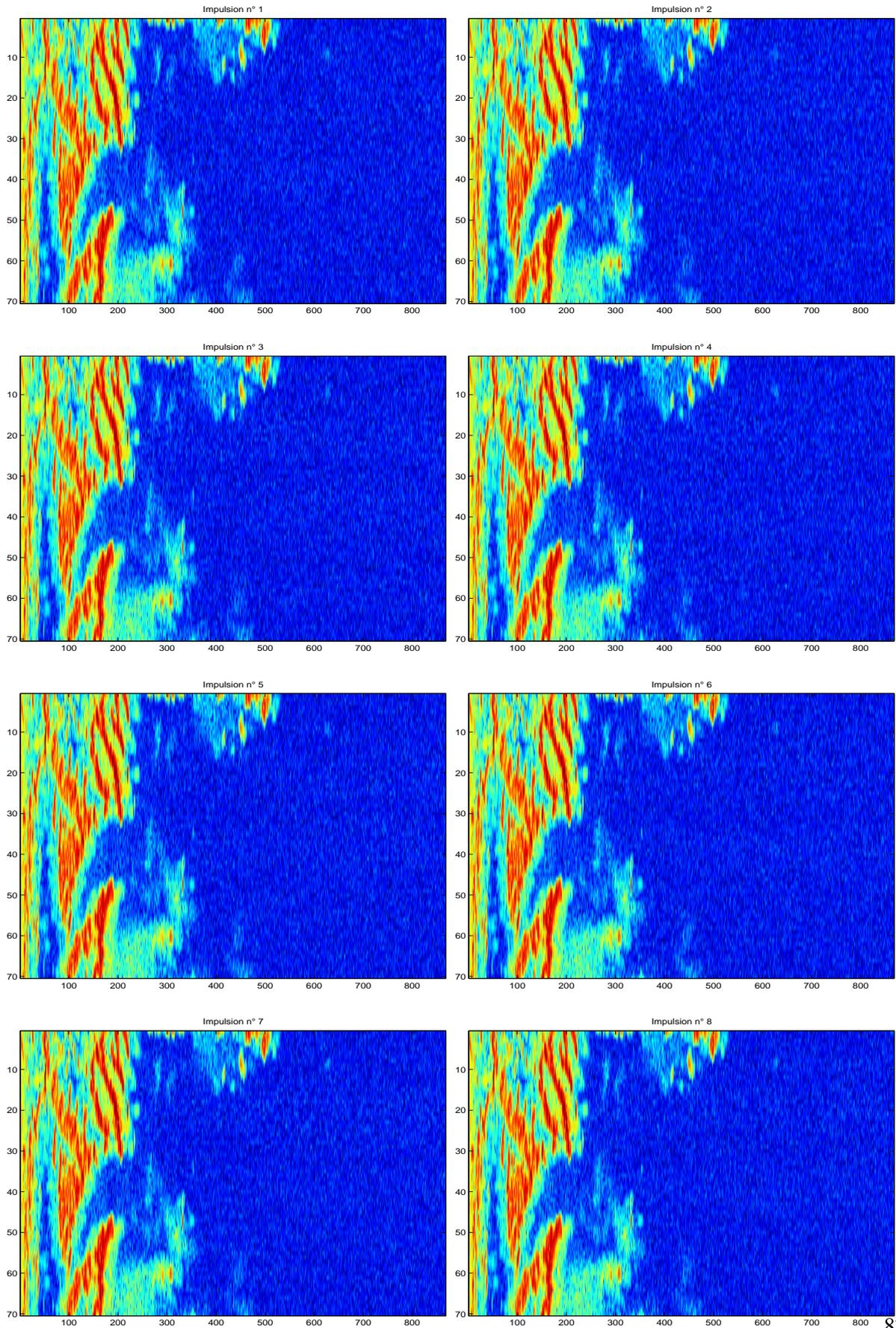


FIG. 5.10 – Échos radar de fouillis de sol : 8 impulsions

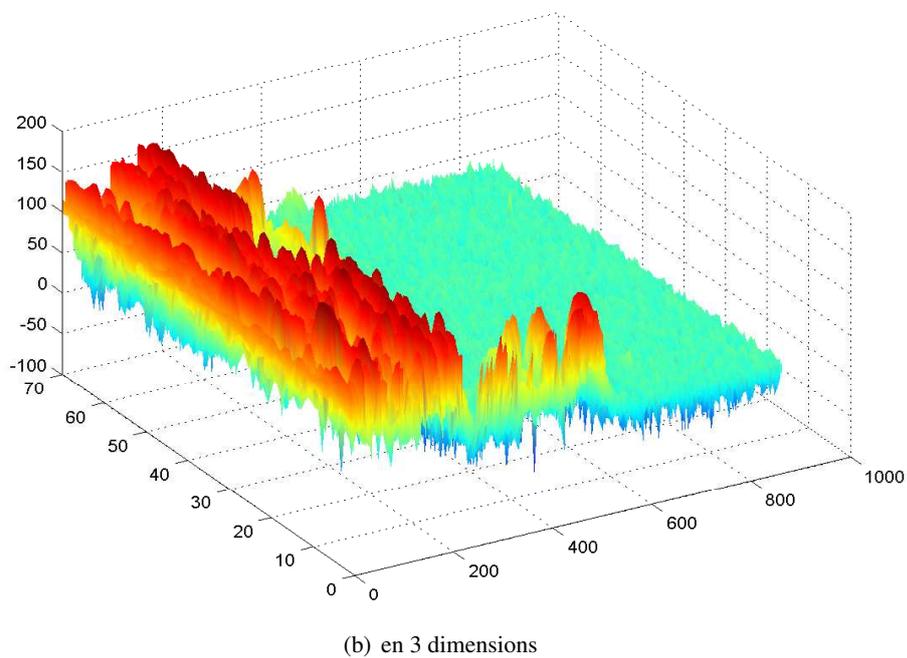
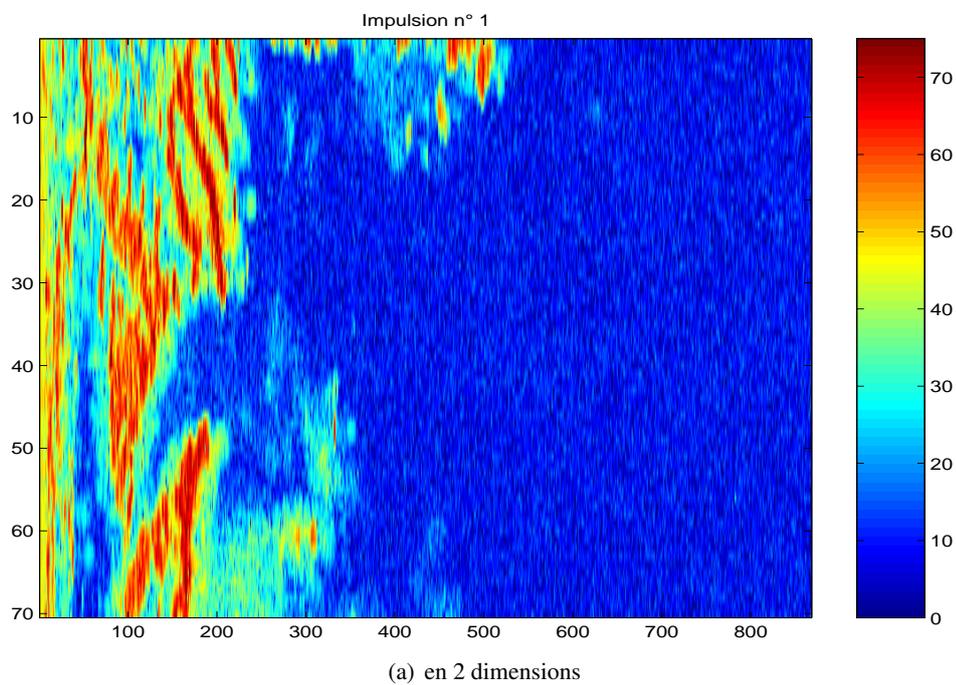
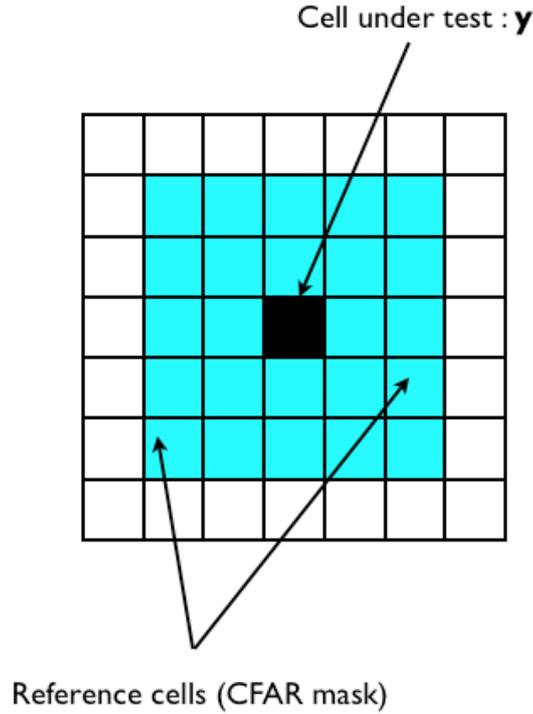


FIG. 5.11 – Échos radar de fouillis de sol : Impulsion n° 1

FIG. 5.12 – Masque CFAR de dimension  $5 \times 5$ 

laire, en croix, en disque ou bien encore en ligne/colonne), selon le nombre de données secondaires ( $N$  peut être plus ou moins grand) ou encore en prenant des cases de gardes autour de la cible. Dans notre problème, les différents essais de masques ont produit des résultats similaires et nous avons choisi, en conséquence, de sélectionner un masque rectangulaire  $5 \times 5$  assez intuitif puisqu'il en prend en compte toutes les données entourant la cible et qu'il est relativement simple à mettre en oeuvre. Notons qu'il n'y avait pas, ici, nécessité de réserver des cases de garde autour de l'observation, puisque par hypothèse, aucune cible n'est présente dans ces données.

Un dernier paramètre est à prendre en compte dans cette analyse, c'est le "steering vector"  $\mathbf{p}$ , qui caractérise complètement la cible. Ici, le "steering vector"  $\mathbf{p}$  utilisé permet de modéliser la fréquence Doppler de la cible, ou de manière équivalente, sa vitesse. Il est défini pour  $k = 1, \dots, 8$ , par

$$\mathbf{p} = \begin{pmatrix} 1 \\ \exp\left(\frac{2i\pi(k-1)}{m}\right) \\ \exp\left(\frac{2i\pi(k-1)2}{m}\right) \\ \vdots \\ \exp\left(\frac{2i\pi(k-1)(m-1)}{m}\right) \end{pmatrix} \quad (5.12)$$

Le cas particulier du vecteur  $\mathbf{p} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$  correspond à une cible de fréquence Doppler nulle, i.e. im-

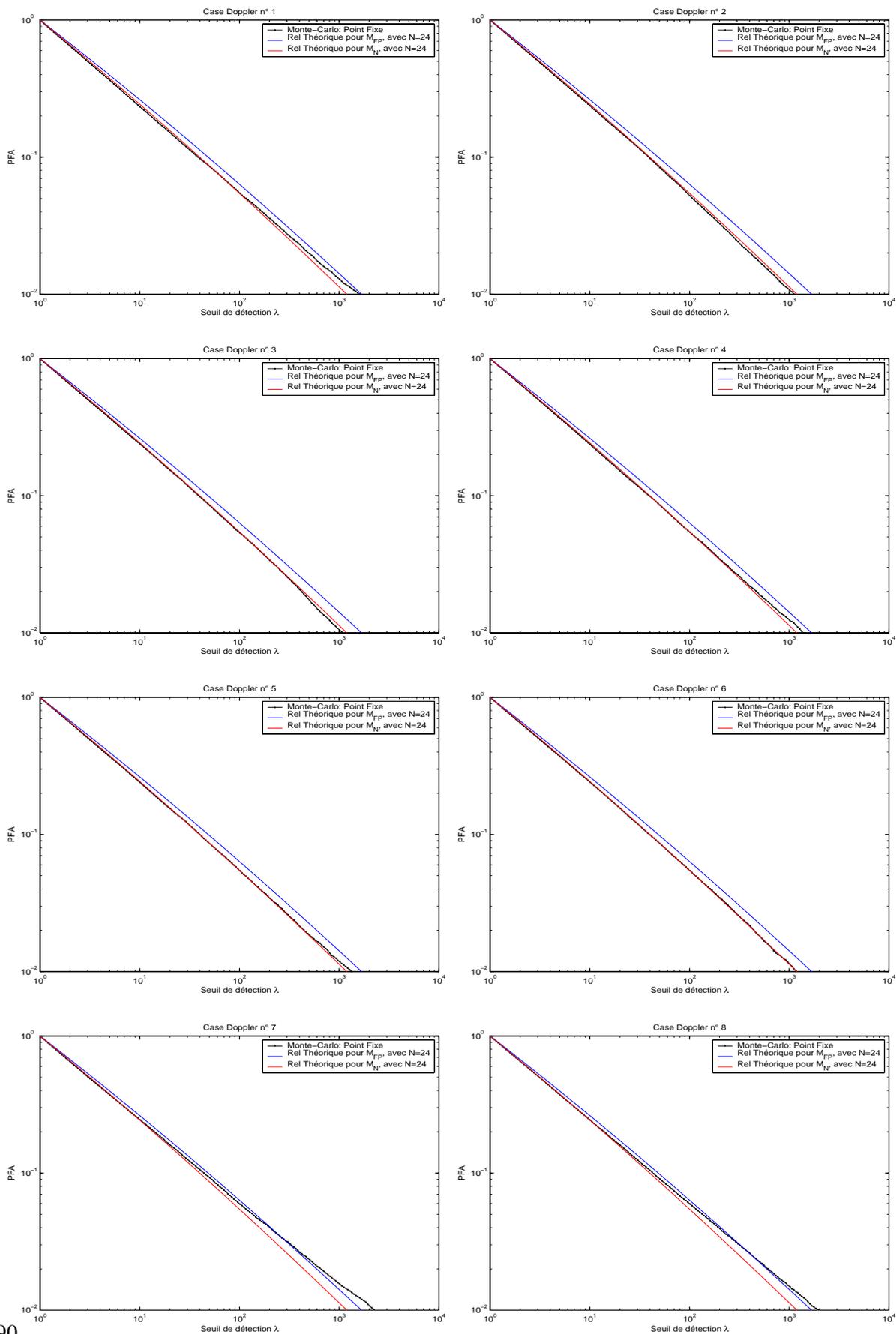


FIG. 5.13 – Comparaison des relations " $P_{fa}$ -seuil" pour le GLRT-LQ construit avec l'estimateur du point fixe, dans les cas théoriques et à partir de données réelles

mobile.

Les 8 "steering vectors" définis par (5.12), vont donc tous être testés dans le détecteur GLRT-LQ, ce qui permettra d'obtenir 8 relations entre la  $P_{fa}$  et le seuil de détection. Ceci n'a pas été le cas pour les simulations Monte-Carlo pour la raison suivante : la loi du GLRT-LQ, construit avec le Point Fixe, est indépendante de la valeur de  $\mathbf{p}$ , comme cela a été montré dans [39]. Cependant, le nombre de données réelles n'étant pas infini, il est intéressant d'étudier toutes les fréquences Doppler.

Analyse de la figure (Fig. 5.13) :

Ceci fait l'objet de la figure (Fig. 5.13) qui représente les relations " $P_{fa}$ -seuil" pour les 8 "steering vectors possibles". La courbe rouge représente la relation théorique obtenue dans le cas d'une matrice de covariance estimée par  $\hat{\mathbf{M}}_N$ , donnée par l'équation (5.3), tandis que la courbe bleue représente la relation théorique obtenue dans le cas d'une matrice de covariance estimée par  $\hat{\mathbf{M}}_{FP}$ , donnée par l'équation 5.6. La courbe noire, représente enfin les résultats obtenus à partir des données réelles. Toutes les courbes ont été tracées pour une  $P_{fa}$  variant de 1 à  $10^{-2}$ . Pour des valeurs plus faibles de la  $P_{fa}$ , les courbes ne sont plus cohérentes, ceci vient du faible nombre de données, environ  $6 * 10^4$ .

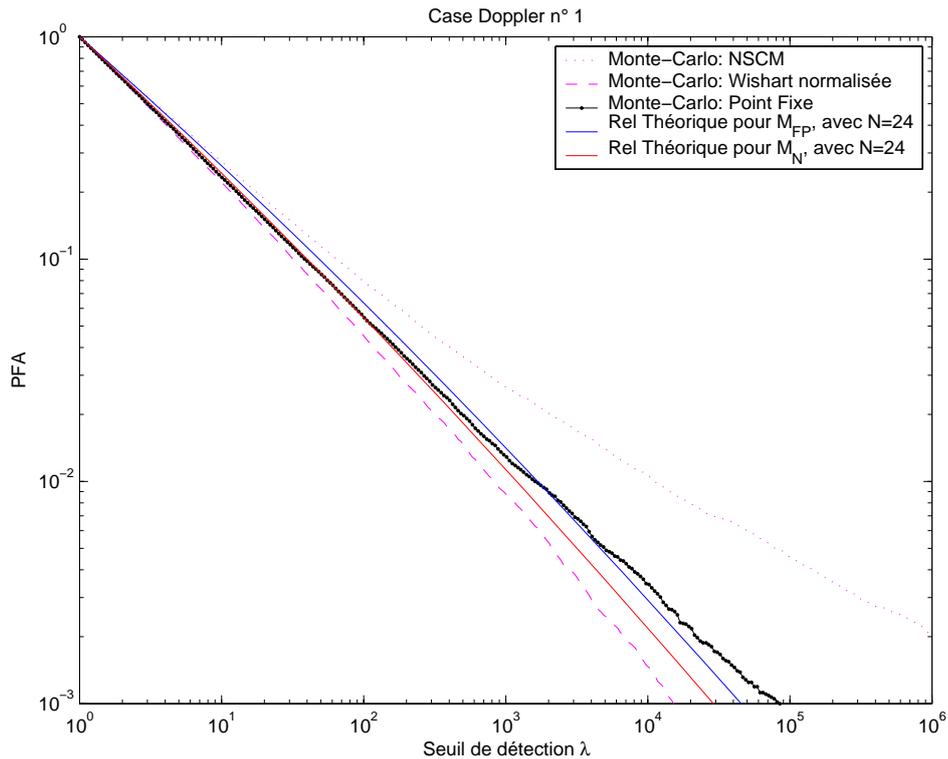


FIG. 5.14 – Comparaison des relations " $P_{fa}$ -seuil" pour le GLRT-LQ construit avec l'estimateur du point fixe et construit avec d'autres estimateurs de  $\mathbf{M}$  ( $\hat{\mathbf{M}}_N$  et  $\hat{\mathbf{M}}_{NSCM}$ ) pour la fréquence Doppler nulle

Les résultats obtenus sur données réelles sont conformes aux simulations présentées précédemment : le nombre de données secondaires utilisées pour l'estimation de la matrice de covariance est relativement faible,  $N = 24$ , ce qui explique que la courbe soit très proche de la relation " $P_{fa}$ -seuil" théorique pour

$\hat{\mathbf{M}}_N$ . De plus, pour les différents "steering vectors", la courbe noire est quasiment la même : elle oscille à l'intérieur du cône formé par les deux courbes bleue et rouge. Cette figure permet donc de conclure à une certaine stabilité du détecteur  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$ , puisqu'avec des données fortement impulsives, la théorie reste valide et donne une idée très précise de la valeur du seuil à régler pour garantir une  $P_{fa}$  donnée.

Analyse de la figure (Fig. 5.14) :

Maintenant, à titre de comparaison, la figure (Fig. 5.13) est complétée par le calcul de régulation de fausses alarmes pour les autres estimateurs présentés dans cette thèse :  $\hat{\mathbf{M}}_N$  et  $\hat{\mathbf{M}}_{NSCM}$  car ce sont les deux seuls autres estimateurs de  $\mathbf{M}$  qui peuvent être utilisés en pratique ; dans le cas de données réelles, nous n'avons évidemment pas accès à la vraie matrice de covariance...

A nouveau, la figure (Fig. 5.14), représentant uniquement la fréquence Doppler nulle, est cohérente avec la théorie. En effet, les autres estimateurs de  $\mathbf{M}$  ne sont pas appropriés à une configuration de données réelles :  $\hat{\Lambda}(\hat{\mathbf{M}}_{NSCM})$  ne possède pas la propriété CFAR-matrice donc le comptage sur l'ensemble de la carte de données n'a aucun sens, car à chaque observation est associée une et une seule matrice de covariance ; il y a autant de matrices de covariance "locales" que d'observations. D'autre part,  $\hat{\Lambda}(\hat{\mathbf{M}}_N)$  ne possède pas la propriété CFAR-texture qui est essentielle pour une carte non homogène en termes de puissance de fouillis. En fait, comme l'indique la remarque 4.2.0.2,  $\hat{\mathbf{M}}_N$  n'est plus, dans cette application, construit à partir des données Gaussiennes, auxquelles on ne peut évidemment pas avoir accès.

### 5.3 Détection dans des zones de transition de fouillis

La dernière section de ce manuscrit peut apparaître comme un "bonus" des travaux précédents. Après une étude théorique du détecteur GLRT-LQ en fonction de l'estimateur  $\hat{\mathbf{M}}$  utilisé, les résultats obtenus ont été validés sur des données réelles de fouillis de terre. Ce travail a permis d'élargir le domaine d'application de ce détecteur, et notamment, dans des zones de transition de fouillis, où  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$  est robuste.

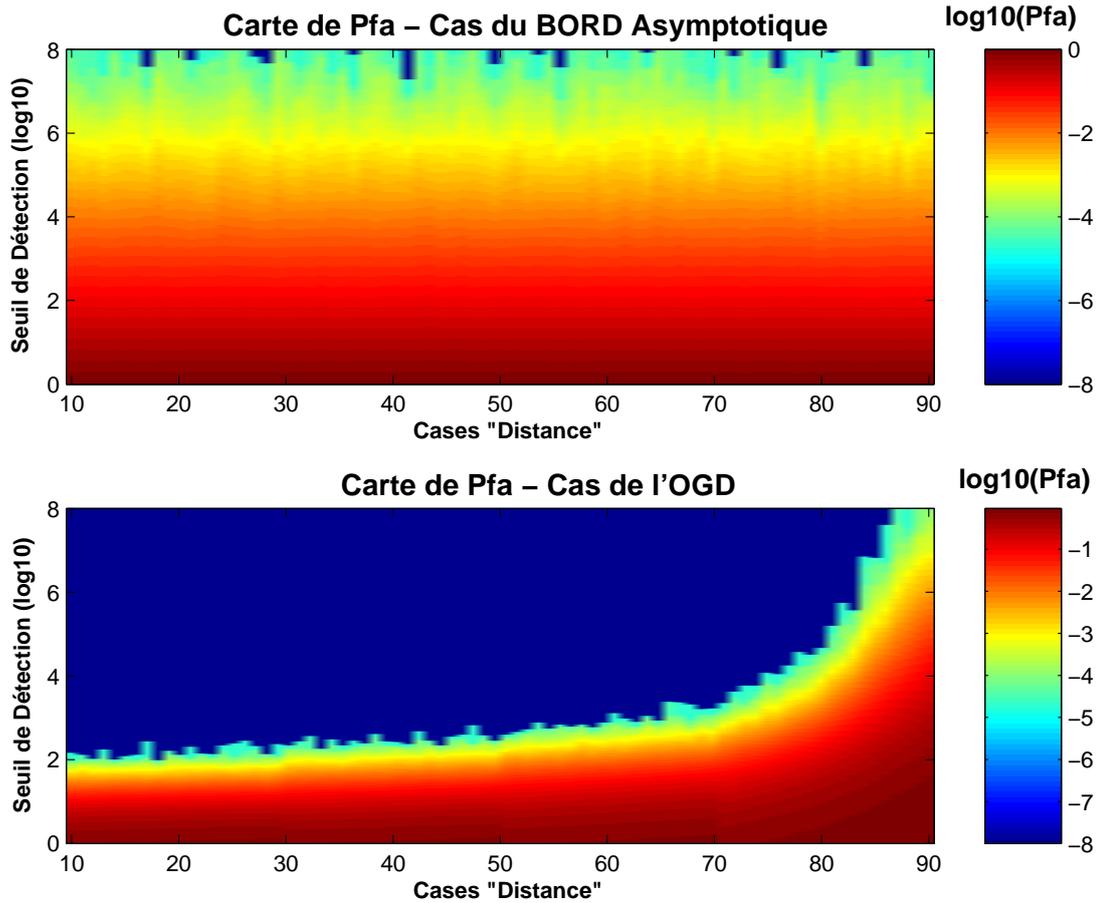


FIG. 5.15 – Régulation du taux de fausses alarmes pour l'OGD et le GLRT-LQ dans 5 différentes zones de fouillis, allant du bruit Gaussien jusqu'à un bruit très impulsif

D'un point de vue analytique, les résultats ont tous été établis dans les chapitres 2 et 3, ce sont les différentes propriétés CFAR : CFAR-texture et CFAR-matrice. Pour illustrer cette robustesse aux transitions de fouillis, des simulations ont été réalisées pour différentes zones de fouillis, allant du bruit Gaussien à un fouillis K-distribué très impulsif. Pour chaque zone, le détecteur  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$  est comparé à l'OGD, détecteur optimal dans du bruit Gaussien et défini au chapitre 1 par (1.10). La première étape consiste toujours à réguler le taux de fausses alarmes, puis, une fois le seuil de détection déterminé, les performances de détection des deux détecteurs sont calculées par une méthode Monte-Carlo, afin de garantir pour chacun d'eux la même probabilité de fausse alarme.

Analyse de la figure (Fig. 5.15) :

La figure (Fig. 5.15) présente l'étape de régulation de fausses alarmes. La taille des vecteurs est  $m = 10$ , tandis que le nombre de données secondaires utilisées pour estimer la matrice de covariance est  $N = 20$ .

L'axe des ordonnées correspond au seuil de détection variant de  $10^0$  à  $10^8$ , tandis que la  $P_{fa}$  est représentée par le code des couleurs noté sur la droite des graphiques, elle prend des valeurs comprises entre  $10^0$  et  $10^{-8}$ . Enfin, sur l'axe des abscisses sont notées 80 cases distance correspondant à différentes zones de fouillis : de 1 à 10, le bruit est Gaussien pour ensuite être K-distribué avec les paramètres d'échelle suivants : de 10 à 30,  $\nu = 2$ , puis de 30 à 50,  $\nu = 1$ , de 50 à 70,  $\nu = 0,5$ , et enfin, de 70 à 80,  $\nu = 0,1$ . En fait, plus le paramètre d'échelle diminue et plus le fouillis devient impulsif. Cette carte " $P_{fa}$ -seuil" a été réalisée dans le cas du GLRT-LQ construit avec l'estimateur du point fixe et dans le cas du détecteur classique Gaussien, construit avec la SCM, comme il est traditionnellement le cas dans la pratique.

