



HAL
open science

Détection en Environnement non Gaussien

Emmanuelle Jay

► **To cite this version:**

Emmanuelle Jay. Détection en Environnement non Gaussien. Traitement du signal et de l'image [eess.SP]. Université de Cergy Pontoise, 2002. Français. NNT: . tel-00174276

HAL Id: tel-00174276

<https://theses.hal.science/tel-00174276>

Submitted on 22 Sep 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.



Université de Cergy Pontoise
Ecole Doctorale Sciences et Ingénierie

École Doctorale ETIS

Détection en Environnement Non-Gaussien

THÈSE

présentée et soutenue publiquement le 14 Juin 2002

pour l'obtention du

Doctorat de l'université de Cergy-Pontoise
(spécialité Traitement du Signal)

par

Emmanuelle JAY

Composition du jury

Président : LACOUME Jean-Louis
Rapporteurs : CASTANIÉ Francis
FARINA Alfonso
Examineurs : DECLERCQ David
DUVAUT Patrick (Directeur de Thèse)
OVARLEZ Jean-Philippe
PICARD Dominique
Invité : DELHOTE Christian

Office National d'Études et de Recherches Aérospatiales

ONERA



Mis en page avec la classe thloria.

Remerciements

Pour avoir effectué ma thèse dans l'unité Traitement du Signal du Département ElectroMagnétisme et Radar de l'ONERA en collaboration avec l'Equipe Traitement des Images et du Signal de l'ENSEA, je tiens avant tout à saluer toutes les personnes de ces deux équipes qui m'ont toujours soutenue tout au long de ce travail. Au-delà de simples remerciements, je leur adresse mes sincères amitiés.

Parmi toutes ces personnes, je tiens à mentionner Patrick Duvaut, mon directeur de thèse, qui m'a attribuée toute sa confiance pour ce travail. Je lui témoigne toute ma reconnaissance.

Selon un fameux proverbe, que je ne citerai pas, sans les idées innovantes et constructives de Jean-Philippe Ovarlez, le sujet de cette thèse n'aurait sûrement jamais vu le jour. Son investissement dans ce travail n'a pas de commune mesure et il m'est difficilement possible de trouver les mots pour lui témoigner tous mes ressentiments. Peut-être tout simplement lui dire merci ?

Dans mon souhait de rendre hommage aux personnes qui m'ont été les plus proches dans mon travail, je souhaiterais également citer David Declercq pour sa disponibilité et sa contribution déterminante dans ce travail. Dans les moments délicats il m'a toujours soutenue dans ma démarche et a toujours cru en l'aboutissement du travail entrepris. Un grand coup de chapeau à toi !

Ce fut un grand honneur pour moi qu'Alfonso Farina accepte de rapporter sur mon travail. Maître incontestable dans le domaine du radar, je ne tarirais pas d'éloges en sa faveur alors je lui adresse un infini grazie mille !

C'est avec toute ma sympathie et toute ma considération que je m'adresse à Francis Castanié et à Jean-Louis Lacoume. Je les remercie vivement de leur présence dans mon jury en qualités respectives de rapporteur et de président du jury. Encore deux grandes références incontournables !

De chaleureuses pensées à Dominique Picard que je remercie de tout coeur d'avoir participé au jury de la thèse. Il était important, à mon sens, qu'elle apporte toute son expérience en qualité de statisticienne reconnue et confirmée dans ce travail. Merci Do.

Par la présence de Christian Delhote dans mon jury, je voudrais remercier Thalès Air Defence pour m'avoir permis de valider les résultats théoriques de la thèse sur signaux expérimentaux. Cet apport a donné toute sa signification à ce travail.

En résumé, je vous remercie tous de m'avoir donné la chance de vivre cette expérience de thèse, qui pour moi s'est révélée être une belle aventure. Les moments de doute mais aussi de joie ont rythmé ces trois dernières années, enrichies par mes participations à divers congrès. Merci à Christèle, Robert, Marc, Christian(s), Luc(s), Christine, Thierry, Anne-Sophie, Eric, Laurent(s), Lola, Goulou, Camille, Nine, André et ma maman ... Papa : ca y'est ! J'ai fini de chercher ... désormais je vais peut-être essayer de trouver ... Amitiés. Manou.

*A Goulou,
De la part de Lola.*

Table des matières

Introduction	1
---------------------	----------

Chapitre 1
État de l'art

1.1	Théorie de la détection radar	5
1.2	Le détecteur classique et l'Optimum Gaussian Detector (OGD)	7
1.3	Techniques d'optimisation du détecteur classique en environnement non-Gaussien	9
1.4	Modélisation de l'environnement non-Gaussien et détection cohérente	11

Chapitre 2
Les processus aléatoires sphériquement invariants : les SIRP

2.1	Modélisation de processus aléatoires non-Gaussiens	15
2.1.1	Théorie des processus Gaussiens composés et propriétés	15
2.1.1.1	Définitions	15
2.1.1.2	Le théorème de représentation d'un SIRV	18
2.1.1.3	Propriété d'invariance des SIRV par transformation linéaire	18
2.1.1.4	Propriétés sur l'addition de deux SIRV	20
2.1.1.5	Caractérisation d'un SIRV blanc	22
2.1.1.6	Caractérisation d'un SIRV par les fonctions h_m	23
2.1.1.7	Détermination des fonctions h_m par la fonction caractéristique	24
2.1.2	Lois jointes de vecteurs non-Gaussiens issus d'un SIRP	25
2.1.2.1	Densités de probabilité pour lesquelles $p(\tau)$ est connue	25
2.1.2.2	Densités de probabilité pour lesquelles $p(\tau)$ n'est pas connue	30
2.1.3	Les SIRV généralisés	32
2.1.3.1	Propriétés des SIRV généralisés	34
2.2	Stratégies de détection optimales pour un environnement SIRV connu a priori	35
2.2.1	Le test du rapport de vraisemblance (RV) pour les SIRV	35
2.2.2	Le rapport de vraisemblance généralisé (RVG)	36
2.2.3	Optimum K Detector : l'OKD	37

2.2.4	Optimum Laplace Detector : OLD	38
2.2.5	Optimum Student-t Detector : OStD	38
2.2.6	Optimum Cauchy Detector : OCD	38
2.2.7	Optimum χ Detector : O χ D	38
2.2.8	Optimum generalized Rayleigh Detector : OgRD	39
2.2.9	Optimum Weibull Detector : OWD	39
2.2.10	Optimum Rice Detector : ORD	39
2.2.11	Optimum Clutter plus Gaussian Detector : l'OCpGD	40

<p>Chapitre 3</p> <p>Stratégies de détection optimales pour un environnement inconnu</p>
--

3.1	Analyse bayésienne du modèle SIRV	43
3.1.1	La règle de Bayes	43
3.1.2	Choix du prior	44
3.1.2.1	Prior non-informatif	45
3.1.2.2	Prior conjugué	45
3.1.2.3	Expression de la loi <i>a posteriori</i> instantanée de la texture	46
3.2	Estimation NSCM de la matrice de covariance	46
3.2.1	Modèle de matrice de covariance utilisé pour les simulations	47
3.2.2	Propriété statistique de la matrice estimée au sens NSCM	48
3.3	Le PEOD : Padé Estimated Optimum (radar) Detector	51
3.3.1	Estimation de la loi de la <i>texture</i> du SIRV par les approximants de Padé	51
3.3.2	Expression du PEOD après formation du rapport de vraisemblance généralisé	52
3.3.3	Avantages et inconvénients de la structure du PEOD	53
3.3.4	Performances de détection du PEOD	54
3.3.4.1	Comparaison à l'optimal	54
3.3.4.2	Environnement inconnu	56
3.3.4.3	Contexte des simulations	56
3.3.5	Synthèse sur le PEOD	61
3.3.5.1	Le PEOD "idéal"	61
3.3.5.2	Le PEOD "non-idéal"	61
3.4	Le BORD : Bayesian Optimum Radar Detector	62
3.4.1	Estimation bayésienne de la loi de la <i>texture</i> du SIRV	62
3.4.2	Expression du détecteur BORD	63
3.4.3	Version adaptative du BORD	64
3.4.4	Propriété TFAC du BORD	64
3.4.5	Performances de détection du BORD	65

3.4.5.1	Contexte des simulations	65
3.4.5.2	Résultats des simulations	66
3.4.6	Robustesse du BORD	70
3.4.7	Comportement asymptotique du BORD	71
3.4.8	Version adaptative du BORD Asymptotique	72
3.4.9	Propriété TFAC du BORD Asymptotique	72
3.4.10	Loi du BORD Asymptotique	74
3.4.10.1	Cas où \mathbf{x} est réel	75
3.4.10.2	Cas où \mathbf{x} est complexe	78
3.4.11	Performances de détection du BORD Asymptotique	80
3.4.12	Influence de la corrélation sur les performances	85
3.5	Synthèse	86
3.5.1	Le BORD	86
3.5.2	Le BORD Asymptotique	86
3.5.3	Validité du seuil théorique du BORD Asymptotique	87

Chapitre 4

Application sur signaux expérimentaux

4.1	Principe opérationnel d'un radar	89
4.1.1	Le traitement théorique optimal	89
4.1.2	Le traitement opérationnel idéal	90
4.1.3	Le traitement opérationnel classique	91
4.2	Mise en place des deux versions du BORD	94
4.2.1	Le BORD	94
4.2.2	Le BORD Asymptotique	96
4.3	Analyse des données expérimentales	97
4.3.1	Description des données	97
4.3.2	Définition des "zones" de données	98
4.3.3	Caractérisation du fouillis lors de la détection	101
4.4	Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique	109
4.4.1	Zones de bruit thermique Gaussien	109
4.4.1.1	"Zone 1"	109
4.4.1.2	"Zone 4"	113
4.4.1.3	"Zone 6"	117
4.4.2	Zones de fouillis	121
4.4.2.1	"Zone 2"	121

4.4.2.2	”Zone 5”	125
4.4.2.3	”Zone 7”	129
4.4.3	Zones transitoires bruit thermique / fouillis	133
4.4.3.1	”Zone 3”	133
4.4.3.2	”Zone 8”	137
4.5	Synthèse des résultats	141
Conclusion		143

Annexe A

Densités de probabilité d’une variable aléatoire et propriétés

A.1	Généralités	149
A.2	Loi uniforme sur $[a, b] \subset \mathbb{R}$, $\mathcal{U}([a, b])$	152
A.3	Lois Gaussienne ou normale $\mathcal{N}(m, \sigma^2)$ définie sur \mathbb{R}	153
A.4	Loi log-normale définie sur \mathbb{R}_+^*	153
A.5	Loi de Cauchy	154
A.6	Loi exponentielle $\mathcal{E}(b)$ définie sur \mathbb{R}_+	154
A.7	Loi de Laplace (ou exp. bilatérale) $\mathcal{L}(b)$ définie sur \mathbb{R}	155
A.8	Loi de Rayleigh de paramètre b , définie sur \mathbb{R}_+	155
A.9	Loi de Rice	156
A.10	Loi gamma $\mathcal{G}(a, b)$ définie sur \mathbb{R}_+	157
A.11	Loi Inverse Gamma $\mathcal{IG}(a, b)$ définie sur \mathbb{R}_+	157
A.12	La loi Beta	158
A.13	La K-distribution	158
A.14	Loi de Weibull	158
A.15	Loi du χ^2 centrée	159
A.16	Loi du χ^2 décentrée	160
A.17	Loi de Fisher F centrée	161
A.18	Loi de Fisher F' décentrée	161
A.19	Loi de Student’s -t	161
A.20	Convergences vers la loi normale	162

Annexe B

Méthodes d’estimation de densités de probabilité

B.1	Estimation empirique par la méthode des histogrammes	165
B.2	Généralisation de la méthode des histogrammes : la méthode des noyaux	166
B.3	Les approximants de Padé	171

B.3.1	L'approximation de Padé à 1 point	171
B.3.2	L'approximation de Padé à 2 points	173
B.3.3	Expressions des fonctions densité de probabilité et de répartition	173
B.3.4	Détermination de $p(x)$ et de $F(x)$ par la méthode des résidus	174
B.3.5	Détermination de l'ordre de l'approximation	175
B.3.6	Qualité de l'estimation sur quelques exemples	175
B.3.6.1	A partir des moments théoriques, pour un ordre M	176
B.3.7	Comparaison des approximations de Padé en fonction des ordres déterminés	176
B.3.8	Qualité de l'approximation de Padé en fonction du nombre d'échantillons disponibles pour l'estimation des moments	176

Annexe C

Les vecteurs Gaussiens

C.1	Rappels de probabilité	185
C.2	Définitions, propriétés	186
C.3	Les vecteurs Gaussiens complexes circulaires	187
C.3.1	Définition	187
C.3.2	Loi d'un vecteur Gaussien complexe circulaire	187
C.4	Formes quadratiques de vecteurs Gaussiens	188
C.4.1	Indépendance de formes quadratiques	188
C.4.2	Lois de formes quadratiques	188
C.4.3	Le théorème de Cochran	189
C.4.4	Formes quadratiques particulières	191
C.4.4.1	Vecteurs Gaussiens réels	191
C.4.4.2	Vecteurs Gaussiens complexes	191
C.4.4.3	SIRV réels	192
C.4.4.4	SIRV complexes	192

Publications

Bibliographie

195

Acronymes

MIT	Massachussets Institute of Technology
ONERA	Office National d'Etudes et de Recherches Aérospatiales
PEA	Projet d'Etude Amont
TAD	THALÉS Air Defence

BLU	Best Linear Unbiased
CML	Censored Maximum Likelihood
ddp	Densité de probabilité
i.i.d.	Indépendant et identiquement distribué
MC	Monte-Carlo
MCMC	Monte-Carlo par Chaîne de Markov
MISE	Mean Integrated Square Error
MV	Maximum de vraisemblance
MAP	Maximum A Poteriori
NSCM (Estimation)	Normalized Structured Covariance Matrix
SCM (Estimation)	Structured Covariance Matrix
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
TCL	Théorème Central Limite
v.a.	Variable aléatoire

ALQ	Adaptive Linear Quadratic
GLRT-LQ	Generalized Likelihood Ratio Test - Linear Quadratic
P_d	Probabilité de détection
P_{fa}	Probabilité de fausse alarme
RCB	Rapport Clutter à Bruit
RSB	Rapport Signal à Bruit
RV / LRT	Rapport de Vraisemblance
RVG / GLRT	Rapport de Vraisemblance Généralisé
TFA	Taux de Fausse Alarme
TFAC / CFAR	Taux de Fausse Alarme Constant / <i>Constant False Alarm Rate</i>

AOS-CFAR	Adaptive Order Statistic-CFAR
CA-CFAR	Cell Averaging-CFAR
GO-CFAR	Greatest-Of-CFAR
GOS-CFAR	Generalized Order Statistic -CFAR
LCOS-CFAR	Linearly-Combined-Order Statistic -CFAR
OS-CFAR	Order Statistic -CFAR
SE-CFAR	Selection-Estimation-CFAR
SO-CFAR	Smallest-Of-CFAR
TM-CFAR	Trimmed-Mean-CFAR
VI-CFAR	Variability-Index-CFAR

OCD	Optimum Cauchy Detector
OCpGD	Optimum Clutter plus Gaussian Detector
$O\chi D$	Optimum χ Detector
OgRD	Optimum generalized-Rayleigh Detector
OGD	Optimum Gaussian Detector
OKD	Optimum K Detector
OKGD	Optimum K plus Gaussian Detector
OLD	Optimum Laplace Detector
ORD	Optimum Rician Detector
OS _t D	Optimum Student-t Detector
OWD	Optimum Weibull Detector

Symboles et notations

$\hat{\theta}$	Valeur estimée de θ
\bar{z}, z^*	Conjugués du nombre complexe z
$ z $	Module du nombre complexe z
$\Re(z)$	Partie réelle de z
$\Im(z)$	Partie imaginaire de z
\dagger	Opérateur de transposition conjugaison
T	Opérateur de transposition
\mathbf{I}_m	Matrice identité de taille $m \times m$
$ \mathbf{M} , \det(\mathbf{M})$	Déterminants de la matrice \mathbf{M}
$Im(E)$	Espace Image de E
$Ker(E)$	Noyau de E
\mathbb{N}^*	$= \mathbb{N}/\{0\}$
$\mathbb{E}(\cdot)$	Espérance mathématique
$p(x/y)$	Densité de probabilité de x "sachant" y
C_n^k	Coefficients du binôme de Newton $= \frac{n!}{k!(n-k)!}, k \leq n$
$res(f, z_0)$	Résidu de la fonction f en z_0
$\delta(\cdot)$	distribution de Dirac ($\delta(x-a) = 1$ si $x = a$, 0 sinon)
$B(a, b)$	Fonction Beta de paramètres a et b
$\Gamma(a)$	Fonction Gamma en a
$J_\nu(\cdot)$	Fonction de Bessel de 1 ^{ère} espèce d'ordre ν
$I_\nu(\cdot)$	Fonction de Bessel de 1 ^{ère} espèce modifiée d'ordre ν
$K_\nu(\cdot)$	Fonction de Bessel de 2 ^{ème} espèce modifiée d'ordre ν
${}_bF_a(\cdot)$	Fonction hypergéométrique F à a et b paramètres en arguments
$P^{[L/M]}(\cdot)$	Approximation de Padé d'ordre $[L/M]$
$\mathcal{B}(a, b)$	Ddp d'une v.a. Beta de paramètres a et b
$\mathcal{N}(m, \sigma^2)$	Loi Normale réelle univariée (moyenne m , variance σ^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 μ , covariance \mathbf{M})
$\mathcal{E}(b)$	Loi Exponentielle univariée de paramètre b
$\mathcal{G}(a, b)$	Loi Gamma univariée de paramètres a et b
$\mathcal{IG}(a, b)$	Loi Inverse Gamma univariée de paramètres a et b
χ_n^2	Loi du χ^2 à n degrés de liberté