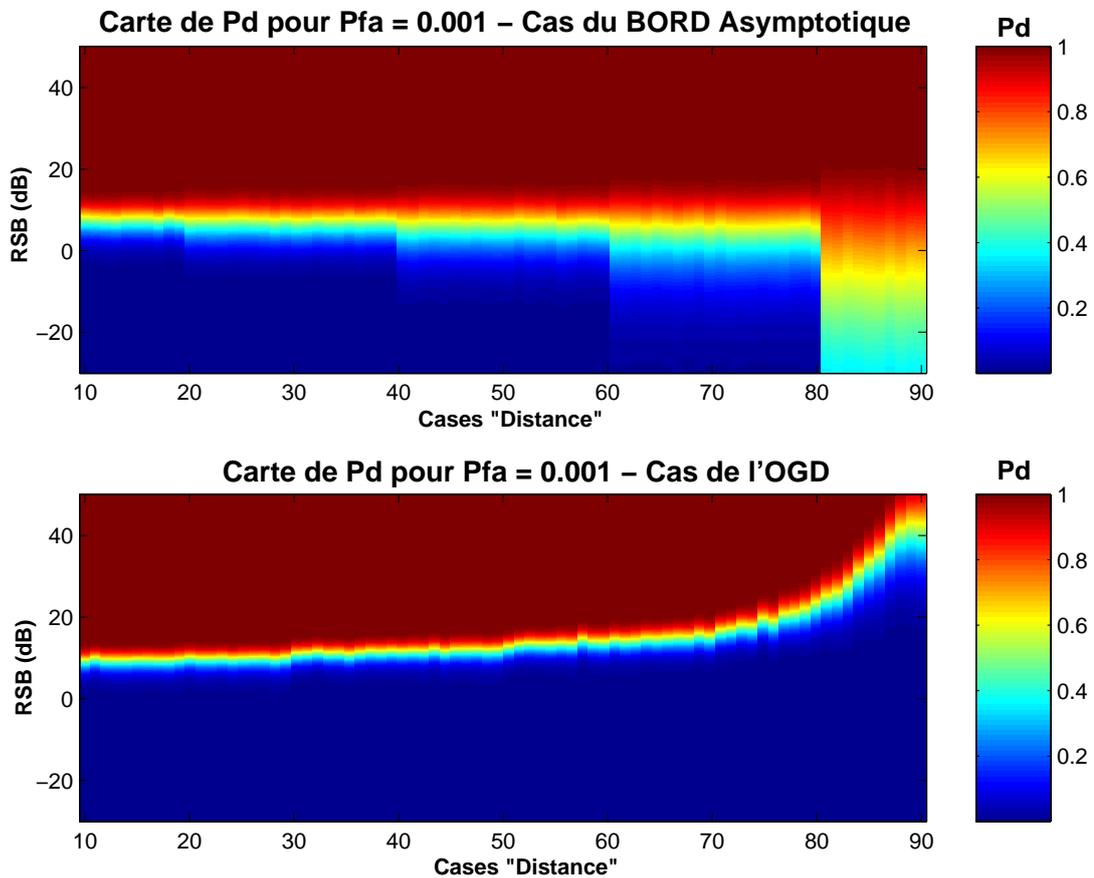


FIG. 5.16 – Performances de détection de l'OGD et du GLRT-LQ dans 5 différentes zones de fouillis, allant du bruit Gaussien jusqu'à un bruit très impulsif, tout en garantissant une  $P_{fa} = 10^{-3}$

Le constat est immédiat,  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$ , possédant la propriété CFAR-texture, garantit la même relation " $P_{fa}$ -seuil" dans toutes les zones de fouillis, il est robuste aux transitions, les zones ne sont plus identifiables. L'OGD a un comportement, quant à lui, totalement différent : le seuil de détection doit être

rehaussé quand le bruit devient plus impulsif pour garantir le même taux de fausses alarmes. Ainsi, pour garantir une  $P_{fa} \simeq 10^{-3}$ , le seuil de détection est approximativement égal à  $10^2$  dans du bruit Gaussien (10 premières cases) puis augmente jusqu'aux environs de  $10^7$  dans les dernières cases. Cette remarque illustre parfaitement les défauts de l'OGD en pratique : quand le bruit devient plus impulsif, l'utilisateur doit rehausser, souvent manuellement, le seuil de détection pour assurer la  $P_{fa}$  exigée.

Analyse de la figure (Fig. 5.16) :

Maintenant, une fois le seuil de détection ajusté, la figure (Fig. 5.16) présente les performances de détection des deux détecteurs. L'axe des ordonnées correspond au rapport signal-à-bruit (RSB) tandis que l'axe des abscisses représente à nouveau les 80 cases distances. Dans ce graphique, le jeu des couleurs est utilisé pour la probabilité de détection  $P_d$ , variant de 0 à 1. Les performances de détection de l'OGD et du GLRT-LQ ont été réalisées pour une  $P_{fa}$  égale à  $10^{-3}$ .

Dans la zone Gaussienne (10 premières cases), l'OGD est optimal mais les performances du GLRT-LQ sont très proches. La seule différence est la vitesse de variation de la  $P_d$  de 0 à 1, cela se fait approximativement de 7 à 14 dB pour l'OGD et de 5 à 15 dB pour le GLRT-LQ. Mais, en progressant dans les transitions de fouillis, on remarque que le GLRT-LQ détecte toujours ( $P_d$  supérieure à 0,9) pour un RSB d'environ 16-17 dB tandis que les performances de l'OGD sont complètement dégradées dans le fouillis impulsif, un RSB de près de 50 dB est nécessaire pour atteindre une  $P_d$  de 0,8.

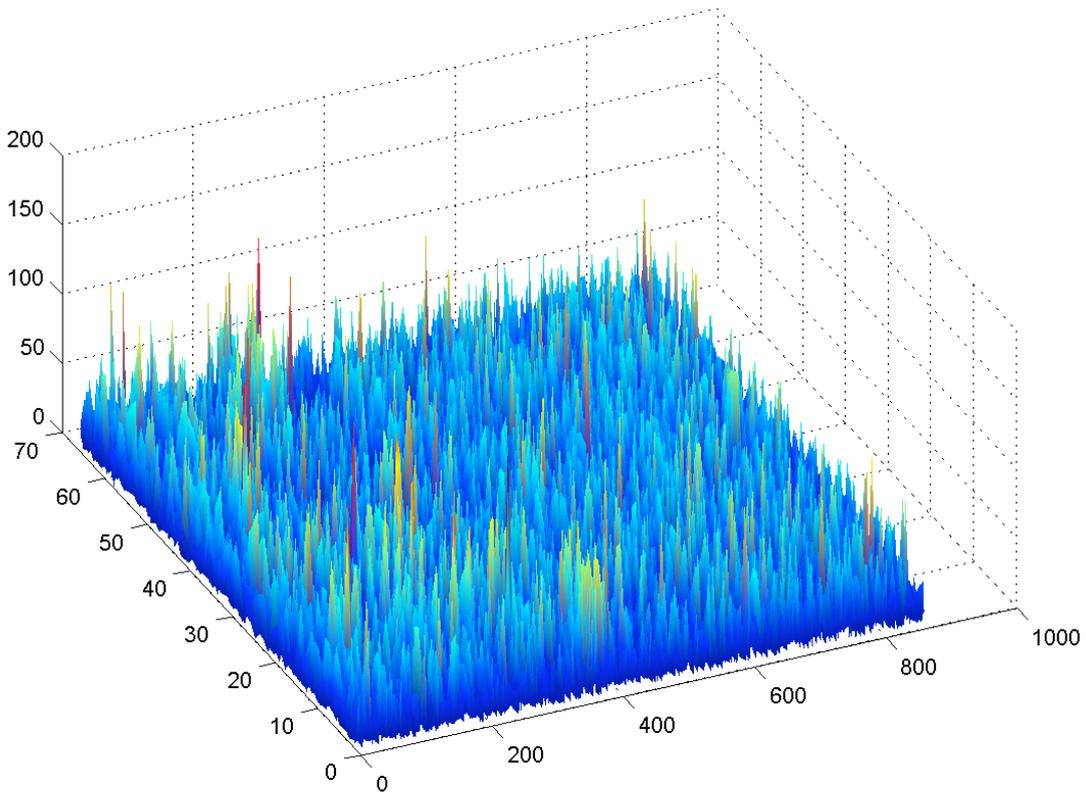


FIG. 5.17 – Carte du GLRT-LQ construit avec l'estimateur du point fixe

Cette figure met en avant un autre aspect des détecteurs : la zone de passage de 0 à 1 pour la  $P_d$  ce qui correspond à la pente de la courbe  $P_d$ -RSB, par exemple sur la figure (Fig. 1.2). Cette pente est très élevée dans le cas de l'OGD, i.e. passage très rapide de la zone "non-détection" à la zone "détection". Ceci est propre au bruit Gaussien, très homogène : intuitivement, la cible est au-dessous du niveau de bruit, elle n'est pas détectée, puis pour une très légère variation du RSB, elle se trouve au-dessus du bruit et, par conséquent, est détectée.

Contrairement à l'OGD, la pente de la courbe de détection du GLRT-LQ est beaucoup plus faible. Il n'y a plus deux zones complètement séparées, "non-détection" et "détection". Ce serait plutôt, "non-détection", "possibilité de cible", "peut-être que oui, peut-être que non", "il y a de grandes chances" et "détection".

Analyse de la figure (Fig. 5.17) :

La deuxième partie de cette section est consacrée à l'analyse des transitions sur les données réelles de fouillis de terre. La figure (Fig. 5.17) représente la carte de détection obtenue à partir des données réelles : pour chaque observation  $y$  de la figure (Fig. 5.10), le détecteur associé est calculé et la figure (Fig. 5.17) présente un tracé de tous ces détecteurs, avec pour troisième dimension, la valeur absolue du détecteur.

Il apparaît rapidement que les deux principales zones (bruit Gaussien et fouillis impulsif) présentes sur les données (Fig. 5.11) ont disparues, i.e. la carte de détection est homogène sur l'ensemble de l'image, il n'y a plus de zones distinctes. Ce résultat corrobore un peu plus la robustesse du GLRT-LQ, construit avec  $\hat{\mathbf{M}}_{FP}$ , face aux transitions de fouillis. Ainsi, les propriétés théoriques de  $\hat{\Lambda}(\hat{\mathbf{M}}_{FP})$  sont exploitées sur données réelles et apportent une amélioration du processus de détection dans des zones de transitions et/ou des zones de fouillis impulsif tout en conservant de très bonnes performances de détection dans du bruit homogène Gaussien.

# Conclusion

Dans le contexte très général de la détection radar, les détecteurs classiques, basés sur l'hypothèse d'un bruit Gaussien, sont souvent mis en défaut dès lors que l'environnement (fouillis de sol, de mer) devient inhomogène, voire impulsif, s'écartant très vite du modèle Gaussien. Des modèles physiques de fouillis basés sur les modèles de bruit composé (SIRP, Compound Gaussian Processes) permettent de mieux représenter la réalité. Ces modèles dépendent cependant de paramètres (matrice de covariance, loi de texture, paramètres de "disturbance") qu'il devient nécessaire d'estimer. Une fois ces paramètres estimés, il est possible de construire des détecteurs radar optimaux (Generalized Likelihood Ratio Test - Linear Quadratic) pour ces environnements. Cette thèse, qui s'appuie sur ces modèles, propose une analyse complète de diverses procédures d'estimation de matrices de covariance, associées à ce problème de détection. Elle décrit également les performances et les propriétés théoriques (SIRV-CFAR) du détecteur GLRT-LQ construits avec ces nouveaux estimateurs. Celles-ci sont analysées sur des données simulées mais également testées sur des données réelles de fouillis de sol. Les principaux résultats élaborés dans ce travail sont, tout d'abord, rappelés de manière détaillée dans les conclusions générales. Plusieurs propositions, non exhaustives, de direction de recherche restant à explorer sont proposées dans le paragraphe des perspectives.

## Conclusions générales

Le premier chapitre est consacré à un état de l'art des bases de la théorie de la détection radar. La méthode optimale de détection pour des paramètres de bruit et de cible connus, sous une hypothèse classique de bruit additif Gaussien, est rappelée par l'intermédiaire de l'Optimum Gaussian Detector ou OGD. Afin de se rapprocher de situations plus réalistes, une modélisation récente de fouillis non-Gaussien est présentée avec les Spherically Invariant Random Vectors ou SIRV. Ces derniers, particulièrement bien adaptés aux environnements de fouillis impulsif et hétérogène ont fait l'objet de nombreuses études qui ont conduit à la mise en place d'un détecteur associé à ce modèle, le GLRT-LQ, s'affranchissant de la variation spatiale de puissance du fouillis, encore appelée texture.

Cependant, beaucoup de travaux s'appuient encore sur l'hypothèse très contraignante et non applicable en pratique d'une connaissance *a priori* de la matrice de covariance  $\mathbf{M}$  du processus SIRP représentant le fouillis. Le travail de cette thèse a donc consisté à proposer un estimateur  $\hat{\mathbf{M}}$  de la matrice de covariance  $\mathbf{M}$  inconnue et d'analyser les performances théoriques du détecteur sur données simulées puis sur des données réelles.

Dans le chapitre 2, après une brève discussion sur les conditions essentielles que doit remplir un estimateur  $\hat{\mathbf{M}}$  de  $\mathbf{M}$ , notamment celle, essentielle, de la normalisation pour des raisons d'identifiabi-

lité, quatre estimateurs sont étudiés : la Sample Covariance Matrix,  $\widehat{\mathbf{M}}_{SCM}$ , classiquement utilisée sous hypothèse Gaussienne et distribuée selon une loi de Wishart,  $\widehat{\mathbf{M}}_N$ , l'acolyte de  $\widehat{\mathbf{M}}_{SCM}$ , la Normalized Sample Covariance Matrix,  $\widehat{\mathbf{M}}_{NSCM}$ , très répandue dans la littérature associée aux modélisations SIRV, et enfin, le Theoretical Maximum Likelihood,  $\widehat{\mathbf{M}}_{TML}$ , estimateur théorique "idéal" puisque sa construction fait appel à la matrice de covariance  $\mathbf{M}$ , supposée inconnue.

Une analyse statistique complète de ces quatre estimateurs est conduite et mène à l'élaboration de plusieurs théorèmes originaux. Les propriétés statistiques ont été étudiées au travers de la consistance, condition indispensable pour un estimateur, du biais, pouvant améliorer (dans le cas d'un biais nul) les performances de détection et la distribution asymptotique (et par conséquent son moment d'ordre 2) donnant une information supplémentaire sur le comportement de l'estimateur.

Le chapitre 3 constitue l'apport majeur de cette thèse. Un estimateur innovant de la matrice de covariance  $\mathbf{M}$  est introduit, l'estimateur du Point Fixe,  $\widehat{\mathbf{M}}_{FP}$ . Il est défini à partir d'une équation ne possédant pas de solution explicite et la première partie de ce chapitre est consacrée à sa mise en place. Dans [40], nous avons montré que  $\widehat{\mathbf{M}}_{FP}$  existe et qu'il est l'unique point fixe d'une fonction  $f_{N,\mathbf{M}}$ . Un algorithme itératif original, essentiel à la construction de  $\widehat{\mathbf{M}}_{FP}$ , est développé. Cet algorithme est convergent et il permet donc d'obtenir, de manière unique, l'estimateur du point fixe quel que soit le jeu de paramètres utilisés dans le problème sous-jacent. Ces résultats sont ensuite validés à l'aide de simulations qui ont permis d'"apprivoiser" son comportement.

La seconde partie de ce chapitre est ensuite consacrée à l'étude de  $\widehat{\mathbf{M}}_{FP}$  : comme pour les estimateurs précédents, une analyse complète (consistance, biais et distribution asymptotique) est effectuée en n'utilisant que la forme particulière de la fonction  $f_{N,\mathbf{M}}$ . Les théorèmes qui en résultent sont démontrés de façon détaillée dans le chapitre 3A. Cette analyse a notamment permis de dégager les bonnes propriétés de  $\widehat{\mathbf{M}}_{FP}$  mais l'a aussi rapproché des quatre estimateurs beaucoup plus maniables étudiés dans le chapitre 2.

Dans une dernière partie, nous nous sommes intéressés à l'utilisation de cet estimateur dans des schémas pratiques de détection radar. Ainsi, deux propriétés du détecteur GLRT-LQ, construit avec  $\widehat{\mathbf{M}}_{FP}$ , ont été établies : la propriété CFAR-texture et la propriété CFAR-matrice, qui rendent la loi du détecteur indépendante de la texture et de la matrice de covariance  $\mathbf{M}$ . Ces propriétés, très intéressantes dans la pratique, permettent, en fait, au détecteur de s'affranchir des deux paramètres d'un SIRV, on dira donc que le détecteur est SIRV-CFAR.

Un chapitre 4 succinct est introduit pour dresser un bilan des propriétés statistiques et "radaristiques" de chaque estimateur étudié dans cette thèse. Il propose, de plus, une comparaison de ces estimateurs afin de sélectionner le plus performant selon le problème. Ce chapitre 4 conclut la partie "Estimation" de ce travail.

Le chapitre 5 a présenté un autre aspect de ces travaux, en intégrant les précédents résultats dans la partie "Détection" du problème : cette partie se décompose en deux phases distinctes. Tout d'abord, le chapitre 5 débute par une approche théorique qui conduit à la détermination du seuil de détection tout en garantissant la probabilité de fausse alarme  $\mathbb{P}_{fa}$  fixée par l'utilisateur. Cet ajustement est établi en établissant la distribution du détecteur GLRT-LQ quand la matrice de covariance  $\mathbf{M}$  est estimée par  $\widehat{\mathbf{M}}_N$  et en donnant une relation théorique entre la  $\mathbb{P}_{fa}$  et le seuil de détection. Cette relation est ensuite étendue pour l'estimateur du point fixe. Ceci se justifie par les comportements asymptotiques très proches de

$\widehat{\mathbf{M}}_{FP}$  et  $\widehat{\mathbf{M}}_N$ .

Ensuite, plusieurs simulations Monte-Carlo sont présentées et permettent de valider les deux relations " $P_{fa}$ -seuil" théoriques. En fait, la validation "finale" de ces résultats intervient après l'expérimentation faite sur des données réelles de fouillis de sol. Ces données constituent une zone hétérogène de fouillis impulsif qui caractérise bien le problème d'environnement non Gaussien. Outre la validation de relations " $P_{fa}$ -seuil" théoriques, cette application montre une fois de plus que la modélisation SIRV est appropriée au problème de fouillis impulsif.

L'analyse de données réelles ayant montré une hétérogénéité des zones de fouillis, il nous a semblé intéressant d'analyser, en simulation, les performances de détection de l'OGD et du GLRT-LQ dans des zones inhomogènes de fouillis. Des simulations Monte-Carlo permettent d'aboutir à deux conclusions : premièrement, les performances de détection du GLRT-LQ, construit avec l'estimateur du point fixe sont très bonnes dans un fouillis fortement impulsif tandis que celles de l'OGD sont très nettement dégradées dans cet environnement non Gaussien. D'autre part, ces simulations, accompagnées des résultats obtenus à partir des signaux réels, mettent en évidence la robustesse du détecteur GLRT-LQ, construit avec  $\widehat{\mathbf{M}}_{FP}$ , sur des zones de transition de fouillis. Ce résultat fondamental provient de la propriété SIRV-CFAR du détecteur.

Un dernier point est à mentionner pour conclure ce travail de thèse, le chapitre 5 propose une conjecture sur la distribution statistique de  $\widehat{\mathbf{M}}_{FP}$ , qui pourrait en fait suivre une loi de Wishart... Cette conjecture est étayée dans la section 5.1.3, mais non démontrée. Elle offre ici une transition naturelle de la conclusion de ces travaux vers les perspectives envisagées.

## Perspectives

Tout au long de cette étude, nous nous sommes concentrés sur le problème de détection dans un environnement non Gaussien, modélisé par des SIRV dans lesquels la texture est une variable aléatoire inconnue et la matrice de covariance  $\mathbf{M}$  est un paramètre déterministe inconnu.

La conjecture sur la distribution de l'estimateur du point fixe n'est pas encore démontrée, il serait donc intéressant de poursuivre dans cette direction afin de caractériser complètement ce nouvel outil statistique ne possédant aucune expression analytique. De plus, dans le cas des SIRV, cet estimateur n'est qu'un estimateur du Maximum de Vraisemblance Approché, qu'en est-il du Maximum de Vraisemblance Exact ? Et dans le cas, où cet estimateur pourra être obtenu, sera-t-il encore "maniable" comme le point fixe ?

Sur le plan de l'estimation de matrices de covariance, aucune borne n'a été mentionnée. En commençant par calculer la borne de Cramer-Rao associée au modèle SIRV pour le paramètre  $\mathbf{M}$ , on pourrait obtenir une notion d'efficacité de chaque estimateur.

D'autre part, ne perdons pas de vue que les problèmes de détection sont des problèmes expérimentaux. Une analyse plus approfondie des performances de détection sur données réelles en présence de cibles permettrait peut-être un jour d'intégrer aux radars opérationnels des outils plus performants que l'OGD et la Sample Covariance Matrix et pour lesquels le seuil de détection est souvent réglé manuellement (comme dans les films américains avec des sous-marins et des avions)...

Plusieurs études de "défrichage" ont été effectuées pendant cette thèse sur des données réelles de fouillis de sol, de fouillis de mer et de fouillis combiné "terre-mer". Des résultats probants ont été obtenus, il faut cependant encore aller plus loin...

Ces techniques de détection ont aussi été appliquées à des données polarimétriques. Il reste cependant beaucoup de choses à régler et/ou à comprendre, par exemple, les vecteurs d'observations ne contiennent plus les différentes impulsions émises par le radar mais les différentes voies polarimétriques (HH, VH et VV). Est-ce que cela a un sens ? D'autre part, dans ce cas précis, que représente le "steering vector" de la cible ? Autant de questions qu'il faudrait aborder pour intégrer les outils développés dans cette thèse afin d'améliorer les techniques de détection dans des images polarimétriques...

Est-il possible d'intégrer ces techniques aux problèmes de détection de points brillants dans les images SAR ?

Enfin, le problème d'estimation de matrice de covariance touche un grand nombre de domaines, les résultats de cette thèse pourraient-ils leur être appliqués ?

En résumé, je pense que l'étude théorique d'outils d'estimation et de détection est une première étape mais qu'il est essentiel, ensuite, d'intégrer ces outils dans des applications réelles. Ceci constitue un travail de longue haleine, assez pénible mais qui en vaut la peine...

FP

# Annexe A

## Généralités

Cette annexe rappelle les densités des principales distributions évoquées dans ce manuscrit.

### A.1 Loi Gaussienne ou normale $\mathcal{N}(m, \sigma^2)$ définie sur $\mathbb{R}$

Une v.a.  $x$ , de moyenne  $m$  et de variance  $\sigma^2$  suit une loi Gaussienne (ou normale)  $\mathcal{N}(m, \sigma^2)$  quand sa densité de probabilité s'écrit,  $\forall x \in \mathbb{R}$  :

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right). \quad (\text{A.1})$$

### A.2 Densité de probabilité du vecteur Gaussien $\mathcal{N}(\mathbf{m}, \mathbf{D})$

Soit  $\mathbf{D}$  une matrice  $d \times d$  inversible, symétrique positive, avec  $\det(\mathbf{D}) \neq 0$ . Soit  $\mathbf{m} \in \mathbb{R}^d$ . La loi du vecteur Gaussien de moyenne  $\mathbf{m}$  et de matrice de dispersion  $\mathbf{D}$  est absolument continue par rapport à la mesure de Lebesgue sur  $\mathbb{R}^d$ , avec pour densité ( $\mathbf{x} \in \mathbb{R}^d$ ) :

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} \sqrt{\det(\mathbf{D})}} \exp\left(-\frac{(\mathbf{x}-\mathbf{m})^H \mathbf{D}^{-1} (\mathbf{x}-\mathbf{m})}{2}\right). \quad (\text{A.2})$$

### A.3 K-distribution

La K-distribution, loi à deux paramètres  $b$  et  $\nu$ , est définie sur  $\mathbb{R}_+$  et doit son nom à la fonction de Bessel modifiée de deuxième espèce,  $K_\nu(x)$ .

$$p(x) = \frac{b^{\nu+1}}{2^{\nu-1} \Gamma(\nu)} x^\nu K_{\nu-1}(bx) \quad (\text{A.3})$$

où

### A.4 Loi de Weibull

La loi de Weibull est une loi à deux paramètres  $a$  et  $b$ , définie sur  $\mathbb{R}_+$  :

$$p(x) = a b x^{b-1} \exp(-ax^b) \quad (\text{A.4})$$

## A.5 Loi Uniforme

La loi uniforme  $\mathcal{U}([a, b])$  est une loi à support compact sur  $\mathbb{R}$ , soit un intervalle  $[a, b] \subset \mathbb{R}$ ,  $a < b$  et a pour expression :

$$p_X(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x),$$

où  $\mathbb{1}_{[a,b]}(x) = 1$  si  $x \in [a, b]$  et 0 sinon.

## A.6 Loi du $\chi^2$ centré

La loi du  $\chi^2$  centrée est une loi à un paramètre  $\nu \in \mathbb{N} - \{0\}$ , noté  $\mathbb{N}^*$ ) déterminant le nombre de degrés de liberté de la loi.

On dit qu'une variable aléatoire  $X$  suit une loi  $\chi_\nu^2$  (loi du  $\chi^2$  à  $\nu$  degrés de liberté), si  $\forall x > 0$ , sa densité est donnée par :

$$p(x) = \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} \exp\left(-\frac{x}{2}\right).$$

La loi du  $\chi_\nu^2$  coïncide avec la loi Gamma de paramètres  $\nu/2$  et 2, soit  $\mathcal{G}\left(\frac{\nu}{2}, 2\right)$ .

## A.7 Loi Beta de première et de seconde espèce

Les densités de variables de loi  $\beta_{a,b}^1$  et  $\beta_{a,b}^2$  sont définies dans [1] par les densités de probabilités suivantes :

$$\beta_{a,b}^1(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \mathbb{1}_{[0,1]}(x),$$

et

$$\beta_{a,b}^2(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{x^{a-1}}{(1+x)^{a+b}}, x > 0$$

où  $\mathbb{1}_{[0,1]}(x)$  est la fonction indicatrice de  $x$  sur l'intervalle  $[0, 1]$ .

## A.8 Loi de Wishart

$\mathbf{A} \sim \mathcal{W}(N, m; \mathbf{M})$  si sa densité est :

$$f(\mathbf{A}) = \frac{\det(\mathbf{A})^{N-m}}{J(\mathbf{M})} \exp(-\text{Tr}(\mathbf{M}^{-1}\mathbf{A}))$$

où  $J(\mathbf{M})$  s'écrit de la manière suivante :

$$J(\mathbf{M}) = \pi^{\frac{1}{2}m(m-1)} \Gamma(N) \Gamma(N-1) \dots \Gamma(N-m+1) \det(\mathbf{M})^N.$$

## A.9 Loi faible des Grands Nombres

Soient  $X_1, \dots, X_n$  une suite de  $n$  variables aléatoires. i.i.d. réelles, de variance finie et d'espérance  $\mathbb{E}(X)$ , alors :

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow +\infty]{Pr.} \mathbb{E}(X).$$

La Loi Forte des Grands Nombres (LFGN) suppose en plus que les variables sont intégrables ( $\mathbb{E}(|X|) < +\infty$ ) mais assure une convergence presque sûre.