Table des figures

3.1	Coefficients d'auto-corrélation des matrices de covariance des SIRV simulés, représentés pour différentes valeurs de σ_0 (l'étalement spectral est $\sigma_f = \sigma_0/(2mT_r)$). $T_r = 10^{-3}$; $P_c = 1$; $f_0 = 0$	48
3.2	Indépendance des estimées NSCM par rapport à la loi de la texture du SIRV, montrée pour un SIRV K-distribué ($\nu = 0.1$) et un SIRV Student-t ($\nu = 2.24$), $m = 10$ et $N = 1000$	49
3.3	Erreur relative entre la loi théorique et son approximation par Padé pour différents paramètres.	55
3.4	Histogrammes et moments des variables de <i>texture</i> : <i>texture</i> théorique (en bleu), <i>texture a posteriori</i> (en rouge); moments théoriques ("Th", ronds rouges), moments empiriques sur <i>texture</i> théorique ("MC-Th", en bleu) et moments empiriques sur <i>texture a posteriori</i> ("MC-AP", en noir) pour un SIRV K-distribué : $\nu = 0.1, 2, 20$; $N_{ref} = 10000$	56
3.5	Histogrammes et moments des variables de <i>texture</i> : <i>texture</i> théorique (en bleu), <i>texture a posteriori</i> (en rouge); moments théoriques ("Th", ronds rouges), moments empiriques sur <i>texture</i> théorique ("MC-Th", en bleu) et moments empiriques sur <i>texture a posteriori</i> ("MC-AP", en noir) pour un SIRV K-distribué : $\nu = 0.1, 2, 20$; $N_{ref} = 1000$	57
3.6	Histogrammes et moments des variables de <i>texture</i> : <i>texture</i> théorique (en bleu), <i>texture a posteriori</i> (en rouge); moments théoriques ("Th", ronds rouges), moments empiriques sur <i>texture</i> théorique ("MC-Th", en bleu) et moments empiriques sur <i>texture a posteriori</i> ("MC-AP", en noir) pour un SIRV K-distribué : $\nu = 0.1, 2, 20$; $N_{ref} = 100$	58
3.7	Performance de détection du PEOD "idéal" pour une cible d'amplitude inconnue dans un environnement K-distribué ($\nu = 0.1; 0.5; 2; 20$). Comparaison avec le détecteur optimal OKD et l'OGD. $P_{fa} = 10^{-2}, 10^{-3}$, $m = 10$, $M_{\nu=0.1} = 9$, $M_{\nu=0.5} = 7$, $M_{\nu=2} = 5$, $M_{\nu=20} = 2$	59
3.8	Performance de détection du PEOD "non-idéal" pour une cible d'amplitude inconnue dans un environnement K-distribué ($\nu = 0.1; 0.5; 2; 20$). Comparaison avec le détecteur optimal OKD et le PEOD "idéal". $P_{fa} = 10^{-3}$, $m = 10$, $M_{\nu=0.1} = 9$, $M_{\nu=2.1} = 6$, $M_{\nu=20} = 3$	60
3.9	Performance de détection du PEOD "idéal" pour une cible d'amplitude inconnue dans un environnement inconnu : la <i>texture</i> suit une loi de Weibull de paramètre $b = 0.658$ et de puissance unité. Comparaison avec l'OKD pour $\nu = 0.1, 0.5, 2$ et l'OGD. $P_{fa} = 10^{-3}$, $m = 10$, $M_{id} = 9$	60

3.10	Propriété TFAC par rapport à la loi de la <i>texture</i> du BORD. Environnements K-distribués : $\nu = 0.1, 0.5, 2, 20$ pour les courbes de gauche. Environnements K-distribué ($\nu = 0.1$), Student-t, Gaussien et inconnu (la loi de la <i>texture</i> suit une loi de Weibull), pour les courbes de droite.	65
3.11	Caractéristiques de la matrice de covariance $\mathbf{M}(m, m)$ utilisée dans les simulations	66
3.12	Performances de détection du BORD comparées aux performances de l'OGD et de l'OKD dans des environnements K-distribués ($\nu = 0.1, 0.5, 2, 20$), avec $m = 10$ et $N_{obs} = 5000$ et $P_{fa} = 10^{-3}$. La matrice de covariance est supposée connue. . .	68
3.13	Performances de détection du BORD comparées aux performances de l'OSTD dans des environnements de Student-t ($\nu = 1.03, 2.3$), avec $m = 10$, $N_{obs} = 1000$, $N_{ref} = 500$ et $P_{fa} = 10^{-3}$. La matrice de covariance est supposée connue.	69
3.14	Performances de détection du BORD dans un environnement SIRV inconnu : la <i>texture</i> suit une loi de Weibull ($b = 0.3, 0.658, 1.763, 2.6$), avec $m = 10$, $N_{obs} = 500$, $N_{ref} = 500$, $P_{fa} = 10^{-3}$. Comparaison avec le PEOD. La matrice est supposée connue.	69
3.15	Robustesse du BORD (en rouge) au nombre de données de référence, évaluée en environnement K-distribué pour $\nu = 0.5$, $m = 10$. Comparaison avec l'OKD (en noir). La $P_{fa} = 10^{-3}$ est respectée. La matrice est supposée connue.	70
3.16	Propriété TFAC par rapport à la loi de la <i>texture</i> du BORD Asymptotique. Environnements K-distribués : $\nu = 0.1, 0.5, 2, 20$ (courbes de gauche) ; environnements K-distribué ($\nu = 0.1$), Gaussien, Student-t($\nu = 0.3$) et dont la <i>texture</i> suit une loi de Weibull (courbes de droite). Comparaison avec les valeurs théoriques (à droite). 73	
3.17	Courbes théoriques déterminant la valeur du seuil de détection en fonction de la P_{fa} (gauche) et en fonction de m pour une P_{fa} fixée.	80
3.18	Adéquation entre les valeurs théoriques du seuil du BORD Asymptotique et les valeurs MC du BORD.	81
3.19	Performances de détection (4 premières courbes) du BORD et du BORD Asymptotique en environnement K-distribué ($\nu = 0.1$). Le seuil théorique est utilisé pour une $P_{fa} = 10^{-3}$. Le nombre de référence utile à l'estimation de la matrice varie : $N_{ref} = 10, 50, 200, 1000$. Erreur relative (courbe gauche en bas) de P_d en fonction de N_{ref} . Evolution associée du taux de fausse alarme constaté (courbe droite en bas) en fonction de N_{ref}	83
3.20	Performances de détection du BORD et du BORD Asymptotique en environnement K-distribué ($\nu = 0.1, 0.5, 2, 20$). Le seuil théorique est utilisé pour une $P_{fa} = 10^{-6}$. Le nombre de référence utile à l'estimation de la matrice est fixé à $N_{ref} = 1000$. Le nombre d'observations est égal à $N_{obs} = 500$	84
3.21	Rang et conditionnement de la matrice de corrélation en fonction de son étalement spectral. Les différentes matrices testées ont un étalement spectral allant de 0.2 (forte corrélation) à 2 (corrélation moyenne).	85
4.1	Récepteur optimal (OGD) en bruit blanc pour une émission d'un train de m impulsions	93
4.2	Mise en place du BORD en mode opérationnel, à partir du filtre adapté classique	95
4.3	Mise en place du BORD Asymptotique en mode opérationnel, à partir du filtre adapté classique	96
4.4	Cartes de fouillis de sol issu d'enregistrement réel par un radar opérationnel. Ces cartes montrent le niveau de fouillis (en dB) pour 70 angles d'azimuts, 868 cases distance et pour 8 PRF.	98

4.5	Evolution du CNR - Clutter to Noise Ratio - par azimuth, comparé à la puissance totale moyenne du fouillis, du bruit thermique et de l'ensemble des données. . . .	99
4.6	Case distance où la rupture définitive fouillis non-Gaussien / bruit thermique a lieu, pour chaque azimuth. Moyenne sur les 8 PRF.	101
4.7	Zone de fouillis : Estimations NSCM des matrices de covariance des références et des observations de la "zone 2" et coefficients d'auto-corrélation correspondants. .	102
4.8	Zone transitoire bruit thermique / fouillis : Estimations NSCM des matrices de covariance des références et des observations de la "zone 8" et coefficients d'auto-corrélation correspondants.	102
4.9	Zone de bruit thermique ("Zone 1") : Amplitude des données de référence (en haut) et d'observation (en bas).	103
4.10	Zone de bruit thermique ("Zone 1") : Histogrammes des 8×2145 phases des données de référence (à gauche) et des 8×1617 données d'observation (à droite).	103
4.11	Zone de bruit thermique ("Zone 4") : Histogrammes empiriques (en bleu), lois $K(\nu, b)$ estimées (en rouge) et approximations de Padé (en vert) de l'enveloppe des observations de bruit thermique pour chaque PRF.	104
4.12	Zone de fouillis ("Zone 2") : Amplitude des données de référence (en haut) et d'observation (en bas).	105
4.13	Zone de fouillis ("Zone 2") : Histogrammes des 8×4867 phases des données de référence (à gauche) et des 8×1441 données d'observation (à droite).	105
4.14	Zone de fouillis ("Zone 5") : Histogrammes empiriques (en bleu), lois $K(\nu, b)$ estimées (en rouge) et approximations de Padé (en vert) de l'enveloppe des observations de fouillis pour chaque PRF.	106
4.15	Zone transitoire bruit thermique / fouillis ("Zone 3") : Amplitude des données de référence (en haut) et d'observation (en bas).	107
4.16	Zone transitoire bruit thermique / fouillis ("Zone 3") : Histogrammes des 8×4720 phases des données de référence (à gauche) et des 8×1703 données d'observation (à droite).	107
4.17	Zone transitoire bruit thermique / fouillis ("Zone 8") : Histogrammes empiriques (en bleu), lois $K(\nu, b)$ estimées (en rouge) et approximations de Padé (en vert) de l'enveloppe des observations pour une zone transitoire bruit thermique / fouillis et pour chaque PRF.	108
4.18	Zone 1 : Bruit thermique Gaussien pour les références (gauche) et les observations (droite). Au total, $N_{ref} = 2145$ et $N_{obs} = 1617$	110
4.19	Zone 1 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).	110
4.20	Zone 1 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du bruit thermique Gaussien	111
4.21	Zone 4 : Bruit thermique Gaussien des observations : $N_{obs} = 3296$	113
4.22	Zone 4 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD), ou au seuil théorique du BORD Asymptotique.	114
4.23	Zone 4 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du bruit thermique Gaussien	115
4.24	Zone 6 : Bruit thermique Gaussien des références (gauche) et des observations (droite) : $N_{obs} = 2408$, $N_{ref} = 2314$	117

4.25	Zone 6 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien ou BORD Asymptotique.	118
4.26	Zone 6 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du bruit thermique Gaussien	119
4.27	Zone 2 : Fouillis pour les références (gauche) et les observations (droite). Au total, $N_{ref} = 4836$ et $N_{obs} = 1300$	121
4.28	Zone 2 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).	122
4.29	Zone 2 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis	123
4.30	Zone 5 : Fouillis sélectionné pour les observations. Au total $N_{obs} = 2660$	125
4.31	Zone 5 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).	126
4.32	Zone 5 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis	127
4.33	Zone 7 : Fouillis pour les références (gauche) et les observations (droite). Au total, $N_{ref} = 2054$ et $N_{obs} = 1550$	129
4.34	Zone 7 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).	130
4.35	Zone 7 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis	131
4.36	Zone 3 : Références (à gauche) et observations (à droite) pour un zone transitoire bruit thermique / fouillis. Au total, $N_{ref} = 4410$ et $N_{obs} = 1560$	133
4.37	Zone 3 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).	134
4.38	Zone 3 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis	135
4.39	Zone 8 : Zone transitoire bruit thermique / fouillis pour les références (gauche) et les observations (droite). Au total, $N_{ref} = 2340$ et $N_{obs} = 2080$	137
4.40	Zone 8 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).	138
4.41	Zone 8 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis	139
B.1	Noyau Triangulaire	167
B.2	Noyau Bipoids	167
B.3	Noyau Gaussien	168
B.4	Noyau d'Epanechnikov	168
B.5	Approximation de Padé d'ordre $M = 10$ pour la K-distribution de paramètre $\nu = 0.1$. Erreur relative correspondante sur la courbe inférieure.	177
B.6	Approximation de Padé d'ordre $M = 7$ pour la K-distribution de paramètre $\nu = 1.67$. Erreur relative correspondante sur la courbe inférieure.	178

B.7	Approximation de Padé d'ordre $M = 10$ pour la loi Gamma de paramètre $a = 0.15$. Erreur relative correspondante sur la courbe inférieure.	179
B.8	Approximation de Padé d'ordre $M = 7$ pour la loi Gamma de paramètre $a = 2.24$. Erreur relative correspondante sur la courbe inférieure.	180
B.9	Approximation de Padé d'ordre $M = 8$ pour la loi de Rayleigh de paramètre $b = 0.76$ et erreur relative correspondante sur la courbe inférieure.	181
B.10	Approximations de Padé d'ordres $M = 2$ à $M = 10$ pour la K-distribution de paramètre $\nu = 0.1$ (en haut) et erreurs relatives associées (en bas).	182
B.11	Approximations de Padé d'ordres $M = 2$ à $M = 7$ pour la loi Gamma de paramètre $a = 2.68$ (en haut) et erreurs relatives associées (en bas).	183
B.12	Erreur quadratique moyenne intégrée calculée entre la fonction de répartition théorique et son approximation de Padé en fonction du nombre $N = 10$ à 5000 d'échantillons K-distribués de paramètre $\nu = 0.1$	184

Introduction

Les radars sont conçus pour détecter la présence d'objets d'intérêt noyés dans un bruit aléatoire ambiant. 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 constituant le sol, le sursol ou encore 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 à du bruit Gaussien, les détecteurs conventionnels, construits sur cette hypothèse Gaussienne, permettent alors une détection optimale.

Dans de nombreuses situations, la statistique du fouillis ne s'apparente plus à une Gaussienne. Ce phénomène se produit, par exemple, lorsqu'un radar à fort pouvoir de résolution est utilisé, ou bien que l'angle de visée du radar est faible. Dans ces cas-là, le signal reçu provenant de chaque cellule d'analyse se compose d'un nombre trop faible de composantes pour les assimiler à une statistique Gaussienne. Ce phénomène est également constaté lorsque la puissance du fouillis est aléatoire, ce qui nécessite de considérer le processus de fouillis comme Gaussien composé. Une bonne description de ce phénomène est donnée dans [Jake76] pour des échos de fouillis de mer. Dans chacune de ces situations l'hypothèse Gaussienne n'est pas 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 ce niveau de fausses alarmes, mais les performances de 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 du fouillis. Ces diverses techniques TFAC, issues d'un traitement non-cohérent sur les données radar, sont pour la plupart adaptatives. Cependant, la stratégie de détection initiale ayant été obtenue sous une hypothèse Gaussienne, l'optimalité des traitements n'est pas atteinte.

La recherche de méthodes visant à déterminer une stratégie optimale de détection a été initiée par les travaux de A. Farina et al. [Fari85], portant, pour la première fois, sur le problème de l'élimination du fouillis non-Gaussien par des traitements cohérents. Les échos du radar sont considérés comme des processus complexes 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 SIRP (Spherically

Invariant Random Process), 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, χ , Rayleigh généralisé, et même de prendre en compte le bruit thermique Gaussien qui s'ajoute à ces perturbations.

L'utilisation des modèles SIRP 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 [Trun70, Gold76, Ward81, Watt85, Cont87, Cont91, Pent92, Gini00c]. Ces modèles se sont révélés être en adéquation avec les mesures de fouillis, ce qui permet de construire des détecteurs adaptés à leur statistique. Cependant, les stratégies optimales de détection, obtenues après modélisation de l'environnement par les SIRP, 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 choisi.

Ce travail a pour objectif de mieux caractériser l'environnement du radar, afin de trouver 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 SIRP. Les SIRP 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 (Synthetic Aperture Radar), est représenté par un processus Gaussien complexe circulaire multidimensionnel, caractéristique des propriétés de "cohérence" du radar. La *texture*, interprétée comme la variance du processus Gaussien, est une variable aléatoire positive dont la densité de probabilité 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.

Un premier travail consiste à estimer, à partir de données de référence, la loi de la *texture* du SIRP. Pour ce faire, j'ai utilisé deux méthodes différentes : une estimation "paramétrique", découlant d'une approximation de Padé de la fonction génératrice de moments de la loi de la *texture*, et un estimateur Bayésien de la loi de la *texture*. Pour chacune de ces méthodes, une expression générale de la loi du processus SIRP peut être déterminée, et deux stratégies de détection sont alors mises en place : le PEOD (Padé Estimated Optimum Detector) et le BORD (pour Bayesian Optimum Radar Detector). Ces deux stratégies optimales de détection s'adaptent automatiquement à l'environnement et sont valables quelle que soit la statistique des données.

Le travail s'articule autour de quatre chapitres, au fil desquels chaque point de la démarche employée est détaillée. 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 le *filtre adapté*, est rappelée. Les diverses méthodes d'optimisation 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, les stratégies optimales de détection, qui ont été obtenues pour un environnement donné et modélisé par un SIRP, sont décrites. 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.

Dans le chapitre 2, une synthèse des résultats théoriques relatifs aux processus SIRP est réalisée. Parmi ces résultats figurent le théorème de représentation de K. Yao, le lemme d'invariance par transformation linéaire et des théorèmes permettant la caractérisation de tels processus. Les expressions de plusieurs lois jointes de vecteurs non-Gaussiens, issus de la famille des SIRP, sont ensuite données, puis utilisées dans le cadre de la détection. L'expression du rapport de vraisemblance généralisé est explicité dans ce cas là, et les stratégies de détection optimales, dont certaines sont décrites au chapitre 1, sont obtenues pour chaque loi de vecteur issu d'un SIRP.

Le chapitre 3 s'inscrit dans la continuité des deux premiers chapitres et constitue l'apport original de ma thèse en détection radar. Deux nouvelles stratégies de détection optimales sont obtenues, valables quelle que soit la statistique de l'environnement : après une analyse bayésienne du modèle des SIRP, la densité de probabilité de la *texture* du fouillis est estimée à partir de données de référence. Les densités estimées admettent des expressions analytiques pour chaque méthode utilisée et permettent d'obtenir ces deux nouvelles structures de détecteurs. La première, nommée PEOD (Padé Estimated Optimum Detector, [Jay00a, Jay01]), découle d'une approximation de Padé de la fonction génératrice de moments de la *texture*. Sa densité de probabilité est ensuite obtenue par inversion de la transformée de Laplace usuelle, et s'exprime sous la forme d'une somme d'exponentielles complexes décroissantes ([Jay98, Jay99a, Jay99b, Ovar99, Jay00b]). Les performances du PEOD, évaluées dans différents types d'environnements, sont comparées aux performances des détecteurs optimaux.

La deuxième stratégie de détection, nommée BORD (Bayesian Optimum Radar Detector, [Jay02a, Jay02b]), découle d'une estimation bayésienne de la loi de la *texture*. Un prior non-informatif est utilisé sur les données de référence et la densité estimée de la *texture* s'interprète comme une mixture de lois inverse Gamma, dont les paramètres dépendent des références et de leurs caractéristiques. L'expression du BORD dépend des mêmes paramètres et s'exprime comme le rapport de deux sommes, fonctions des références et des observations.

Il conserve structurellement la propriété TFAC par rapport à la loi de la *texture*, ce qui signifie, que quelle que soit la statistique du SIRP, la loi du BORD reste inchangée. D'après son expression, une analyse asymptotique est réalisée. Lorsque le nombre de données de référence devient grand, le BORD converge en loi vers un autre détecteur, appelé le BORD Asymptotique. Ce dernier coïncide avec l'expression de détecteurs déjà connus, obtenus soit sous l'hypothèse Gaussienne, soit en estimant la variable de *texture* au sens du maximum de vraisemblance. Ce détecteur vérifie également la propriété TFAC par rapport à la loi de la *texture*, et peut donc s'exprimer en fonction du *speckle* Gaussien qui compose le SIRP. La loi du BORD Asymptotique est alors obtenue en utilisant un théorème sur la loi et l'indépendance des formes quadratiques de vecteurs Gaussiens, le théorème de Cochran. Des analyses de performances de détection du BORD et de sa version Asymptotique sont réalisées pour tester leur optimalité ainsi que leur propriété TFAC.

Le chapitre 4 est consacré à la validation des résultats de simulation de plusieurs détecteurs sur des données expérimentales. Elles représentent des données de fouillis de sol, enregistrées par un radar opérationnel de THALÉS TAD (Thalès Air Defence) à faible incidence. L'étude se concentre principalement sur l'analyse des performances du BORD et du BORD Asymptotique, comparées à celles du détecteur classique (l'OGD - Optimum Gaussian Detector) et du détecteur

optimal pour une K-distribution (l'OKD - Optimum K Detector).

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

Chapitre 1

État de l'art

Ce chapitre fait état des résultats principaux concernant la détection radar en environnement non-Gaussien. Les éléments essentiels de la théorie de la détection radar sont d'abord rappelés, puis une synthèse des résultats est effectuée. Dans un premier temps, les différentes techniques d'optimisation du détecteur classique sont présentées, suivies d'une synthèse des résultats déjà obtenus en détection cohérente après modélisation du non-Gaussien par les SIRP.

1.1 Théorie de la détection radar

Le schéma classiquement utilisé en détection radar est le suivant :

$$\mathbf{y} = \zeta \mathbf{s} + \mathbf{b} \quad (1.1)$$

$$H_0 : \zeta = 0 \quad (1.2)$$

$$H_1 : \zeta \neq 0 \quad (1.3)$$

Sous l'hypothèse H_0 , le signal complexe \mathbf{y} reçu par le radar est supposé ne contenir que les échos indésirables et nuisibles à une bonne détection, provenant des réflexions sur les différents éléments de l'environnement. Ces parasites sont supposés admettre une certaine densité de probabilité, notée $p_{\mathbf{y}}(\mathbf{y}/H_0) = p_{\mathbf{b}}(\mathbf{y}/H_0)$.

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

Dans la suite du document, seules les hypothèses H_0 et H_1 seront utilisées pour représenter la présence ou l'absence du signal cible. Le paramètre ζ nous permet juste de décrire le signal reçu par le radar dans l'espace des paramètres d'intérêt.

L'objectif de la détection est de déterminer laquelle des deux hypothèses est la plus vraisemblable, tout en minimisant les erreurs possibles. Ces erreurs sont de la forme :

- Décider H_0 alors que H_1 est vraie. Dans ce cas, on parle de *non-détection*, avec la probabilité $P_{nd} = p(H_0/H_1) = 1 - p(H_1/H_1) = 1 - P_d$,
- Décider H_1 alors que H_0 est vraie. Dans ce cas, on parle de *fausse alarme*, avec la probabilité $P_{fa} = p(H_1/H_0)$.

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 (décider H_1 lorsque H_1 est vraie) pour un taux de fausses alarmes $P_{fa} = \alpha$ fixé (dans la pratique, le plus faible possible). Il est montré que le test du rapport de vraisemblance (RV), donné par :

$$\Lambda(\mathbf{y}) = \frac{p(\mathbf{y}/H_1)}{p(\mathbf{y}/H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (1.4)$$

est alors le test optimal pour le critère de Neymann-Pearson. Cela s'obtient facilement ([Kay98]) par la méthode des multiplicateurs de Lagrange en maximisant la fonction $L = P_d + \lambda (P_{fa} - \alpha)$.

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

$$P_{fa} = \mathbb{P}(\Lambda(\mathbf{y}; H_0) > \eta) = \alpha \quad (1.5)$$

$$P_{fa} = \int_{\mathcal{D}_0} p(\mathbf{y}/H_0) d\mathbf{y}, \quad (1.6)$$

où \mathcal{D}_0 est l'ensemble des \mathbf{y} contenus dans le domaine de décision de l'hypothèse H_0 . Il est souvent difficile d'obtenir une expression analytique de ces équations, et le calcul de η s'effectue soit en résolvant numériquement (1.5) ou (1.6), soit par Monte-Carlo.

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

$$P_d = \mathbb{P}(\Lambda(\mathbf{y}; H_1) > \eta) \quad (1.7)$$

$$P_d = \int_{\mathcal{D}_1} p(\mathbf{y}/H_1) d\mathbf{y}, \quad (1.8)$$

où \mathcal{D}_1 est l'ensemble des \mathbf{y} contenus dans le domaine de décision de l'hypothèse H_1 . Comme pour le calcul du seuil de détection à P_{fa} fixée, il est très rare d'obtenir une expression analytique de P_d .

L'expression du RV donné par (1.4) 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", toujours supposé déterministe, est inconnu, le rapport de vraisemblance généralisé (RVG) est alors mis en place. Sa qualification *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 estimées ont remplacé les vrais paramètres dans l'expression (1.4) du RV.

L'ensemble θ 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 τ du signal (indicateur de la distance de la cible) et de la direction de visée du radar. Le RV devient alors une fonction de θ 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 θ , au sens du MV s'écrit :

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \Lambda(\mathbf{y}, \theta) = \underset{\theta}{\operatorname{argmax}} p(\mathbf{y} - f(\theta)/H_0), \quad (1.9)$$

qui se résout difficilement dans son intégralité [Alte79].

Dans la suite, le signal "cible" \mathbf{s} est noté $\mathbf{s} = A \mathbf{p}$, où \mathbf{p} est le "steering vector" désigné par les tests, comprenant les valeurs estimées des paramètres autres que l'amplitude A . Le RVG s'écrit donc :

$$\Lambda(\mathbf{y}, \hat{A}_{mv}) = \frac{p(\mathbf{y}, \hat{A}_{mv}/H_1)}{p(\mathbf{y}/H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (1.10)$$

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 .

Par exemple, si le vecteur \mathbf{b} est un vecteur Gaussien complexe, centré, de matrice de covariance $2\mathbf{M}$, le logarithme du rapport de vraisemblance (RV), donné en (1.4), s'écrit en fonction de A comme suit :

$$\log(\Lambda(\mathbf{y}, A)) = \frac{1}{2} \left(A \mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{p} + A^* \mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{y} - |A|^2 \mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p} \right) \underset{H_0}{\overset{H_1}{\gtrless}} \log(\eta), \quad (1.11)$$

où \dagger désigne le transposé conjugué, et $*$ le conjugué.

La résolution de (1.9) en A s'effectue alors en maximisant (1.11) sur A , ce qui revient à trouver les zéros de sa dérivée. On rappelle, que pour un nombre complexe $z = x + jy$, les dérivées de toute fonction $f(x, y)$ par rapport à z et à son conjugué \bar{z} s'écrivent (voir par exemple [Dolb90]) :

$$\frac{\partial}{\partial z} f(x, y) = \frac{1}{2} \left(\frac{\partial}{\partial x} f(x, y) - j \frac{\partial}{\partial y} f(x, y) \right) \quad (1.12)$$

$$\frac{\partial}{\partial \bar{z}} f(x, y) = \frac{1}{2} \left(\frac{\partial}{\partial x} f(x, y) + j \frac{\partial}{\partial y} f(x, y) \right) \quad (1.13)$$

On obtient alors :

$$\frac{d}{dA} [\log(\Lambda(\mathbf{y}, A))] = \frac{1}{2} \left(\mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{p} - A^* \mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p} \right), \quad (1.14)$$

et l'estimation de A au sens du MV admet pour expression :

$$\hat{A}_{mv} = \left(\frac{\mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{p}}{\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}} \right)^* = \frac{\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{y}}{\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}}. \quad (1.15)$$

Il suffit de remplacer cette expression dans (1.10) pour obtenir l'expression du RVG dans le cas d'un bruit complexe Gaussien.

1.2 Le détecteur classique et l'Optimum Gaussian Detector (OGD)

Les premiers résultats obtenus étaient basés sur une hypothèse de bruit Gaussien. Cette hypothèse était vérifiée de par la conception des systèmes radar et simplifiait considérablement l'ensemble des calculs. En effet, le bruit est alors complètement déterminé par ses propriétés du second ordre, sa covariance, et dans le cas stationnaire, par sa fonction de corrélation. Le test

optimal découlant du RV coïncide dans ce cas-là avec le filtre adapté classique. Pour une observation y et un signal cible connu, le détecteur classique s'écrit :

$$\Re(y s^*) \underset{H_0}{\overset{H_1}{\gtrless}} \sigma^2 \ln(\eta) + \frac{|s|^2}{2} \quad (1.16)$$

où σ^2 est la puissance du bruit Gaussien et \Re désigne la partie réelle. Lorsque l'amplitude a du signal $s = a q$ est estimée au sens du MV, le filtre adapté est suivi d'une étape quadratique, ce qui donne :

$$|y q^*|^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln(\eta) |q|^2 \quad (1.17)$$

Le calcul du seuil de détection est alors donné par :

$$\eta = \sqrt{-2\sigma^2 \log(P_{fa})}, \quad (1.18)$$

Dans le cas où un traitement cohérent est effectué, le signal observé est un vecteur complexe. Le bruit Gaussien est alors déterminé par sa matrice de covariance \mathbf{M} normalisée par sa puissance $2\sigma^2$. Le détecteur, appelé OGD, se comporte toujours comme un filtre adapté, suivi d'une étape quadratique si l'amplitude du signal est estimée au sens du MV. Les deux expressions deviennent alors :

$$\Re(\mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{s}) \underset{H_0}{\overset{H_1}{\gtrless}} \sigma^2 \ln(\eta) + \frac{\mathbf{s}^\dagger \mathbf{M}^{-1} \mathbf{s}}{2} \quad (1.19)$$

et :

$$|\mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{p}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln(\eta) \mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}. \quad (1.20)$$

Une forme bien connue et bien particulière du RVG est donnée par E.J. Kelly dans [Kell86]. Ce dernier suppose observer $K+1$ vecteurs de taille m sous chacune des hypothèses, composés de K vecteurs de bruit adjacents à la cellule d'observation. Sous l'hypothèse H_0 , les $K+1$ vecteurs ne comportent donc que du bruit Gaussien, alors que sous H_1 , seuls K vecteurs ne représentent que du bruit. Les lois jointes des $K+1$ vecteurs sous chaque hypothèse s'expriment alors comme le produit des $K+1$ densités Gaussiennes multivariées de chacun des vecteurs, en distinguant l'observation des K références. Dans un premier temps, la matrice de covariance \mathbf{M} des données est remplacée, sous chacune des hypothèses, par son estimée au sens du MV $\widehat{\mathbf{M}}$, puis l'amplitude inconnue de la cible est remplacée à son tour par son estimée au sens du MV, qui coïncide avec (1.15).

Le RVG de Kelly s'écrit alors :

$$|\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{p} \left[1 + \frac{1}{K} \mathbf{y}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y} \right], \quad (1.21)$$

où $\lambda = K(1 - \eta^{-1/(K+1)})$, et où $\widehat{\mathbf{M}}$ est donnée par $\widehat{\mathbf{M}} = K^{-1} \sum_{k=1}^K \mathbf{y}_k \mathbf{y}_k^\dagger$.

Si l'on compare (1.21) à (1.20), la différence majeure résulte en la présence du terme additionnel $K^{-1} \mathbf{y}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}$. Cette différence peut s'interpréter comme une mesure de l'erreur commise lors de l'estimation de la matrice de covariance à partir des K vecteurs de référence. En particulier, lorsque K devient grand ($K \rightarrow \infty$), ce terme tend vers 0, $\widehat{\mathbf{M}}$ tend vers sa vraie valeur \mathbf{M} et le RVG de Kelly tend alors vers le filtre adapté (1.20).

Les performances du détecteur classique sont évidemment optimales en environnement Gaussien. Cependant, de nombreux résultats expérimentaux [Trun70, Gold76, Schl76, Jake80, Ward81, Jao84, Watt85, Cont87, Cont91, Bill93, Ovar98, Shni99, Gini00c, Bill02] ont montré que les bruits parasites réels ne sont ni Gaussiens ni stationnaires, notamment lorsqu'il s'agit de fouillis de mer ou de terre, éclairé à faible incidence par les systèmes radar actuels. Dans ces environnements impulsionnels, les performances du détecteur classique sont dégradées et le taux de fausses alarmes augmente considérablement.

A la suite de ces constatations, deux axes de recherche ont été explorés : l'objectif du premier est d'optimiser le détecteur classique par des techniques de mise en œuvre adaptatives. Cela a pour effet d'adapter le seuil de détection à la puissance réelle du fouillis et de conserver ainsi un taux de fausses alarmes constant tout au long du traitement. Le deuxième axe de recherche porte plutôt sur la modélisation de l'environnement non-Gaussien et la mise en place de structures de détecteurs adaptés à la statistique de celui-ci.

1.3 Techniques d'optimisation du détecteur classique en environnement non-Gaussien

Pour mettre en place le détecteur classique, il est nécessaire de calculer la puissance du bruit et éventuellement sa matrice de covariance dans le cas de traitements cohérents, pour déterminer la valeur du seuil de détection. Lorsque la statistique du bruit est connue, ces paramètres peuvent être calculés de manière optimale, ou tout du moins s'approcher de leurs valeurs théoriques, en utilisant par exemple la méthode des moments pour la puissance [Schl76, Watt87, Cont91, Anas95, Anas99, Shni99], ou les statistiques d'ordre supérieur pour la corrélation [Sadl94, Gini97b].

Dans le but majeur de conserver un taux de fausses alarmes constant tout au long du traitement non-cohérent, des techniques appelées TFAC (Taux de Fausse Alarme Constant ou CFAR en anglais) ont été conçues.

La technique du TFAC à moyennage de cellules (ou Cell-Averaging-CFAR en anglais), est la plus répandue [Watt85, Watt87, Ragh95, Shni95, Watt96, Bark99]. Le seuil de détection est calculé "par blocs", c'est-à-dire d'après la puissance moyenne d'une "fenêtre" constituée de quelques cellules de fouillis, adjacentes à la cellule testée. La fenêtre est ainsi déplacée sur toutes les cellules, ce qui a pour but d'adapter le seuil de détection à la puissance locale du fouillis. Cette procédure est optimale dans des environnements de fouillis Gaussien homogène, et lorsque la taille de la fenêtre augmente, les performances du CA-CFAR tendent vers les performances du détecteur classique à seuil fixe, calculé sur l'ensemble des cellules de référence bruit. Une comparaison des deux approches (seuil fixe, seuil adaptatif) est par exemple effectuée dans [Watt85] pour un environnement K-distribué. Dans le cas où cet environnement est corrélé, l'auteur montre que les fluctuations locales sont mieux appréhendées par un calcul adaptatif du seuil. Cependant, cette technique souffre d'une augmentation de la probabilité de fausse alarme lors des transitions du fouillis dans des environnements non-homogènes.

De nombreuses variantes à ce détecteur ont été proposées. Les deux plus proches sont le Greatest-Of-CFAR (ou GO-CFAR) et le Smallest-Of-CFAR (ou SO-CFAR), qui utilisent plusieurs fenêtres glissantes, pour déterminer non pas la moyenne des puissances, mais respective-

ment la plus grande et la plus petite des valeurs. Dans le cas où deux fenêtres sont utilisées, elles constituent par exemple les cellules précédant et succédant la cellule testée. Le GO-CFAR offre de meilleures performances de détection que le CA-CFAR en présence de fouillis présentant des ruptures alors que ses performances se dégradent en environnement homogène. Par opposition au GO-CFAR, le SO-CFAR permet une meilleure discrimination en présence de cibles multiples, alors que ses performances se dégradent en environnement présentant des ruptures.

Les techniques, dites de "censoring", procèdent à une sélection de quelques cellules pour déterminer la valeur du seuil. Par exemple, l'Order-Statistic-CFAR (ou OS-CFAR, [Gand88]) ordonne, dans la fenêtre, les cellules de référence d'après leur puissance, et leur attribue un "rang", dont le plus grand est noté M . Le seuil de détection est ensuite calculé d'après la statistique d'ordre d'un rang préselectionné. Ce détecteur est plus robuste que le CA-CFAR en environnement non-homogène et en situations multi-cibles. Ses pertes TFAC sont minimisées lorsque la plus grande valeur M est utilisée. Il en existe une version adaptative, le AOS-CFAR, [Gand89]), qui fonctionne comme le OS-CFAR à un test supplémentaire d'hypothèses près, utilisé afin de déterminer la présence de rupture dans les cellules de référence. De même, le Selection-Estimation-CFAR (ou SE-CFAR), présenté dans [Visw92], est lui aussi adaptatif et présente de meilleures performances que le OS-CFAR aux transitions du fouillis.

Une autre technique proche du OS-CFAR est appelée Trimmed-Mean-CFAR (TM-CFAR, [Gand88]). Les cellules sont ordonnées selon leurs puissances croissantes et un certain nombre (prédéfini à l'avance) des plus grandes valeurs est éliminé pour ensuite procéder comme un CA-CFAR. Cette procédure est robuste en situation multi-cibles et en environnement de fouillis impulsionnel, très similairement à l'OS-CFAR.

La difficulté à trouver un traitement qui s'accommode de toutes les variétés d'environnements rencontrés dans la pratique, mène au développement de procédures composées. Ainsi, R.Rifkin, [Rifk94], combine le OS-CFAR et le GO-CFAR pour évaluer les performances de la procédure dans du fouillis de Weibull.

Dans [Smit00], les CA, SO et GO-CFAR sont combinées pour donner le VI-CFAR (Variability Index-CFAR). Cette procédure détermine un indice de variabilité (VI) des données et classe les cellules selon l'homogénéité du fouillis. Cela permet de conserver les performances TFAC en environnement homogène (par exemple en environnement Gaussien) et de garantir une bonne robustesse du détecteur en environnement non homogène, incluant éventuellement la détection de cibles multiples.