## A.10 Théorème Centrale Limite

Soient  $X_1, \dots, X_n$  une suite de  $n$  variables aléatoires. i.i.d. réelles. On suppose que  $\mathbb{E}(X^2) < +\infty$  et soit  $m = \mathbb{E}(X)$ ,  $\sigma^2 = \text{Var}(X)$ , alors :

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - m) \xrightarrow[n \rightarrow +\infty]{dist.} \mathcal{N}(0, \sigma^2).$$



## Annexe B

### Lemme B.0.0.1

#### Lemme B.0.0.1

Soit  $\mathbf{A}$  une matrice Hermitienne et soit  $\mathbf{U}$  une matrice unitaire, alors

$$\mathbf{A} = \mathbf{U}\mathbf{A}\mathbf{U}^H, \forall \mathbf{U} \Leftrightarrow \mathbf{A} = \alpha\mathbf{I}, \forall \alpha \in \mathbb{R} \quad (\text{B.1})$$

#### Preuve B.0.0.1

- $\mathbf{A} = \alpha\mathbf{I}$  or  $\mathbf{A} = \mathbf{0}$  implique évidemment que  $\mathbf{A} = \mathbf{U}\mathbf{A}\mathbf{U}^H$ .
- Maintenant, supposons que pour une matrice diagonale  $\mathbf{A}$  et que pour une matrice unitaire  $\mathbf{U}$ , on ait  $\mathbf{A} = \mathbf{U}\mathbf{A}\mathbf{U}^H$ . Soit  $\mathbf{V}$  la matrice des vecteurs propres de  $\mathbf{A}$  et  $\Lambda$  la matrice diagonale des valeurs propres de  $\mathbf{A}$ , alors  $\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^H$ .  
Si  $\mathbf{U} = \mathbf{V}^H$ , on a  $\mathbf{V}\mathbf{D}\mathbf{V}^H = \mathbf{D}$ . Ceci implique que  $\mathbf{A}$  est une matrice diagonale. En prenant maintenant pour  $\mathbf{U}$  la matrice des permutations qui transforme le  $i^{\text{ème}}$  élément de  $\mathbf{A}$ ,  $A_{ii}$  en le  $(i + 1)^{\text{ème}}$ , on arrive à la conclusion.



## **Annexe C**

# **"First and Second Order Moments of the Normalized Sample Covariance Matrix of Spherically Invariant Random Vectors"**

# First and Second Order Moments of the Normalized Sample Covariance Matrix of Spherically Invariant Random Vectors

Sébastien Bausson, Frédéric Pascal, Philippe Forster, Jean-Philippe Ovarlez and Pascal Larzabal

**Abstract**—Under Gaussian assumptions, the Sample Covariance Matrix (SCM) is encountered in many covariance based processing algorithms. In case of impulsive noise, this estimate is no more appropriate. This is the reason why when the noise is modeled by Spherically Invariant Random Vectors (SIRV), a natural extension of the SCM is extensively used in the literature: the well-known Normalized Sample Covariance Matrix (NSCM) which estimates the covariance of SIRV. Indeed, this estimate gets rid of a fluctuating noise power and is widely used in radar applications. The aim of this paper is to derive closed-form expressions of the first and second order moments of the NSCM.

**Index Terms**—SIRV, NSCM, estimation, performance analysis.

## I. INTRODUCTION

Given independent identically distributed observations of a zero-mean complex Gaussian random vector, the Sample Covariance Matrix (SCM) is the Maximum Likelihood estimate of the data covariance matrix. It is well-known that the SCM is complex Wishart distributed, unbiased, and its second order moments have simple expressions [1]. The full statistical characterization of the SCM allows performance analysis of numerous algorithms relying on this estimate. However, this widespread estimate is no more appropriate when observations are not Gaussian. This is for instance the case for radar clutter returns [2], [3], radio fading analysis [4] or sonar interferences [5]. In these contexts Spherically Invariant Random Vectors (SIRV) have been appropriately used in modeling non-Gaussian problems. A SIRV is a complex compound Gaussian process with random power. More precisely a SIRV  $\mathbf{c}$  [6] is the product of the square root of a positive random variable  $\tau$ , called the *texture*, and a  $m$ -dimensional independent zero mean complex Gaussian vector  $\mathbf{x}$  with covariance matrix  $\Sigma$  normalized according to  $\text{tr}(\Sigma) = m$ , where  $\text{tr}(\cdot)$  is the trace of a matrix,

$$\mathbf{c} = \sqrt{\tau}\mathbf{x}.$$

S. Bausson and P. Forster are with the Groupe d'Électromagnétisme Appliqué (GEA), Université Paris X, 92410 Ville d'Avray, France (email: sebastien.bausson, philippe.forster@u-paris10.fr) (tel: 00 33 1 47 09 45 42, fax: 00 33 1 47 50 66 55).

F. Pascal and J-P. Ovarlez are with the Office National d'Études et de Recherches Aérospatiales (ONERA), DEMR/TSI, BP 72, 92322 Chatillon Cedex, France (email: frederic.pascal, jean-philippe.ovarlez@onera.fr).

P. Larzabal is with SATIE, École Normale Supérieure de Cachan, UMR CNRS 8029, 94235 Cachan Cedex, France (email: larzabal@satie.ens-cachan.fr).

The notation  $\mathbf{x} \sim \mathbb{CN}(\mathbf{0}, \Sigma)$  means that  $\mathbf{x}$  is a zero mean complex Gaussian vector with covariance matrix  $\Sigma$ . In this paper, we consider the estimation scheme of  $\Sigma$  from  $N$  independent SIRV observations,  $\mathbf{c}_k = \sqrt{\tau_k}\mathbf{x}_k$ , for  $k = 1, \dots, N$ . In this context, we analyze the statistical properties of the well-known Normalized Sample Covariance Matrix (NSCM), introduced in [7], and defined by

$$\hat{\Sigma} = \frac{m}{N} \sum_{k=1}^N \frac{\mathbf{c}_k \mathbf{c}_k^H}{\mathbf{c}_k^H \mathbf{c}_k} = \frac{m}{N} \sum_{k=1}^N \frac{\mathbf{c}_k \mathbf{c}_k^H}{\|\mathbf{c}_k\|^2} = \frac{m}{N} \sum_{k=1}^N \frac{\mathbf{x}_k \mathbf{x}_k^H}{\|\mathbf{x}_k\|^2}, \quad (1)$$

where  $^H$  denotes the transpose conjugate operator. Notice that the NSCM does not depend on the texture. The Central Limit Theorem ensures that the NSCM is asymptotically Gaussian but first and second order moments of this estimate never appeared in the literature. Thus the goal of this paper is to fill these gaps when the  $\Sigma$ -eigenvalues are distinct, i.e. the most common and realistic case.

## II. FIRST AND SECOND ORDER MOMENTS OF THE NSCM

In this section we present the main results while computational details are provided in Appendix.

Let us introduce the eigenvalue decomposition of  $\Sigma$

$$\Sigma = \mathbf{U} \Lambda \mathbf{U}^H = \sum_{k=1}^m \lambda_k \mathbf{u}_k \mathbf{u}_k^H, \quad (2)$$

- $\Lambda$  is the diagonal matrix of the  $\Sigma$ -eigenvalues,  $\lambda_1 > \dots > \lambda_m > 0$ ,
- $\mathbf{U}$  is the unitary matrix of the  $\Sigma$ -eigenvectors.

Notice that we assume that all eigenvalues  $\lambda_1, \dots, \lambda_m$ , are strictly positive and different, i.e. their multiplicity order is 1. We note  $\mathbb{E}[\cdot]$  the statistical mean.

**Theorem 1** *The first order moment of the NSCM is given by*

$$\mathbb{E}[\hat{\Sigma}] = m \mathbf{U} \Delta \mathbf{U}^H, \quad (3)$$

where

$$\delta_k = \sum_{\substack{n=1 \\ n \neq k}}^m d_n \lambda_k \left( \frac{\log \lambda_n - \log \lambda_k}{\lambda_n - \lambda_k} - \frac{1}{\lambda_n} \right), \quad (4)$$

$$d_n = \prod_{\substack{p=1 \\ p \neq n}}^m (1 - \lambda_p / \lambda_n)^{-1}, \quad (5)$$

and where  $\Delta$  is the diagonal matrix of the  $\delta_k$ 's, with  $\delta_1 > \dots > \delta_m > 0$ .

*Proof:* See Appendix I. ■

**Remark 1** This theorem provides as a by-product the eigen-decomposition of  $\mathbb{E}[\widehat{\Sigma}]$ . It shows also that  $\mathbb{E}[\widehat{\Sigma}]$  and  $\Sigma$  share the same eigenvectors but have different eigenvalues. Consequently, the NSCM is a biased estimate of  $\Sigma$ .

**Remark 2** The NSCM preserves the ordering of the eigenvectors.

Let us denote  $\text{vec}(\cdot)$  the operator which reshapes a  $m \times n$  matrix elements into a  $mn$  column vector. Let us note  $\mathbf{v} = \text{vec}(\widehat{\Sigma})$  and introduce the two matrices

$$\mathbf{V}_1 = \mathbb{E}[\mathbf{v}\mathbf{v}^H] \quad \text{and} \quad \mathbf{V}_2 = \mathbb{E}[\mathbf{v}\mathbf{v}^T], \quad (6)$$

from which the covariances of the real and imaginary parts of the NSCM are straightforwardly derived.

**Theorem 2** The NSCM is asymptotically Gaussian and

$$\mathbf{V}_1 = \frac{m^2}{N} \sum_{p=1}^m \sum_{k=1}^m (w_{pk} + (N-1)\delta_p\delta_k) \text{vec}(\mathbf{u}_p\mathbf{u}_p^H) \text{vec}(\mathbf{u}_k\mathbf{u}_k^H)^H + \frac{m^2}{N} \sum_{p=1}^m \sum_{\substack{k=1 \\ k \neq p}}^m w_{pk} \text{vec}(\mathbf{u}_p\mathbf{u}_k^H) \text{vec}(\mathbf{u}_p\mathbf{u}_k^H)^H, \quad (7)$$

$$\mathbf{V}_2 = \frac{m^2}{N} \sum_{p=1}^m \sum_{k=1}^m (w_{pk} + (N-1)\delta_p\delta_k) \text{vec}(\mathbf{u}_p\mathbf{u}_p^H) \text{vec}(\mathbf{u}_k\mathbf{u}_k^H)^T + \frac{m^2}{N} \sum_{p=1}^m \sum_{\substack{k=1 \\ k \neq p}}^m w_{pk} \text{vec}(\mathbf{u}_p\mathbf{u}_k^H) \text{vec}(\mathbf{u}_k\mathbf{u}_p^H)^T, \quad (8)$$

where

$$w_{kk} = \sum_{\substack{n=1 \\ n \neq k}}^m d_n \lambda_k \left( \frac{2\lambda_k \log(\lambda_k/\lambda_n)}{(\lambda_k - \lambda_n)^2} - \frac{1}{\lambda_n} \frac{\lambda_k + \lambda_n}{\lambda_k - \lambda_n} \right), \quad (9)$$

$$w_{pk} = \sum_{\substack{n=1 \\ n \neq p \\ n \neq k}}^m d_n \tilde{w}_{pkn}, \quad \text{for } p \neq k, \quad (10)$$

with

$$\tilde{w}_{pkn} = \lambda_p \lambda_k \left\{ \frac{\lambda_n(\lambda_p + \lambda_k) - 2\lambda_p\lambda_k}{\lambda_n^2(\lambda_p - \lambda_k)^2} - \frac{\log(\lambda_n/\lambda_p)}{(\lambda_n - \lambda_p)(\lambda_n - \lambda_k)} + \left[ \frac{\lambda_k(\lambda_n - \lambda_p) \log(\lambda_k/\lambda_p)}{\lambda_n^2(\lambda_n - \lambda_k)(\lambda_p - \lambda_k)^3} \right] (2\lambda_p\lambda_n - \lambda_p\lambda_k - \lambda_k^2) \right\}, \quad (11)$$

where  $\delta_n$  and  $d_n$  are respectively defined in (4) and (5).

*Proof:* See Appendix II, III, IV. ■

### III. CONCLUSION

The closed-form expressions of the first and second order moments of the NSCM for SIRV modeling have been provided in this paper with full detailed proofs. These analytical equations are essential for analyzing performance of signal processing methods based on NSCM: detection schemes in radar applications, direction of arrivals estimation in array processing.

### APPENDIX I PROOF OF THEOREM 1

Using the eigen-decomposition of (2), let us whiten  $\mathbf{x}$  according to  $\mathbf{y} = \mathbf{\Lambda}^{-1/2} \mathbf{U}^H \mathbf{x}$ . Hence  $\mathbf{y} \sim \mathbb{CN}(\mathbf{0}, \mathbf{I})$  and

$$\frac{\mathbf{x}\mathbf{x}^H}{\|\mathbf{x}\|^2} = \mathbf{U} \mathbf{\Lambda}^{1/2} \frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H \mathbf{\Lambda} \mathbf{y}} \mathbf{\Lambda}^{1/2} \mathbf{U}^H.$$

The NSCM (1) statistical mean can be rewritten as

$$\mathbb{E}[\widehat{\Sigma}] = m \mathbf{U} \mathbf{\Lambda}^{1/2} \mathbb{E} \left[ \frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H \mathbf{\Lambda} \mathbf{y}} \right] \mathbf{\Lambda}^{1/2} \mathbf{U}^H. \quad (I.12)$$

Each component  $y_k$  of  $\mathbf{y}$  is a zero-mean unit variance circular complex Gaussian variable and can be expressed as:

$$y_k = \sqrt{\frac{1}{2}} \chi_k^2 \exp(i\theta_k),$$

where  $\chi_k^2$  is Chi-squared-distributed with 2 degrees of freedom,  $\theta_k$  is uniformly distributed on  $[0, 2\pi]$ . All the  $\chi_k^2$ 's and  $\theta_k$ 's are two-by-two independent. It follows that (I.12) yields

$$\mathbb{E}[\widehat{\Sigma}] = m \sum_{k=1}^m \lambda_k \mathbb{E} \left[ \chi_k^2 / \sum_{n=1}^m \lambda_n \chi_n^2 \right] \mathbf{u}_k \mathbf{u}_k^H.$$

Let us set

$$\delta_k = \mathbb{E} \left[ \lambda_k \chi_k^2 / \sum_{n=1}^m \lambda_n \chi_n^2 \right] = \mathbb{E} \left[ \frac{1}{1 + X_2/X_1} \right], \quad (I.13)$$

where  $X_1 = \lambda_k \chi_k^2$  and  $X_2 = \sum_{\substack{n=1 \\ n \neq k}}^m \lambda_n \chi_n^2$ .

The PDF of  $X_2$  has to be derived to complete the proof. Since all  $\chi_k^2$ 's are independent, the characteristic function of  $X_2$  is

$$\phi_{X_2}(u) = \prod_{\substack{n=1 \\ n \neq k}}^m (1 - 2i\lambda_n u)^{-1} = \sum_{\substack{n=1 \\ n \neq k}}^m \frac{c_n}{1 - 2i\lambda_n u},$$

where  $c_n = \prod_{\substack{p=1 \\ p \neq n \\ p \neq k}}^m \left( 1 - \frac{\lambda_p}{\lambda_n} \right)^{-1}$ . Thus, the PDF of  $X_2$  follows

$$p_{X_2}(x) = \frac{1}{2} \sum_{\substack{n=1 \\ n \neq k}}^m \frac{c_n}{\lambda_n} \exp\left(-\frac{x}{2\lambda_n}\right), \quad x \geq 0. \quad (I.14)$$

So, the density of  $X_2$  is obtained by the weighted sum of the densities of  $\lambda_n \chi_n^2$  by the coefficient  $c_n$ . Now, the PDF of the ratio  $X_2/X_1$  is a weighted sum of F laws (Fisher-Snedecor)

$$p_{X_2/X_1}(x) = \sum_{\substack{n=1 \\ n \neq k}}^m c_n \frac{\lambda_k}{\lambda_n} \left( 1 + \frac{\lambda_k}{\lambda_n} x \right)^{-2}, \quad x \geq 0, \quad (I.15)$$

and after some manipulations, (I.13) yields

$$\delta_k = \sum_{\substack{n=1 \\ n \neq k}}^m c_n \left( \frac{\lambda_n/\lambda_k}{(1 - \lambda_n/\lambda_k)^2} \log(\lambda_n/\lambda_k) + \frac{1}{1 - \lambda_n/\lambda_k} \right).$$

It remains to show that  $\delta_1 > \dots > \delta_m > 0$ . First, the  $\delta_k$ 's, defined in (I.13), are strictly positive. Now, let us consider the following function for  $x > 0$  and  $y > 0$ :

$$f_w(x, y) = \frac{1}{4} \int_{\mathbb{R}_+^2} \frac{xu}{xu + yv + w} \exp\left(-\frac{u+v}{2}\right) du dv.$$

It follows from (I.13) that we have  $\delta_k = \mathbb{E}_w [f_w(\lambda_k, \lambda_p)]$  and

$$\delta_p = \mathbb{E}_w [f_w(\lambda_p, \lambda_k)], \text{ for } w = \sum_{\substack{n=1 \\ n \neq k \\ n \neq p}}^m \lambda_n \chi_n^2, \text{ and where } \mathbb{E}_w [\cdot]$$

stands for the statistical mean related to  $w$ . To show that  $\delta_k < \delta_p$ , we prove that  $f_w(\lambda_k, \lambda_p) < f_w(\lambda_p, \lambda_k)$  for all  $w$ , assuming  $\lambda_k < \lambda_p$ . Let us define the functions

$$\begin{aligned} f_1(t) &= f_w((1-t)\lambda_p + t\lambda_k, \lambda_p), \\ f_2(t) &= f_w(\lambda_p, (1-t)\lambda_p + t\lambda_k), \end{aligned}$$

which verify  $f_1(0) = f_2(0)$ ,  $f_1(1) = \delta_k$  and  $f_2(1) = \delta_p$ . To demonstrate that  $\delta_k < \delta_p$ , we show hereafter that  $f_1$  and  $f_2$  are respectively strictly decreasing and strictly increasing functions of  $t$  on the interval  $[0, 1]$ . We have

$$\begin{aligned} \frac{\partial f_w}{\partial x}(x, y) &= \frac{1}{4} \int_{\mathbb{R}_+^2} \frac{u(yv+w)}{(xu+yv+w)^2} e^{-(u+v)/2} dudv > 0, \\ \frac{\partial f_w}{\partial y}(x, y) &= -\frac{1}{4} \int_{\mathbb{R}_+^2} \frac{vXu}{(xu+yv+w)^2} e^{-(u+v)/2} dudv < 0, \end{aligned}$$

from which we obtain:

$$\begin{aligned} \frac{df_1}{dt} &= \left( \frac{\partial f_w}{\partial x} \right)_{((1-t)\lambda_p + t\lambda_k, \lambda_p)} (\lambda_k - \lambda_p) < 0, \\ \frac{df_2}{dt} &= \left( \frac{\partial f_w}{\partial y} \right)_{(\lambda_p, (1-t)\lambda_p + t\lambda_k)} (\lambda_k - \lambda_p) > 0. \end{aligned}$$

In summary,  $\delta_k < \delta_p$  for any  $k, p$  such that  $\lambda_k < \lambda_p$ . This completes the proof of Theorem 1.

## APPENDIX II

### PROOF OF (7), (8) AND (9) OF THEOREM 2

By expressing the variance of the NSCM as a linear combination of functions of the  $\Sigma$ -eigenvectors, we compute the statistical means of the coefficients. Eqs. (1), (3), (6) and (I.12) leads to

$$\begin{aligned} \mathbf{V}_1 &= \frac{m^2}{N^2} \sum_{k=1}^m \sum_{p=1}^m \mathbb{E} \left[ \text{vec} \left( \frac{\mathbf{x}_k \mathbf{x}_k^H}{\|\mathbf{x}_k\|^2} \right) \text{vec} \left( \frac{\mathbf{x}_p \mathbf{x}_p^H}{\|\mathbf{x}_p\|^2} \right)^H \right], \\ &= \frac{m^2}{N} \sum_{p,j,n,k} \{ [\omega_{p j n k} + (N-1)\delta_p \delta_n \delta(p-j)\delta(n-k)] \\ &\quad \text{vec}(\mathbf{u}_p \mathbf{u}_j^H) \text{vec}(\mathbf{u}_n \mathbf{u}_k^H)^H \}, \end{aligned}$$

where

$$\omega_{p j n k} = \mathbb{E} \left[ \frac{\sqrt{\lambda_p \lambda_j \lambda_n \lambda_k \chi_p^2 \chi_j^2 \chi_n^2 \chi_k^2}}{(\sum_{t=1}^m \lambda_t \chi_t^2)^2} \right] \mathbb{E} \left[ e^{i(\theta_p - \theta_j + \theta_k - \theta_n)} \right],$$

and  $\delta(\cdot)$  is the Kronecker delta. The  $\theta$ 's being independent uniform variables, the last term of previous equations is zero unless  $p = j, k = n$  or  $p = n, k = j$ , which leads to

$$\begin{aligned} \mathbf{V}_1 &= \frac{m^2}{N} \sum_{p,k} [w_{pk} + (N-1)\delta_p \delta_k] \text{vec}(\mathbf{u}_p \mathbf{u}_p^H) \text{vec}(\mathbf{u}_k \mathbf{u}_k^H)^H \\ &\quad + \frac{m^2}{N} \sum_{p=1}^m \sum_{\substack{k=1 \\ k \neq p}}^m w_{pk} \text{vec}(\mathbf{u}_p \mathbf{u}_k^H) \text{vec}(\mathbf{u}_p \mathbf{u}_k^H)^H \end{aligned}$$

where

$$w_{pk} = \lambda_p \lambda_k \mathbb{E} \left[ \chi_p^2 \chi_k^2 / \left( \sum_{n=1}^m \lambda_n \chi_n^2 \right)^2 \right], \quad (\text{II.16})$$

This is (7) of Theorem 2 and (8) is derived from the same reasoning. Concerning (9), one has, from (II.16), for  $p = k$ ,  $w_{kk} = \mathbb{E} \left[ (1 + X_2/X_1)^{-2} \right]$ , where  $X_1$  and  $X_2$  are defined in (I.13). Thus  $w_{kk} = \int_0^{+\infty} (1+x)^{-2} p_{X_2/X_1}(x) dx$ . Eq. (10) is derived in Appendix IV. The proof needs some results related to exponential integrals introduced in Appendix III.

## APPENDIX III

### EXPONENTIAL INTEGRALS AND RELATED FUNCTIONS

This section contains some mathematical tools used in Appendix IV. From pp. 228 of [8], let us recall the definition of the exponential integral

$$\begin{aligned} E_n(z) &= \int_1^{+\infty} \frac{e^{-zt}}{t^n} dt, \quad n \in \mathbb{N}, \Re(z) > 0, \\ E_1(z) &= \int_1^{+\infty} \frac{e^{-zt}}{t} dt = \int_z^{+\infty} \frac{e^{-t}}{t} dt, \\ &= -\gamma - \ln z - \sum_{n=1}^{+\infty} \frac{(-1)^n z^n}{n n!}, \end{aligned} \quad (\text{III.17})$$

where  $\Re(z)$  denotes the real part of  $z$  and  $\gamma$  is Euler's gamma constant. It is assumed that the path of integration on the complex plane excludes the origin and does not cross the negative real axis [8]. Analytic continuation yields multi-valued functions with branch points at  $z = 0$  and  $z = \infty$ . Let us introduce the real function

$$F_n(a, x) = \int_x^{+\infty} t^n e^{-at} E_1(t) dt, \quad n \in \mathbb{N}, x > 0, a > -1. \quad (\text{III.18})$$

Let us show that the integral involved in the definition of  $F_n(a, x)$  is well defined for  $x > 0$  and  $a > -1$ . From [8] pp. 229 we have  $e^x E_1(x) < \log(1 + 1/x)$  for  $x > 0$  which leads to

$$0 \leq F_n(a, x) \leq \log(1 + 1/x) x^{n+1} \alpha_n([a+1]x),$$

where the function  $\alpha_n(y) = \int_1^{+\infty} t^n e^{-yt} dt$ ,  $n \in \mathbb{N}$ , is defined for  $y > 0$ , see [8] pp. 228. In conclusion, function  $F_n(a, x)$  is well defined for  $x > 0$  and  $(a+1)x > 0$ , i.e. for  $x > 0$  and  $a > -1$ .

We are interested in the limiting values of  $F_n(a, x)$  when  $x$  tends to zero. Integration by parts leads to

$$F_n(a, x) = \frac{x^n}{a} e^{-ax} E_1(x) + \frac{n}{a} F_{n-1}(a, x) - \frac{x^n}{a} \alpha_{n-1}([a+1]x), \quad (\text{III.19})$$

where  $\alpha_n(y)$  is given by [8] pp. 228

$$\alpha_n(y) = n! y^{-n-1} e^{-y} \left( 1 + y + \frac{y^2}{2!} + \dots + \frac{y^n}{n!} \right).$$

Eq. (III.19) combined with  $\lim_{x \rightarrow 0} x^n E_1(x) = 0$  for  $n \geq 1$  which results from the series expansion (III.17), with

$\lim_{x \rightarrow 0} F_0(a, x) = \int_{\mathbb{R}_+} e^{-at} E_1(t) dt = \ln(1+a)/a$ , see [8] pp. 230, and with the above expression of  $\alpha_n(y)$ , leads to

$$\begin{aligned} \lim_{x \rightarrow 0} F_1(a, x) &= F_1(a, 0) = \frac{\ln(1+a)}{a^2} - \frac{1}{a(1+a)}, \\ \lim_{x \rightarrow 0} F_2(a, x) &= F_2(a, 0) = \frac{2\ln(1+a)}{a^3} - \frac{3a+2}{a^2(1+a)^2}. \end{aligned} \quad (\text{III.20})$$

#### APPENDIX IV

END OF PROOF OF THEOREM 2 ( $w_{pk}$  FOR  $p \neq k$ , SEE (10))

At the end of Appendix II, it remained to compute (II.16) to complete the proof of Theorem 2.

Let us write  $w_{pk} = \mathbb{E} \left[ (1 + \tilde{X}_2/\tilde{X}_1)^{-1} \right]$  with  $\tilde{X}_1 = \lambda_k \chi_k^2 + \lambda_p \chi_p^2$  and  $\tilde{X}_2 = \sum_{\substack{n=1 \\ n \neq k \\ n \neq p}}^m \lambda_n \chi_n^2$ . A PDF decomposition similar to (I.14), but for  $\tilde{X}_2$ , provides

$$w_{pk} = \sum_{\substack{n=1 \\ n \neq p \\ n \neq k}}^m \delta_{pkn} \prod_{\substack{j=1 \\ j \neq n \\ j \neq p \\ j \neq k}}^m \left( 1 - \frac{\lambda_j}{\lambda_n} \right)^{-1}, \quad (\text{IV.21})$$

where  $\delta_{pkn} = \lambda_p \lambda_k \mathbb{E} \left[ \chi_p^2 \chi_k^2 / (\lambda_p \chi_p^2 + \lambda_k \chi_k^2 + \lambda_n \chi_n^2)^2 \right]$  is

$$\delta_{pkn} = \frac{\lambda_p \lambda_k}{8} \int_{\mathbb{R}_+^3} \frac{x_p x_k e^{-(x_p+x_k+x_n)/2}}{(\lambda_p x_p + \lambda_k x_k + \lambda_n x_n)^2} dx_p dx_k dx_n.$$

An analytic expression of  $\delta_{pkn}$  is obtained by computing the above integral. The previous equation is rewritten as

$$\delta_{pkn} = \frac{\lambda_p \lambda_k}{8} \int_{\mathbb{R}_+^2} t_1 x_p x_k e^{-(x_p+x_k)/2} dx_p dx_k, \quad (\text{IV.22})$$

where  $t_1 = \int_0^{+\infty} e^{-x_n/2} (\lambda_p x_p + \lambda_k x_k + \lambda_n x_n)^{-2} dx_n$ .

Then, by setting  $C = \lambda_p x_p + \lambda_k x_k$ ,  $t_1$  is rewritten as

$$\begin{aligned} t_1 &= \frac{1}{\lambda_n C} \exp\left(\frac{C}{2\lambda_n}\right) E_2\left(\frac{C}{2\lambda_n}\right), \\ &= \frac{1}{\lambda_n C} \left[ 1 - \frac{C}{2\lambda_n} \exp\left(\frac{C}{2\lambda_n}\right) E_1\left(\frac{C}{2\lambda_n}\right) \right], \end{aligned}$$

where  $E_1$  and  $E_2$  are defined in (III.17). Now, by replacing  $t_1$  in (IV.22), we obtain

$$\delta_{pkn} = \frac{\lambda_p \lambda_k}{8 \lambda_n} \left( t_2 - \frac{1}{2\lambda_n} t_3 \right), \quad (\text{IV.23})$$

$$t_2 = \int_{\mathbb{R}_+^2} \frac{x_p x_k}{\lambda_p x_p + \lambda_k x_k} e^{-(x_p+x_k)/2} dx_p dx_k,$$

$$t_3 = \int_{\mathbb{R}_+^2} x_p x_k \exp\left(-\frac{1}{2} \left( x_p + x_k - \frac{\lambda_p x_p + \lambda_k x_k}{\lambda_n} \right)\right) E_1\left(\frac{\lambda_p x_p + \lambda_k x_k}{2\lambda_n}\right) dx_p dx_k.$$

Integrating firstly along  $x_k$  in  $t_2$  allow to rewrite  $t_2$  as

$$t_2 = \frac{8}{\lambda_k} - \frac{8\lambda_k}{\lambda_p^2} F_2\left(\frac{\lambda_k - \lambda_p}{\lambda_p}, 0\right), \quad (\text{IV.24})$$

where the function  $F_2(\cdot)$  is defined in (III.20). Now, let us compute  $t_3$  as

$$t_3 = \int_0^{+\infty} x_p \exp\left(-\frac{x_p}{2} \left(1 - \frac{\lambda_p}{\lambda_n}\right)\right) t_4 dx_p,$$

with  $t_4 = \int_{\mathbb{R}_+} x_k e^{-x_k(1-\lambda_k/\lambda_n)/2} E_1\left(\frac{\lambda_p x_p + \lambda_k x_k}{2\lambda_n}\right) dx_k$ .

By a change of variable,  $t_4$  is rewritten as

$$\begin{aligned} t_4 &= \frac{2\lambda_n}{\lambda_k^2} \int_{\frac{\lambda_p x_p}{2\lambda_n}}^{+\infty} (2\lambda_n t - \lambda_p x_p) e^{-\frac{2\lambda_n t - \lambda_p x_p}{2\lambda_k} (1 - \frac{\lambda_k}{\lambda_n})} E_1(t) dt, \\ &= \frac{2\lambda_n}{\lambda_k^2} e^{\lambda_p x_p / 2 / \lambda_k (1 - \lambda_k / \lambda_n)} \left[ 2\lambda_n F_1\left(\frac{\lambda_n}{\lambda_k} - 1, \frac{\lambda_p x_p}{2\lambda_n}\right) - \lambda_p x_p F_0\left(\frac{\lambda_n}{\lambda_k} - 1, \frac{\lambda_p x_p}{2\lambda_n}\right) \right]. \end{aligned}$$

and can be simplified with (III.19) and with  $F_0(a, x) = e^{-a} E_1(x)/a - E_1([a+1]x)/a$ . The simplified expression of  $t_4$  allows to rewrite  $t_3$  as

$$\begin{aligned} t_3 &= \frac{16\lambda_n^2}{\lambda_n - \lambda_k} \left( \frac{1}{\lambda_p^2 (\lambda_n - \lambda_k)} [\lambda_n^2 F_1(b_n, 0) - \lambda_k^2 F_1(b_k, 0)] \right. \\ &\quad \left. - \frac{1}{\lambda_n} + \frac{\lambda_k^2}{\lambda_n \lambda_p^2} F_2(b_k, 0) \right) \end{aligned}$$

where  $b_j = (\lambda_j - \lambda_p)/\lambda_p$ , for  $j = k, n$ . Finally, combining the previous result with (III.20), (IV.23) and (IV.24), one has

$$\begin{aligned} \delta_{pkn} &= \lambda_p \lambda_k \left[ \frac{\lambda_n (\lambda_p + \lambda_k) - 2\lambda_p \lambda_k}{(\lambda_n - \lambda_p)(\lambda_n - \lambda_k)(\lambda_p - \lambda_k)^2} \right. \\ &\quad - \frac{\lambda_n^2 \log \lambda_n}{(\lambda_n - \lambda_p)^2 (\lambda_n - \lambda_k)^2} + \frac{\lambda_p \lambda_k \log \lambda_k}{(\lambda_p - \lambda_k)^3} \frac{2\lambda_p \lambda_n - \lambda_p \lambda_k - \lambda_k^2}{\lambda_p (\lambda_n - \lambda_k)^2} \\ &\quad \left. - \frac{\lambda_p \lambda_k}{(\lambda_p - \lambda_k)^3} \frac{2\lambda_k \lambda_n - \lambda_p \lambda_k - \lambda_p^2}{\lambda_k (\lambda_n - \lambda_p)^2} \log \lambda_p \right]. \end{aligned}$$

Thanks to (IV.21), the previous equation provides (10) and (11). This concludes the proof of Theorem 2.

#### REFERENCES

- [1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley & Sons, ISBN 0-471-36091-0.
- [2] E. Conte and A. De Maio and G. Ricci, "Recursive Estimation of the Covariance Matrix of a Compound-Gaussian Process and Its Application to Adaptive CFAR Detection", *IEEE Trans. on SP*, vol. 50, no. 8, pp. 1908-1915, Aug. 2002.
- [3] E. Conte, A. De Maio and G. Ricci, "Covariance Matrix Estimation for Adaptive CFAR Detection in Compound-Gaussian Clutter", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 38, no. 2, pp. 415-426, Apr. 2002.
- [4] K. Yao, M.K. Simon and E. Biglieri, "A Unified Theory on Wireless Communication Fading Statistics based on SIRV", *Fifth IEEE workshop on SP advances in Wireless Communications*, Lisboa Portugal, Jul. 2004.
- [5] T. J. Barnard and F. Khan, "Statistical Normalization of Spherically Invariant Non-Gaussian Clutter", *IEEE Journal of Oceanic Engineering*, vol. 29, no. 2, pp. 303-309, Apr. 2004.
- [6] K. Yao, "A representation theorem and its application to spherically invariant random processes", *IEEE Trans. on IT*, vol. 19, pp. 600-608, Jul. 1973.
- [7] E. Conte, M. Lops and G. Ricci, "Adaptive Radar Detection in Compound-Gaussian Clutter", *Proceedings of Eusipco'94*, Edinburgh, Scotland, pp. 526-529, Sep. 1994.
- [8] M. Abramowitz, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Edited by Milton Abramowitz and Irene A. Stegun, ISBN 04866127724.



## Annexe D

# **Preuves des théorèmes 3.1.3.1 et 3.1.3.2 : "Covariance Structure Maximum Likelihood Estimates in Compound Gaussian Noise : Existence and Algorithm Analysis"**

Afin de bien comprendre le détail des preuves, l'article les contenant est joint dans sa totalité. En effet, de nombreuses notations ont été introduites ainsi que plusieurs lemmes essentiels aux preuves d'existence et d'unicité de l'estimateur du point fixe.

# Covariance Structure Maximum Likelihood Estimates in Compound Gaussian Noise : Existence and Algorithm Analysis

Frédéric Pascal, Yacine Chitour, Jean-Philippe Ovarlez, Philippe Forster and Pascal Larzabal

**Abstract**—Recently, a new adaptive scheme [1], [2] has been introduced for covariance structure matrix estimation in the context of adaptive radar detection under non Gaussian noise. This latter has been modelled by compound Gaussian noise, which is the product  $\mathbf{c}$  of the square root of a positive unknown variable  $\tau$  (deterministic or random) and an independent Gaussian vector  $\mathbf{x}$ ,  $\mathbf{c} = \sqrt{\tau} \mathbf{x}$ .

Because of the implicit algebraic structure of the equation to solve, we called the corresponding solution, the Fixed Point (FP) estimate. When  $\tau$  is assumed deterministic and unknown, the FP is the exact Maximum Likelihood (ML) estimate of the noise covariance structure, while when  $\tau$  is a positive random variable, the FP is an Approximate Maximum Likelihood (AML).

This estimate has been already used for its excellent statistical properties without proofs of its existence and uniqueness. The major contribution of this paper is to fill these gaps. Our derivation is based on some Likelihood functions general properties like homogeneity and can be easily adapted to other recursive contexts. Moreover, the corresponding iterative algorithm used for the FP estimate practical determination is also analyzed and we show the convergence of this recursive scheme, ensured whatever the initialization.

**Index Terms**—Compound-Gaussian, SIRV, Maximum likelihood estimate, adaptive detection, CFAR detector.

## I. INTRODUCTION

THE basic problem of detecting a complex signal embedded in an additive Gaussian noise has been extensively studied these last decades. In these contexts, adaptive detection schemes required an estimate of the noise covariance matrix generally obtained from signal free data traditionally called secondary data or reference data. The resulting adaptive detectors, as those proposed by [7] and [8], are all based on the Gaussian assumption for which the Maximum Likelihood (ML) estimate of the covariance matrix is given by the sample

F. Pascal is with the Office National d'Etudes et de Recherches Aéropatiales, DEMR/TSI, BP 72, 92322 Chatillon Cedex, France (e-mail: frederic.pascal@onera.fr).

Y. Chitour is with the Laboratoire des Signaux et Systèmes, Supélec, 3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France (e-mail: yacine.chitour@lss.supelec.fr)

J. P. Ovarlez is with the Office National d'Etudes et de Recherches Aéropatiales, DEMR/TSI, BP 72, 92322 Chatillon Cedex, France (e-mail: jean-philippe.ovarlez@onera.fr).

P. Forster is with the Groupe d'Electromagnétisme Appliqué (GEA), Institut Universitaire de Technologie de Ville d'Avray, 92410 Ville d'Avray, France (e-mail: philippe.forster@cva.u-paris10.fr).

P. Larzabal is with the IUT de Cachan, C.R.I.P., Université Paris Sud, 94234 Cachan Cedex, France, and also with the SATIE, ENS Cachan, UMR CNRS 8029, 94235 Cachan Cedex, France (e-mail: larzabal@satie.ens-cachan.fr).

covariance matrix. However, these detectors may exhibit poor performance when the additive noise is no more Gaussian [6].

This is the case in radar detection problems where the additive noise is due to the superposition of unwanted echoes reflected by the environment and traditionally called the clutter. Indeed, experimental radar clutter measurements showed that these data are non-Gaussian. This fact arises for example when the illuminated area is non-homogeneous or when the number of scatterers is small. This kind of non-Gaussian noises is usually described by distributions such as K-distribution, Weibull, ... Therefore, this non-Gaussian noise characterization has gained a lot of interest in the radar detection community.

One of the most general and elegant non-Gaussian noise model is provided by the compound-Gaussian process which includes the so-called *Spherically Invariant Random Vectors* (SIRV). These processes encompass a large number of non-Gaussian distributions mentioned above and include of course Gaussian processes. They have been recently introduced, in radar detection, to model clutter for solving the basic problem of detecting a known signal. This approach resulted in the adaptive detectors development such as the Generalized Likelihood Ratio Test-Linear Quadratic (GLRT-LQ) in [1], [2] or the Bayesian Optimum Radar Detector (BORD) in [3], [4]. These detectors require an estimate of the SIRV covariance matrix. In this context, ML estimates based on secondary data have been introduced in [11], [12], together with a numerical procedure supposed to obtain them. However, as noticed in [12] p.1852, "existence of the ML estimate and convergence of iteration [...] is still an open problem".

To the best of our knowledge, the proofs of existence, uniqueness of the ML estimate and convergence of the algorithm proposed in [1] have never been established. The main purpose of this paper is to fill these gaps.

The paper is organized as follows. In the Section II, we present the two main models of interest in our ML estimation framework. Both models lead to ML estimates which are solution of a transcendental equation. Section IV presents the main results of this paper while a proofs outline is given in Section V: for presentation clarity, full demonstrations are provided in Appendices. Finally, Section VI gives some simulations results which confirm the theoretical analysis.

## II. STATE OF THE ART AND PROBLEM FORMULATION

A compound-Gaussian process  $\mathbf{c}$  is the product of the square root of a positive scalar quantity  $\tau$  called the texture and a

$m$ -dimensional zero mean complex Gaussian vector  $\mathbf{x}$  with covariance matrix  $\mathbf{M} = \mathbb{E}(\mathbf{x}\mathbf{x}^H)$  usually normalized according to  $\text{Tr}(\mathbf{M}) = m$ , where  $H$  denotes the conjugate transpose operator and  $\text{Tr}(\cdot)$  stands for the trace operator:

$$\mathbf{c} = \sqrt{\tau} \mathbf{x}. \quad (1)$$

This general model leads to two distinct approaches: the well-known SIRV modeling where the texture is considered random and the case where the texture is treated as an unknown nuisance parameter.

Generally, the covariance matrix  $\mathbf{M}$  is not known and an estimate  $\hat{\mathbf{M}}$  is required for the Likelihood Ratio (LR) computation. Classically, such an estimate  $\hat{\mathbf{M}}$  is obtained from Maximum Likelihood (ML) theory, well known for its good statistical properties. In this problem, estimation of  $\mathbf{M}$  must respect the previous  $\mathbf{M}$ -normalization,  $\text{Tr}(\hat{\mathbf{M}}) = m$ . This estimate  $\hat{\mathbf{M}}$  will be built using  $N$  independent realizations of  $\mathbf{c}$  denoted  $\mathbf{c}_i = \sqrt{\tau_i} \mathbf{x}_i$  for  $i = 1, \dots, N$ .

It straightforwardly appears that the Likelihood will depend on the assumption relative to texture. The two most often met cases are presented in the two following subsections.

#### A. SIRV case

Let us recap that a SIRV [5] is the product of the square root of a positive random variable  $\tau$  (*texture*) and a  $m$ -dimensional independent complex Gaussian vector  $\mathbf{x}$  (*speckle*) with zero mean normalized covariance matrix  $\mathbf{M}$ . This model led to many investigations [1], [2], [3], [4].

To obtain the ML estimate of  $\mathbf{M}$ , with no proofs of existence and uniqueness, Gini *et al.* derived in [12] an Approximate Maximum Likelihood (AML) estimate  $\hat{\mathbf{M}}$  as the solution of the following equation

$$\hat{\mathbf{M}} = f(\hat{\mathbf{M}}), \quad (2)$$

where  $f$  is given by

$$f(\hat{\mathbf{M}}) = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{c}_i \mathbf{c}_i^H}{\mathbf{c}_i^H \hat{\mathbf{M}}^{-1} \mathbf{c}_i}. \quad (3)$$

#### B. Unknown deterministic $\tau$ case

This approach has been developed in [13] where the  $\tau_i$ 's are assumed to be unknown deterministic quantities. The corresponding Likelihood function to maximize with respect to  $\mathbf{M}$  and  $\tau_i$ 's, is given by

$$p_C(\mathbf{c}_1, \dots, \mathbf{c}_N; \mathbf{M}, \tau_1, \dots, \tau_N) = \frac{1}{(\pi)^{mN} |\mathbf{M}|^N} \times \prod_{i=1}^N \frac{1}{\tau_i^m} \exp\left(-\frac{\mathbf{c}_i^\dagger \mathbf{M}^{-1} \mathbf{c}_i}{\tau_i}\right). \quad (4)$$

where  $|\mathbf{M}|$  denotes the determinant of matrix  $\mathbf{M}$ .

Maximization with respect to  $\tau_i$ 's, for a given  $\mathbf{M}$ , leads to  $\hat{\tau}_i = \frac{\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i}{m}$ , and then by replacing the  $\tau_i$ 's in (4) by their ML estimates  $\hat{\tau}_i$ 's, we obtain the reduced likelihood function

$$\hat{p}_C(\mathbf{c}_1, \dots, \mathbf{c}_N; \mathbf{M}) = \frac{1}{(\pi)^{mN} |\mathbf{M}|^N} \prod_{i=1}^N \frac{m^m \exp(-m)}{(\mathbf{c}_i^H \mathbf{M}^{-1} \mathbf{c}_i)^m}.$$

Finally, maximizing  $\hat{p}_C(\mathbf{c}_1, \dots, \mathbf{c}_N; \mathbf{M})$  with respect to  $\mathbf{M}$  is equivalent to maximize the following function  $F$ , written in terms of  $\mathbf{x}_i$ 's and  $\tau_i$ 's thanks to (1)

$$F(\mathbf{M}) = \frac{1}{|\mathbf{M}|^N} \prod_{i=1}^N \frac{1}{\tau_i^m (\mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i)^m}. \quad (5)$$

By cancelling the gradient of  $F$  with respect to  $\mathbf{M}$ , we obtain the following equation

$$\hat{\mathbf{M}} = f(\hat{\mathbf{M}}), \quad (6)$$

where  $f$  is given again by (3).

Note that  $f$  can be rewritten from (1) as

$$f(\hat{\mathbf{M}}) = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \hat{\mathbf{M}}^{-1} \mathbf{x}_i}. \quad (7)$$

Equation (7) shows that  $f(\hat{\mathbf{M}})$  does not depend on the texture  $\tau$  but only on the Gaussian vectors  $\mathbf{x}_i$ 's.

#### C. Problem Formulation

It has been shown in [12], [13] that estimation schemes developed under both the stochastic case (Section II-A) and the deterministic case (Section II-B) lead to the analysis of the same equation ((2) and (6)), whose solution is a fixed point of  $f$  (7). A first contribution of this paper is to establish the existence and the uniqueness, up to a scalar factor, of this fixed point  $\hat{\mathbf{M}}_{FP}$  which is the Approximate Maximum Likelihood (AML) estimate under the stochastic assumption and the exact ML under the deterministic assumption.

Moreover, a second contribution is to analyze an algorithm based on the key equation (6), which defines  $\hat{\mathbf{M}}_{FP}$ . The convergence of this algorithm will be established. Then, numerical results of Section VI will illustrate the computational efficiency of the algorithm for obtaining the FP estimate.

Finally, the complete statistical properties investigation of the corresponding ML estimate will be addressed in a forthcoming paper.

### III. STATEMENT OF THE MAIN RESULT

We first provide some notations. Let  $m$  and  $N$  be positive integers such that  $m < N$ . We use  $\mathbb{R}^{+*}$  to denote the set of strictly positive real scalars,  $M_m(\mathbb{C})$  to denote the set of  $m \times m$  complex matrices, and  $\mathcal{G}$ , the subset of  $M_m(\mathbb{C})$  defined by the positive definite Hermitian matrices. For  $\mathbf{M} \in M_m(\mathbb{C})$ ,  $\|\mathbf{M}\| := \text{Tr}(\mathbf{M}^H \mathbf{M})^{1/2}$  the Frobenius norm of  $\mathbf{M}$  which is the norm associated to an inner product on  $M_m(\mathbb{C})$ . Moreover, from the statistical independence hypothesis of the  $N$  complex  $m$ -vectors  $\mathbf{x}_i$ , it is natural to assume the following

(H): Let us set  $\mathbf{x}_i = \mathbf{x}_i^{(1)} + j\mathbf{x}_i^{(2)}$ . Any  $2m$  distinct vectors

taken in

$$\left\{ \begin{pmatrix} \mathbf{x}_1^{(1)} \\ \mathbf{x}_1^{(2)} \end{pmatrix}, \dots, \begin{pmatrix} \mathbf{x}_N^{(1)} \\ \mathbf{x}_N^{(2)} \end{pmatrix}, \begin{pmatrix} -\mathbf{x}_1^{(2)} \\ \mathbf{x}_1^{(1)} \end{pmatrix}, \dots, \begin{pmatrix} -\mathbf{x}_N^{(2)} \\ \mathbf{x}_N^{(1)} \end{pmatrix} \right\}$$

are linearly independent.

From (5) and (7), one has

$$F : \mathcal{G} \longrightarrow \mathbb{R}^{+*}$$

$$\mathbf{M} \longrightarrow F(\mathbf{M}) = \frac{1}{|\mathbf{M}|^N} \prod_{i=1}^N \frac{1}{\tau_i^m (\mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i)^m},$$

and

$$f : \mathcal{G} \longrightarrow \mathcal{G}$$

$$\mathbf{M} \longrightarrow f(\mathbf{M}) = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i}.$$

### Theorem III.1

- (i) *There exists  $\widehat{\mathbf{M}}_{FP} \in \mathcal{G}$  with unit norm such that, for every  $\alpha > 0$ ,  $f$  admits a unique fixed point of norm  $\alpha > 0$  equal to  $\alpha \widehat{\mathbf{M}}_{FP}$ . Moreover,  $F$  reaches its maximum over  $\mathcal{G}$  only on  $\mathcal{L}_{\widehat{\mathbf{M}}_{FP}}$ , the open half-line spanned by  $\widehat{\mathbf{M}}_{FP}$ .*
- (ii) *Let  $(S)_{dis}$  be the discrete dynamical system defined on  $\mathcal{D}$  by*

$$(S)_{dis} : \mathbf{M}_{k+1} = f(\mathbf{M}_k). \quad (8)$$

*Then, for every initial condition  $\mathbf{M}_0 \in \mathcal{G}$ , the resulting sequence  $(\mathbf{M}_k)_{k \geq 0}$  converges to a fixed point of  $f$ , i.e. to a point where  $F$  reaches its maximum;*

- (iii) *Let  $(S)_{cont}$  be the continuous dynamical system defined on  $\mathcal{G}$  by*

$$(S)_{cont} : \dot{\mathbf{M}} = \nabla F(\mathbf{M}). \quad (9)$$

*Then, for every initial condition  $\mathbf{M}(0) = \mathbf{M}_0 \in \mathcal{G}$ , the resulting trajectory  $\mathbf{M}(t)$ ,  $t \geq 0$ , converges when  $t$  tends to  $+\infty$ , to the point  $\|\mathbf{M}_0\| \widehat{\mathbf{M}}_{FP}$ , i.e. to a point where  $F$  reaches its maximum.*

Consequently to (i),  $\widehat{\mathbf{M}}_{FP}$  is the unique positive definite  $m \times m$  matrix of norm one satisfying

$$\widehat{\mathbf{M}}_{FP} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{M}}_{FP}^{-1} \mathbf{x}_i}. \quad (10)$$

*Proof:* The same problem and the same result can be formulated with real numbers instead of complex numbers and symmetric matrices instead of hermitian matrices, while hypothesis (H) becomes hypothesis (H2) stated below (just before Remark IV.1). The proof of Theorem III.1 breaks up into two stages. We first show in Appendix I how to derive Theorem III.1 from the corresponding real results. Then, the rest of the paper is devoted to the study of the real case. ■

## IV. NOTATIONS AND STATEMENTS OF THE RESULTS IN THE REAL CASE

### A. Notations

In this paragraph, we introduce the main notations of the paper for the real case. Notations already defined in the complex case are translated in the real one. Moreover, real results will be valid for every integer  $m$ . For every positive integer  $n$ ,  $\llbracket 1, n \rrbracket$  denotes the set of integers  $\{1, \dots, n\}$ . For vectors of  $\mathbb{R}^m$ , the norm used is the Euclidean one. Throughout the paper, we will use several basic results on square matrices, especially regarding diagonalization of real symmetric and orthogonal matrices. We refer to [14] for such standard results.

We use  $M_m(\mathbb{R})$  to denote the set of  $m \times m$  real matrices,  $SO(m)$  to denote the set of  $m \times m$  orthogonal matrices and  $\mathbf{M}^\top$ , the transpose of  $\mathbf{M}$ . We denote the identity matrix of  $M_m(\mathbb{R})$  by  $\mathbf{I}_m$ .

We next define and list the several sets of matrices used in the sequel:

- \*  $\mathcal{D}$ , the subset of  $M_m(\mathbb{R})$  defined by the symmetric positive definite matrices;
- \*  $\overline{\mathcal{D}}$ , the closure of  $\mathcal{D}$  in  $M_m(\mathbb{R})$ , i.e. the subset of  $M_m(\mathbb{R})$  defined by the symmetric non negative matrices;
- \* For every  $\alpha > 0$ ,  $\begin{cases} \mathcal{D}(\alpha) = \{\mathbf{M} \in \mathcal{D} \mid \|\mathbf{M}\| = \alpha\} \\ \overline{\mathcal{D}}(\alpha) = \{\mathbf{M} \in \overline{\mathcal{D}} \mid \|\mathbf{M}\| = \alpha\} \end{cases}$ .

It is obvious that  $\overline{\mathcal{D}}(\alpha)$  is compact in  $M_m(\mathbb{R})$ .

For  $\mathbf{M} \in \mathcal{D}$ , we use  $\mathcal{L}_{\mathbf{M}}$  to denote the open-half line spanned by  $\mathbf{M}$  in the cone  $\mathcal{D}$ , i.e. the set of points  $\lambda \mathbf{M}$ , with  $\lambda > 0$ . Recall that the order associated with the cone structure of  $\mathcal{D}$  is called the Loewner order for symmetric matrices of  $M_m(\mathbb{R})$  and is defined as follows. Let  $\mathbf{A}, \mathbf{B}$  be two symmetric  $m \times m$  real matrices. Then  $\mathbf{A} \leq \mathbf{B}$  ( $\mathbf{A} < \mathbf{B}$  respectively) means that the quadratic form defined by  $\mathbf{B} - \mathbf{A}$  is non negative (positive definite respectively), i.e., for every non zero  $\mathbf{x} \in \mathbb{R}^m$ ,  $\mathbf{x}^\top (\mathbf{A} - \mathbf{B}) \mathbf{x} \leq 0$ , ( $> 0$  respectively). Using that order, one has  $\mathbf{M} \in \mathcal{D}$  ( $\in \overline{\mathcal{D}}$  respectively) if and only if  $\mathbf{M} > \mathbf{0}$  ( $\mathbf{M} \geq \mathbf{0}$  respectively).