Une autre procédure, le GOS-CFAR (Generalized OS-CFAR), combine le OS-CFAR avec le TM-CFAR, [Kim95]. En fait, le calcul d'un certain coefficient détermine l'utilisation de l'une ou l'autre des procédures.

Dans [Nagl95], les auteurs étudient les performances du détecteur LCOS-CFAR (Linearly Combined Order Statistic-CFAR) après avoir effectué une estimation robuste et efficace du seuil de détection. L'environnement est supposé de loi exponentielle et deux méthodes d'optimisation sont utilisées. La première utilise les estimées efficaces du paramètre d'échelle que sont les estimées CML (Censored Maximum Likelihood) et BLU (Best Linear Unbiased). Avec ces deux estimées, les performances de détection sont meilleures que les performances des détecteurs OS et TM-CFAR. L'autre procédure d'optimisation détermine des coefficients de pondération par maximisation de la probabilité de détection (MPL-CFAR, Most Powerful Linear-CFAR) pour des cibles fluctuant selon une loi de Swerling 2 [Swer54].

Dans la série des techniques de "censoring", on trouve également une méthode basée sur un test de corrélation. Cette méthode utilise la transformée de Fisher pour situer les ruptures dans le fouillis et déterminer si la cellule testée se trouve dans du fouillis corrélé ou du fouillis non corrélé, [Solt97]. Il est montré que la probabilité de détecter les ruptures augmente avec le nombre de cellules et le coefficient de corrélation du fouillis. Un détecteur "M-out-of-L" est proposé, qui permet de détecter la cible dans M cellules parmi L observées.

Dans la majorité des cas étudiés, les performances du détecteur classique sont évaluées par Monte-Carlo. La non-linéarité du test de détection ne permet effectivement pas d'obtenir la loi des observations sous l'hypothèse H_1 , dont se déduit l'expression de la probabilité de détection. Cependant, certaines méthodes permettent de résoudre ce problème numériquement. Ainsi, dans [Hou87], des expressions théoriques de la probabilité de fausse alarme et de la probabilité de détection sont données après avoir supposé que la fonction caractéristique des données pouvait s'écrire sous la forme d'une fraction rationnelle. Dans [Hels84], les performances de détection sont évaluées par la méthode d'intégration de la plus grande pente (steepest integration descent), avec pour point de départ le point de selle (ou saddlepoint). Enfin, dans [Bird95, Jay00b], l'utilisation de la fonction caractéristique cohérente radiale du fouillis permet d'obtenir les expressions de la probabilité de fausse alarme et de la probabilité de détection à partir de la fonction génératrice de moments. Dans [Amin91, Amin94a, Amin94b, Jay00b], une approximation de Padé est utilisée pour estimer les densités de probabilité du fouillis. Les expressions résultantes permettent d'obtenir des expressions simplifiées de la probabilité de fausse alarme et de la probabilité de détection.

Malgré la diversité de toutes ces techniques, l'optimalité du détecteur n'est jamais atteinte, dans des environnements de fouillis qui ne lui sont pas adaptés. Ces techniques permettent en fait de conserver la structure simple du détecteur classique en réduisant partiellement le taux de fausses alarmes.

Le deuxième axe de recherche, exploré plus récemment après la publication des travaux de A. Farina et al. [Fari85], se base sur des traitements cohérents, pour lesquels l'information de la phase est conservée. En utilisant des modèles statistiques de processus non-Gaussiens, tels les SIRP (Spherically Invariant Random Process), la loi jointe des vecteurs d'observation est déterminée et des structures optimales de détection découlent du RV.

1.4 Modélisation de l'environnement non-Gaussien et détection cohérente

Le regain d'intérêt pour la détection radar cohérente trouve en fait sa source dans l'utilisation des processus composés pour modéliser l'environnement non-Gaussien du radar. 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 peuvent se représenter par les processus aléatoires sphériquement invariants (ou SIRP, [Yao73, Gold76, Cont87, Rang91, Rang93, Rang95, Barn96]) dont la loi de probabilité est connue dans de nombreux cas.

A partir de cette modélisation, le rapport de vraisemblance 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 SIRP, ou plus simplement la *loi de la texture*. A chaque SIRP lui correspond sa loi de texture, connue analytiquement ou non. Par exemple, un SIRP issu d'une K-distribution possède une texture de loi Gamma. En revanche, un SIRP issu d'une loi de Weibull ne peut pas se caractériser par la loi de sa texture puisque son expression n'existe pas sous forme analytique simple.

Les stratégies de détection optimales, obtenues après formation du rapport de vraisemblance, ont été étudiées selon la statistique de l'environnement considéré. De nombreuses études se rapportent à un fouillis strictement K-distribué ou de Weibull [Cont87, Fari87, Cont91, Pent92, Cont94, Cont98, Sang99], en effectuant des comparaisons entre les deux types de fouillis. On retrouve également du fouillis Gaussien, avec un signal cible aléatoire [Ragh95] ou plus réaliste, du fouillis K-distribué plus du bruit Gaussien, représentant le bruit thermique, très souvent négligé dans les analyses de performances [Gini97a, Gini97b, Gini98, Grec01b]. Les auteurs montrent ainsi que la prise en compte, dans la modélisation, du bruit thermique, en plus du fouillis K-distribué, permet encore d'améliorer les performances de détection par rapport à un détecteur basé sur une hypothèse de fouillis K-distribué uniquement. De manière générale, les performances des détecteurs SIRP dépassent très largement les performances du détecteur classique cohérent, l'OGD.

Avec de telles stratégies, l'optimalité est atteinte dans l'environnement désigné. L'inconvénient majeur réside dans la structure même de ces détecteurs qui ne permet pas d'obtenir des expressions simples pour le calcul du seuil ou de la probabilité de détection. Ainsi, des études asymptotiques ont été menées pour tenter de simplifier ces expressions, en estimant, par exemple, la variable de texture au sens du MV ou au sens du Maximum A Posteriori (MAP). Ces approches se retrouvent dans [Kell86, Cont95, Gini99, Sang99, Gini00c] et les détecteurs sous-optimaux ainsi obtenus coïncident avec des détecteurs obtenus sous l'hypothèse d'un bruit Gaussien.

Les préoccupations pratiques restant les mêmes, les propriétés TFAC et le caractère adaptatif sont également recherchés pour ces détecteurs optimaux. La propriété TFAC se caractérise par une loi du test indépendante de la statistique de l'environnement et ne peut être montrée que pour les détecteurs sous-optimaux. En revanche, dès que les paramètres du fouillis sont estimés à partir des données, les détecteurs sont déclarés adaptatifs. Par exemple, le RVG et l'ALQ (Adaptive Linear Quadratic detector) sont adaptatifs et se différencient l'un de l'autre uniquement par la méthode d'estimation de la matrice de covariance. Pour le premier, l'estimation est réalisée au sens du MV, alors que pour le second, une estimation NSCM (Normalized Structured Covariance Matrix, [Gini00b]) est utilisée.

La modélisation de l'environnement par les SIRP a pris une part importante en détection cohérente. Ceci dit, d'autres modèles encore plus généraux existent, comme les processus alpha-stables, ou $S\alpha S$ [Tsih97, Sza01], isotropiques et circulairement symétriques. Ces processus permettent de modéliser des processus non-Gaussiens avec 3 ou 4 degrés de liberté. Pour ces modèles, la variance du Gaussien n'est pas seulement une variable aléatoire, mais un processus stable positif, et de ces processus se déduisent certains SIRP.

Cette étude bibliographique nous a permis d'identifier la nécessité d'adapter les détecteurs radar aux environnements de fouillis non-Gaussiens, modélisés par des processus SIRP. Notre objectif est d'y parvenir sans connaissance *a priori* sur la statistique du fouillis, en estimant la

loi de la *texture* du SIRP uniquement à partir des données reçues par le radar.

Pour bien comprendre le cheminement du travail, le chapitre qui suit est dédié entièrement aux SIRP. Sa première partie rapporte les éléments principaux de la théorie des SIRP et des processus composés, avec des propriétés et des théorèmes très importants sur ces processus. La seconde partie applique la modélisation de l'environnement radar à la détection cohérente et les principales stratégies optimales sont données.

Chapitre 2

Les processus aléatoires sphériquement invariants : les SIRP

Ce chapitre décrit les processus aléatoires sphériquement invariants, ou SIRP, caractérisés par les vecteurs qui les composent, les SIRV (V pour vecteur). Des propriétés et théorèmes sont d'abord énoncés, puis ces processus sont utilisés dans le cadre de la modélisation du fouillis radar non-Gaussien pour obtenir des stratégies de détection optimales dans des environnements de statistique connue a priori.

2.1 Modélisation de processus aléatoires non-Gaussiens

Les processus aléatoires sphériquement invariants (d'acronyme anglo-saxon SIRP) permettent la représentation des processus aléatoires non-Gaussiens. Leurs formes composées donnent naturellement les lois jointes de vecteurs non-Gaussiens, issus de ces processus, nécessaires en théorie de la détection cohérente.

2.1.1 Théorie des processus Gaussiens composés et propriétés

2.1.1.1 Définitions

Un processus Gaussien composé est un processus aléatoire, tel que tout vecteur issu de ce processus est un vecteur aléatoire Gaussien composé. Il en va de même pour les processus aléatoires sphériquement invariants (pour lesquels les Vecteurs de même nature sont des SIRV) et les deux familles de processus se définissent à partir des vecteurs qui les composent.

Les deux premières définitions se rapportent aux processus Gaussiens composés, caractérisés par la forme particulière des fonctions caractéristiques et densités de probabilité des vecteurs

qui les composent.

Définition 1 (Les vecteurs aléatoires réels Gaussiens composés) sont des vecteurs aléatoires réels de taille m dont la densité de probabilité est donnée par

$$p_m(\mathbf{y}) = \int_0^{+\infty} g_m(\mathbf{y}, \tau) p(\tau) d\tau,$$

ou bien, dont la fonction caractéristique est donnée par

$$\Phi_m(\mathbf{u}) = \int_0^{+\infty} \exp\left(-\frac{\tau \mathbf{u}^T \mathbf{M} \mathbf{u}}{2}\right) p(\tau) d\tau.$$

$g_m(\mathbf{y}, \tau)$ est une densité de probabilité Gaussienne multivariée (à m composantes), de matrice de covariance $\tau \mathbf{M}$ avec τ une variable aléatoire positive de loi $p(\tau)$, d'expression :

$$g_m(\mathbf{y}, \tau) = \frac{1}{(2\pi \tau)^{m/2} \sqrt{|\mathbf{M}|}} \exp\left(-\frac{\mathbf{y}^T \mathbf{M}^{-1} \mathbf{y}}{2\tau}\right).$$

Définition 2 (Les vecteurs aléatoires complexes Gaussiens composés) sont des vecteurs aléatoires complexes de taille m dont la densité de probabilité est donnée par

$$p_m(\mathbf{y}) = \int_0^{+\infty} g_{c_m}(\mathbf{y}, \tau) p(\tau) d\tau,$$

ou bien, dont la fonction caractéristique est donnée par

$$\Phi_m(\mathbf{u}) = \int_0^{+\infty} \exp\left(-\frac{\tau \mathbf{u}^\dagger \mathbf{M} \mathbf{u}}{2}\right) p(\tau) d\tau.$$

$g_{c_m}(\mathbf{y}, \tau)$ est une densité de probabilité Gaussienne multivariée (à m composantes), de matrice de covariance $2\tau \mathbf{M}$ avec τ une variable aléatoire positive de loi $p(\tau)$, d'expression :

$$g_{c_m}(\mathbf{y}, \tau) = \frac{1}{(2\pi \tau)^m |\mathbf{M}|} \exp\left(-\frac{\mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y}}{2\tau}\right).$$

De manière identique aux processus Gaussiens composés, les SIRP se définissent à partir des caractéristiques de leurs vecteurs, les SIRV, comme l'indique la définition qui suit :

Définition 3 (Les vecteurs aléatoires sphériquement invariants ou SIRV) sont des vecteurs dont la fonction caractéristique s'écrit sous la forme :

$$\Phi_m(\mathbf{u}) = f\left(\frac{\mathbf{u}^T \mathbf{M} \mathbf{u}}{2}\right),$$

fonction de la forme quadratique définie positive $\mathbf{u}^T \mathbf{M} \mathbf{u}$. La fonction $f(v)$ admet la représentation intégrale

$$f(v) = \int_0^{+\infty} e^{-\tau v} p(\tau) d\tau.$$

La dénomination *sphériquement invariant* est liée au fait que, par une transformation linéaire du vecteur aléatoire \mathbf{y} transformant \mathbf{M} en la matrice identité, la fonction caractéristique $\Phi_m(\mathbf{u})$ devient $f(\mathbf{u}^T \mathbf{u}/2) = g(\|\mathbf{u}\|^2)$, qui possède bien une invariance sphérique.

Des deux définitions des classes de vecteurs, il découle que tout vecteur aléatoire sphériquement invariant réel ou complexe est nécessairement de type Gaussien composé, respectivement réel ou complexe. En effet, la fonction caractéristique du SIRV coïncide parfaitement avec la fonction caractéristique du vecteur Gaussien composé réel ou complexe. Une conséquence de ce résultat est l'interprétation d'un SIRV \mathbf{y} , réel ou complexe, sous la forme :

$$\mathbf{y} = \sqrt{\tau} \mathbf{x}, \quad (2.1)$$

produit d'une variable aléatoire τ positive par un vecteur réel ou complexe \mathbf{x} Gaussien indépendant de τ et de matrice de covariance \mathbf{M} dans le cas réel. La matrice de covariance de \mathbf{y} est donc égale à $\mathbb{E}(\tau) \mathbf{M}$, où $\mathbb{E}(\cdot)$ désigne l'espérance mathématique.

Dans le cas complexe, le vecteur $\mathbf{x} = \mathbf{a} + j\mathbf{b}$ est Gaussien complexe circulaire. Cette condition de circularité d'ordre 2, nécessaire et explicitée dans l'annexe C (C.6), nous donne $\mathbf{M}_x = 2(\mathbf{M}_a + j\mathbf{M}_{ab})$. La matrice de covariance de \mathbf{y} s'écrit donc $\mathbf{M}_y = 2\mathbb{E}(\tau) (\mathbf{M}_a + j\mathbf{M}_{ab})$.

Remarque : Ce cas de SIRV complexe se rencontre dans les problèmes appliqués au radar. Les composantes réelle et complexe de \mathbf{y} sont alors appelées les **composantes en quadrature** (ou speckle) du récepteur radar (respectivement, voie I pour la partie réelle et voie Q pour la partie imaginaire), et sont indépendantes, identiquement distribuées, de moyenne nulle, de même variance et de même matrice de covariance. Dans pareil cas, on a $\mathbf{M}_y = 2\mathbb{E}(\tau) \mathbf{M}$, où \mathbf{M} est la matrice de covariance commune des composantes en quadrature de \mathbf{x} . La densité de probabilité de \mathbf{y} s'écrit alors :

$$p(\mathbf{y}) = \int_0^{+\infty} p(\mathbf{y}/\tau) p(\tau) d\tau = \frac{1}{(2\pi)^m |\mathbf{M}|} h_m(\alpha) \quad (2.2)$$

où

$$h_m(\alpha) = \int_0^{+\infty} \tau^{-m} \exp\left(-\frac{\alpha}{2\tau}\right) p(\tau) d\tau, \quad (2.3)$$

avec $\alpha = \mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y}$ une forme quadratique définie positive, où † désigne le transposé conjugué.

Dans le jargon des radaristes, la variable aléatoire τ est appelée **texture** du processus et sa densité de probabilité caractérise la loi du SIRV. Dans la littérature, la densité de probabilité de τ est aussi appelée **densité de probabilité caractéristique** du SIRV, mais pour éviter toute confusion avec la fonction caractéristique, nous l'appellerons tout simplement la **loi de la texture**.

L'équation (2.2) montre que la densité de probabilité d'un SIRV est complètement déterminée par la spécification du vecteur moyenne (nul dans ce cas), de la matrice de covariance \mathbf{M} et de la loi de la texture $p(\tau)$. De plus, la densité de probabilité d'un SIRV est une fonction de α , une forme quadratique définie positive qui contient toutes les informations nécessaires pour identifier la densité de probabilité du SIRV. Ces particularités sont proches de celles d'un vecteur aléatoire Gaussien et toutes les propriétés relatives aux SIRV peuvent être considérées comme des extensions des propriétés des vecteurs Gaussiens. Il est à noter que le vecteur aléatoire Gaussien est un cas particulier des SIRV et s'obtient quand $p(\tau) = \delta(\tau-1)$, où $\delta(\cdot)$ est l'impulsion de Dirac.

2.1.1.2 Le théorème de représentation d'un SIRV

Le théorème de représentation donné par K. Yao dans [Yao73] précise le lien entre un SIRV et un vecteur aléatoire réel Gaussien :

Théorème 1 (Théorème de représentation - 1) *Une condition nécessaire et suffisante pour chaque classe de densités de probabilité $\{p_m(\mathbf{X}), m \geq 1\}$ d'appartenir à la classe des densités de probabilité des SIRV est que :*

$$p_m(\mathbf{X}) = (2\pi)^{-m/2} |\mathbf{M}|^{-1/2} h_m(\alpha), \quad (2.4)$$

où $\alpha = \mathbf{X}^T \mathbf{M}^{-1} \mathbf{X}$, \mathbf{M} est la matrice de covariance du vecteur \mathbf{X} et $h_m(r)$ est donnée comme une intégrale de Lebesgue-Stieljes pour $0 < r < +\infty$ par :

$$h_m(r) = \int_0^{+\infty} v^{-m} \exp(-r/2v^2) p(v) dv. \quad (2.5)$$

$h_m(r)$ est définie par continuité en $r = 0$ et $p(v)$ est une densité de probabilité arbitraire sur \mathbb{R}_+ .