As explained in Appendix I, we will study in this section the applications  $F$  and  $f$  (same notations as in the complex case) defined as follows:

$$F : \mathcal{D} \longrightarrow \mathbb{R}^{+*}$$

$$\mathbf{M} \longrightarrow \frac{1}{|\mathbf{M}|^N} \prod_{i=1}^N \frac{1}{\tau_i^m (\mathbf{x}_i^\top \mathbf{M}^{-1} \mathbf{x}_i)^m},$$

and

$$f : \mathcal{D} \longrightarrow \mathcal{D}$$

$$\mathbf{M} \longrightarrow \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^\top}{\mathbf{x}_i^\top \mathbf{M}^{-1} \mathbf{x}_i}.$$

Henceforth,  $F$  and  $f$  stay for the real formulation. In the above, the vectors  $(\mathbf{x}_i)$ ,  $1 \leq i \leq N$ , belong to  $\mathbb{R}^m$  and verify the next two hypothesis:

- (H1) :  $\|\mathbf{x}_i\| = 1, 1 \leq i \leq N$ ;
- (H2) : For any  $m$  two by two distinct indices  $i(1) < \dots < i(m)$  chosen in  $\llbracket 1, N \rrbracket$ , the vectors  $\mathbf{x}_{i(1)}, \dots, \mathbf{x}_{i(m)}$  are linearly independent.

Consequently, the vectors  $\mathbf{c}_1, \dots, \mathbf{c}_m$  verify (H2).

Hypothesis (H1) stems from the fact that function  $f$  does not depend on  $\mathbf{x}_i$ 's norm.

Let us already emphasize that hypothesis (H2) is the key assumption for getting all our subsequent results. Hypothesis (H2) has the following trivial but fundamental consequence that we state as a remark.

**Remark IV.1**

For every  $n$  vectors  $\mathbf{x}_{i(1)}, \dots, \mathbf{x}_{i(n)}$  (respectively  $\mathbf{c}_{i(1)}, \dots, \mathbf{c}_{i(n)}$ ) with  $1 \leq n \leq m, 1 \leq i \leq N$ , the vector space generated by  $\mathbf{x}_{i(1)}, \dots, \mathbf{x}_{i(n)}$  (respectively  $\mathbf{c}_{i(1)}, \dots, \mathbf{c}_{i(n)}$ ) has dimension  $n$ .

In the sequel, we use  $f^n, n \geq 1$ , to denote the  $n$ -th iterate of  $f$  i.e.,  $f^n := f \circ \dots \circ f$ , where  $f$  is repeated  $n$  times. We also adopt the following standard convention  $f^0 := Id_{\mathcal{D}}$ .

The two functions  $F$  and  $f$  are related by the following relation, which is obtained after an easy computation. For every  $\mathbf{M} \in \mathcal{D}$ , let  $\nabla F(\mathbf{M})$  be the gradient of  $F$  at  $\mathbf{M} \in \mathcal{D}$  i.e. the unique symmetric matrix verifying, for every matrix  $M \in \mathcal{S}$ ,

$$\nabla F(\mathbf{M}) = N F(\mathbf{M}) \mathbf{M}^{-1} (f(\mathbf{M}) - \mathbf{M}) \mathbf{M}^{-1}.$$

Clearly  $\mathbf{M}$  is a fixed point of  $f$  if and only if  $M$  is a critical point of the vector field defined by  $\nabla F$  on  $\mathcal{D}$ .

**B. Statements of the results**

The goal of this paper is to establish the following theorems whose proofs are outlined in the next Section.

**Theorem IV.1**

There exists  $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}$  with unit norm such that, for every  $\alpha > 0$ ,  $f$  admits a unique fixed point of norm  $\alpha > 0$  equal to  $\alpha \widehat{\mathbf{M}}_{FP}$ . Moreover,  $F$  reaches its maximum over  $\mathcal{D}$  only on  $\mathcal{L}_{\widehat{\mathbf{M}}_{FP}}$ , the open half-line spanned by  $\widehat{\mathbf{M}}_{FP}$ .

Consequently,  $\widehat{\mathbf{M}}_{FP}$  is the unique positive definite  $m \times m$  matrix of norm one satisfying

$$\widehat{\mathbf{M}}_{FP} = \frac{m}{N} \sum_{i=1}^N \frac{\mathbf{x}_i \mathbf{x}_i^\top}{\mathbf{x}_i^\top \widehat{\mathbf{M}}_{FP}^{-1} \mathbf{x}_i}. \quad (11)$$

**Remark IV.2**

Theorem IV.1 relies on the fact that  $F$  reaches its maximum on  $\mathcal{D}$ . Roughly speaking, that issue is proved as follows. The function  $F$  is continuously extended by the zero function on the boundary of  $\mathcal{D}$ , excepted on the zero matrix. Since  $F$  is positive and bounded on  $\mathcal{D}$ , we conclude. Complete argument is provided in Appendix II.

As a consequence of Theorem IV.1, one obtains the next result.

**Theorem IV.2**

• Let  $(S)_{dis}$  be the discrete dynamical system defined on  $\mathcal{D}$  by

$$(S)_{dis} : \mathbf{M}_{k+1} = f(\mathbf{M}_k). \quad (12)$$

Then, for every initial condition  $\mathbf{M}_0 \in \mathcal{D}$ , the resulting sequence  $(\mathbf{M}_k)_{k \geq 0}$  converges to a fixed point of  $f$ , i.e. to a

point where  $F$  reaches its maximum;

• Let  $(S)_{cont}$  be the continuous dynamical system defined on  $\mathcal{D}$  by

$$(S)_{cont} : \dot{\mathbf{M}} = -\nabla F(\mathbf{M}). \quad (13)$$

Then, for every initial condition  $\mathbf{M}(0) = \mathbf{M}_0 \in \mathcal{D}$ , the resulting trajectory  $\mathbf{M}(t), t \geq 0$ , converges, when  $t$  tends to  $+\infty$ , to the point  $\|\mathbf{M}_0\| \widehat{\mathbf{M}}_{FP}$ , i.e. to a point where  $F$  reaches its maximum.

The last theorem can be used to characterize numerically the points where  $F$  reaches its maximum and the value of that maximum.

Notice that algorithm defined by (12) does not allow the control of the FP norm. Therefore, for practical convenient, we propose a slightly modified algorithm in which the  $\mathbf{M}$ -normalization is applied at each iteration. This is summarized in the following corollary:

**Corollary IV.1**

The following scheme

$$\mathbf{M}'_{k+1} = \frac{f(\mathbf{M}'_k)}{\text{Tr}(f(\mathbf{M}'_k))}. \quad (14)$$

yields the matrices sequence  $\{\mathbf{M}'_0, \dots, \mathbf{M}'_k\}$  which is related to the matrices sequence  $\{\mathbf{M}_0, \dots, \mathbf{M}_k\}$ , provided by (12), by, for  $1 \leq i \leq k$ ,

$$\mathbf{M}'_i = \frac{\mathbf{M}_i}{\text{Tr}(\mathbf{M}_i)}.$$

This algorithm converges to  $\widehat{\mathbf{M}}_{FP}$  up to a scaling factor which is:  $\frac{1}{\text{Tr}(\widehat{\mathbf{M}}_{FP})}$ .

As a consequence of Theorem IV.1, we can prove a matrix inequality which is interesting on its own. It simply expresses that the Hessian computed at a critical point of  $F$  is non positive. We also provide an example showing that, in general, the Hessian is not definite negative. Therefore, in general, the convergence rate to the critical points of  $F$  for the dynamical systems  $(S)_{dis}$  and  $(S)_{cont}$  is not exponential.

**Proposition IV.1**

Let  $m, N$  be two positive integers with  $m < N$  and  $\mathbf{x}_1, \dots, \mathbf{x}_N$  be unit vectors of  $\mathbb{R}^m$  subject to (H2) and such that

$$\frac{m}{N} \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^\top = \mathbf{I}_m. \quad (15)$$

Then, for every matrix  $\mathbf{M}$  of  $M_m(\mathbb{R})$ , we have

$$\frac{m}{N} \sum_{i=1}^N (\mathbf{x}_i^\top \mathbf{M} \mathbf{x}_i)^2 \leq \|\mathbf{M}\|^2. \quad (16)$$

Assuming Theorem IV.1, the proof of the proposition is short enough to be provided next.

We may assume  $\mathbf{M}$  to be symmetric since it is enough to prove the result for  $(\mathbf{M} + \mathbf{M}^T)/2$ , the symmetric part of  $\mathbf{M}$ . Applying Theorem IV.1, it is clear that the function  $F$  associated to the  $\mathbf{x}_i$ 's reaches its maximum over  $\mathcal{D}$  at  $\mathbf{I}_m$ . The

expression of  $H_{\mathbf{I}_m}$ , the Hessian of  $F$  at  $\mathbf{I}_m$  is the following. For every symmetric matrix  $\mathbf{M}$ , we have

$$H_{\mathbf{I}_m}(\mathbf{M}, \mathbf{M}) = N F(\mathbf{I}_m) \left( \frac{m}{N} \sum_{i=1}^N (\mathbf{x}_i^\top \mathbf{M} \mathbf{x}_i)^2 - \|\mathbf{M}\|^2 \right).$$

Since  $H_{\mathbf{I}_m}$  is non positive, (16) follows. Note that a similar formula can be given if, instead of (15), the  $\mathbf{x}_i$ 's verify the more general equation (11).

Because of the homogeneity properties of  $F$  and  $f$  and in order to prove that the rates of convergence of both  $(S)_{dis}$  and  $(S)_{cont}$  are not exponential, one must prove that the Hessian  $H_{\mathbf{I}_m}$  is not negative definite on the orthogonal to  $\mathbf{I}_m$  in the set of all symmetric matrices. The latter is simply the set of symmetric matrices with null trace. We next provide a numerical example describing that situation. Here,  $m = 3$ ,  $N = 4$  and

$$\mathbf{x}_1 = \begin{pmatrix} 2\sqrt{2} \\ 3 \\ 0 \\ 1 \\ 3 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} -\sqrt{2} \\ 3 \\ \sqrt{3} \\ 1 \\ 3 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} -\sqrt{2} \\ 3 \\ -\sqrt{3} \\ 1 \\ 3 \end{pmatrix}, \quad \mathbf{x}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Then, hypotheses (H1), (H2) and (15) are satisfied. Moreover, it is easy to see that, for every diagonal matrix  $\mathbf{D}$ , we have equality in (16).

## V. PROOFS OUTLINE

In that Section, we give Theorem IV.1 proof and Theorem IV.2 one. Each proof is decomposed in a sequence of lemmas and propositions whose arguments are postponed in the Appendices.

### A. Proof of Theorem IV.1

Theorem conclusions are the consequences of several propositions whose statements are listed below.

First of all, it is clear that  $F$  is homogeneous of degree zero and  $f$  is homogeneous of degree one, i.e., for every  $\lambda > 0$  and  $\mathbf{M} \in \mathcal{D}$ , one has

$$F(\lambda \mathbf{M}) = F(\mathbf{M}), \quad f(\lambda \mathbf{M}) = \lambda f(\mathbf{M}).$$

The first proposition is the following.

#### Proposition V.1

The supremum of  $F$  over  $\mathcal{D}$  is finite and is reached at a point  $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}$  with  $\|\widehat{\mathbf{M}}_{FP}\| = 1$ . Therefore,  $f$  admits the open-half line  $\mathcal{L}_{\widehat{\mathbf{M}}_{FP}}$  as fixed points.

*Proof:* See Appendix II ■

It remains to show that there are no other fixed points of  $f$  except  $\mathcal{L}_{\widehat{\mathbf{M}}_{FP}}$ . For that purpose, one must study the function  $f$ . We first establish the following result.

#### Proposition V.2

The function  $f$  verifies the following properties.

- (P1) : For every  $\mathbf{M}, \mathbf{Q} \in \mathcal{D}$ , if  $\mathbf{M} \leq \mathbf{Q}$ , then  $f(\mathbf{M}) \leq f(\mathbf{Q})$  (also true with strict inequalities);

- (P2) : for every  $\mathbf{M}, \mathbf{Q} \in \mathcal{D}$ , then

$$f(\mathbf{M} + \mathbf{Q}) \geq f(\mathbf{M}) + f(\mathbf{Q}), \quad (17)$$

and equality occurs if and only if  $\mathbf{M}$  and  $\mathbf{Q}$  are colinear.

*Proof:* See Appendix III ■

The property of  $f$  described in the next proposition turns out to be basic for the proofs of both theorems.

#### Proposition V.3

The function  $f$  is eventually strictly increasing, i.e. for every  $\mathbf{Q}, \mathbf{P} \in \mathcal{D}$  such that  $\mathbf{Q} \geq \mathbf{P}$  and  $\mathbf{Q} \neq \mathbf{P}$ , then  $f^m(\mathbf{Q}) > f^m(\mathbf{P})$ .

*Proof:* See Appendix IV ■

We next proceed by establishing another property of  $f$ , which can be seen as an intermediary step towards the conclusion.

Recall that the orbit of  $f$  associated to  $\mathbf{M} \in \mathcal{D}$  is the trajectory of  $(S)_{dis}$  (12) starting at  $\mathbf{M}$ .

#### Proposition V.4

The following statements are equivalent.

- (A)  $f$  admits a fixed point;
- (B)  $f$  has one bounded orbit in  $\mathcal{D}$ ;
- (C) every orbit of  $f$  is bounded in  $\mathcal{D}$ .

*Proof:* See Appendix V ■

From proposition V.1,  $f$  admits a fixed point. Thus, proposition V.4 ensures that every orbit of  $f$  is bounded in  $\mathcal{D}$ .

Finally, using Proposition V.3, we get the following corollary, which concludes the proof of Theorem IV.1.

#### Corollary V.1

Assume that every orbit of  $f$  is bounded in  $\mathcal{D}$ . The following holds true.

- (C1) : Let  $\mathbf{P} \in \mathcal{D}$  and  $n \geq 1$  such that  $\mathbf{P}$  can be compared with  $f^n(\mathbf{P})$ , i.e.  $\mathbf{P} \geq f^n(\mathbf{P})$  or  $\mathbf{P} \leq f^n(\mathbf{P})$ . Then,  $\mathbf{P} = f^n(\mathbf{P})$ . In particular, if  $\mathbf{P} \geq f(\mathbf{P})$  or  $\mathbf{P} \leq f(\mathbf{P})$ , then  $\mathbf{P}$  is a fixed point of  $f$ ;
- (C2) : All the fixed points of  $f$  are colinear.

*Proof:* See Appendix VI ■

To summarize, proposition V.1 establishes the existence of a fixed point while corollary V.1 ensures the uniqueness of the unit norm fixed point.

### B. Proof of Theorem IV.2

1) *Convergence results for  $(S)_{dis}$ :* In the previous Section, we already proved several important facts relative to the trajectories of  $(S)_{dis}$  defined by (12), i.e. the orbits of  $f$ . Indeed, since  $f$  has fixed points, then all the orbits of  $f$  are bounded in  $\mathcal{D}$ . It remains to show now that each of them is convergent to a fixed point of  $f$ .

For that purpose, we consider, for every  $\mathbf{M} \in \mathcal{D}$ , the positive limit set  $\omega(\mathbf{M})$  associated to  $\mathbf{M}$ , i.e., the set made of the cluster points of the sequence  $(\mathbf{M}_k)_{k \geq 0}$ , where  $\mathbf{M}_{k+1} = f(\mathbf{M}_k)$  with  $\mathbf{M}_0 = \mathbf{M}$ . Since the orbit of  $f$  associated to  $\mathbf{M}$  is bounded in  $\mathcal{D}$ , the set  $\omega(\mathbf{M})$  is a compact of  $\mathcal{D}$  and is invariant by  $f$ : for every  $\mathbf{P} \in \omega(\mathbf{M})$ ,  $f(\mathbf{P}) \in \omega(\mathbf{M})$ . It is clear that the sequence

$(\mathbf{M}_k)_{k \geq 0}$  converges if and only if  $\omega(\mathbf{M})$  reduces to a single point.

The last part of the proof is divided into two lemmas, whose statements are given below.

**Lemma V.1**

For every  $\mathbf{M} \in \mathcal{D}$ ,  $\omega(\mathbf{M})$  contains a periodic orbit of  $f$  (i.e. contain a finite number of points).

*Proof:* See Appendix VII ■

**Lemma V.2**

Let  $\mathbf{M}_1$  and  $\mathbf{M}_2 \in \mathcal{D}$  be such that their respective orbits are periodic. Then  $\mathbf{M}_1$  and  $\mathbf{M}_2$  are colinear and are both fixed points of  $f$ .

*Proof:* See Appendix VIII ■

We now complete the proof of theorem IV.2 in the discrete case.

Let  $\mathbf{M} \in \mathcal{D}$ . Using both lemmas, it is easy to deduce that  $\omega(\mathbf{M})$  contains a fixed point of  $f$ , which will be denoted by  $\mathbf{Q}$ . Notice that there exists a compact  $\mathcal{K}$  containing both the orbit of  $f$  associated to  $\mathbf{M}$  and  $\omega(\mathbf{M})$ . We next prove that, for every  $\varepsilon > 0$ , there exists a positive integer  $n_\varepsilon > 0$  such that

$$(1 - \varepsilon) \mathbf{Q} \leq f^{n_\varepsilon}(\mathbf{M}) \leq (1 + \varepsilon) \mathbf{Q}. \quad (18)$$

Indeed, since  $\mathbf{Q} \in \omega(\mathbf{M})$ , for every  $\varepsilon > 0$ , there exists a positive integer  $n_\varepsilon > 0$  such that

$$\|f^{n_\varepsilon}(\mathbf{M}) - \mathbf{Q}\| \leq \varepsilon.$$

After standard computations, one can see that there exists a constant  $K > 0$ , only depending on the compact  $\mathcal{K}$ , such that, for  $\varepsilon > 0$  small enough,

$$(1 - K\varepsilon) \mathbf{Q} \leq f^{n_\varepsilon}(\mathbf{M}) \leq (1 + K\varepsilon) \mathbf{Q}.$$

The previous inequality implies at once (18).

Applying  $f^l$ ,  $l \geq 0$ , to (18), and taking into account that  $\mathbf{Q}$  is a fixed point of  $f$ , one deduces that

$$(1 - \varepsilon) \mathbf{Q} \leq f^{l+n_\varepsilon}(\mathbf{M}) \leq (1 + \varepsilon) \mathbf{Q}.$$

This is nothing else but the definition of the convergence of the sequence  $(f^l(\mathbf{M}))_{l \geq 0}$  to  $\mathbf{Q}$ . ■

2) *Convergence results for  $(S)_{cont}$ :* Let  $t \rightarrow \mathbf{M}(t)$ ,  $t \geq 0$ , be a trajectory of  $(S)_{cont}$  with initial condition  $\mathbf{M}_0 \in \mathcal{D}$ .

Thanks to equation (II.26) which appears in the proof of proposition V.1 in Appendix II, we have for every trajectory  $\mathbf{M}(t)$  of  $(S)_{cont}$

$$\frac{d}{dt} \|\mathbf{M}\|^2 = 2 \operatorname{Tr}(\mathbf{M}\dot{\mathbf{M}}) = 2 \operatorname{Tr}(\nabla F(\mathbf{M}) \cdot \mathbf{M}) = 0.$$

Then, for every  $t \geq 0$ ,  $\mathbf{M}(t)$  keeps a constant norm equal to  $\|\mathbf{M}_0\|$ . Moreover, one has for every  $t \geq 0$

$$F(\mathbf{M}(t)) - F(\mathbf{M}(0)) = \int_0^t \frac{d}{dt} F(\mathbf{M}) = \int_0^t \|\nabla F(\mathbf{M})\|^2 > 0.$$

Since  $F$  is bounded over  $\mathcal{D}(\|\mathbf{M}_0\|)$ , we deduce that

$$\int_0^{+\infty} \|\nabla F(\mathbf{M})\|^2 < +\infty. \quad (19)$$

In addition, since  $t \rightarrow F(\mathbf{M}(t))$  is an increasing function, then  $\mathbf{M}(t)$  remains in a compact subset  $\mathcal{K}$  of  $\mathcal{D}(\|\mathbf{M}_0\|)$  which is independent of the time  $t$ . As  $\mathcal{D}(\|\mathbf{M}_0\|)$  contains a unique equilibrium point of  $(S)_{cont}$ , we proceed by proving theorem IV.2 in the continuous case

$$\forall \mathbf{M}_0 \in \mathcal{D}, \mathbf{M}(t) \xrightarrow[t \rightarrow +\infty]{} \|\mathbf{M}_0\| \widehat{\mathbf{M}}_{FP}. \quad (20)$$

Without loss of generality, we assume that  $\|\mathbf{M}_0\| = 1$ . Let  $F_0$  be the limit of  $F(\mathbf{M}(t))$  as  $t$  tends to  $+\infty$ . Thanks to Theorem IV.1 and the fact that  $\|\mathbf{M}(t)\|$  is constant, it is easy to see that (20) follows if one can show that  $F_0 = F(\widehat{\mathbf{M}}_{FP})$ . We assume the contrary and will reach a contradiction.

Indeed, if we assume that  $F_0 < F(\widehat{\mathbf{M}}_{FP})$  then there exists  $\varepsilon_0$  such that  $\|\mathbf{M}(t) - \widehat{\mathbf{M}}_{FP}\| \geq \varepsilon_0$ , for every  $t \geq 0$ . This implies together with the fact that  $\widehat{\mathbf{M}}_{FP}$  is the unique fixed point of  $f$  in  $\mathcal{D}(1)$  and  $\|\nabla F(\mathbf{M})\|$  is continuous, that there exists  $C_0$  such that  $\|\nabla F(\mathbf{M})\| \geq C_0$ , for every  $t \geq 0$ . Then,  $\int_{t_0}^{+\infty} \|\nabla F(\mathbf{M})\|^2 = +\infty$ , which contradicts (19). Therefore, (20) holds true. ■

## VI. SIMULATIONS

The main purpose of this section is to give some tools for computing of the FP estimate regardless of its statistical properties; in particular, we investigate the numerical accuracy and the algorithm convergence in different contexts for the complex case.

The two algorithms presented in section IV will be compared:

- the discrete case algorithm of theorem IV.2, called algorithm 1 in the sequel, defined by (12) and whose convergence to the FP estimate has been proved in Section V;
- the normalized algorithm, called algorithm 2 in the sequel, defined by (14).

The first purpose of simulations is to compare the two algorithms in order to choose the best one in terms of convergence speed.

Secondly, we study the parameters influence in the retained algorithm: the order  $m$  of matrix  $\mathbf{M}$ , the number  $N$  of reference data  $(\mathbf{c}_1, \dots, \mathbf{c}_N)$  and the algorithm starting point. Note that the distribution of the  $\mathbf{c}_i$ 's has no influence on the simulations because of the independence of equation (3) (which completely defines the FP estimate) with respect to the distribution of the  $\tau_i$ 's. Thus, without loss of generality, the Gaussian distribution will be used in the sequel.

Convergence will be analyzed by evaluating the widely used criterion  $C$

$$C(k) = \frac{\|\widehat{\mathbf{M}}_{k+1} - \widehat{\mathbf{M}}_k\|}{\|\widehat{\mathbf{M}}_k\|} \quad (21)$$

as a function of algorithm iteration  $k$ . The numerical limit of  $C$  (when algorithm has converged) is called the floor level.

The first subsection compares algorithms 1 and 2 while the second subsection studies parameters influence.

### A. Comparison of the two Algorithms

This section is devoted to the comparison of Algorithm 1 and 2 for Toeplitz matrices which are met when the processes are stationary. We will use the set of Toeplitz matrices  $\mathbf{M}$  defined by the following widely used structure:

$$M_{ij} = \rho^{|i-j|}, \quad (22)$$

for  $1 \leq i, j \leq m$  and for  $0 < \rho < 1$ . Notice that the covariance matrix  $\mathbf{M}$  is fully defined by the parameter  $\rho$ , which characterizes the correlation of the data.

1) *Convergence behavior for different values of  $\rho$* : Fig. 1 displays the criterion  $C(k)$  versus the iterations number  $k$  for the following set of parameters:  $m = 10$ ,  $N = 20$  and the starting point  $\mathbf{M}_0 = \mathbf{I}_m$ . Three typical cases are investigated: weak correlation ( $\rho = 10^{-5}$ , Fig. 1.a), medium correlation ( $\rho = 0.9$ , Fig. 1.b) and strong correlation ( $\rho = 1 - 10^{-5}$ , Fig. 1.c).

Fig. 1 leads to four main comments.

- For a given of  $\rho$ , both algorithms numerical convergence occurs for the same iteration number. Moreover, algorithm 2 always presents a better accuracy (in terms of floor level).
- Higher the  $\rho$ , faster the convergence is; for  $\rho = 10^{-5}$ , convergence is reached around 90 iterations, for  $\rho = 0.9$ , 60 iterations are enough and for  $\rho = 1 - 10^{-5}$ , only 20 iterations are required.
- Stronger the correlation, lower the limit accuracy is.
- The improvement of algorithm 2 in term of accuracy increases with  $\rho$ .

With this first analysis, we infer that algorithm 2 is better than algorithm 1.

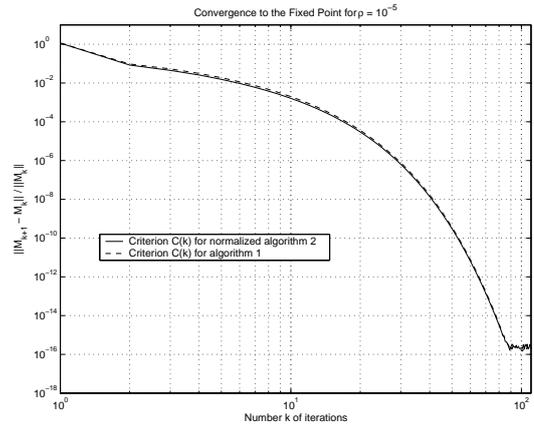
On Fig. 2, we have plotted the criterion  $C$  versus  $\rho$  when the convergence has occurred. Floor level is evaluated at the 150<sup>th</sup> iteration. Both algorithms exhibit the same behavior: the floor level gets worth when correlation parameter  $\rho$  increases. Floor level is always better for the normalized algorithm than for the algorithm 1. Moreover, the distance between the two curves increases with  $\rho$ .

Fig. 3 shows the required iteration number  $k$  to achieve a relative error  $C$  equal to  $10^{-5}$ . Plots are given as a function of correlation parameter  $\rho$ . Algorithm 1 is quite insensitive to the correlation parameter influence. The number of iteration  $k$  is always close to 21. Conversely, for algorithm 2, the iteration number  $k$  decreases with  $\rho$ , starting at  $k = 20$  for small  $\rho$  and ending at  $k = 8$  for  $\rho$  close to 1. Surprisingly, more the data are correlated, faster the convergence is (but according to Fig. 1.c, the floor level gets worse).

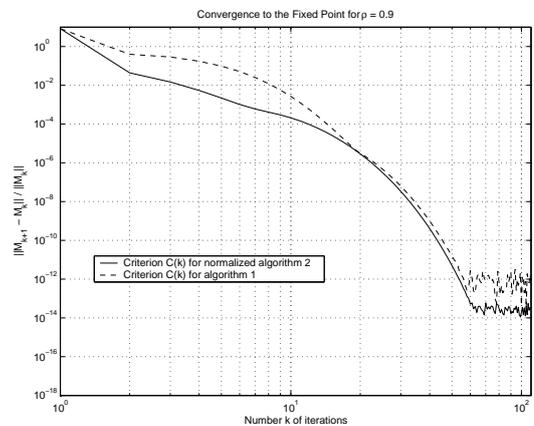
These results allow to conclude that algorithm 2 (normalized algorithm) is the best in all situations. That is why, in the sequel, we will study parameters influence on the normalized algorithm.

### B. Parameters influence

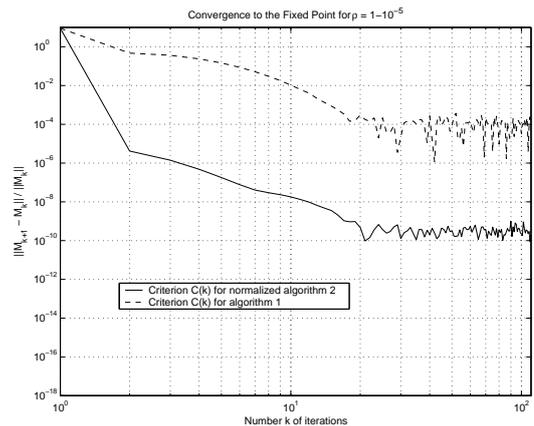
This section studies the influence on the normalized algorithm of the starting point  $\mathbf{M}_0$  and the number  $N$  of reference data.



(a)  $\rho = 10^{-5}$



(b)  $\rho = 0.9$



(c)  $\rho = 1 - 10^{-5}$

Fig. 1. Convergence to the FP for three different  $\rho$ . a)  $\rho = 10^{-5}$ , b)  $\rho = 0.9$ , c)  $\rho = 1 - 10^{-5}$

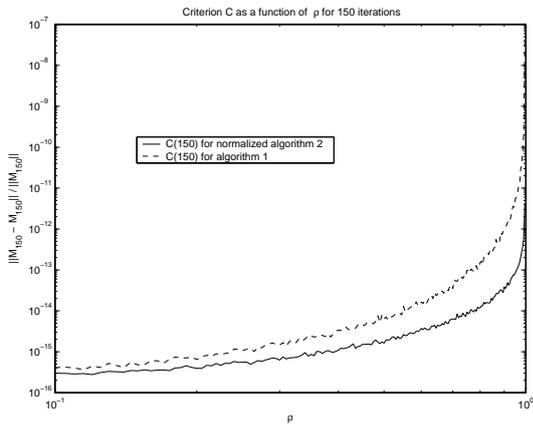
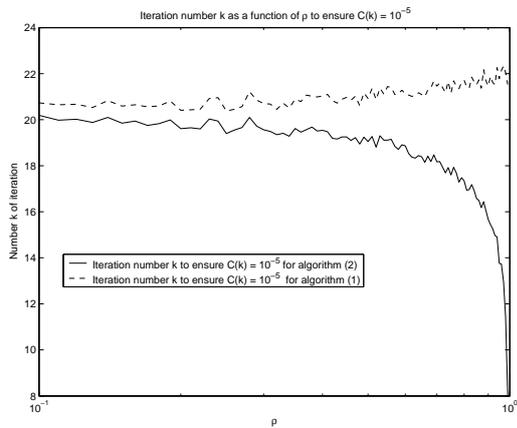
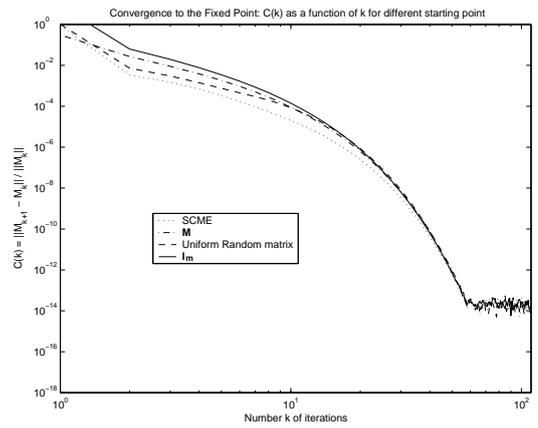
Fig. 2. Floor level,  $C(150)$ , against  $\rho$ Fig. 3. Required iteration number  $k$  to achieve the relative error  $C = 10^{-5}$ 

Fig. 4.a shows the criterion  $C(k)$  for four different initial conditions  $\mathbf{M}_0$  and a medium correlation parameter  $\rho = 0.9$ : the well known Sample Covariance Matrix Estimate (SCME), the true covariance matrix  $\mathbf{M}$ , a random matrix whose elements are uniformly distributed and the identity matrix  $\mathbf{I}_m$ . Floor level and convergence speed are independent of the algorithm initialization, after 10 iterations, all the curves merge. Fig. 4.b represents  $C(k)$  for various values of  $N$ : 20, 200, 2000 and 4000. Notice that convergence speed increases with  $N$ , while the floor level is almost independent of  $N$ .

## VII. CONCLUSION

In this work we have considered the problem of covariance matrix estimation for adaptive radar detection in compound-Gaussian clutter. The corresponding ML estimate of the covariance matrix built with secondary data is known to be the solution (if such a solution exists and is unique) of an equation for which no closed form solution is available. We have established in this paper a sound demonstration of the existence and uniqueness of this ML estimate, called FPE (Fixed Point Estimator). We have also derived two algorithms for obtaining the FPE. The convergence of each algorithm has been theoretically proved and emphasized by extensive



(a) Error influence with starting points

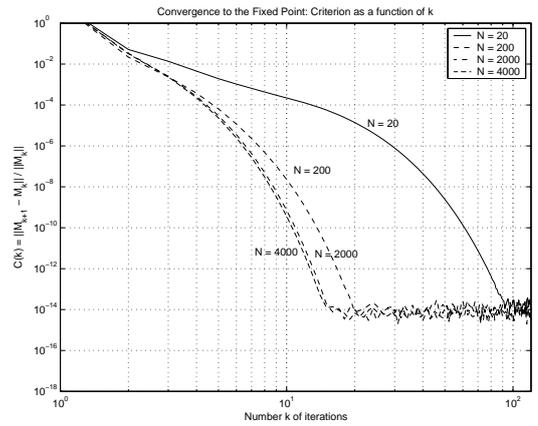
(b) Error influence with the number  $N$  of secondary data

Fig. 4. Convergence to the fixed point. a)  $C(k)$  as a function of  $k$  for different starting points  $\mathbf{M}_0$ . b)  $C(k)$  as a function of  $k$  for various values of  $N$ : 20, 200, 2000 and 4000

simulations which have shown the superiority of one of them, the so-called normalized algorithm. The numerical behavior of the two algorithms in realistic scenario has been also investigated as a function of main parameters, correlation and number of reference data, highlighting their fast convergence and therefore their great practical interests. These important results will allow the use of the FPE in real radar detection scheme [15]. It remains now to analyze the statistical behavior of the FPE, preliminary results in that direction have been already obtained in [16].

## APPENDIX I

### REDUCTION OF THE COMPLEX CASE TO THE REAL CASE

Let  $\mathcal{G}$  be the set of  $m \times m$  definite positive Hermitian matrices and  $\mathcal{S}$  the set of  $2m \times 2m$  symmetric matrices. Let us define the function  $g$  by

$$g : \mathcal{G} \longrightarrow \tilde{\mathcal{D}} = g(\mathcal{G}) \subset \mathcal{S}$$

$$\mathbf{M} \longrightarrow g(\mathbf{M}) = \begin{pmatrix} \mathbf{M}^{(1)} & -\mathbf{M}^{(2)} \\ \mathbf{M}^{(2)} & \mathbf{M}^{(1)} \end{pmatrix},$$

where  $\mathbf{M} = \mathbf{M}^{(1)} + j\mathbf{M}^{(2)}$  with  $\mathbf{M}^{(1)}$ , symmetric matrix, the real part of  $\mathbf{M}$  and  $\mathbf{M}^{(2)}$ , antisymmetric matrix, the imaginary part. It is obvious that  $g$  is a bijection between  $\mathcal{G}$  and  $\mathcal{D}$ . Moreover, we have the following proposition

**Proposition I.1**

$$\forall \mathbf{M} \in \mathcal{G}, g(f(\mathbf{M})) = f_R(g(\mathbf{M})),$$

where  $f$  is given by (7) and  $f_R$  by

$$f_R(\mathbf{M}_r) = \frac{m}{N} \sum_{i=1}^{2N} \frac{\mathbf{w}_i \mathbf{w}_i^\top}{\mathbf{w}_i^\top \mathbf{M}_r^{-1} \mathbf{w}_i},$$

with  $\mathbf{M}_r \in \tilde{\mathcal{D}}$ , and the  $2m$ -vectors  $\mathbf{w}_1, \dots, \mathbf{w}_{2N}$  are defined by

- for the  $N$  first vectors  $\mathbf{w}_1, \dots, \mathbf{w}_N$  (called  $\mathbf{u}_i$  for clarity),  
 $\mathbf{w}_i = \mathbf{u}_i = \begin{pmatrix} \mathbf{x}_i^{(1)} \\ \mathbf{x}_i^{(2)} \end{pmatrix},$
- for the  $N$  last vectors  $\mathbf{w}_{N+1}, \dots, \mathbf{w}_{2N}$  (called  $\mathbf{v}_i$ ),  
 $\mathbf{w}_{N+i} = \mathbf{v}_i = \begin{pmatrix} -\mathbf{x}_i^{(2)} \\ \mathbf{x}_i^{(1)} \end{pmatrix}.$

*Proof:* We have

$$g(f(\mathbf{M})) = \frac{m}{N} \sum_{i=1}^N \frac{g(\mathbf{x}_i \mathbf{x}_i^H)}{\mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i}.$$

Thanks to the following results:  $g(\mathbf{M}^{-1}) = (g(\mathbf{M}))^{-1}$ ,  $g(\mathbf{x}_i \mathbf{x}_i^H) = \mathbf{u}_i \mathbf{u}_i^\top + \mathbf{v}_i \mathbf{v}_i^\top$  and  $\mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i = \mathbf{u}_i^\top g(\mathbf{M}^{-1}) \mathbf{u}_i + \mathbf{v}_i^\top g(\mathbf{M}^{-1}) \mathbf{v}_i$ , Proposition I.1 follows straightforwardly. ■

Hypothesis (H) of Section III implies hypothesis (H2) (just before Remark IV.1) of linear independence for the real problem just defined in  $\mathbb{R}^{2m}$ . Thanks to Theorem IV.1, there exists a unique fixed point  $\mathbf{M}_R^{FP}$  (up to scalar factor) in  $\mathcal{S}$ . Thus, it remains to show that  $\mathbf{M}_R^{FP}$  belongs to  $\tilde{\mathcal{D}}$ . Thanks to Proposition I.1, if initialization of algorithm defined in Theorem IV.2, Eqn. (12), belongs to  $\tilde{\mathcal{D}}$ , the resulting sequence  $\{\mathbf{M}_0, \dots, \mathbf{M}_k\}$  obviously belongs to  $\tilde{\mathcal{D}}$ . Since this sequence converges in  $\mathcal{S}$ , by elementary topological considerations, the limit belongs to  $\tilde{\mathcal{D}}$ .

Now, since  $f_R$  admits a unique fixed point  $\mathbf{M}_R^{FP}$  (up to a scalar factor) in  $\tilde{\mathcal{D}}$ , the proof of Theorem III.1 is completed. Indeed, there exists a unique matrix  $\mathbf{M}^{FP}$  (up to a scalar factor) which verifies

$$\begin{aligned} \mathbf{M}^{FP} &= g^{-1}(\mathbf{M}_R^{FP}) = g^{-1}(f_R(\mathbf{M}_R^{FP})) \\ &= g^{-1}(f_R(g(\mathbf{M}^{FP}))) = g^{-1}(g(f(\mathbf{M}^{FP}))) = f(\mathbf{M}^{FP}). \end{aligned}$$

## APPENDIX II

### PROOF OF PROPOSITION V.1

If such a  $\widehat{\mathbf{M}}_{FP}$  exists, then for every  $\lambda > 0$ ,  $\lambda \widehat{\mathbf{M}}_{FP}$  is also a fixed point of  $f$ , since  $f$  is homogeneous of degree one. We start by demonstrating the following lemma.

**Lemma II.1**

The function  $F$  can be extended as a continuous function of  $\overline{\mathcal{D}} \setminus \{\mathbf{0}\}$  so that, for every non invertible  $\mathbf{M} \in \overline{\mathcal{D}} \setminus \{\mathbf{0}\}$ ,  $F(\mathbf{M}) = \mathbf{0}$ .

*Proof:* It is enough to show that, for every non invertible  $\mathbf{M} \in \overline{\mathcal{D}} \setminus \{\mathbf{0}\}$ , and every sequence  $(\mathbf{Q}^{(k)})_{k \geq 0}$  in  $\overline{\mathcal{D}}$  converging to zero and so that  $\mathbf{M} + \mathbf{Q}^{(k)}$  is invertible, we have

$$\forall \mathbf{Q} \in \mathcal{D}, \lim_{k \rightarrow \infty} F(\mathbf{M} + \mathbf{Q}^{(k)}) = 0.$$

Since  $F$  is smooth, we may assume that  $\mathbf{Q}^{(k)} \in \mathcal{D}$  for every  $k \geq 0$ . We introduce the notation  $F^{\mathbf{c}}$  for the function  $F$  in order to emphasize the dependence of  $F$  with respect to the  $N$ -tuple  $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_N)$ . If  $\mathbf{R}$  is an invertible matrix, let  $\mathbf{R.c}$  be the  $N$ -tuple  $\mathbf{R.c} := (\mathbf{Rc}_1, \dots, \mathbf{Rc}_N)$ . Clearly one has for every  $\mathbf{M} \in \mathcal{D}$ ,

$$F^{\mathbf{c}}(\mathbf{M}) = |\mathbf{R}|^{2N} F^{\mathbf{R.c}}(\mathbf{RMR}^T).$$

Fix now a symmetric matrix  $\mathbf{M}$  such that  $\mathbf{M} \geq 0$  and the rank of  $\mathbf{M}$ ,  $\text{rk}(\mathbf{M})$ , is equal to  $l$ , with  $0 < l < m$ . Thanks to the previous equation, we may assume that  $\mathbf{M} = \mathbf{J}_l$ , with  $\mathbf{J}_l := \text{diag}(\mathbf{I}_l \mathbf{0} \dots \mathbf{0})$ , where  $\mathbf{0}$  is repeated  $m - l$  times. For  $i \in [1, N]$ , we write  $\mathbf{c}_i$  as

$$\mathbf{c}_i = \begin{pmatrix} \mathbf{c}_i^1 \\ \mathbf{c}_i^2 \end{pmatrix}, \text{ with } \mathbf{c}_i^1 \in \text{Im}(\mathbf{J}_l) \text{ and } \mathbf{c}_i^2 \in \text{Ker}(\mathbf{J}_l).$$

According to that orthogonal decomposition, we write  $\mathbf{Q}^{(k)}$  by blocks,

$$\mathbf{Q}^{(k)} = \begin{pmatrix} \mathbf{Q}_1^{(k)} & \mathbf{Q}_2^{(k)T} \\ \mathbf{Q}_2^{(k)} & \mathbf{Q}_3^{(k)} \end{pmatrix}.$$

Then,

$$\mathbf{M} + \mathbf{Q}^{(k)} = \begin{pmatrix} \mathbf{I}_l + \mathbf{Q}_1^{(k)} & \mathbf{Q}_2^{(k)T} \\ \mathbf{Q}_2^{(k)} & \mathbf{Q}_3^{(k)} \end{pmatrix}.$$

For every  $k \geq 0$ , set  $\mathbf{P}_k := (\mathbf{I}_l + \mathbf{Q}_1^{(k)})^{-1}$ , and  $\mathbf{R}_k := \mathbf{Q}_3^{(k)} - \mathbf{Q}_2^{(k)} \mathbf{P}_k (\mathbf{Q}_2^{(k)})^T$ . Then, for every  $k \geq 0$ , one has, after standard computations using the Schur complement formula (cf. [14] for instance), that

$$\begin{aligned} (\mathbf{M} + \mathbf{Q}^{(k)})^{-1} &= \\ &= \begin{pmatrix} \mathbf{P}_k + \mathbf{P}_k (\mathbf{Q}_2^{(k)})^T \mathbf{R}_k^{-1} \mathbf{Q}_2^{(k)} \mathbf{P}_k & -\mathbf{P}_k (\mathbf{Q}_2^{(k)})^T \mathbf{R}_k^{-1} \\ -\mathbf{R}_k^{-1} \mathbf{Q}_2^{(k)} \mathbf{P}_k & \mathbf{R}_k^{-1} \end{pmatrix}, \end{aligned}$$

and  $|\mathbf{M} + \mathbf{Q}^{(k)}| = |\mathbf{I}_l + \mathbf{Q}_1^{(k)}| |\mathbf{R}_k|$ .

We next compute  $\mathbf{c}_i^T (\mathbf{M} + \mathbf{Q}^{(k)})^{-1} \mathbf{c}_i$  for  $i \in [1, N]$  and  $k \geq 0$ . We get

$$\begin{aligned} \mathbf{c}_i^T (\mathbf{M} + \mathbf{Q}^{(k)})^{-1} \mathbf{c}_i &= (\mathbf{c}_i^1)^T (\mathbf{P}_k + \mathbf{P}_k (\mathbf{Q}_2^{(k)})^T \mathbf{R}_k^{-1} \mathbf{Q}_2^{(k)} \mathbf{P}_k) \mathbf{c}_i^1 \\ &\quad - 2(\mathbf{c}_i^1)^T \mathbf{P}_k (\mathbf{Q}_2^{(k)})^T \mathbf{R}_k^{-1} \mathbf{c}_i^2 \\ &\quad + (\mathbf{c}_i^2)^T \mathbf{R}_k^{-1} \mathbf{c}_i^2. \end{aligned}$$

**Lemma II.2**

With the above notations, we have

$$\lim_{k \rightarrow \infty} \mathbf{P}_k + \mathbf{P}_k (\mathbf{Q}_2^{(k)})^T \mathbf{R}_k^{-1} \mathbf{Q}_2^{(k)} \mathbf{P}_k = \mathbf{I}_l,$$

and, if  $\mathbf{c}_i^2 \neq \mathbf{0}$ , then,

$$\lim_{k \rightarrow \infty} \frac{\mathbf{c}_i^T (\mathbf{M} + \mathbf{Q}^{(k)})^{-1} \mathbf{c}_i}{\mathbf{c}_i^2 \mathbf{R}_k^{-1} \mathbf{c}_i^2} = 1. \quad (\text{II.23})$$

*Proof:* Both results are a consequence of the following fact,

$$\lim_{k \rightarrow \infty} \mathbf{P}_k (\mathbf{Q}_2^{(k)})^T \mathbf{R}_k^{-1} \mathbf{Q}_2^{(k)} \mathbf{P}_k = 0. \quad (\text{II.24})$$

To see that, first recall that  $\mathbf{S}_k := \mathbf{Q}_3^{(k)} - \mathbf{Q}_2^{(k)} (\mathbf{Q}_1^{(k)})^{-1} (\mathbf{Q}_2^{(k)})^T$  is definite positive since  $\mathbf{Q}^{(k)}$  is positive definite. Next, we write

$$\begin{aligned} \mathbf{R}_k &= \mathbf{S}_k + \mathbf{Q}_2^{(k)} (\mathbf{Q}_1^{(k)})^{-1} (\mathbf{Q}_2^{(k)})^T - \mathbf{Q}_2^{(k)} \mathbf{P}_k (\mathbf{Q}_2^{(k)})^T \\ &= \mathbf{Q}_2^{(k)} (\mathbf{Q}_1^{(k)})^{-1} \mathbf{P}_k (\mathbf{Q}_2^{(k)})^T, \end{aligned}$$

and we then have

$$\begin{aligned} &\mathbf{P}_k (\mathbf{Q}_2^{(k)})^T \mathbf{R}_k^{-1} \mathbf{Q}_2^{(k)} \mathbf{P}_k = \\ &\mathbf{P}_k^{1/2} (\mathbf{Q}_1^{(k)})^{1/2} \mathbf{B}_k^T (\mathbf{I}_l + \mathbf{B}_k \mathbf{B}_k^T)^{-1} \mathbf{B}_k (\mathbf{Q}_1^{(k)})^{1/2} \mathbf{P}_k^{1/2}, \end{aligned}$$

where  $\mathbf{B}_k := \mathbf{S}_k^{-1/2} \mathbf{Q}_2^{(k)} (\mathbf{Q}_1^{(k)})^{-1/2} \mathbf{P}_k^{1/2}$ .

It is now clear that (II.24) holds true if the  $l \times l$  symmetric non negative matrix  $\mathbf{B}_k^T (\mathbf{I}_l + \mathbf{B}_k \mathbf{B}_k^T)^{-1} \mathbf{B}_k$  is bounded. Computing the norm, we end up with

$$\|\mathbf{B}_k^T (\mathbf{I}_l + \mathbf{B}_k \mathbf{B}_k^T)^{-1} \mathbf{B}_k\|^2 = \|(\mathbf{I}_l + \mathbf{T}_k)^{-1} \mathbf{T}_k\|^2,$$

where  $\mathbf{T}_k := \mathbf{B}_k \mathbf{B}_k^T \in \overline{\mathcal{D}}$ . Since  $(\mathbf{I}_l + \mathbf{T}_k)^{-1} \mathbf{T}_k \leq \mathbf{I}_l$ , we conclude the proof of Lemma II.2. ■

We next consider the diagonalization of  $\mathbf{R}_k$  in an orthonormal basis, given by

$$\mathbf{R}_k = \mathbf{U}_k^T \mathbf{D}_k \mathbf{U}_k, \text{ for } k \geq 0,$$

with  $\mathbf{U}_k \in SO(m-l)$  and  $\mathbf{D}_k = \text{diag}(\varepsilon_k^{(l+1)}, \dots, \varepsilon_k^{(m)})$ . By definition,  $\lim_{k \rightarrow \infty} \varepsilon_k^{(j)} = 0^+$ , for every  $j \in [l+1, m]$ ,

and, with no loss of generality, we will assume that  $\varepsilon_k^{(m)} = \min_{l+1 \leq j \leq m} \varepsilon_k^{(j)}$  and  $\lim_{k \rightarrow \infty} \mathbf{U}_k = \mathbf{U} \in SO(m-l)$ .

We next establish the following lemma.

### Lemma II.3

Let  $\mathbf{E}_m = (0 \cdots 0 \ 1)^T$  with 0 repeated  $m-l-1$  times. With the previous notations, there exist  $C > 0$  and  $i^* \in [1, N]$  such that, for  $k \geq 0$  large enough, we have

$$|\mathbf{E}_m^T \mathbf{U}_k \mathbf{c}_{i^*}^2| \geq C. \quad (\text{II.25})$$

*Proof:* By a continuity argument, it is enough to show the existence of an index  $i^*$  so that  $\mathbf{E}_m^T \mathbf{U} \mathbf{c}_{i^*}^2 \neq 0$ . Moreover, according to hypothesis (H2), it is not possible to find  $m$  vectors  $\mathbf{c}_{i(1)}, \dots, \mathbf{c}_{i(m)}$  linearly independent such that

$$\mathbf{e}_m^T \overline{\mathbf{U}} \mathbf{c}_i = \mathbf{E}_m^T \mathbf{U} \mathbf{c}_i^2 = 0,$$

where  $\mathbf{e}_m = (0 \cdots 0 \ 1)^T \in \mathbb{R}^m$  and  $\overline{\mathbf{U}} = \text{diag}(\mathbf{I}_l, \mathbf{U})$ . (Otherwise, there exist  $m$  vectors  $\mathbf{c}_{i(1)}, \dots, \mathbf{c}_{i(m)}$  linearly independent belonging to the orthogonal of  $\overline{\mathbf{U}} \mathbf{e}_m$ , which has dimension  $m-1$ .)

By a simple counting argument, the index  $i^*$  therefore exists. Indeed, otherwise the  $N$  vectors  $\mathbf{c}_i$ 's, with  $i \notin S$ , verify  $\mathbf{e}_m^T \overline{\mathbf{U}} \mathbf{c}_i = 0$ , meaning that all the vectors  $\mathbf{c}_i$ ,  $1 \leq i \leq N$ , are orthogonal to  $\overline{\mathbf{U}}^T \mathbf{e}_m$ , which is impossible. The proof of Lemma II.3 is complete. ■

We can now finish the proof of Lemma II.1. Let  $\mathbf{c}^*$  be the  $(N-1)$ -tuple made of the  $\mathbf{c}_i$ 's for  $i \in [1, N]/\{i^*\}$ . For every  $k \geq 0$ , we have

$$\begin{aligned} F^{\mathbf{c}}(\mathbf{M} + \mathbf{Q}^{(k)}) &= \frac{1}{|\mathbf{M} + \mathbf{Q}^{(k)}|} \left( \frac{1}{\mathbf{c}_{i^*}^T (\mathbf{M} + \mathbf{Q}^{(k)})^{-1} \mathbf{c}_{i^*}} \right)^m \\ &\times F^{\mathbf{c}^*}(\mathbf{M} + \mathbf{Q}^{(k)}). \end{aligned}$$

Since  $N-1 \geq m$ , we apply the result of [13] which states that the supremum of  $F^{\mathbf{c}^*}$  over  $\mathcal{D}$  is finite, i.e., there exists a positive constant  $C^*$  such that, for every  $\mathbf{R} \in \mathcal{D}$ ,  $F^{\mathbf{c}^*}(\mathbf{R}) \leq C^*$ . Therefore, the conclusion holds true if

$$\lim_{k \rightarrow \infty} \frac{1}{|\mathbf{M} + \mathbf{Q}^{(k)}|} \left( \frac{1}{\mathbf{c}_{i^*}^T (\mathbf{M} + \mathbf{Q}^{(k)})^{-1} \mathbf{c}_{i^*}} \right)^m = 0.$$

Thanks to (II.23), that amounts to show that

$$\lim_{k \rightarrow \infty} \frac{1}{|\mathbf{D}_k|} \left( \frac{1}{(\mathbf{c}_{i^*}^2)^T (\mathbf{R}_k)^{-1} \mathbf{c}_{i^*}^2} \right)^m = 0.$$

It is clear that  $|\mathbf{D}_k| \geq (\varepsilon_k^{(m)})^{m-l}$ . In addition, by using Lemma II.3, we can write

$$(\mathbf{c}_{i^*}^2)^T (\mathbf{R}_k)^{-1} \mathbf{c}_{i^*}^2 = (\mathbf{U}_k \mathbf{c}_{i^*}^2)^T (\mathbf{D}_k)^{-1} \mathbf{U}_k \mathbf{c}_{i^*}^2 = \xi_k \frac{(\mathbf{E}_m^T \mathbf{U}_k \mathbf{c}_{i^*}^2)^2}{\varepsilon_k^{(m)}},$$

where  $\xi_k$  is bounded below and above by positive constants independent on  $k$ . We finally get that

$$\frac{1}{|\mathbf{D}_k|} \left( \frac{1}{(\mathbf{c}_{i^*}^2)^T (\mathbf{R}_k)^{-1} \mathbf{c}_{i^*}^2} \right)^m \leq C (\varepsilon_k^{(m)})^l,$$

with a positive constant  $C$  independent of  $k$ . By letting  $k$  go to infinity, we conclude the proof of Lemma II.1. ■

End of the proof of Proposition V.1:

Recall that  $\overline{\mathcal{D}}(1)$  is a compact subset of  $\overline{\mathcal{D}} \setminus \{\mathbf{0}\}$ . Then  $F$  is well-defined on  $\overline{\mathcal{D}}(1)$  and is continuous. The application  $F$  reaches its maximum over  $\overline{\mathcal{D}}(1)$  at a point  $\widehat{\mathbf{M}}_{FP}$ . Since  $F$  is strictly positive on  $\mathcal{D}(1)$  and equal to zero on  $\overline{\mathcal{D}}(1) \setminus \mathcal{D}(1)$ , then  $F(\widehat{\mathbf{M}}_{FP}) > 0$ , implying that  $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}(1)$ . We complete the proof of Proposition V.1 by establishing the next lemma.