Ce théorème peut s'étendre au cas des SIRV complexes en introduisant la contrainte de circularité d'ordre 2 sur le vecteur \mathbf{X} , qui modifie la densité de probabilité donnée en (2.4). De manière plus générale, ce théorème s'énonce réciproquement pour des SIRV réels ou complexes :

Théorème 2 (Théorème de représentation - 2) *Si un vecteur aléatoire est un SIRV, alors il existe une variable aléatoire non négative τ telle que la densité de probabilité du vecteur aléatoire conditionné par τ est une densité de probabilité Gaussienne multivariée.*

Remarque : *Le parallèle effectué entre les définitions des SIRV réels et des SIRV complexes permet d'énoncer les prochains résultats dans le cas réel. L'extension au cas complexe s'effectue en considérant la densité de probabilité donnée en (2.2) associée au vecteur Gaussien complexe circulaire.*

2.1.1.3 Propriété d'invariance des SIRV par transformation linéaire

Par souci de clarté dans le texte, on notera un SIRV \mathbf{X} de dimension m , de vecteur moyenne μ , de matrice de covariance \mathbf{M} et de loi de texture $p(\cdot)$ par $\mathbf{X}(m, \mu, \mathbf{M}, p)$.

Lemme 1 (Invariance d'un SIRV par transformation linéaire) *Soit \mathbf{X} un SIRV (m, μ, \mathbf{M}, p) . Soit \mathbf{A} une matrice de taille $n \times m$ telle que $\mathbf{A}\mathbf{A}^T$ est non singulière et soit \mathbf{b} un vecteur de taille n . Alors : $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ est un SIRV $(n, \mathbf{A}\mu + \mathbf{b}, \mathbf{A}\mathbf{M}\mathbf{A}^T, p)$.*

Remarques :

- (1) Ce lemme permet d'étendre tous les résultats donnés pour les SIRV centrés aux SIRV de moyenne non nulle. Par conséquent, seuls les résultats sur les SIRV centrés sont énoncés dans ce chapitre.
- (2) Ce lemme est très pratique d'un point de vue simulation numérique d'un SIRV, car, à partir d'un SIRV blanc $\mathbf{B}(m, \mathbf{0}, \mathbf{Id}, p)$, on obtient aisément un SIRV $\mathbf{X}(m, \mu, \mathbf{M}, p)$ par la transformation linéaire $X = \mathbf{M}^{1/2} \mathbf{B} + \mu$.
- (3) D'autre part, le lemme d'invariance par transformation linéaire permet d'établir l'invariance de la forme quadratique contenue dans l'expression de la densité de probabilité d'un SIRV. En effet, considérons un SIRV \mathbf{X} quelconque de matrice de covariance \mathbf{M}_x dont la forme quadratique Q_x associée s'écrit $Q_x = \mathbf{X}^\dagger \mathbf{M}_x^{-1} \mathbf{X}$. Soit un autre SIRV \mathbf{Y} de même loi de texture que \mathbf{X} et dont la forme quadratique Q_y associée est $Q_y = \mathbf{Y}^\dagger \mathbf{M}_y^{-1} \mathbf{Y}$. D'après le lemme ci-dessus, on peut trouver une matrice \mathbf{A} telle que $\mathbf{Y} = \mathbf{A} \mathbf{X}$, et de ce fait, $\mathbf{M}_y = \mathbf{A} \mathbf{M}_x \mathbf{A}^T$. La forme quadratique Q_y s'écrit alors : $Q_y = (\mathbf{A} \mathbf{X})^\dagger (\mathbf{A} \mathbf{M}_x \mathbf{A}^T)^{-1} (\mathbf{A} \mathbf{X}) = \mathbf{X}^\dagger \mathbf{M}_x^{-1} \mathbf{X} = Q_x$.

L'invariance de la forme quadratique donnée en (3) a une implication statistique : quel que soit le SIRV considéré, la densité de probabilité de la forme quadratique est toujours la même (les détails des calculs se trouvent dans l'annexe C (C.19)). Sa statistique dépend de la loi de la texture du SIRV et s'écrit, pour un SIRV réel \mathbf{y} , de taille m , de matrice de covariance \mathbf{M}_x et de moyenne nulle :

$$p_\alpha(\alpha) = \int_0^{+\infty} \frac{\alpha^{m/2-1}}{2^{m/2} \Gamma(m/2)} \tau^{-m/2} \exp\left(-\frac{\alpha}{2\tau}\right) p(\tau) d\tau, \quad (2.6)$$

où $\alpha = \mathbf{y}^T \mathbf{M}^{-1} \mathbf{y} = \tau \mathbf{x}^T \mathbf{M}^{-1} \mathbf{x}$ d'après le modèle (2.1) pour un vecteur réel. Dans le cas particulier où $p(\tau) = \delta(\tau - 1)$, à savoir où le vecteur \mathbf{y} est Gaussien, réel, de taille m , la forme quadratique α est alors une somme de m variables aléatoires $\mathcal{N}(0, 1)$ au carré, soit une variable du χ_m^2 . Sa densité de probabilité coïncide bien avec (2.6) pour $p(\tau) = \delta(\tau - 1)$. La connaissance de la densité de probabilité de la forme quadratique du SIRV (réel ou complexe) est suffisante pour identifier la loi jointe de ce même SIRV [Rang93].

Pour un SIRV complexe, $\mathbf{M}_x = 2\mathbf{M}$ si \mathbf{M} est la matrice de covariance des composantes en quadrature de \mathbf{x} . La forme quadratique s'écrit donc

$$\alpha' = \mathbf{y}^\dagger \mathbf{M}_x^{-1} \mathbf{y} = \mathbf{y}^\dagger (2\mathbf{M})^{-1} \mathbf{y} = \frac{\mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y}}{2},$$

et représente, dans $(\mathbb{R}^2)^m$, la somme de $2m$ composantes (au lieu de m dans le cas réel). Sa densité de probabilité peut se déduire directement de (2.6) en remplaçant m par $2m$ et où $\alpha' = \alpha/2$:

$$p_{\alpha'}(\alpha') = p_\alpha(\alpha' = \frac{\alpha}{2}) = \int_0^{+\infty} \frac{\alpha'^{m-1}}{\Gamma(m)} \tau^{-m} \exp\left(-\frac{\alpha'}{\tau}\right) p(\tau) d\tau. \quad (2.7)$$

Remarque : La loi jointe d'un SIRV complexe a été précédemment exprimée en fonction de \mathbf{M} et non pas de $\mathbf{M}_x = 2\mathbf{M}$. La forme quadratique présente dans (2.3) est $\alpha = \tau \mathbf{x}^T \mathbf{M}^{-1} \mathbf{x}$, dont la loi se déduit de (2.6) pour $2m$ au lieu de m et dont l'expression est alors :

$$p_\alpha(\alpha) = \int_0^{+\infty} \frac{\alpha^{m-1}}{2^m \Gamma(m)} \tau^{-m} \exp\left(-\frac{\alpha}{2\tau}\right) p(\tau) d\tau. \quad (2.8)$$

2.1.1.4 Propriétés sur l'addition de deux SIRV

De même que pour les vecteurs Gaussiens, le résultat de l'addition de deux SIRV est encore un SIRV. Il n'est cependant pas toujours possible de caractériser le SIRV résultant et deux cas particuliers sont donnés ci-dessous.

La première propriété concerne deux SIRV de même loi de texture, mais dont les caractéristiques du premier et second ordre (moyenne et matrice de covariance) sont différentes.

Propriété 1 Soient $\mathbf{X}(m, \mu_x, \mathbf{M}_x, p)$ et $\mathbf{Y}(m, \mu_y, \mathbf{M}_y, p)$ deux SIRV statistiquement indépendants et dont la variance de la texture est bornée.

Alors, $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ est un SIRV $(m, \mu_x + \mu_y, \mathbf{M}_x + \mathbf{M}_y, p)$ si et seulement si $p(\tau) = \delta(\tau - a)$ où $0 < a < +\infty$ et ainsi \mathbf{X}, \mathbf{Y} et \mathbf{Z} sont des vecteurs aléatoires Gaussiens.

La propriété suivante détermine les caractéristiques d'un SIRV résultant de l'addition de deux autres SIRV, de même matrice de covariance et de même vecteur moyenne, mais dont toutes les autres composantes sont différentes :

Propriété 2 Soient $\mathbf{X}(m, \mu_x, \mathbf{M}, p_x)$ et $\mathbf{Y}(m, \mu_y, \mathbf{M}, p_y)$ deux SIRV statistiquement indépendants et dont les lois de texture sont des fonctions absolument continues.

Alors, $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ est un SIRV $(m, \mu_x + \mu_y, k\mathbf{M}, p_z)$ où :

$$p_z(\tau) = \sqrt{k} \int_0^{\tau^2} \frac{\tau}{\sqrt{\tau^2 - t^2}} p_x\left(\sqrt{\tau^2 - t^2}\right) p_y(t) dt, \quad (2.9)$$

$$k = \left[\frac{\int_0^{+\infty} x^2 p_x(x) dx + \int_0^{+\infty} y^2 p_y(y) dy}{\int_0^{+\infty} \int_0^r \frac{r^3}{\sqrt{r^2 - t^2}} p_x\left(\sqrt{r^2 - t^2}\right) p_y(t) dt dr} \right]^{2/3}, \quad (2.10)$$

avec la condition d'avoir k borné.

Ces propriétés sont résumées dans TAB. 2.1 et les démonstrations peuvent se trouver dans [Yao73].

	Taille du SIRV	Vecteur moyenne	Matrice de covariance	Loi de la texture τ	Condition \Rightarrow Conséquence
\mathbf{X}	m	$\boldsymbol{\mu}$	\mathbf{M}	p	
$\mathbf{AX} + \mathbf{b}$	n	$\mathbf{A}\boldsymbol{\mu} + \mathbf{b}$	$\mathbf{A}\mathbf{M}\mathbf{A}^T$	p	\mathbf{A} de taille $n \times m$ \mathbf{b} de taille n et $\mathbf{A}\mathbf{A}^T$ non singulière
\mathbf{X}	m	μ_x	\mathbf{M}_x	p	$\mathbb{E}(\tau^2) < \infty$
\mathbf{Y}	m	μ_y	\mathbf{M}_y	p	$\mathbb{E}(\tau^2) < \infty$
$\mathbf{Z} = \mathbf{X} + \mathbf{Y}$	m	$\mu_x + \mu_y$	$\mathbf{M}_x + \mathbf{M}_y$	p	ssi $p(\tau) = \delta(\tau - a)$, $a > 0$ $\Rightarrow \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ Gaussiens
\mathbf{X}	m	μ_x	\mathbf{M}	p_x	p_x abs.continue
\mathbf{Y}	m	μ_y	\mathbf{M}	p_y	p_y abs.continue
$\mathbf{Z} = \mathbf{X} + \mathbf{Y}$	m	$\mu_x + \mu_y$	$k\mathbf{M}$	p_z	

TAB. 2.1 – Quelques propriétés des SIRV

2.1.1.5 Caractérisation d'un SIRV blanc

Pour caractériser un SIRV blanc de moyenne nulle et de matrice de covariance identité, il est possible d'utiliser le théorème suivant, que l'on retrouve dans [Gold76] :

Théorème 3 (Caractérisation d'un SIRV blanc) *Un vecteur aléatoire $\mathbf{Y} = [Y_1, \dots, Y_m]^T$ est un SIRV blanc de moyenne nulle et de matrice de covariance identité si et seulement si il existe des variables aléatoires $R \in \mathbb{R}_+$, $\theta \in [0, 2\pi]$ et $\phi_k \in [0, \pi], k = 1, \dots, m-2$ telles que quand les composantes de \mathbf{Y} sont exprimées en coordonnées sphériques généralisées comme suit :*

$$\begin{aligned} Y_1 &= R \cos(\phi_1) \\ Y_k &= R \cos(\phi_k) \prod_{i=1}^{k-1} \sin(\phi_i) \quad (1 < k \leq m-2) \\ Y_{m-1} &= R \cos(\theta) \prod_{i=1}^{m-2} \sin(\phi_i) \\ Y_m &= R \sin(\theta) \prod_{i=1}^{m-2} \sin(\phi_i), \end{aligned} \quad (2.11)$$

alors les variables aléatoires R, θ et ϕ sont mutuellement et statistiquement indépendantes de densités de probabilité de la forme :

$$p_R(r) = \frac{r^{m-1}}{2^{m/2-1} \Gamma(m/2)} f_m(r^2) u(r) \quad (2.12)$$

$$p_{\phi_k} = \frac{\Gamma((m-k-1)/2)}{\sqrt{\pi} \Gamma((m-k)/2)} (\sin(\phi_k))^{m-k-1} \times [u(\phi_k) - u(\phi_k - \pi)] \quad (2.13)$$

$$p_{\theta}(\theta) = (2\pi)^{-1} [u(\theta) - u(\theta - 2\pi)] \quad (2.14)$$

où $\Gamma(\cdot)$ est la fonction Gamma-Eulérienne, $u(t)$ est la distribution d'Heaviside et où la fonction $f_m(t)$ est donnée par :

$$f_m(t) = \int_0^{+\infty} \tau^{-m} \exp\left(-\frac{t}{2\tau^2}\right) p(\tau) d\tau. \quad (2.15)$$

Par ce théorème, tout SIRV blanc de vecteur moyenne nul et de matrice de covariance identité peut être représenté en coordonnées sphériques généralisées, mutuellement et statistiquement indépendantes. De plus, les densités de probabilité des variables θ et ϕ_k étant fonctionnellement indépendantes du SIRV blanc considéré, seule la loi de la variable R change d'un SIRV à l'autre.

On peut noter que R , représentant la norme du SIRV ($R^2 = \sum_{k=1}^m Y_k^2 = \mathbf{Y}^T \mathbf{Y}$), est en fait la racine carrée de la forme quadratique associée à \mathbf{Y} . De par l'invariance de la forme quadratique d'un SIRV par transformation linéaire, ce théorème reste valable pour tout SIRV. Les densités de probabilité des variables θ , ϕ_k et R restent inchangées, quelle que soit la matrice de covariance du SIRV.

Remarque : *D'un point de vue simulation numérique, ce théorème permet de générer un SIRV blanc de moyenne nulle et de loi de texture quelconque à partir d'un vecteur aléatoire Gaussien, admettant lui aussi cette représentation. En effet, si on note par \mathbf{Z} un SIRV blanc*

Gaussien de taille m , de norme R_g , alors les composantes du vecteur s'écrivent comme indiqué par (2.11). Les densités de probabilité des variables ϕ_k et θ étant invariantes selon le choix du SIRV, on a alors :

$$\frac{Z_k}{R_g} = \frac{Y_k}{R} \quad k = 1, \dots, m, \quad (2.16)$$

où R est la norme du SIRV blanc \mathbf{Y} . Le lemme d'invariance des SIRV par transformation linéaire permet ensuite d'obtenir le SIRV voulu.

Cette procédure de simulation d'un SIRV à partir d'un SIRV blanc Gaussien de moyenne nulle a été présentée par M. Rangaswami, D.D. Weiner et A. Öztürk dans [Rang95]. La difficulté réside dans la simulation des variables de l'enveloppe R qui suivent la loi donnée par (2.12). En effet, si la loi de la texture n'est pas connue analytiquement, et si la procédure classique d'inversion de la fonction de répartition est difficile à utiliser alors il faut recourir à des méthodes de rejection classiques ou généralisées.

Remarque (sur la remarque) : Lorsque la variable de texture peut être simulée directement à partir de sa loi, la simulation de N SIRV quelconques réels ou complexes de taille m est immédiate : à partir de la représentation donnée en (2.1), il suffit de générer N SIRV Gaussiens réels ou complexes de moyenne et de matrice de covariance voulues (obtenus après transformation linéaire de N SIRV Gaussiens blancs réels ou complexes), puis de multiplier chaque SIRV par une réalisation de la variable de texture.

2.1.1.6 Caractérisation d'un SIRV par les fonctions h_m

Il existe une autre méthode pour caractériser un SIRV dont la loi de la texture n'est pas connue analytiquement. Cette méthode découle de la relation de récurrence obtenue après dérivations successives des fonctions h_m de (2.3). On la retrouve dans [Rang93] dans le cas d'un SIRV réel, dont la texture est non pas $\sqrt{\tau}$, mais s . La relation diffère un peu, car dans pareil cas, il faut distinguer les ordres pairs et impairs des fonction h_m .

Dans la suite, les fonctions h_m sont définies par (2.3). En calculant la dérivée première de h_m , la relation de récurrence est immédiate, soit :

$$h_{m+1}(w) = (-2) \frac{dh_m(w)}{dw}. \quad (2.17)$$

Par récurrence, on obtient alors :

$$h_{m+1}(w) = (-2)^m \frac{d^m h_1(w)}{dw^m}. \quad (2.18)$$

Pour que la densité de probabilité multivariée obtenue de cette manière existe et reste une densité de probabilité, il est nécessaire que la fonction h_1 soit positive, monotone décroissante.

D'après la relation (2.18), il est nécessaire de connaître h_1 pour déduire les ordres supérieurs de la fonction h_m . Selon la statistique du SIRV, la fonction h_1 s'exprime soit en fonction de la loi marginale d'une composante du SIRV complexe, soit en fonction de la loi de l'enveloppe d'une

composante.

Si la loi marginale connue est la densité de probabilité marginale d'une composante du vecteur complexe \mathbf{Y} , elle est directement reliée à h_1 par (en posant $m = 1$ dans (2.3)) :

$$f_{Y_k}(Y_k) = \frac{1}{2\pi\sigma^2} h_1(\alpha). \quad (2.19)$$

Dans ce cas précis, $\alpha = |Y_k|^2$ et σ^2 est la variance commune des composantes en quadrature de \mathbf{X}_k . De plus, la loi (2.19) correspond à la loi jointe des composantes quadrature Y_{kI} et Y_{kQ} de Y_k .

Si la loi marginale connue est la densité de probabilité marginale de l'enveloppe de Y_k , la relation entre h_1 et cette loi marginale se déduit de (2.19). En effet, l'enveloppe et la phase des composantes de Y_k sont données par

$$R_k = \sqrt{Y_{kI}^2 + Y_{kQ}^2}$$

et

$$\phi_k = \arctan\left(\frac{Y_{kQ}}{Y_{kI}}\right),$$

où la phase est supposée uniformément distribuée sur $[0, 2\pi]$. En partant de la loi du couple (Y_{kI}, Y_{kQ}) , en passant en coordonnées polaires (R_k, ϕ_k) on obtient la densité de probabilité marginale de l'enveloppe R_k (en intégrant sur la loi de la phase ϕ_k) qui s'exprime en fonction de h_1 comme suit :

$$p_{R_k}(r) = \frac{r}{\sigma^2} h_1\left(\frac{r^2}{\sigma^2}\right). \quad (2.20)$$

La fonction h_1 ainsi définie, la relation (2.18) permet d'obtenir les lois jointes de tout SIRV multi-dimensionnel et ce, quelle que soit la nature du SIRV.

Remarque : *Il faut bien distinguer la loi jointe d'un vecteur en coordonnées cartésiennes et la loi jointe d'un vecteur SIRV qui est toujours sous forme radiale. L'appellation "sphériquement invariant" provient de cette forme radiale, et par exemple, si la matrice de covariance du vecteur est l'identité, alors la forme quadratique α , présente dans l'expression de la loi jointe, représente l'équation de la sphère dans \mathbb{R}^m (ou \mathbb{R}^{2m} pour un vecteur complexe), d'où l'invariance sphérique.*

2.1.1.7 Détermination des fonctions h_m par la fonction caractéristique

La fonction caractéristique usuelle d'un SIRV (transformée de Fourier multidimensionnelle) est une fonction sphériquement invariante à partir de laquelle il est possible de trouver h_m .

Comme précédemment, la démarche consiste à exprimer h_1 en fonction de la fonction caractéristique des composantes en quadrature Y_{kI} et Y_{kQ} de \mathbf{Y}_k , puis d'utiliser la relation de récurrence (2.17) pour en déduire h_m .

La fonction caractéristique bidimensionnelle des composantes quadrature Y_{kI} et Y_{kQ} de \mathbf{Y}_k s'écrit :

$$\Phi_{Y_{kI}, Y_{kQ}}(w_1, w_2) = \mathbb{E} [\exp(j w_1 Y_{kI} + j w_2 Y_{kQ})]. \quad (2.21)$$

C'est une fonction de $\sqrt{w_1^2 + w_2^2}$ (elle est sphériquement invariante) et en la notant $\Phi(w)$ on a la relation suivante entre $h_m(\alpha)$ et $\Phi(w)$:

$$h_m(\alpha) = (\sqrt{\alpha})^{1-m} \int_0^{+\infty} w^m \Phi(w) J_{m-1}(\sqrt{\alpha} w) dw, \quad (2.22)$$

où $J_\nu(x)$ est la fonction de Bessel de première espèce d'ordre ν . Toute fonction caractéristique marginale paire et absolument intégrable peut être choisie pour $\Phi(w)$.

2.1.2 Lois jointes de vecteurs non-Gaussiens issus d'un SIRP

Les expressions des lois données ci-après et récapitulées dans les tableaux (TAB. 2.2) et (TAB. 2.3) représentent les lois jointes du vecteur complexe \mathbf{y} tel qu'il est décrit en (2.1). Les densités de probabilité sont séparées en deux catégories : celles pour lesquelles la loi de la texture est connue et celles pour lesquelles la loi de la texture est inconnue. Dans chacun des cas, les expressions de h_m et de la fonction caractéristique (quand elle existe) sont données.

2.1.2.1 Densités de probabilité pour lesquelles $p(\tau)$ est connue

Loi Gaussienne

La loi Gaussienne s'obtient avec $p(\tau) = \delta(\tau - 1) = 1$ si $\tau = 1$, 0 sinon, et la fonction $h_m(\alpha)$ est égale à $\exp(-\alpha/2)$. La loi marginale est définie sur \mathbb{R} par $p_{y_{kI}}(y) = \sqrt{2\pi}^{-1} \exp(-y^2/2)$ et la loi jointe d'un vecteur SIRV Gaussien complexe de moyenne nulle est donnée par :

$$p(\mathbf{y}) = \frac{1}{(2\pi)^m |\mathbf{M}|} \exp\left(-\frac{\alpha}{2}\right). \quad (2.23)$$

La matrice \mathbf{M} est la matrice de covariance des composantes en quadrature de \mathbf{x} ; dans ce cas-là, la matrice de covariance \mathbf{M}_y de \mathbf{y} est égale à \mathbf{M} . Pour déterminer (2.23), il est également possible de passer par la fonction caractéristique radiale et d'utiliser (2.22).

Pour un vecteur Gaussien complexe, on obtient h_m puis $p(\mathbf{y})$ à partir de la fonction caractéristique radiale qui s'écrit :

$$\Phi(w) = \exp\left(-\frac{w^2}{2}\right). \quad (2.24)$$

Loi de Laplace

La loi de Laplace de paramètre b s'obtient avec τ qui suit une loi exponentielle de paramètre $b^2/2$ (notée $\mathcal{E}(b^2/2)$) définie sur \mathbb{R}_+ et donnée par :

$$p(\tau) = \frac{b^2}{2} \exp\left(-\frac{b^2 \tau}{2}\right). \quad (2.25)$$

La loi marginale est définie sur \mathbb{R} par $p_{y_{kI}}(y) = \frac{b}{2} \exp(-b|y|)$. La fonction $h_m(\alpha)$ dépend de la fonction de Bessel modifiée de deuxième espèce d'ordre $m-1$, $K_{m-1}(\cdot)$ par la relation suivante :

$$h_m(\alpha) = \frac{b^{m+1}}{\sqrt{\alpha}^{m-1}} K_{m-1}(b\sqrt{\alpha}), \quad (2.26)$$

et la loi jointe du vecteur SIRV complexe admet donc l'expression suivante :

$$p(\mathbf{y}) = \frac{b^{m+1}}{(2\pi)^m |\mathbf{M}| \sqrt{\alpha}^{m-1}} K_{m-1}(b\sqrt{\alpha}). \quad (2.27)$$

Loi de Cauchy

La loi de Cauchy de paramètre b s'obtient quand τ suit une loi Inverse Gamma de paramètres 1 et $2/b^2$ (notée $\mathcal{IG}(1, 2/b^2)$) définie sur \mathbb{R}_+ et donnée par :

$$p(\tau) = \frac{b^2}{2} \tau^{-2} \exp\left(-\frac{b^2}{2\tau}\right). \quad (2.28)$$

(La loi Inverse Gamma de paramètres a et b est la loi de l'inverse d'une variable aléatoire de loi Gamma, de paramètres a et b .)

La loi marginale est définie sur \mathbb{R} par :

$$p_{y_{kI}}(y) = \frac{b}{\pi(b^2 + y^2)}, \quad (2.29)$$

et la fonction $h_m(\alpha)$ a pour expression :

$$h_m(\alpha) = \frac{2^m b \Gamma(m + 1/2)}{\sqrt{\pi}} (b^2 + \alpha)^{-m-1/2}, \quad (2.30)$$

et la loi jointe du vecteur SIRV complexe devient :

$$p(\mathbf{y}) = \frac{2^m b \Gamma(m + 1/2)}{\sqrt{\pi} (2\pi)^m |\mathbf{M}|} (b^2 + \alpha)^{-m-1/2}. \quad (2.31)$$

La K-distribution

La K-distribution est une loi à deux paramètres ν et b . Le paramètre ν est appelé le paramètre de forme car il détermine la forme de la loi. Quand ν est petit la densité de probabilité est très "piquée" alors que quand ν tend vers l'infini, la densité de probabilité tend vers une densité

de probabilité Gaussienne. La loi de la texture pour la K-distribution est une loi Gamma de paramètres ν et $2/b^2$ (notée $\mathcal{G}(\nu, 2/b^2)$) dont l'expression est la suivante :

$$p(\tau) = \left(\frac{b^2}{2}\right)^\nu \frac{\tau^{\nu-1}}{\Gamma(\nu)} \exp\left(-\frac{b^2\tau}{2}\right). \quad (2.32)$$

La loi marginale indiquée pour la K-distribution est la loi de l'enveloppe R_k des composantes quadrature définie sur \mathbb{R}_+ et est donnée par :

$$p_{R_k}(r) = \frac{2b}{\Gamma(\nu)} \left(\frac{br}{2}\right)^\nu K_{\nu-1}(br). \quad (2.33)$$

De là, en utilisant (2.20) et (2.17) on obtient :

$$h_m(\alpha) = \frac{b^{2m}}{\Gamma(\nu)} \frac{(b\sqrt{\alpha})^{\nu-m}}{2^{\nu-1}} K_{m-\nu}(b\sqrt{\alpha}), \quad (2.34)$$

et la loi jointe du vecteur SIRV complexe s'écrit :

$$p(\mathbf{y}) = \frac{b^{2m}}{\Gamma(\nu)\sqrt{\pi}(2\pi)^m |\mathbf{M}|} \frac{(b\sqrt{\alpha})^{\nu-m}}{2^{\nu-1}} K_{m-\nu}(b\sqrt{\alpha}). \quad (2.35)$$

La loi du SIRV dont la marginale de l'enveloppe est une K-distribution peut également s'obtenir à partir de la fonction caractéristique radiale définie sur \mathbb{R}_+ par :

$$\Phi(w) = \left(1 + \frac{w^2}{b^2}\right)^{-\nu}. \quad (2.36)$$

La loi de Student-t

La loi de Student-t est une loi à deux paramètres ν et b . La loi de la texture du SIRV la représentant est une loi $\mathcal{IG}(\nu, 2/b^2)$ dont l'expression est la suivante :

$$p(\tau) = \left(\frac{b^2}{2}\right)^\nu \frac{\tau^{-\nu-1}}{\Gamma(\nu)} \exp\left(-\frac{b^2}{2\tau}\right). \quad (2.37)$$

La loi marginale est définie sur \mathbb{R} par :

$$p_{y_{kI}}(y) = \frac{\Gamma(\nu + 1/2)}{b\sqrt{\pi}\Gamma(\nu)} \left(1 + \frac{y^2}{b^2}\right)^{-\nu-1/2}, \quad (2.38)$$

et la fonction $h_m(\alpha)$ a pour expression :

$$h_m(\alpha) = \frac{2^m b^{2\nu} \Gamma(\nu + m)}{\Gamma(\nu)(b^2 + \alpha)^{m+\nu}}, \quad (2.39)$$

ce qui donne lieu à la loi jointe suivante du SIRV complexe :

$$p(\mathbf{y}) = \frac{2^m b^{2\nu} \Gamma(\nu + m)}{(2\pi)^m |\mathbf{M}| \Gamma(\nu)(b^2 + \alpha)^{m+\nu}}. \quad (2.40)$$

Cette expression s'obtient également à partir de la fonction caractéristique radiale :

$$\Phi(w) = \frac{(bw)^\nu K_\nu(bw)}{2^{\nu-1} \Gamma_\nu}. \quad (2.41)$$

Remarque : la loi de Cauchy marginale (2.29) se déduit de la loi de Student-t pour $\nu = 1/2$.

	SIRV dont la texture est de loi connue		$p(\tau)$ du SIRV	
Densité marginale	$f_Y(y)$	$h_m(\alpha)$	$f_\tau(\tau)$	$E(\tau)$
Gauss	$\frac{1}{\sqrt{2\pi}} \exp(-y^2/2)$	$\exp(-\alpha/2)$	$\delta(\tau - 1)$	1
Laplace	$\frac{b}{2} \exp(-b y)$	$b^{2m} (b\sqrt{\alpha})^{1-m} K_{m-1}(b\sqrt{\alpha})$	$\frac{b^2}{2} \exp\left(-\frac{b^2 \tau}{2}\right) : \text{Exp}\left(\frac{b^2}{2}\right)$	$\frac{2}{b^2}$
Cauchy	$\frac{b}{\pi(b^2 + y^2)}$	$\frac{2^m b \Gamma(m + 1/2)}{\sqrt{\pi} (b^2 + \alpha)^{m+1/2}}$	$\frac{b^2 \tau^{-2}}{2} \exp\left(-\frac{b^2}{2\tau}\right) : \text{IG}\left(1, \frac{2}{b^2}\right)$	∞
K-distribution	$\frac{2b}{\Gamma(\nu)} \left(\frac{by}{2}\right)^\nu K_{\nu-1}(by)$	$\frac{b^{2m}}{\Gamma(\nu)} \frac{(b\sqrt{\alpha})^{\nu-m}}{2^{\nu-1}} K_{m-\nu}(b\sqrt{\alpha})$	$\left(\frac{b^2}{2}\right)^\nu \frac{\tau^{\nu-1}}{\Gamma(\nu)} \exp\left(-\frac{b^2}{2} \tau\right) : \mathcal{G}\left(\nu, \frac{2}{b^2}\right)$	$\frac{2\nu}{b^2}$
Student-t	$\frac{\Gamma(\nu + 1/2)}{b \sqrt{\pi} \Gamma(\nu)} \left(1 + \frac{y^2}{b^2}\right)^{-\nu-1/2}$	$\frac{2^m b^{2\nu} \Gamma(\nu + m)}{\Gamma(\nu) (b^2 + \alpha)^{m+\nu}}$	$\left(\frac{b^2}{2}\right)^\nu \frac{\tau^{-\nu-1}}{\Gamma(\nu)} \exp\left(-\frac{b^2}{2\tau}\right) : \text{IG}\left(\nu, \frac{2}{b^2}\right)$	$\frac{b^2}{2(\nu - 1)}$

TAB. 2.2 – Tableau récapitulatif : densité de probabilité des SIRV dont la texture est de loi connue

Densité marginale	$f_{\mathbf{Y}}(y)$	$h_m(\alpha)$
Chi	$\frac{2b}{\Gamma(\nu)} (by)^{2\nu-1} \exp(-b^2 y^2)$	$\frac{(-2)^m}{\Gamma(\nu)} (b\sigma)^{2\nu} \sum_{k=1}^m G_k \alpha^{\nu-k} \exp(-b^2 \sigma^2 \alpha)$ $G_k = (-1)^{k-1} C_{m-1}^{k-1} (b^2 \sigma^2)^{k-1} \prod_{i=1}^{m-1} (\nu - i) \quad \nu \leq 1$
Rayleigh	$\frac{\eta y}{\beta^2 \Gamma(2/\eta)} \exp\left(-\left(\frac{y}{\beta}\right)^\eta\right)$	$\frac{(-2)^{m-1} \eta \sigma^2 \exp\left(-\left(\frac{\sigma}{\beta}\right)^\eta \alpha^{\eta/2}\right)}{\beta^2 \Gamma(2/\eta)} \sum_{k=1}^{m-1} \frac{D_k}{k!} \left(\frac{\sigma}{\beta}\right)^{\eta k} \alpha^{1-m+k\eta/2}$ $D_k = \sum_{n=1}^k (-1)^n C_k^n \prod_{i=0}^{n-2} (n\eta/2 - i) \quad \eta \leq 2$
Weibull	$a b y^{b-1} \exp(-a y^b)$	$(-2)^m \sum_{k=1}^m B_k \frac{(a\sigma^b)^k}{k!} \alpha^{kb/2-m} \exp(-a\sigma^b \alpha^{kb/2})$ $B_k = \sum_{n=1}^k (-1)^n C_k^n \prod_{i=0}^{n-1} (nb/2 - i) \quad b \leq 2$
Rice	$\frac{y \exp\left(-\frac{y^2}{2(1-\rho^2)}\right)}{\sqrt{1-\rho^2}} I_0\left(\frac{\rho y^2}{2(1-\rho^2)}\right)$	$\frac{\sigma^{2m} \exp\left(-\frac{\alpha \sigma^2}{2(1-\rho^2)}\right)}{(1-\rho^2)^{m-1/2}} \sum_{k=0}^{m-1} (-1)^k C_{m-1}^k (\rho/2)^k B_k$ $B_k = \sum_{n=0}^k C_k^n I_{k-2n}\left(\rho \frac{\alpha \sigma^2}{2(1-\rho^2)}\right)$

TAB. 2.3 – Tableau récapitulatif : densité de probabilité des SIRV dont la texture est de loi inconnue

2.1.2.2 Densités de probabilité pour lesquelles $p(\tau)$ n'est pas connue

Lorsque $p(\tau)$ n'est pas connue analytiquement, il est nécessaire d'utiliser les fonctions h_m pour déterminer la loi du SIRV correspondant. Ces lois s'obtiennent à partir de la loi marginale de l'enveloppe des composantes quadrature du vecteur complexe \mathbf{y} . Pour simuler de tels vecteurs, il est nécessaire d'utiliser la représentation en coordonnées sphériques donnée par (2.11).

La loi χ

La loi χ est à différencier de la loi couramment appelée loi du χ^2 . La loi marginale de l'enveloppe de la loi du χ_n^2 (n degrés de liberté) est connue pour être la loi de la somme du carré de n variables aléatoires $\mathcal{N}(0, 1)$ alors que la loi du χ est la loi de la racine carrée de la somme du carré de n variables aléatoires $\mathcal{N}(0, 1)$.

Dans le cas d'un SIRV dont la loi marginale de l'enveloppe des composantes en quadrature suit une loi χ à ν degrés de liberté et de paramètre b donnée sur \mathbb{R}_+ par :

$$p_{R_k}(r) = \frac{2b}{\Gamma\nu} (br)^{2\nu-1} \exp(-b^2 r^2), \quad (2.42)$$

la fonction h_m se déduit de (2.20) et de (2.17). En utilisant (2.20) on peut écrire :

$$h_1(\alpha) = \frac{2(b\sigma)^{2\nu}}{\Gamma(\nu)} \alpha^{\nu-1} \exp(-b^2 \sigma^2 \alpha), \quad (2.43)$$

et en utilisant (2.17) on a :

$$h_m(\alpha) = (-2)^{m-1} \frac{d^{m-1} h_1(\alpha)}{d\alpha^{m-1}} = \frac{(-2)^{m-1}}{\Gamma(\nu)} 2(b\sigma)^{2\nu} \frac{d^{m-1}}{d\alpha^{m-1}} \left[\alpha^{\nu-1} e^{-b^2 \sigma^2 \alpha} \right]. \quad (2.44)$$

L'expression de $h_m(\alpha)$ qui en résulte est donnée dans (TAB. 2.3) et n'est valable que pour $\nu \leq 1$. Cette condition est imposée par le fait que h_1 doit être une fonction monotone décroissante, ce qui est vrai uniquement pour $\nu \leq 1$.

La loi de \mathbf{y} s'obtient en utilisant (2.2) et donne :

$$p(\mathbf{y}) = \frac{(-1)^{m-1}}{(\pi)^m |\mathbf{M}| \Gamma(\nu)} (b\sigma)^{2\nu} \sum_{k=1}^m (-1)^{k-1} C_{m-1}^{k-1} (b^2 \sigma^2)^{k-1} D_k(\alpha, \nu, b, \sigma), \quad (2.45)$$

avec

$$D_k(\alpha, \nu, b, \sigma) = \prod_{i=1}^{m-1} (\nu - i) \alpha^{\nu-k} \exp(-b^2 \sigma^2 \alpha), \quad (2.46)$$

$$C_{m-1}^{k-1} = \frac{(m-1)!}{(k-1)! (m-k)!} \quad (2.47)$$

La loi de Rayleigh généralisée

La loi marginale de l'enveloppe des composantes quadrature de \mathbf{y} est définie sur \mathbb{R}_+ est une loi de Rayleigh généralisée de paramètres η et β donnée par :

$$p_{R_k}(r) = \frac{\eta r}{\beta^2 \Gamma(2/\eta)} \cdot \exp\left(-\left(\frac{r}{\beta}\right)^\eta\right) \quad (2.48)$$

Quand $\eta = 2$, on retrouve la densité de probabilité de Rayleigh usuelle.