### Lemma II.4

Let  $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}(1)$  be defined as previously. Then,  $\nabla F(\widehat{\mathbf{M}}_{FP}) = \mathbf{0}$ , which implies that  $\widehat{\mathbf{M}}_{FP}$  is a fixed point of  $f$ .

*Proof:* By definition of  $\widehat{\mathbf{M}}_{FP}$ , one has  $F(\widehat{\mathbf{M}}_{FP}) = \max_{\mathbf{M} \in \mathcal{D}(1)} F(\mathbf{M})$ . By standard calculus, it results that  $\nabla F(\widehat{\mathbf{M}}_{FP})$  and  $\nabla \mathcal{N}(\widehat{\mathbf{M}}_{FP})$  are colinear, where  $\mathcal{N}(\mathbf{M}) = \|\mathbf{M}\|^2$  for every  $\mathbf{M} \in \mathcal{D}$ . Since  $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}$ , there exists a real number  $\mu$  such that  $\nabla F(\widehat{\mathbf{M}}_{FP}) = \mu \widehat{\mathbf{M}}_{FP}$ . Recall that, since  $F$  is homogeneous of degree zero, then,

$$\forall \mathbf{M} \in \mathcal{D}, \quad \nabla(\mathbf{M}) \cdot \mathbf{M} = 0. \quad (\text{II.26})$$

One deduces that  $\mu = \mu \|\widehat{\mathbf{M}}_{FP}\|^2 = \nabla F(\widehat{\mathbf{M}}_{FP}) \cdot \widehat{\mathbf{M}}_{FP} = 0$ . The proof of Lemma II.4 is complete. ■

APPENDIX III  
 PROOF OF PROPOSITION V.2

We start by establishing (P1). Let  $\mathbf{M}, \mathbf{Q} \in \mathcal{D}$  with  $\mathbf{M} \leq \mathbf{Q}$ . Then,  $\mathbf{M}^{-1} \geq \mathbf{Q}^{-1}$  and, for every  $1 \leq i \leq N$ , we have

$$\frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i} \leq \frac{1}{\mathbf{c}_i^\top \mathbf{Q}^{-1} \mathbf{c}_i}.$$

The reasoning for the case with strict inequalities is identical. Then, clearly, (P1) follows.

We next turn to the proof of (P2). We first recall that, for every unit vector  $\mathbf{c} \in \mathbb{R}^m$ ,  $\|\mathbf{c}\| = 1$  and  $\mathbf{M} \in \mathcal{D}$ , then

$$\frac{1}{\mathbf{c}^\top \mathbf{M}^{-1} \mathbf{c}} = \inf_{\mathbf{z}^\top \mathbf{c} \neq 0} \frac{\mathbf{z}^\top \mathbf{M} \mathbf{z}}{(\mathbf{c}^\top \mathbf{z})^2}, \quad (\text{III.27})$$

and the infimum is reached only on the line generated by  $\mathbf{M}^{-1} \mathbf{c}$ .

Let  $\mathbf{M}, \mathbf{Q} \in \mathcal{D}$ . Then, one has

$$\begin{aligned} f(\mathbf{M} + \mathbf{Q}) &= \frac{m}{N} \sum_{i=1}^N \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \frac{\mathbf{z}^\top (\mathbf{M} + \mathbf{Q}) \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} \\ &= \frac{m}{N} \sum_{i=1}^N \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \left( \frac{\mathbf{z}^\top \mathbf{M} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} + \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} \right). \end{aligned}$$

More generally, the following holds true,

$$\min_{\mathbf{z} \in \mathcal{A}} (f_1(\mathbf{z}) + f_2(\mathbf{z})) \geq \min_{\mathbf{z} \in \mathcal{A}} f_1(\mathbf{z}) + \min_{\mathbf{z} \in \mathcal{A}} f_2(\mathbf{z}),$$

for every functions  $f_1, f_2$  and set  $\mathcal{A}$  giving a sense to the previous inequality. Then, (P2) clearly holds true. It remains to study when equality occurs in (P2). That happens if and only if, for every  $1 \leq i \leq N$ , one has

$$\begin{aligned} \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \left( \frac{\mathbf{z}^\top \mathbf{M} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} + \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} \right) &= \\ \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \frac{\mathbf{z}^\top \mathbf{M} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} + \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2}. \end{aligned} \quad (\text{III.28})$$

Let us first show that equality occurs in (III.28) if and only if there exists some  $\mu_i > 0$  such that

$$\mathbf{M}^{-1} \mathbf{c}_i = \frac{1}{\mu_i} \mathbf{Q}^{-1} \mathbf{c}_i. \quad (\text{III.29})$$

Indeed, for every vector  $\mathbf{z} \in \mathbb{R}^m$  with  $\mathbf{z}^\top \mathbf{c}_i \neq 0$ , we have

$$\frac{\mathbf{z}^\top (\mathbf{M} + \mathbf{Q}) \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} \geq \frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i} + \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2}.$$

Choosing  $\mathbf{z} = (\mathbf{M} + \mathbf{Q})^{-1} \mathbf{c}_i$  yields

$$\begin{aligned} \frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i} + \frac{1}{\mathbf{c}_i^\top \mathbf{Q}^{-1} \mathbf{c}_i} &= \frac{1}{\mathbf{c}_i^\top (\mathbf{M} + \mathbf{Q})^{-1} \mathbf{c}_i} \\ &\geq \frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i} \\ &\quad + \frac{\mathbf{c}_i^\top (\mathbf{M} + \mathbf{Q})^{-1} \mathbf{Q} (\mathbf{M} + \mathbf{Q})^{-1} \mathbf{c}_i}{(\mathbf{c}_i^\top (\mathbf{M} + \mathbf{Q})^{-1} \mathbf{c}_i)^2}. \end{aligned}$$

Therefore, the function of  $z$  given by  $\frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2}$  reaches its

minimum value  $\frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i}$  at  $\mathbf{z} = (\mathbf{M} + \mathbf{Q})^{-1} \mathbf{c}_i$ . Using (III.27), we get that  $(\mathbf{M} + \mathbf{Q})^{-1} \mathbf{c}_i$  is colinear to  $\mathbf{Q}^{-1} \mathbf{c}_i$ . Exchanging  $\mathbf{M}$  and  $\mathbf{Q}$  and proceeding as above yields that  $(\mathbf{M} + \mathbf{Q})^{-1} \mathbf{c}_i$  is also colinear to  $\mathbf{M}^{-1} \mathbf{c}_i$ , which finally implies that  $\mathbf{M}^{-1} \mathbf{c}_i$  and  $\mathbf{Q}^{-1} \mathbf{c}_i$  are themselves colinear. (III.29) is proved.

To finish the proof, one must show that all the  $(\mu_i)$ 's,  $1 \leq i \leq N$ , as defined in (III.29), are equal.

Set  $\mathbf{D} := \text{diag}(\mu_1, \dots, \mu_m)$  for the first  $m$  indices of  $\llbracket 1, N \rrbracket$ . Since  $(\mathbf{c}_1, \dots, \mathbf{c}_m)$  is a basis of  $\mathbb{R}^n$  and  $\mathbf{M}^{-1} - \mathbf{D}^{-1} \mathbf{Q}^{-1}$  is equal to  $\mathbf{0}$  on that basis, we deduce that  $\mathbf{M}^{-1} = \mathbf{D}^{-1} \mathbf{Q}^{-1}$ . Consider now another basis of  $\mathbb{R}^m$  defined by  $(\mathbf{c}_2, \dots, \mathbf{c}_{m+1})$  and set  $\tilde{\mathbf{D}} = \text{diag}(\mu_2, \dots, \mu_{m+1})$ . Reasoning as previously, we obtain that  $\mathbf{M}^{-1} = \tilde{\mathbf{D}}^{-1} \mathbf{Q}^{-1}$ , which firstly implies that  $\tilde{\mathbf{D}} = \mathbf{D}$  and, secondly, that  $\mu_1 = \mu_2, \mu_2 = \mu_3, \dots, \mu_m = \mu_{m+1}$ . Repeating that reasoning for any pair of  $m$ -tuples of distinct indices  $(i_1, \dots, i_m)$  of  $\llbracket 1, N \rrbracket$ , we get that, for every  $i \in \llbracket 1, N \rrbracket$ ,  $\mu_i = \mu$ , yielding  $\mathbf{D} = \mu \mathbf{I}_m$ . ■

 APPENDIX IV  
 PROOF OF PROPOSITION V.3

We first establish the following fact. For every  $\mathbf{Q}, \mathbf{P} \in \mathcal{D}$ , we have

$$\text{If } \mathbf{Q} \geq \mathbf{P} \text{ and } f(\mathbf{Q}) = f(\mathbf{P}), \text{ then } \mathbf{Q} = \mathbf{P}. \quad (\text{IV.30})$$

Indeed, it is clear that  $\mathbf{Q} \geq \mathbf{P}$  implies that  $\mathbf{P}^{-1} - \mathbf{Q}^{-1} \geq \mathbf{0}$ . Therefore, for every  $1 \leq i \leq N$ , we have

$$\frac{1}{\mathbf{c}_i^\top \mathbf{Q}^{-1} \mathbf{c}_i} \geq \frac{1}{\mathbf{c}_i^\top \mathbf{P}^{-1} \mathbf{c}_i}.$$

Assuming  $f(\mathbf{Q}) = f(\mathbf{P})$  implies that, for every  $1 \leq i \leq N$ , we have  $\mathbf{c}_i^\top \mathbf{Q}^{-1} \mathbf{c}_i = \mathbf{c}_i^\top \mathbf{P}^{-1} \mathbf{c}_i$  i.e.

$$\mathbf{c}_i^\top (\mathbf{P}^{-1} - \mathbf{Q}^{-1}) \mathbf{c}_i = 0.$$

Since  $\mathbf{P}^{-1} - \mathbf{Q}^{-1} \geq \mathbf{0}$ , the previous equality says that  $(\mathbf{P}^{-1} - \mathbf{Q}^{-1}) \mathbf{c}_i = \mathbf{0}$ , for every  $1 \leq i \leq N$ . By (H2), the claim (IV.30) is proved.

We now turn to the proof of Proposition V.3. We consider  $\mathbf{Q}, \mathbf{P} \in \mathcal{D}$  such that  $\mathbf{Q} \geq \mathbf{P}$  and  $\mathbf{Q} \neq \mathbf{P}$ . From what precedes, we also have that  $f(\mathbf{Q}) \geq f(\mathbf{P})$  and  $f(\mathbf{Q}) \neq f(\mathbf{P})$ . That implies the existence of an index  $i_0 \in \llbracket 1, N \rrbracket$  such that

$$\xi_{i_0} := \frac{m}{N} \left( \frac{1}{\mathbf{c}_{i_0}^\top \mathbf{Q}^{-1} \mathbf{c}_{i_0}} - \frac{1}{\mathbf{c}_{i_0}^\top \mathbf{P}^{-1} \mathbf{c}_{i_0}} \right) > 0.$$

Up to a relabel, we may assume that  $i_0 = 1$ . We then have

$$f(\mathbf{Q}) \geq f(\mathbf{P}) + \xi_1 \mathbf{c}_1 \mathbf{c}_1^\top. \quad (\text{IV.31})$$

Next, we will show by induction on the index  $l \leq m$  that there exist  $l$  positive real numbers  $\xi_k, 1 \leq k \leq l$ , so that

$$f^l(\mathbf{Q}) \geq f^l(\mathbf{P}) + \sum_{k=1}^l \xi_k \mathbf{c}_k \mathbf{c}_k^\top \quad (\text{IV.32})$$

In the previous equation, the vectors  $(\mathbf{c}_k)_{1 \leq k \leq l}$  only need to be two by two distinct among all the vectors  $(\mathbf{c}_i)_{1 \leq i \leq N}$ . At each step of the induction, we will have the possibility to relabel the indices in  $\llbracket l+1, N \rrbracket$  in such a way to get (IV.32). The induction starts for  $l = 1$  and, in this case, (IV.32) reduces to (IV.31). Therefore the induction is initialized. We then assume that (IV.32) holds true for some index  $l \leq m-1$  and proceed in showing the same for the index  $l+1$ . It is clear that it will be a consequence of the next lemma.

**Lemma IV.1**

Let  $1 \leq l \leq m-1$ ,  $\mathbf{Q}, \mathbf{P} \in \mathcal{D}$  such that

$$\mathbf{Q} \geq \mathbf{P} + \sum_{k=1}^l \xi_k \mathbf{c}_k \mathbf{c}_k^\top, \quad \xi_k > 0. \quad (\text{IV.33})$$

Then, there exists a vector of  $\{\mathbf{c}_{l+1}, \dots, \mathbf{c}_N\}$  (to be set equal to  $\mathbf{c}_{l+1}$ , up to a relabelling of  $\{\mathbf{c}_{l+1}, \dots, \mathbf{c}_N\}$ ) and a positive real number  $\xi_{l+1} > 0$  such that

$$f(\mathbf{Q}) \geq f(\mathbf{P}) + \sum_{k=1}^{l+1} \xi_k \mathbf{c}_k \mathbf{c}_k^\top. \quad (\text{IV.34})$$

*Proof:* Using (IV.33), we have for every  $j \in \llbracket 1, N \rrbracket$ ,

$$\frac{1}{\mathbf{c}_j^\top \mathbf{Q}^{-1} \mathbf{c}_j} = \min_{\mathbf{z}^\top \mathbf{c}_j \neq 0} \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{z}^\top \mathbf{c}_j)^2},$$

and

$$\min_{\mathbf{z}^\top \mathbf{c}_j \neq 0} \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{z}^\top \mathbf{c}_j)^2} \geq \min_{\mathbf{z}^\top \mathbf{c}_j \neq 0} \left( \frac{\mathbf{z}^\top \mathbf{P} \mathbf{z}}{(\mathbf{z}^\top \mathbf{c}_j)^2} + \sum_{k=1}^l \xi_k \frac{(\mathbf{z}^\top \mathbf{c}_k)^2}{(\mathbf{z}^\top \mathbf{c}_j)^2} \right). \quad (\text{IV.35})$$

Using the induction hypothesis, we also have for every  $1 \leq j \leq l$ , that

$$\frac{1}{\mathbf{c}_j^\top \mathbf{Q}^{-1} \mathbf{c}_j} \geq \frac{1}{\mathbf{c}_j^\top \mathbf{P}^{-1} \mathbf{c}_j} + \xi_j.$$

We next show the following claim

(C1) there exists two indices, one index  $j_0 \in \llbracket l+1, N \rrbracket$  and another one  $k_0 \in \llbracket 1, l \rrbracket$  such that  $\mathbf{c}_{k_0}^\top \mathbf{Q}^{-1} \mathbf{c}_{j_0} \neq 0$ .

Claim (C1) is proved reasoning by contradiction. Therefore, let us assume that  $\mathbf{c}_k^\top \mathbf{Q}^{-1} \mathbf{c}_j = 0$ , for every  $1 \leq k \leq l$  and  $l+1 \leq j \leq N$ . Since  $l < m$ , the vectors  $(\mathbf{Q}^{-1} \mathbf{c}_k)$ ,  $1 \leq k \leq l$  generate a vector space  $\mathbf{V}_l$  of dimension  $l$ , we deduce that, for every  $j \in \llbracket l+1, N \rrbracket$ ,  $\mathbf{c}_j$  is orthogonal to  $\mathbf{V}_l$  and, therefore, belongs to an  $m-l$ -dimensional vector space of  $\mathbb{R}^m$ . But there are  $N-l$  indices  $j$  verifying the previous fact. According to (H2), these vectors  $(\mathbf{c}_j)_{l+1 \leq j \leq N}$  generate a vector space of dimension  $\min(N-l, m)$  in  $\mathbb{R}^m$ . We finally get that  $\min(N-l, m) \leq m-l$ . This is impossible because  $N > m$  and claim (C1) is proved.

We now finish the proof of Lemma IV.1. Choosing in (IV.35)  $\mathbf{z} = \mathbf{Q}^{-1} \mathbf{c}_{j_0}$ , we get

$$\begin{aligned} \frac{1}{\mathbf{c}_{j_0}^\top \mathbf{Q}^{-1} \mathbf{c}_{j_0}} &\geq \frac{(\mathbf{c}_{j_0}^\top \mathbf{Q}^{-1}) \mathbf{P} (\mathbf{Q}^{-1} \mathbf{c}_{j_0})}{(\mathbf{c}_{j_0}^\top \mathbf{Q}^{-1} \mathbf{c}_{j_0})^2} + \xi_{k_0} \frac{(\mathbf{c}_{j_0}^\top \mathbf{Q}^{-1} \mathbf{c}_{k_0})^2}{(\mathbf{c}_{j_0}^\top \mathbf{Q}^{-1} \mathbf{c}_{j_0})^2} \\ &\geq \frac{1}{\mathbf{c}_{j_0}^\top \mathbf{P}^{-1} \mathbf{c}_{j_0}} + \xi_{j_0} \end{aligned}$$

with  $\xi_{j_0} > 0$ , thanks to claim (C1). It is clear that  $\mathbf{c}_{j_0}$  is the vector of  $\{\mathbf{c}_{l+1}, \dots, \mathbf{c}_N\}$  needed with  $\xi_{j_0}$  so that, up to relabelling, yields (IV.34). Proofs of Lemma IV.1 and Proposition V.3 are now complete.  $\blacksquare$

APPENDIX V

PROOF OF PROPOSITION V.4

We first need to make precise a definition. An orbit  $(\mathbf{M}_k)_{k \geq 0}$  is bounded in  $\mathcal{D}$  if it is contained in a compact subset of  $\mathcal{D}$ , i.e., there exists  $\mathbf{M}, \mathbf{P} \in \mathcal{D}$  such that, for every  $k \geq 0$ ,  $\mathbf{M} \leq \mathbf{M}_k \leq \mathbf{P}$ .

We will show the following chain of implications  $(A) \Rightarrow (B) \Rightarrow (C) \Rightarrow (A)$ .

$(A) \Rightarrow (B)$ : Trivial (simply  $\mathbf{M}_0 = \mathbf{P}$ ).

$(B) \Rightarrow (C)$ : Assume that  $f$  has a bounded orbit in  $\mathcal{D}$ , starting at  $\mathbf{M}$ . Then, there exists  $\mu, \mu' > 0$  such that, for every  $k \geq 0$ ,  $\mu \mathbf{M} \leq \mathbf{M}_k \leq \mu' \mathbf{M}$ , for every  $k \geq 0$ .

Let  $\mathbf{Q}$  be an arbitrary matrix of  $\mathcal{D}$ . Then, there exists  $\lambda, \lambda' > 0$  such that  $\lambda \mathbf{M} \leq \mathbf{Q} \leq \lambda' \mathbf{M}$ . Using the homogeneity of degree one of  $f$ , property (P1) and the definition of an orbit of  $f$ , we get, after a trivial induction, that  $\lambda \mu \mathbf{M} \leq \lambda \mathbf{M}_k \leq \mathbf{Q}_k \leq \lambda' \mathbf{M}_k \leq \lambda' \mu' \mathbf{M}$ , for every  $k \geq 0$ . Then, the orbit associated to  $\mathbf{Q}$  is bounded in  $\mathcal{D}$ .

$(C) \Rightarrow (A)$ : Consider an orbit  $(\mathbf{M}_k)_{k \geq 0}$  of  $f$  starting at  $\mathbf{M} \in \mathcal{D}$  and bounded in  $\mathcal{D}$ . It is then contained in a compact  $\mathcal{K}$  of  $\mathcal{D}$ . For  $l \geq 1$ , set

$$\mathbf{Q}_l := \frac{1}{l} \sum_{i=1}^l \mathbf{M}_i.$$

Then, the sequence  $(\mathbf{Q}_l)_{l \geq 1}$  is bounded in  $\mathcal{D}$  because every point  $\mathbf{Q}_l$  belongs to the convex hull of  $\mathcal{K}$ , which is itself a compact subset of  $\mathcal{D}$ . For every  $l \geq 1$ , we have by using Proposition V.2 that

$$\begin{aligned} f(\mathbf{Q}_l) &= \frac{1}{l} f \left( \sum_{i=1}^l \mathbf{M}_i \right) \geq \frac{1}{l} \sum_{i=2}^{l+1} \mathbf{M}_i \\ &\geq \frac{1}{l} \left( \sum_{i=1}^{l+1} \mathbf{M}_i - \mathbf{M}_1 \right) = \mathbf{Q}_l + \frac{\mathbf{M}_{l+1} - \mathbf{M}_1}{l}. \end{aligned}$$

Since  $(\mathbf{Q}_l)_{l \geq 1}$  is bounded in  $\mathcal{D}$ , we have that, up to extracting a sub-sequence, that the sequence  $(\mathbf{Q}_l)$  converges to  $\bar{\mathbf{Q}}$ , with  $\bar{\mathbf{Q}} \in \mathcal{D}$ , as  $l$  tends to  $+\infty$ . From the last equation, it follows that  $f(\bar{\mathbf{Q}}) \geq \bar{\mathbf{Q}}$ .

We now consider the orbit of  $f$  starting at  $\bar{\mathbf{Q}}$ . It defines an increasing, bounded in  $\mathcal{D}$  sequence. It is therefore converging in  $\mathcal{D}$  to a fixed point of  $f$ .  $\blacksquare$

APPENDIX VI

PROOF OF COROLLARY V.1

The proof of (C1) goes by contradiction. Let  $\mathbf{P} \in \mathcal{D}$  with  $f^l(\mathbf{P}) \geq \mathbf{P}$  and  $f^l(\mathbf{P}) \neq \mathbf{P}$  for some positive integer  $l \geq 1$ . According to Proposition V.3, we have

$$f^n(f^l(\mathbf{P})) > f^n(\mathbf{P}) \Leftrightarrow f^l(f^n(\mathbf{P})) > f^n(\mathbf{P}).$$

Set  $\mathbf{Q} := f^n(\mathbf{P})$  and  $g := f^l$ . It is clear that  $g$  is a function from  $\mathcal{D}$  to  $\mathcal{D}$ , homogeneous of degree one and verifies properties (P1) and (P2) of Proposition V.2. We will show that the orbit of  $g$  associated to  $\mathbf{Q}$  is not bounded, which will be the desired contradiction.

We have  $g(\mathbf{Q}) > \mathbf{Q}$  which is equivalent to  $g(\mathbf{Q}) - \mathbf{Q}$  being positive definite. By a simple continuity argument, there exists  $\varepsilon_{\mathbf{Q}} > 0$  such that

$$\varepsilon_{\mathbf{Q}} \mathbf{Q} \leq f(\mathbf{Q}) - \mathbf{Q} \Leftrightarrow f(\mathbf{Q}) \geq (1 + \varepsilon_{\mathbf{Q}}) \mathbf{Q}.$$

By a trivial induction, we have  $f^k(\mathbf{Q}) \geq (1 + \varepsilon_{\mathbf{Q}})^k \mathbf{Q}$ , for every  $k \geq 0$ , with the right-hand side of the above inequality tending to  $+\infty$  as  $k$  tends to  $\infty$ . Therefore, the orbit of  $f$  associated to  $\mathbf{M}$  is not bounded.

We now prove statement (C2). Let  $\widehat{\mathbf{M}}_{FP}$  and  $\mathbf{P}_2$  be two fixed points of  $f$ . Applying (P2), we have

$$f(\widehat{\mathbf{M}}_{FP} + \mathbf{P}_2) \geq f(\widehat{\mathbf{M}}_{FP}) + f(\mathbf{P}_2) = \widehat{\mathbf{M}}_{FP} + \mathbf{P}_2,$$

According to (C1) above, we have that  $\widehat{\mathbf{M}}_{FP} + \mathbf{P}_2$  is also a fixed point of  $f$  and therefore, we have equality in (P2). It implies that  $\widehat{\mathbf{M}}_{FP}$  and  $\mathbf{P}_2$  are colinear. The proofs of Corollary V.1 is complete and it concludes the argument of Theorem IV.1.

## APPENDIX VII

### PROOF OF LEMMA V.1

The argument goes by contradiction. We thus assume that  $\omega(\mathbf{M})$  does not contain any periodic orbit. Let  $\mathcal{K}$  be a compact subset of  $\mathcal{D}$  containing both the orbit associated to  $\mathbf{M}$  and  $\omega(\mathbf{M})$ .

Let  $\mathbf{Q} \in \omega(\mathbf{M})$ . Then, there exists a sequence  $(f^{n_j}(\mathbf{M}))_{j \geq 0}$  converging to  $\mathbf{Q}$ , as  $j$  tends to  $+\infty$ , with  $(n_j)_{j \geq 0}$  a strictly increasing sequence of integers tending to  $+\infty$ .

Let  $\varepsilon \in (0, 1)$  small enough and  $n_{j_0} \in \mathbb{N}$  such that  $\|f^{n_{j_0}}(\mathbf{M}) - \mathbf{Q}\| \leq \varepsilon$ . It is easy to see that there exists a constant  $K$  only depending on  $\mathcal{K}$  such that  $(1 - K\varepsilon)\mathbf{Q} \leq f^{n_{j_0}}(\mathbf{M}) \leq (1 + K\varepsilon)\mathbf{Q}$ . Using Proposition V.2, we have for every  $p \geq 0$ ,

$$(1 - K\varepsilon)f^p(\mathbf{Q}) \leq f^{n_0+p}(\mathbf{M}) \leq (1 + K\varepsilon)f^p(\mathbf{Q}). \quad (\text{VII.36})$$

Since  $\mathbf{Q}$  is a cluster point for the orbit associated to  $\mathbf{M}$ , there exists  $n_{j_1} \geq 0$  such that

$$\left(1 - \frac{K\varepsilon}{4}\right) \mathbf{Q} \leq f^{n_1}(\mathbf{M}) \leq \left(1 + \frac{K\varepsilon}{4}\right) \mathbf{Q}.$$

Using (VII.36) and the previous equation, there exists  $p$  large enough such that

$$\left(1 - \frac{K\varepsilon}{2}\right) \mathbf{Q} \leq f^p(\mathbf{Q}) \leq \left(1 + \frac{K\varepsilon}{2}\right) \mathbf{Q}. \quad (\text{VII.37})$$

We set  $\mathbf{Q}_0 := \mathbf{Q}$  and  $\varepsilon_0$  "maximal" with respect to (VII.37), i.e.,  $\varepsilon_0$  is the smallest positive real number so that  $(1 - \varepsilon_0)\mathbf{Q}_0 \leq f^p(\mathbf{Q}_0) \leq (1 + \varepsilon_0)\mathbf{Q}_0$  holds true. Then,  $\varepsilon_0 \leq \frac{K\varepsilon}{2}$  and one of the two previous inequalities is not strict, by maximality of  $\varepsilon_0$ . Moreover,  $\varepsilon_0 > 0$ . Indeed, if it were not the case, then  $\mathbf{Q}_0$  and  $f^p(\mathbf{Q}_0)$  would be comparable

and, according to Corollary V.1, the orbit associated to  $\mathbf{Q}_0$  would be periodic. We now consider the subset  $V$  of  $\omega(\mathbf{M})$ , made of the matrices  $\mathbf{P}$  such that there exists  $\varepsilon(\mathbf{P}) > 0$  such that

$$(1 - \varepsilon(\mathbf{P})) \mathbf{P} \leq f^p(\mathbf{P}) \leq (1 + \varepsilon(\mathbf{P})) \mathbf{P}, \quad (\text{VII.38})$$

and  $\varepsilon(\mathbf{P})$  is "maximal" with respect to (VII.38).