La fonction h_m , valable uniquement pour $\eta \leq 2$ est donnée dans (TAB. 2.3) et la loi du SIRV correspondant se déduit de (2.2) et donne :

$$p(\mathbf{y}) = \frac{(-1)^{m-1} \eta \sigma^2 \exp\left(-\left(\frac{\sigma}{\beta}\right)^\eta \alpha^{\eta/2}\right)}{2\pi^m |\mathbf{M}| \beta^2 \Gamma(2/\eta)} \sum_{k=1}^{m-1} \frac{1}{k!} D_k(\eta) \left(\frac{\sigma}{\beta}\right)^{\eta k} \alpha^{1-m+k\eta/2}, \quad (2.49)$$

avec

$$D_k(\eta) = \sum_{j=1}^k (-1)^j C_k^j \prod_{i=0}^{j-1} (j\eta/2 - i). \quad (2.50)$$

La loi de Weibull

La loi de Weibull est une loi à deux paramètres a et b . La loi marginale de l'enveloppe des composantes quadrature de \mathbf{y} est définie sur \mathbb{R}_+ par :

$$p_{R_k}(r) = a b r^{b-1} \exp(-a r^b). \quad (2.51)$$

En utilisant là encore (2.20) et (2.17), on obtient la fonction h_m donnée dans (TAB. 2.3), valable pour $0 \leq b \leq 2$. La loi résultante du SIRV donne :

$$p(\mathbf{y}) = \frac{(-1)^m}{\pi^m |\mathbf{M}|} \sum_{k=1}^m \frac{1}{k!} D_k(b) (a\sigma^b)^k \alpha^{kb/2-m} \exp(-a\sigma^b \alpha^{kb/2}), \quad (2.52)$$

avec

$$D_k(b) = \sum_{j=1}^k (-1)^j C_k^j \prod_{i=0}^{j-1} (mb/2 - i). \quad (2.53)$$

Remarque : La loi de la variable τ n'est pas connue analytiquement, mais admet cependant une représentation intégrale. Cette expression est reliée aux fonctions G de Meijer et peut se retrouver dans [Cont98]. Le lecteur pourra juger par lui-même de la complexité de l'intégrale à calculer :

$$p(\tau) = \frac{1}{2\pi\tau^2} \int_0^{+\infty} \exp\left[-a r^{b/2} \cos\left(\frac{\pi b}{4}\right)\right] \cos\left[a r^{b/r^{b/2}} \sin\left(\frac{\pi b}{4} - \frac{r}{2\tau}\right)\right] dr. \quad (2.54)$$

La loi de Rice

La loi marginale de l'enveloppe des composantes quadrature de \mathbf{y} est définie sur \mathbb{R}_+ et dépend de la fonction de Bessel modifiée de première espèce et d'ordre 0, $I_0(\cdot)$:

$$p_{R_k}(r) = \frac{r}{\sqrt{1-\rho^2}} \exp\left(-\frac{r^2}{2(1-\rho^2)}\right) I_0\left(\frac{\rho r^2}{2(1-\rho^2)}\right), \quad (2.55)$$

pour $0 < \rho \leq 1$. En posant

$$A = \frac{\alpha \sigma^2}{2(1-\rho^2)},$$

on obtient $h_1(\alpha)$:

$$h_1(\alpha) = \frac{\sigma^2}{\sqrt{1-\rho^2}} \exp(-A) I_0(\rho A). \quad (2.56)$$

La fonction h_m résultante est donnée dans (TAB. 2.3) et la loi du SIRV devient :

$$p(\mathbf{y}) = \frac{\sigma^{2m} \exp\left(-\frac{\alpha\sigma^2}{2(1-\rho^2)}\right)}{(2\pi)^m |\mathbf{M}| (1-\rho^2)^{m-1/2}} \sum_{k=1}^m (-1)^k C_k^{m-1} (\rho/2)^k D_k(\alpha, \sigma, \rho), \quad (2.57)$$

avec

$$D_k(\alpha, \sigma, \rho) = \sum_{m=0}^k C_k^m I_{k-2m} \left(\rho \frac{\alpha\sigma^2}{2(1-\rho^2)} \right). \quad (2.58)$$

2.1.3 Les SIRV généralisés

La modélisation de processus non-Gaussiens par les SIRP généralisés a été étudié par T.J. Barnard pour ses travaux de thèse, dont on retrouve les principaux résultats dans [Barn96].

Dans le cas des SIRV traditionnels, le vecteur Gaussien modélise les fluctuations rapides du fouillis, la matrice de covariance détermine la corrélation de ces fluctuations, et la texture caractérise les effets environnementaux associés à la cellule observée. Les réalisations successives de la variable de texture sont indépendantes, ce qui signifie que les échos provenant des différentes cases distance sont statistiquement indépendants, alors que cela peut s'avérer ne pas être le cas.

Deux limitations aux SIRV traditionnels sont soulevées dans [Barn96]. La première concerne la dépendance des composantes du SIRV malgré une possible décorrélation. En effet, même si les composantes du vecteur Gaussien sont indépendantes (si la matrice de corrélation est diagonale), la densité de probabilité du vecteur SIRV résultant donnée en (2.2) ne peut pas se factoriser en produit de m fonctions, dépendant respectivement de chaque composante du SIRV. La deuxième limitation concerne l'indépendance des réalisations successives du SIRV. Cela est dû aussi bien à l'indépendance des réalisations de la texture, qu'à l'indépendance des réalisations du vecteur Gaussien. Physiquement, cela signifie que le modèle SIRV traditionnel traite les échos des cellules voisines comme étant indépendants des échos de la cellule observée.

Généraliser les SIRV traditionnels permet d'outrepasser ces limitations et suppose que la loi de la texture est différente pour chaque vecteur observé. Ce modèle s'écrit comme suit, en considérant N observations de taille m pour lesquelles les vecteurs Gaussiens sont supposés indépendants :

$$\begin{aligned} \mathbf{Y}_1 &= \sqrt{\tau_1} \mathbf{X}_1 \\ &\vdots \\ \mathbf{Y}_N &= \sqrt{\tau_N} \mathbf{X}_N. \end{aligned} \quad (2.59)$$

$$(2.60)$$

Ainsi, chaque vecteur \mathbf{Y}_i est un SIRV différent des autres. Les vecteurs Gaussiens \mathbf{X}_i étant indépendants, il en va de même pour les vecteurs conditionnels \mathbf{Y}_i/τ_i et la loi jointe du vecteur

$(\mathbf{Y}_1/\tau_1, \dots, \mathbf{Y}_N/\tau_N)$ s'écrit :

$$p(\mathbf{Y}_1, \dots, \mathbf{Y}_N/\tau_1, \dots, \tau_N) = \prod_{i=1}^N p(\mathbf{Y}_i/\tau_i), \quad (2.61)$$

où

$$p(\mathbf{Y}_i/\tau_i) = \frac{1}{(2\pi)^m |\mathbf{M}_i| \tau_i^m} \exp\left(-\frac{\mathbf{Y}_i^\dagger \mathbf{M}_i^{-1} \mathbf{Y}_i}{2\tau_i}\right). \quad (2.62)$$

La loi jointe du vecteur $(\mathbf{Y}_1, \dots, \mathbf{Y}_N)$ résultant s'écrit alors :

$$\begin{aligned} p(\mathbf{Y}_1, \dots, \mathbf{Y}_N) &= \int_0^{+\infty} \dots \int_0^{+\infty} p(\mathbf{Y}_1, \dots, \mathbf{Y}_N/\tau_1, \dots, \tau_N) p(\tau_1, \dots, \tau_N) d\tau_1 \dots d\tau_N \\ &= (2\pi)^{-mN} \left(\prod_{i=1}^N |\mathbf{M}_i|^{-1} \right) h_{mN}(\alpha_1, \dots, \alpha_N), \end{aligned} \quad (2.63)$$

où $\alpha_i = \mathbf{Y}_i^\dagger \mathbf{M}_i^{-1} \mathbf{Y}_i$, et

$$h_{mN}(\alpha_1, \dots, \alpha_N) = \int_0^{+\infty} \dots \int_0^{+\infty} \left[\prod_{i=1}^N \tau_i^{-m} \exp\left(-\frac{\alpha_i}{2\tau_i}\right) \right] p(\tau_1, \dots, \tau_N) d\tau_1 \dots d\tau_N. \quad (2.64)$$

D'après cette expression, l'indépendance des composantes d'un SIRV généralisé peut être réalisée en choisissant la taille voulue pour N et en posant $m = 1$.

On peut cependant faire une remarque sur les SIRV généralisés : les vecteurs Gaussiens \mathbf{X}_i étant de moyenne nulle, les vecteurs \mathbf{Y}_i sont décorrélés entre eux, et ceci quelle que soit la corrélation des τ_i . Pour prouver ce résultat, il suffit de calculer l'intercovariance entre deux vecteurs SIRV généralisés, qui s'écrit :

$$\mathbf{A}_{Y_i, Y_j} = \mathbb{E}(\mathbf{Y}_i \mathbf{Y}_j^\dagger) - \mathbb{E}(\mathbf{Y}_i) \mathbb{E}(\mathbf{Y}_j)^\dagger.$$

L'indépendance des \mathbf{Y}_i avec les τ_i et des \mathbf{X}_i entre eux, on trouve alors que $\mathbf{A}_{Y_i, Y_j} = \mathbf{0}$, tandis que $\mathbf{A}_{Y_i, Y_i} = \mathbf{M}_i$.

Utiliser le modèle généralisé permet de prendre en compte l'éventuelle non-homogénéité des échos de l'environnement du radar. Supposons, par exemple, que les données enregistrées se présentent, pour une case distance et un azimuth donnés, comme une mixture de trois lois. Dans ce cas là, le modèle de SIRV généralisé s'appliquerait en considérant les données segmentées en trois classes, avec trois lois de texture différentes. Le modèle SIRV traditionnel ne tiendrait pas compte de cette non-homogénéité et la loi de la texture serait une "moyenne" des trois lois présentes. Ceci dit, il faudrait être capable de segmenter les données en plusieurs classes puis d'identifier la statistique associée à chaque classe pour ensuite obtenir une loi globale sur toutes les classes. L'alternative avec les SIRV traditionnels serait de se placer dans chacune des classes pour lesquelles une statistique "SIRV traditionnel" serait associée.

2.1.3.1 Propriétés des SIRV généralisés

Les SIRV généralisés admettent certaines propriétés, qui peuvent être vues comme des extensions des propriétés des SIRV traditionnels.

La première est la *propriété d'invariance par transformation linéaire* où, à la $i^{\text{ème}}$ transformation sont associés une matrice \mathbf{A}_i et un vecteur \mathbf{b}_i . Le vecteur formé des vecteurs $\mathbf{z}_i = \mathbf{A}_i \mathbf{y}_i + \mathbf{b}_i$, où $i = 1, \dots, N$ est encore un SIRV généralisé, dont la loi jointe est donnée par :

$$p(\mathbf{z}_1, \dots, \mathbf{z}_N) = (2\pi)^{-mN} \left(\prod_{i=1}^N |\mathbf{A}_i^\dagger \mathbf{M}_i \mathbf{A}_i|^{-1} \right) h_{mN}(\alpha_{z_1}, \dots, \alpha_{z_N}), \quad (2.65)$$

où $\alpha_{z_i} = (\mathbf{A}_i \mathbf{z}_i + \mathbf{b}_i)^\dagger \mathbf{M}_i^{-1} (\mathbf{A}_i \mathbf{z}_i + \mathbf{b}_i)$ et où $h_{mN}(\alpha_{z_1}, \dots, \alpha_{z_N})$ est donnée par (2.64) en $(\alpha_{z_1}, \dots, \alpha_{z_N})$.

La deuxième propriété concerne les fonctions h_{mN} des SIRV généralisés. De la même manière que pour les SIRV traditionnels, les fonctions h_{mN} admettent une relation de récurrence qui découle du calcul de la dérivée partielle de h_{mN} selon le vecteur $(\alpha_1, \dots, \alpha_N)$:

$$h_{(m+1)N}(\alpha_1, \dots, \alpha_N) = (-2)^N \frac{\partial^N h_{mN}(\alpha_1, \dots, \alpha_N)}{\partial \alpha_1 \dots \partial \alpha_N}. \quad (2.66)$$

L'équation (2.66) est appelée le *théorème de "bootstrap"* des SIRV généralisés. À partir de h_{1N} , il est possible d'obtenir les SIRV généralisés d'ordres supérieurs.

De plus, en utilisant la définition suivante :

Définition 4 Une fonction $f(x_1, \dots, x_N)$ est monotone décroissante, si

$$f(x_1 + \delta_1, \dots, x_N + \delta_N) \leq f(x_1, \dots, x_N),$$

pour toutes les valeurs (x_1, \dots, x_N) et pour toutes valeurs positives de $(\delta_1, \dots, \delta_N)$.

on montre facilement que les fonctions h_{mN} sont monotones décroissantes. Par application de (2.66), les dérivées successives sont donc monotones, alternativement croissantes et décroissantes.

Enfin, T.J. Barnard montre dans sa thèse, que la transformation de Laplace simplifie l'expression des fonctions h_{mN} . Ce résultat se trouve dans [Barn96], où l'expression simplifiée de h_{mN} en fonction de la transformation de Laplace découle d'une expression particulière de la loi jointe de (τ_1, \dots, τ_N) . Cette dernière expression est un produit de lois Inverse Gamma multipliées par une somme de fonctions g dépendantes de l'inverse des τ_i . L'expression de h_{mN} fait alors apparaître une somme normalisée de dérivées $m^{\text{ème}}$ de la transformée de Laplace de g , dont les arguments sont les α_i . Pour cette représentation des h_{mN} , le théorème de bootstrap ne s'applique pas car la loi jointe des textures dépend de m . La relation de récurrence donnée par (2.66) n'est donc plus directe. De plus, avec cette représentation, il est possible de trouver une paire de transformées de Laplace (i.e. une fonction associée à sa transformée) telle que la loi jointe des textures ne soit pas une densité, alors que la loi jointe du SIRV générée à partir des h_{mN} est une densité.

2.2 Stratégies de détection optimales pour un environnement SIRV connu a priori

Dans le cadre de la détection radar selon le critère de Neymann-Pearson explicité dans le chapitre 1, il est nécessaire de connaître la statistique des enregistrements effectués par le récepteur radar. Les données reçues dans chaque case distance après émission de chaque impulsion radar sont des vecteurs complexes et représentent les composantes en quadrature (ou *speckle*). La modélisation de ces données par les SIRV permet d'obtenir les expressions des lois jointes des composantes sous chacune des hypothèses du critère de Neymann-Pearson, nécessaires à l'élaboration du test du rapport de vraisemblance (RV). Dans cette section, l'expression du RV est donnée sous une forme générale valable pour tout SIRV puis des exemples de détecteurs sont présentés pour des environnements de fouillis donnés.

2.2.1 Le test du rapport de vraisemblance (RV) pour les SIRV

L'environnement non-Gaussien du radar est représenté par des vecteurs complexes reçus sous les hypothèses H_0 et H_1 du test de détection, modélisés par des SIRV. Le test de détection s'écrit alors :

$$\mathbf{y} = \zeta \mathbf{s} + \mathbf{x} \sqrt{\tau} \quad (2.67)$$

$$H_0 : \zeta = 0 \quad (2.68)$$

$$H_1 : \zeta \neq 0. \quad (2.69)$$

Sous les deux hypothèses, le vecteur complexe \mathbf{y} est de taille m et sa matrice de covariance \mathbf{M}_y peut s'écrire en fonction de la matrice de covariance \mathbf{M} du *speckle* comme $\mathbf{M}_y = 2 \mathbb{E}(\tau) \mathbf{M}$. La densité de probabilité sous l'hypothèse H_0 du vecteur \mathbf{y} est déterminée par (2.2) et s'écrit :

$$p(\mathbf{y}/H_0) = \frac{1}{(2\pi)^m |\mathbf{M}|} \int_0^{+\infty} \tau^{-m} \exp\left(-\frac{q_0(\mathbf{y})}{2\tau}\right) p(\tau) d\tau, \quad (2.70)$$

où $q_0(\mathbf{y}) = \mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y}$ est une forme quadratique non négative.

Dans le cas où le vecteur \mathbf{s} est connu, la densité de probabilité de \mathbf{y} sous l'hypothèse H_1 se déduit facilement de $p(\mathbf{y}/H_0)$ par :

$$p(\mathbf{y}/H_1) = p(\mathbf{y} - \mathbf{s}/H_0) = \frac{1}{(2\pi)^m |\mathbf{M}|} \int_0^{+\infty} \tau^{-m} \exp\left(-\frac{q_1(\mathbf{y})}{2\tau}\right) p(\tau) d\tau, \quad (2.71)$$

où $q_1(\mathbf{y}) = q_0(\mathbf{y} - \mathbf{s}) = (\mathbf{y} - \mathbf{s})^\dagger \mathbf{M}^{-1} (\mathbf{y} - \mathbf{s})$.

Avec ces deux expressions, le RV peut être donné pour tout SIRV, et la condition à vérifier pour décider H_1 ou H_0 est :

$$\Lambda(\mathbf{y}) = \frac{p(\mathbf{y}/H_1)}{p(\mathbf{y}/H_0)} = \frac{\int_0^{+\infty} \tau^{-m} \exp\left(-\frac{q_1(\mathbf{y})}{2\tau}\right) p(\tau) d\tau}{\int_0^{+\infty} \tau^{-m} \exp\left(-\frac{q_0(\mathbf{y})}{2\tau}\right) p(\tau) d\tau} \underset{H_0}{\overset{H_1}{>}} \eta, \quad (2.72)$$

où le seuil de détection η est défini tel que :

$$P_{fa} = \mathbb{P}(\Lambda(\mathbf{y}) \stackrel{H_0}{>} \eta) = \int_{\{\mathbf{y}:\Lambda(\mathbf{y})>\eta\}} p(\mathbf{y}/H_0)d\mathbf{y} = \beta. \quad (2.73)$$

La région définie pour le domaine d'intégration correspond à l'ensemble de décision $R_1 = \{\mathbf{y} : \text{décider } H_1 \text{ ou rejeter } H_0\}$, le complémentaire de cet ensemble étant $R_0 = \{\mathbf{y} : \text{décider } H_0 \text{ ou rejeter } H_1\}$.