We showed previously that  $V$  is not empty since  $\mathbf{Q} \in V$ . We next show that  $\bar{\varepsilon} = \inf_{\mathbf{P} \in V} \varepsilon(\mathbf{P}) = 0$ .

By definition of  $\bar{\varepsilon}$ , there exists two sequences  $(\mathbf{Q}^{(j)})_{j \geq 0}$  and  $(\varepsilon(\mathbf{Q}^{(j)}))_{j \geq 0}$  such that  $(\varepsilon(\mathbf{Q}^{(j)}))_{j \geq 0}$  converges to  $\bar{\varepsilon}$ , as  $j$  tends to  $+\infty$ . Up to considering a subsequence in the compact  $\omega(\mathbf{M})$ , we may assume that  $(\mathbf{Q}^{(j)})_{j \geq 0}$  converges to some  $\bar{\mathbf{Q}} \in \omega(\mathbf{M})$ . Passing to the limit in (VII.38), we get

$$(1 - \bar{\varepsilon}) \bar{\mathbf{Q}} \leq f^p(\bar{\mathbf{Q}}) \leq (1 + \bar{\varepsilon}) \bar{\mathbf{Q}}. \quad (\text{VII.39})$$

If  $\bar{\varepsilon} > 0$ , then necessarily  $\bar{\mathbf{Q}} \in V$  and  $\bar{\varepsilon}$  is "maximal" with respect to (VII.39). Since  $f$  is eventually strictly increasing, we get  $(1 - \bar{\varepsilon}) f^n(\bar{\mathbf{Q}}) \leq f^p(f^n(\bar{\mathbf{Q}})) < (1 + \bar{\varepsilon}) f^n(\bar{\mathbf{Q}})$ . Setting  $\tilde{\mathbf{Q}} := f^n(\bar{\mathbf{Q}})$ , then  $\tilde{\mathbf{Q}}$  belongs to  $\omega(\mathbf{M})$  since the latter is an invariant set with respect to  $f$ . Choosing  $\tilde{\varepsilon}$  "maximal" with respect to

$$(1 - \tilde{\varepsilon}) \tilde{\mathbf{Q}} \leq f^p(\tilde{\mathbf{Q}}) \leq (1 + \tilde{\varepsilon}) \tilde{\mathbf{Q}},$$

we first have that  $\tilde{\varepsilon} > 0$  (otherwise we would have a periodic orbit) and  $\tilde{\varepsilon} < \bar{\varepsilon}$ . We finally proved that  $\tilde{\mathbf{Q}} \in V$  with  $0 < \tilde{\varepsilon} = \varepsilon(\tilde{\mathbf{Q}}) < \bar{\varepsilon}$ . This is a contradiction with the minimality of  $\bar{\varepsilon}$ . Therefore,  $\bar{\varepsilon} = 0$ , which implies that  $\bar{\mathbf{Q}} = f^p(\bar{\mathbf{Q}})$ , i.e.  $\omega(\mathbf{M})$  contains a periodic orbit. Lemma V.2 is proved. ■

## APPENDIX VIII

### PROOF OF LEMMA V.2

Let  $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{D}$  whose associated orbits are periodic, with respective (positive) periods  $l_1$  and  $l_2$ .

We first show that  $\mathbf{M}_1$  and  $\mathbf{M}_2$  are colinear, which will imply that  $l_1 = l_2$ .

For  $i = 1, 2$ , the orbit associated to  $\mathbf{M}_i$  is the set  $\{\mathbf{M}_i, f(\mathbf{M}_i), \dots, f^{l_i-1}(\mathbf{M}_i)\}$ . Consider  $\mathbf{M} := \mathbf{M}_1 + \mathbf{M}_2$  and  $l := l_1 l_2$ . Then,  $f(\mathbf{M}) = f(\mathbf{M}_1 + \mathbf{M}_2) \geq f(\mathbf{M}_1) + f(\mathbf{M}_2)$  and, for every  $k \geq 0$ , we have

$$f^k(\mathbf{M}) \geq f^k(\mathbf{M}_1) + f^k(\mathbf{M}_2).$$

It implies that  $f^l(\mathbf{M}) \geq f^l(\mathbf{M}_1) + f^l(\mathbf{M}_2) = \mathbf{M}_1 + \mathbf{M}_2 = \mathbf{M}$ . By Corollary V.1, we get that  $f^l(\mathbf{M}) = \mathbf{M}$ . It implies that all the previous inequalities must be in fact equalities and, in particular, we have  $f(\mathbf{M}) = f(\mathbf{M}_1) + f(\mathbf{M}_2)$ . By (P2), we deduce that  $\mathbf{M}_1$  and  $\mathbf{M}_2$  are colinear. It remains to show that a periodic orbit reduces to a single point.

Consider  $\mathbf{M} \in \mathcal{D}$  such that

$$\begin{cases} l \geq 1, & f^l(\mathbf{M}) = \mathbf{M}, \\ \text{(if } l = 1, \text{ no condition)} & f^{l-1}(\mathbf{M}) \neq \mathbf{M}. \end{cases}$$

We have to prove that  $l = 1$ .

Since the orbit associated to every  $f^j(\mathbf{M})$ ,  $0 \leq j \leq l$ , is again  $\omega(\mathbf{M})$  and thus finite, we deduce that  $f^j(\mathbf{M})$  must be colinear to  $\mathbf{M}$ , according to what precedes. Then, for every

$0 \leq j \leq l - 1$ , we have  $f^j(\mathbf{M}) = \lambda_j \mathbf{M}$ , for some  $\lambda_j > 0$ . Obviously,  $\lambda_0 = \lambda_l = 1$ . In particular, we have  $f(\mathbf{M}) = \lambda_1 \mathbf{M}$ , implying that, either  $f(\mathbf{M}) \leq \mathbf{M}$  or  $f(\mathbf{M}) \geq \mathbf{M}$ . By (C1) of Corollary V.1, we get that  $\mathbf{M}$  is a fixed point of  $f$ . The proof of Lemma V.1 is complete. ■

#### REFERENCES

- [1] E. Conte, M. Lops and G. Ricci, "Asymptotically optimum radar detection in compound-Gaussian clutter", *IEEE Trans. Aerosp. Electron. System*, vol. 31, no. 2, pp. 617-625, Apr. 1995.
- [2] F. Gini, "Sub-optimum coherent radar detection in a mixture of K-distributed and Gaussian clutter", *IEE Proc. Radar, Sonar and Navigation*, vol. 144, no. 1, pp. 39-48, Feb. 1997.
- [3] E. Jay, J. P. Ovarlez, D. Declercq and P. Duvaut, "BORD : bayesian optimum radar detector", *Signal Processing*, vol. 83, no. 6, pp. 1151-1162, Jun. 2003.
- [4] E. Jay, Détection en environnement non-Gaussien, *Ph.D. Thesis*, University of Cergy-Pontoise / ONERA, France, Jun. 2002.
- [5] K. Yao, "A representation theorem and its applications to spherically invariant random processes", *IEEE Trans. Inform. Theory*, vol. 19, no. 5, pp. 600-608, Sep. 1973.
- [6] J.B. Billingsley, Ground Clutter Measurements for Surface-Sited Radar, *Technical Report 780*, MIT, February 1993.
- [7] E. J. Kelly "An adaptive detection algorithm", *IEEE Trans. Aerosp. Electron. System*, vol. 23, no. 1, pp. 115-127, Nov. 1986.
- [8] F. C. Robey, D. R. Fuhrmann, E. J. Kelly and R. Nitzberg, "A CFAR adaptive matched filter detector", *Trans. Aerosp. Electron. System*, vol. 23, no. 1, pp. 208 - 216, Jan. 1992.
- [9] E. Conte, M. Lops and G. Ricci, "Adaptive radar detection in compound-Gaussian clutter", *Proc. of the European Signal Processing Conf.*, Edinburgh, Scotland, Sep. 1994.
- [10] F. Gini, M. V. Greco and L. Verrazzani, "Detection problem in mixed clutter environment as a Gaussian problem by adaptive pre-processing", *Electronics Letters*, vol. 31, no. 14, pp. 1189-1190, Jul. 1995.
- [11] R. S. Raghavan and N. B. Pulsone, "A generalization of the adaptive matched filter receiver for array detection in a class of a non-Gaussian interference", *Proc. of the Adaptive Sensor Array Processing (ASAP) Workshop*, Lexington, MA, pp. 499-517, Mar. 1996.
- [12] F. Gini and M. V Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter", *Signal Processing*, special section on Signal Processing with Heavy Tailed Distributions, vol. 82, no. 12, pp. 1847-1859, Dec. 2002.
- [13] E. Conte, A. De Maio and G. Ricci, "Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection", *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 1908-1915, Aug. 2002.
- [14] R. A. Horn and Ch. R. Johnson, "Matrix analysis", Cambridge University Press, Cambridge, U.K., 1985.
- [15] F. Pascal, J. P. Ovarlez, P. Forster and P. Larzabal, "Constant false alarm rate detection in spherically invariant random processes", *Proc. of the European Signal Processing Conf.*, Vienna, Austria, pp. 2143-2146, Sep. 2004.
- [16] F. Pascal, P. Forster, J. P. Ovarlez and P. Larzabal, "Theoretical analysis of an improved covariance matrix estimator in non-Gaussian noise", *Proc. IEEE-ICASSP*, Philadelphia, Pennsylvania, USA, vol. IV, pp. 69-72, Mar. 2005.



# Publications

## Reuves à comité de lecture

[J1] F. Pascal, Y. Chitour, J-P. Ovarlez, P. Forster and P. Larzabal, "Covariance Structure Maximum Likelihood Estimates in Compound Gaussian Noise : Existence and Algorithm Analysis", *IEEE Trans. on Signal Processing*. (accepté)

[J2] F. Pascal, P. Forster, J-P. Ovarlez and P. Larzabal, "Performance Analysis of Covariance Matrix Estimates in Impulsive Noise", *IEEE Trans. on Signal Processing*. (soumis en octobre 2006)

[J3] S. Bausson, F. Pascal, P. Forster, J-P. Ovarlez and P. Larzabal, "First and Second Order Moments of the Normalized Sample Covariance Matrix of Spherically Invariant Random Vectors", *IEEE Signal Processing Letters*. (accepté)

[J4] F. Pascal, P. Forster, J-P. Ovarlez and P. Larzabal, "An Improved Estimation Scheme for the GLRT Detection in Non Gaussian Noise", *IEEE Trans. on AES*. (en rédaction)

## Conférences avec actes et à comité de lecture

[C1] F. Pascal, J.-P. Ovarlez, P. Forster and P. Larzabal, "Constant False Alarm Rate Detection in Spherically Invariant Random Processes", *Proc. of the European Signal Processing Conference, EUSIPCO-04*, Vienna, pp. 2143-2146, Sep. 2004.

[C2] F. Pascal, J.-P. Ovarlez, P. Forster and P. Larzabal, "Radar Detection in Compound-Gaussian Clutter", *Proc. of the IEEE-RADAR 2004*, Toulouse, Oct. 2004.

[C3] J.-P. Ovarlez, E. Jay and F. Pascal, "Bayesian Optimum Radar Detector Performance Against Ground Data", *Proc. of the IEEE-RADAR 2004*, Toulouse, Oct. 2004.

[C4] F. Pascal, P. Forster, J.-P. Ovarlez and P. Larzabal, "Theoretical Analysis of an Improved Covariance Matrix Estimator in Non-Gaussian Noise", *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-05*, Philadelphia, vol. IV, pp. 69-72, Mar. 2005.

[C5] F. Pascal, J.-P. Ovarlez, P. Forster and P. Larzabal, "Propriété CFAR-Matrice du Détecteur BORD - Application Radar sur Signaux Expérimentaux non Gaussien", *Actes du Vingtième Colloque GRETSI 2005*, Louvain-La-Neuve, Belgium, Sep. 2005.

[C6] F. Pascal, J.-P. Ovarlez, P. Forster and P. Larzabal, "On a SIRV-CFAR detector with radar experimentations in impulsive noise", *Accepted to the European Signal Processing Conference, EUSIPCO-06*, Florence, Italy, Sep. 2006.

## **Séminaires**

[S1] "Utilisation de Techniques non Gaussiennes pour la Détection sur des Images Polarimétriques", Journée scientifique "Identification Radar", ONERA, Palaiseau, Novembre 2005.

[S2] "Détection et Estimation en Environnement Non Gaussien - Application sur données réelles", Journées des doctorants, ONERA, Chatillon, Mai 2005 et Janvier 2006.

[S3] "Détection Radar en Environnement de Fouillis Non-Gaussien", Comité d'Evaluation et d'Oriention (CEO), ONERA, Palaiseau, Mars 2004.

[S4] "Nouvel estimateur de la matrice de covariance pour la détection en environnement non Gaussien", GEA, Ville d'Avray, Décembre 2004.

## **Rapport**

[R1] F. Pascal, "Détection en environnement non Gaussien : loi d'un rapport de vraisemblance", Rapport de stage de DEA, Septembre 2003.

# Bibliographie

- [1] M. Abramowitz and I.A. Stegun, "Handbook of Mathematical Functions", *National Bureau of Standard*, AMS 55, June 1964.
- [2] R.A. Altes, "Target position estimation in radar and sonar, and generalized ambiguity analysis for maximum likelihood parameter estimation", *IEEE-Proc.*, vol. 67, no. 6, pp. 920-930, June 1979.
- [3] T.W. Anderson, "An Introduction to Multivariate Statistical Analysis", *John Wiley & Sons*, New York, 1984.
- [4] T.J. Barnard and D.D. Weiner, "Non-Gaussian clutter modeling with generalized spherically invariant random vectors", *IEEE Trans.-SP*, vol. 44, no. 10, pp. 2384-2390, October 1996.
- [5] S. Bausson, F. Pascal, P. Forster, J.-P. Ovarlez and P. Larzabal, "First and Second Order Moments of the Normalized Sample Covariance Matrix of Spherically Invariant Random Vectors", *accepted to IEEE SP Letters in june 2006*.
- [6] J.B. Billingsley, "Ground Clutter Measurements for Surface-Sited Radar", *Technical Report 780, MIT*, February 1993.
- [7] E. Conte and M. Longo, "Characterization of radar clutter as a spherically invariant random process", *IEE Proc.-Pt.F*, vol. 134, no. 2, pp. 191-197, April 1987.
- [8] E. Conte, M. Longo, M. Lops, "Modelling and simulation of non-Rayleigh radar clutter", *IEE Proc.-Pt.F*, vol. 138, no. 2, pp. 121-138, April 1991.
- [9] E. Conte and G. Ricci, "Performance prediction in compound-Gaussian clutter", *IEEE Trans.-AES*, vol. 30, no. 2, pp. 611-616, April 1994.
- [10] E. Conte, M. Lops and G. Ricci, "Asymptotically optimum radar detection in compound-Gaussian clutter", *IEEE Trans.-AES*, vol. 31, no. 2, pp. 617-625, April 1995.
- [11] E. Conte and G. Ricci, "Sensitivity study of GLRT detection in compound-Gaussian clutter", *IEEE Trans.-AES*, vol. 34, no. 1, pp. 308-316, January 1998.
- [12] E. Conte, M. Lops and G. Ricci, "Incoherent radar detection in compound-Gaussian clutter", *IEEE Trans.-AES*, vol. 35, no. 3, pp. 790-800, July 1999.
- [13] E. Conte, A. De Maio and C. Galdi, "Signal detection in compound-Gaussian noise : Neymann-Pearson and CFAR detectors", *IEEE Trans.-SP*, vol. 48, no. 2, pp. 419-428, February 2000.
- [14] E. Conte, A. De Maio and G. Ricci, "Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection", *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 1908-1915, August 2002.
- [15] A. Farina, A. Russo and F.A. Studer, "Advanced models of target and disturbances and related processors", *Proc. CCIR*, pp. 151-158, Washington D.C, May 1985.
- [16] A. Farina, A. Russo, F. Scannapieco and S. Barbarossa, "Theory of radar detection in coherent Weibull clutter", *IEE Communication, Radar and Signal Processing, Pt-F*, vol. 134, no. 2, pp. 174-190, April 1987.

- [17] F. Gini, "Sub-optimum coherent radar detection in a mixture of K-distributed and Gaussian clutter", *IEE Proc.Radar, Sonar Navig.*, vol. 144, no. 1, pp. 39-48, February 1997.
- [18] F. Gini, "A cumulant-based adaptive technique for coherent radar detection in a mixture of K-distributed clutter and Gaussian disturbance", *IEEE Trans.-SP*, vol. 45, no. 6, pp. 1507-1519, June 1997.
- [19] F. Gini, M.V. Greco, A. Farina and P. Lombardo, "Optimum and mismatched detection against K-distributed plus Gaussian clutter", *IEEE Trans.-AES*, vol. 34, no. 3, pp. 860-876, July 1998.
- [20] F. Gini, M.V. Greco and A. Farina, "Clairvoyant and adaptive signal detection in non-Gaussian clutter : a data-dependent threshold interpretation", *IEEE Trans.-SP*, vol. 47, no. 6, pp. 1522-1531, June 1999.
- [21] F. Gini and M.V. Greco, "Sub-optimum approach to adaptive coherent radar detection in compound-Gaussian clutter", *IEEE Trans.-AES*, vol. 35, no. 3, pp. 1095-1103, July 1999.
- [22] F. Gini, "Estimation of chirp radar signals in compound Gaussian clutter : a cyclostationary approach", *IEEE Trans.-SP*, vol. 48, no. 4, pp. 1029-1039, April 2000.
- [23] F. Gini, "Performance analysis of two structured covariance matrix estimators in compound-Gaussian clutter", *Signal Processing*, vol. 80, pp. 365-371, 2000.
- [24] F. Gini, M.V. Greco, M. Diani, L. Verrazzani, "Performance analysis of two adaptive radar detectors against non-Gaussian real sea clutter data", *IEEE Trans.-AES*, vol. 36, no. 4, pp. 1429-1439, October 2000.
- [25] F. Gini, M.V. Greco, A. Farina and P. Lombardo, Note on "Optimum and mismatched detection against K-distributed plus Gaussian clutter", *IEEE Trans.-AES*, vol. 37, no. 1, pp. 296-297, January 2001.
- [26] F. Gini, A. Farina and M.V. Greco, "Selected list of references on radar signal processing", *IEEE Trans.-AES*, vol. 37, no. 1, pp. 329-359, January 2001.
- [27] F. Gini and M. V Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter", *Signal Processing*, special section on Signal Processing with Heavy Tailed Distributions, vol. 82, no. 12, pp. 1847-1859, December 2002.
- [28] J. Goldman, "Detection in the Presence of Spherically Symmetric Random Vectors", *IEEE Trans.-IT*, vol. 22, no. 1, pp. 52-59, January 1976.
- [29] A.K. Gupta and D.K. Nagar, "Matrix Variate Distributions", *Chapman & Hall/CRC*, 2000.
- [30] E. Jakeman and P.N. Pusey, "A model for non-Rayleigh sea echo", *IEEE-Trans.-AP*, vol. 24, no. 6, pp. 806-814, November 1976.
- [31] E. Jay, "Détection en Environnement Non-Gaussien", *Ph.D. Thesis*, University of Cergy-Pontoise / ONERA, France, June 2002.
- [32] E. Jay, J.-P. Ovarlez, D. Declercq and P. Duvaut, "BORD : Bayesian Optimum Radar Detector", *Signal Processing*, vol. 83, no. 6, pp. 1151-1162, Jun. 2003.
- [33] S.M. Kay, "Fundamentals of statistical signal processing - Detection theory - Vol.2", *Prentice Hall PTR*, 1998.
- [34] J.T. Kent and D.E. Tyler, "Redescending  $M$ -Estimates of Multivariate Location and Scatter", *Annals of Statistics*, vol. 19, no. 4, pp. 2102-2119, December 1991.
- [35] V.A. Korado, "Optimum detection of signals with random parameters against the background of noise of unknown intensity under conditions of constant false alarm probability", *Radio Engineering and Electronic Physics*, vol. 13, 1968.

- 
- [36] S. Kraut, L.L. Scharf and L.T. McWhorter, "Adaptive Subspace Detectors", *IEEE Trans.-SP*, vol. 49, no. 1, pp. 1-16, January 2001.
- [37] R.A. Maronna, "Robust  $M$ -Estimators of Multivariate Location and Scatter", *Annals of Statistics*, vol. 4, no. 1, pp. 51-67, January 1976.
- [38] F. Pascal, P. Forster, J.-P. Ovarlez and P. Larzabal, "Theoretical Analysis of an Improved Covariance Matrix Estimator in Non-Gaussian Noise", *Proc. IEEE-ICASSP*, Philadelphia, vol. IV, pp. 69-72, March 2005.
- [39] F. Pascal, J.-P. Ovarlez, P. Forster and P. Larzabal, "On a SIRV-CFAR Detector with Radar Experimentations in Impulsive Noise", *Proc. of the European Signal Processing Conf.*, Florence, Italy, September 2006.
- [40] F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster and P. Larzabal, "Covariance Structure Maximum Likelihood Estimates in Compound Gaussian Noise : Existence and Algorithm Analysis", *IEEE Trans. on Signal Processing*. (Accepted)
- [41] F.A. Pentini, A. Farina and F. Zirilli, "Radar detection of targets located in a coherent K distributed clutter background", *IEE Proc.-F*, vol.139, no.3, pp.238-245, June 1992.
- [42] B. Picinbono and G. Vezzosi, "Détection d'un signal certain dans un bruit non stationnaire et non Gaussien", *Annales des Télécomm.*, vol. 25, pp. 433-439, 1970.
- [43] R.S. Raghavan, H.F. Qiu and D.J. McLaughlin, "CFAR detection in clutter with unknown correlation properties", *IEEE Trans.-AES*, vol. 31, no. 2, pp. 647-656, April 1995.
- [44] M. Rangaswamy, D.D. Weiner and A. Öztürk, "Non-Gaussian Random Vector Identification Using Spherically Invariant Random Processes", *IEEE Trans.-AES*, vol. 29, no. 1, pp. 111-123, January 1993.
- [45] C.P. Robert, "L'analyse Statistique Bayésienne", *Economica*, Paris, 1992.
- [46] C.P. Robert, "The Bayesian Choice : a Decision-Theoretic Motivation", *Springer-Verlag*, New York, 1994.
- [47] C.P. Robert, "Méthodes de Monte Carlo par chaînes de Markov", *Economica*, Paris, 1996.
- [48] K.J. Sangston, F. Gini, M. V. Greco and A. Farina, "Structures for radar detection in compound Gaussian clutter", *IEEE Trans.-AES*, vol. 35, no. 2, pp. 445-458, April 1999.
- [49] M.A. Tanner, "Tools for Statistical Inference : Methods for the Exploration of Posterior Distributions and Likelihood Functions", *Springer-Verlag*, 3rd ed., 1996.
- [50] G.V. Trunk and S.F. George, "Detection of targets in non-Gaussian sea clutter", *IEEE Trans.-AES*, vol. 6, no. 8, pp. 620-628, September 1970.
- [51] A.W. van der Vaart, "Asymptotic Statistics", *Cambridge University Press*, 1998.
- [52] K.D. Ward, "Compound representation of high resolution sea clutter", *Electronics letters*, vol. 17, no. 16, pp. 561-563, August 1981.
- [53] S. Watts, "Radar detection prediction in sea clutter using the compound K-distribution model", *IEE Proc.Pt-F*, vol. 132, no. 7, pp. 613-620, December 1985.
- [54] K. Yao, "A Representation Theorem and its Applications to Spherically Invariant Random Processes", *IEEE Trans.-IT*, vol. 19, no. 5, pp. 600-608, September 1973.



## Résumé

Dans le contexte très général de la détection radar, les détecteurs classiques, basés sur l'hypothèse d'un bruit Gaussien, sont souvent mis en défaut dès lors que l'environnement (fouillis de sol, de mer) devient inhomogène, voire impulsif, s'écartant très vite du modèle Gaussien. Des modèles physiques de fouillis basés sur les modèles de bruit composé (SIRP, Compound Gaussian Processes) permettent de mieux représenter la réalité (variations spatiales de puissance et nature de fouillis, transitions, ...). Ces modèles dépendent cependant de paramètres (matrice de covariance, loi de texture, paramètres de "disturbance") qu'il devient nécessaire d'estimer. Une fois ces paramètres estimés, il est possible de construire des détecteurs radar optimaux (Generalized Likelihood Ratio Test - Linear Quadratic) pour ces environnements. Cette thèse, qui s'appuie sur ces modèles, propose une analyse complète de diverses procédures d'estimation de matrices de covariance, associées à ce problème de détection. Une étude statistique des principaux estimateurs de matrice de covariance, utilisés actuellement, est réalisée. De plus, un nouvel estimateur est proposé : l'estimateur du point fixe, très attractif grâce à ses bonnes propriétés statistiques et "radaristiques".

Elle décrit également les performances et les propriétés théoriques (SIRV-CFAR) du détecteur GLRT-LQ construits avec ces nouveaux estimateurs. En particulier, on montre l'invariance du détecteur à la loi de la texture mais également à la matrice de covariance régissant les propriétés spectrales du fouillis. Ces nouveaux détecteurs sont ensuite analysés sur des données simulées mais également testés sur des données réelles de fouillis de sol.

**Mots-clés:** Détection Radar, Estimation de paramètres, Fouillis non-Gaussien, Détecteur GLRT-LQ, SIRV, Matrice de covariance, Maximum de Vraisemblance.

## Abstract

This thesis deals with radar detection in impulsive noise contexts. Indeed, under Gaussian assumptions, classical detectors, like Optimum Gaussian Detector, present several failures when the background scatterers are heterogeneous and non Gaussian, which is the case for ground or sea clutter. Clutter physical models based on compound noise modeling (SIRP, Compound Gaussian Processes) allow to correctly describe the reality (range power variations or clutter transitions areas). However, these models depend on several unknown parameters (covariance matrix, statistical distribution of the texture, disturbance parameters) which have to be estimated. When this estimation scheme is carried out, it is possible to build optimum radar detectors (Generalized Likelihood Ratio Test - Linear Quadratic) associated to this non Gaussian background. Based on these noise models, this thesis presents a complete analysis of several estimation schemes of the noise covariance matrix, associated to the detection problem. A statistical study of the main covariance matrix estimates which are used in the literature, is performed. Moreover, an improved estimate is proposed : the Fixed Point estimate, very attractive thanks to its good statistical and detection properties.

This thesis also describes detection performance and theoretical properties (texture-CFAR and matrix-CFAR) of the GLRT-LQ detector built with the studied covariance matrix estimates. In particular, the detector invariance to the texture distribution and to the covariance matrix structure is shown. Finally, these detectors are analyzed with simulated data and then, experimented on real ground clutter data.

**Keywords:** Radar Detection, Parameters Estimation, Non-Gaussian Noise, GLRT-LQ, SIRV, Covariance Matrix, Maximum Likelihood.