Remarque : La loi du vecteur SIRV \mathbf{y} est fonction de h_m comme décrit en (2.2). Le RV s'écrit donc simplement comme étant le rapport des fonctions $h_m(q_1(\mathbf{y}))$ et $h_m(q_0(\mathbf{y}))$, soit :

$$\Lambda(\mathbf{y}) = \frac{h_m(q_1(\mathbf{y}))}{h_m(q_0(\mathbf{y}))} \stackrel{H_1}{\underset{H_0}{>}} \eta. \quad (2.74)$$

Extension aux SIRV généralisés

Si le fouillis est modélisé par un SIRV généralisé, une expression générale du RV est donnée, non plus pour un vecteur d'observation \mathbf{y} reçu sous chacune des hypothèses, mais pour les N vecteurs d'observations $(\mathbf{y}_1, \dots, \mathbf{y}_N)$.

Dans le cas d'un signal cible \mathbf{s} connu on a :

$$\Lambda(\mathbf{y}_1, \dots, \mathbf{y}_N) = \frac{h_{mN}(q_0(\mathbf{y}_1 - \mathbf{s}), \dots, q_0(\mathbf{y}_N - \mathbf{s}))}{h_{mN}(q_0(\mathbf{y}_1), \dots, q_0(\mathbf{y}_N))}, \quad (2.75)$$

où l'expression de h_{mN} est donnée par (2.64).

2.2.2 Le rapport de vraisemblance généralisé (RVG)

Quand le signal cible \mathbf{s} n'est pas connu, le rapport de vraisemblance généralisé (RVG) est mis en place, après estimation au sens du MV de l'ensemble des paramètres θ inconnus du signal cible. Comme au chapitre 1, l'estimation de θ s'effectue comme suit, à partir, cette fois, de (2.72) :

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \Lambda(\mathbf{y}) = \underset{\theta}{\operatorname{argmax}} p(\mathbf{y}, \theta/H_1). \quad (2.76)$$

La quantité $p(\mathbf{y}/H_0)$ n'apparaît pas dans l'expression à maximiser, puisqu'elle est indépendante de θ .

Si l'amplitude A du signal cible est la seule inconnue du problème ($\mathbf{s} = A\mathbf{p}$ où \mathbf{p} est fonction des autres paramètres), le RVG, pour les SIRV, admet la même expression que (2.72), où désormais :

$$q_1(\mathbf{y}) = q_0(\mathbf{y}) - \frac{|\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{y}|^2}{\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}}. \quad (2.77)$$

Cette expression est obtenue en remplaçant A par son estimée \hat{A}_{mv} , dont l'expression, donnée au chapitre 1 par (1.15), reste la même que dans le cas Gaussien. En effet, les intégrales (2.70)

et (2.71) étant positives et uniformément convergentes, le théorème de dérivation des fonctions intégrales s'applique. En d'autres termes, annuler la dérivée de (2.71) revient à annuler la dérivée de son intégrand, qui n'est rien d'autre que la dérivée de la fonction exponentielle.

La procédure d'estimation reste la même pour les autres paramètres du signal cible ; en revanche, lorsque tous les paramètres doivent être estimés simultanément, un problème d'optimisation complexe se pose, pour lequel il n'existe pas de solution analytique. Dans ce cas, on peut avoir recours à des méthodes numériques d'optimisation comme la méthode du gradient (méthode du premier ordre, quand il n'existe pas de minima locaux), la méthode de Newton (méthode utilisant le second ordre mais s'il n'y a pas de minima locaux), les méthodes dites heuristiques comme le recuit simulé et les algorithmes génétiques, ou encore les méthodes MCMC (Monte Carlo par Chaîne de Markov).

Dans les exemples de détecteurs qui suivent, seule l'amplitude du signal cible est estimée et les formes quadratiques sont données par $q_0(\mathbf{y}) = \mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y}$ et par (2.77) pour $q_1(\mathbf{y})$.

Les expressions des détecteurs optimaux sont données pour les environnements de fouillis non-Gaussiens dont les expressions des lois jointes sont données dans (TAB. 2.2) et (TAB. 2.3). Les noms des détecteurs sont attribués en fonction de la nature de l'environnement comme par exemple l'OKD pour Optimum K Detector (détecteur optimal pour un environnement K-distribué) ou encore l'OCD pour Optimum Cauchy Detector (optimal pour un environnement de Cauchy), et ainsi de suite. Le dernier détecteur présenté (OCpGD pour l'Optimum Clutter plus Gaussian Detector) est un cas général pour lequel l'expression analytique ne peut pas être déterminée : l'environnement est supposé être un SIRV de statistique quelconque ($p(\tau)$ quelconque) auquel s'ajoute le bruit thermique engendré par l'électronique du système radar et connu pour être un vecteur de bruit Gaussien.

2.2.3 Optimum K Detector : l'OKD

Lorsque la statistique du fouillis radar est connue ou supposée être K-distribuée (de paramètres ν et b), la loi de la texture est une loi Gamma de paramètres ν et $2/b^2$ (cf. (2.32)) d'expression :

$$p(\tau) = \frac{b^{2\nu}}{2^\nu \Gamma(\nu)} \exp\left(-\frac{2\tau}{b^2}\right). \quad (2.78)$$

La loi du test sous l'hypothèse H_0 est donnée par (2.35) pour le vecteur \mathbf{y} et sa forme quadratique $q_0(\mathbf{y})$:

$$p(\mathbf{y}/H_0) = \frac{b^{2m}}{\Gamma(\nu)\sqrt{\pi}(2\pi)^m |\mathbf{M}|} \frac{(b\sqrt{q_0(\mathbf{y})})^{\nu-m}}{2^{\nu-1}} K_{m-\nu}(b\sqrt{q_0(\mathbf{y})}). \quad (2.79)$$

Pour un signal cible d'amplitude inconnue et estimée selon (1.15), le RVG devient $\forall m \geq 2$:

$$\left(\frac{q_1(\mathbf{y})}{q_0(\mathbf{y})}\right)^{\frac{\nu-m}{2}} \cdot \frac{K_{\nu-m}(b\sqrt{q_1(\mathbf{y})})}{K_{\nu-m}(b\sqrt{q_0(\mathbf{y})})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (2.80)$$

Pour $m = 1$, $q_0(y) = |y|^2$ et l'expression devient :

$$|y|^{\nu-1} K_{\nu-1}(b|y|) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{(2\beta)^{\frac{\nu-1}{2}} \Gamma(\nu)}{2\nu\eta}. \quad (2.81)$$

2.2.4 Optimum Laplace Detector : OLD

Ce détecteur peut être considéré comme un cas particulier de l'OKD puisqu'il coïncide avec ce dernier pour $\nu = 1$. Avec l'expression de la loi jointe donnée dans TAB.(2.2) correspondant à $p(\mathbf{y}/H_0)$ où α devient $q_0(\mathbf{y})$, le RVG s'écrit pour $m \geq 2$:

$$\left(\frac{q_1(\mathbf{y})}{q_0(\mathbf{y})} \right)^{\frac{1-m}{2}} \cdot \frac{K_{m-1}\left(b\sqrt{q_1(\mathbf{y})}\right)}{K_{m-1}\left(b\sqrt{q_0(\mathbf{y})}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (2.82)$$

2.2.5 Optimum Student-t Detector : OStD

D'après TAB. 2.2, le RVG s'écrit dans le cas où la loi marginale de \mathbf{y} est une loi de Student-t :

$$\left(\frac{b^2 + q_0(\mathbf{y})}{b^2 + q_1(\mathbf{y})} \right)^{m+\nu} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (2.83)$$

2.2.6 Optimum Cauchy Detector : OCD

La loi de Cauchy est un cas particulier de la loi de Student-t lorsque $\nu = 1/2$. Le RVG devient donc :

$$\left(\frac{b^2 + q_0(\mathbf{y})}{b^2 + q_1(\mathbf{y})} \right)^{m+1/2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (2.84)$$

2.2.7 Optimum χ Detector : O χ D

Lorsque la loi marginale de l'enveloppe des composantes en quadrature de \mathbf{y} est une loi χ (cf. TAB. 2.3), le RVG s'écrit :

$$\frac{\sum_{k=1}^m (-1)^{k-1} C_{m-1}^{k-1} (b^2 \sigma^2)^{k-1} D_k(q_1(\mathbf{y}), \nu, b, \sigma)}{\sum_{k=1}^m (-1)^{k-1} C_{m-1}^{k-1} (b^2 \sigma^2)^{k-1} D_k(q_0(\mathbf{y}), \nu, b, \sigma)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (2.85)$$

où, pour $j = 0, 1$: $D_k(q_j(\mathbf{y}), \nu, b, \sigma) = \prod_{i=1}^{m-1} (\nu-i) (q_j(\mathbf{y}))^{\nu-k} e^{-b^2 \sigma^2 q_j(\mathbf{y})}$, et $C_{m-1}^{k-1} = \frac{(m-1)!}{(k-1)! (m-k)!}$.

2.2.8 Optimum generalized Rayleigh Detector : OgRD

Lorsque la loi marginale de l'enveloppe des composantes en quadrature de \mathbf{y} est une loi de Rayleigh généralisée (cf. TAB. 2.3), le RVG s'écrit pour $\eta \leq 2$:

$$\left(\frac{q_1(\mathbf{y})}{q_0(\mathbf{y})} \right)^{1-m} \exp \left(\frac{\sigma \eta}{\beta \eta} [q_0(\mathbf{y})^{\eta/2} - q_1(\mathbf{y})^{\eta/2}] \right) \frac{\sum_{k=1}^{m-1} \frac{D_k(\eta)}{k!} \left(\frac{\sigma \sqrt{q_1(\mathbf{y})}}{\beta} \right)^{\eta k}}{\sum_{k=1}^{m-1} \frac{D_k(\eta)}{k!} \left(\frac{\sigma \sqrt{q_0(\mathbf{y})}}{\beta} \right)^{\eta k}} \underset{H_0}{\overset{H_1}{\gtrless}} \xi, \quad (2.86)$$

où ξ est le nouveau seuil, $D_k(\eta) = \sum_{j=1}^k (-1)^j C_k^j \prod_{i=0}^{j-1} \left(\frac{j\eta}{2} - i \right)$, et $C_k^j = \frac{(k)!}{(j)! (k-j)!}$.

2.2.9 Optimum Weibull Detector : OWD

Lorsque la loi marginale de l'enveloppe des composantes en quadrature de \mathbf{y} est une loi de Weibull (cf. TAB. 2.3), le RVG s'écrit pour $0 \leq b \leq 2$:

$$\left(\frac{q_0(\mathbf{y})}{q_1(\mathbf{y})} \right)^m \frac{\sum_{k=1}^{m-1} \frac{D_k(b)}{k!} (a\sigma^b)^k \sqrt{q_1(\mathbf{y})}^{kb} \exp(-a\sigma^b \sqrt{q_1(\mathbf{y})}^{kb})}{\sum_{k=1}^{m-1} \frac{D_k(b)}{k!} (a\sigma^b)^k \sqrt{q_0(\mathbf{y})}^{kb} \exp(-a\sigma^b \sqrt{q_0(\mathbf{y})}^{kb})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (2.87)$$

où $D_k(b) = \sum_{j=1}^k (-1)^j C_k^j \prod_{i=0}^{j-1} \left(\frac{j b}{2} - i \right)$, et $C_k^j = \frac{(k)!}{(j)! (k-j)!}$.

2.2.10 Optimum Rice Detector : ORD

Lorsque la loi marginale de l'enveloppe des composantes en quadrature de \mathbf{y} est une loi de Rice (cf. TAB. 2.3), le RVG s'écrit pour $0 < \rho \leq 1$:

$$\exp [K(q_0(\mathbf{y}) - q_1(\mathbf{y}))] \frac{\sum_{k=1}^m (-1)^k C_{m-1}^k (\rho/2)^k D_k(q_1(\mathbf{y}), \sigma, \rho)}{\sum_{k=1}^m (-1)^k C_{m-1}^k (\rho/2)^k D_k(q_0(\mathbf{y}), \sigma, \rho)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (2.88)$$

où $K = \frac{\sigma^2}{2(1-\rho^2)}$ et pour $j = 0, 1$: $D_k(q_j(\mathbf{y})) = \sum_{l=0}^k C_k^l I_{k-2l}(K \rho q_j(\mathbf{y}))$.

$I_\nu(\cdot)$ est la fonction de Bessel modifiée de première espèce d'ordre ν , et $C_k^l = \frac{(k)!}{(l)!(k-l)!}$.

2.2.11 Optimum Clutter plus Gaussian Detector : l'OCpGD

Quand le bruit thermique \mathbf{b} est pris en compte dans les données reçues par le radar, on a alors, sous l'hypothèse H_0 :

$$H_0 : \mathbf{y} = \mathbf{x} \sqrt{\tau} + \mathbf{b}, \quad (2.89)$$

où \mathbf{b} est un vecteur aléatoire complexe circulaire de variables aléatoires Gaussiennes décorréllées, de moyenne nulle et de variance $2\sigma_b^2$, c'est-à-dire que $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}, 2\sigma_b^2 \mathbf{I})$.

Ainsi la matrice de covariance de \mathbf{y} conditionnellement à τ s'écrit :

$$\mathbf{M}_{\mathbf{y}/\tau} = \frac{1}{2} \mathbb{E}(\mathbf{y}\mathbf{y}^\dagger / \tau) = \tau \mathbf{M}_x + \sigma_b^2 \mathbf{I}, \quad (2.90)$$

où \mathbf{I} désigne la matrice identité, et la densité de probabilité de \mathbf{y}/H_0 devient :

$$p(\mathbf{y}/H_0) = \frac{1}{(2\pi)^m} \int_0^{+\infty} \frac{1}{|\tau \mathbf{M}_x + \sigma_b^2 \mathbf{I}|} \exp\left(-\frac{\mathbf{y}^\dagger (\tau \mathbf{M}_x + \sigma_b^2 \mathbf{I})^{-1} \mathbf{y}}{2}\right) p(\tau) d\tau. \quad (2.91)$$

Le RV donné en (2.72) sous forme générale admet dans ce cas l'expression suivante :

$$\Lambda(\mathbf{y}) = \frac{p(\mathbf{y}/H_1)}{p(\mathbf{y}/H_0)} = \frac{\int_0^{+\infty} \frac{1}{|\tau \mathbf{M}_x + \sigma_b^2 \mathbf{I}|} \exp\left(-\frac{q_1(\mathbf{y}, \tau)}{2}\right) p(\tau) d\tau}{\int_0^{+\infty} \frac{1}{|\tau \mathbf{M}_x + \sigma_b^2 \mathbf{I}|} \exp\left(-\frac{q_0(\mathbf{y}, \tau)}{2}\right) p(\tau) d\tau} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (2.92)$$

où désormais

$$q_0(\mathbf{y}, \tau) = \mathbf{y}^\dagger (\tau \mathbf{M}_x + \sigma_b^2 \mathbf{I})^{-1} \mathbf{y}, \quad (2.93)$$

et $q_1(\mathbf{y}, \tau)$ se déduit de $q_0(\mathbf{y}, \tau)$ comme dans les cas précédents selon le degré de connaissance sur le signal cible (cf. (2.77)).

Cette nouvelle expression du RV ne se calcule pas analytiquement, et ceci quelque soit l'expression de $p(\tau)$. F. Gini, M.V. Greco, A. Farina et P. Lombardo ont étudié ce cas précis en considérant un fouillis K-distribué [Gini98]. Les auteurs montrent que si des détecteurs tels l'OGD ou l'OKD sont mis en place alors que du bruit thermique est présent dans les données, alors les performances de ces détecteurs sont dégradées en comparaison des performances obtenues par l'OKGD.

Une solution alternative, sous-jacente à la propriété d'invariance des SIRV après addition, serait de considérer la somme des deux vecteurs de bruit (parfois plus que deux) comme un seul et unique SIRV. Cependant, lorsque les composantes principales du SIRV (matrice de covariance

et loi de texture) sont différentes, il n'existe pas, d'après les propriétés énoncées, d'expression pour la matrice de covariance et la loi de la texture du SIRV résultant. On peut donc envisager d'estimer ces quantités inconnues à partir des données reçues du radar en considérant être en présence d'un seul et unique SIRV. L'expression du détecteur obtenue sous une telle considération serait donc valable quelle que soit la nature statistique des données.

Dans le chapitre 3, deux méthodes d'estimation de la loi de la texture sont proposées, après avoir mené une étude bayésienne du problème posé. Ces deux méthodes sont dans leur ordre de présentation, la méthode des approximants de Padé (méthode basée sur une approximation sous forme de fraction rationnelle de la fonction génératrice de moments) et une estimation bayésienne à la loi de la texture. Le problème de l'estimation de la matrice de covariance est juste évoqué dans ce rapport dans le but d'étudier l'influence de la corrélation sur les performances des détecteurs obtenus et en prenant pour référence principale les travaux de F.Gini [Gini99, Gini00b, Gini00c].

Chapitre 3

Stratégies de détection optimales pour un environnement inconnu

Dans ce chapitre, deux nouvelles stratégies optimales de détection sont établies, après modélisation de l'environnement par un SIRP. Elles constituent un apport nouveau et important en détection radar puisqu'aucune connaissance a priori n'est nécessaire sur la statistique de l'environnement. Suite à une analyse bayésienne du modèle SIRP, la loi de la texture est estimée à partir de données de référence, selon deux méthodes différentes. La première utilise une approximation de Padé de la fonction génératrice de moments de la texture, ce qui permet d'obtenir le détecteur PEOD, ou Padé Estimated Optimum Detector. La seconde méthode fournit un estimateur Monte Carlo-bayésien à la loi de la texture, sous forme analytique, qui nous conduit au détecteur BORD, ou Bayesian Optimum Radar Detector. Les performances de ces deux détecteurs sont analysées dans des environnements SIRV de natures différentes, et des résultats très importants sont établis, notamment pour le BORD.

3.1 Analyse bayésienne du modèle SIRV

3.1.1 La règle de Bayes

L'objectif d'une telle analyse est de déterminer *a posteriori* la loi de la texture du modèle SIRV de l'environnement du radar à partir de N_r mesures de vecteurs référence bruit. Soit $\mathbf{R} = [\mathbf{r}_1, \dots, \mathbf{r}_{N_r}]^t$ une collection de N_r SIRV de taille m tel que, $\forall 1 \leq i \leq N_r$, $\mathbf{r}_i = [r_i(1), \dots, r_i(m)]$.

D'après le modèle SIRV, on a : $\mathbf{r}_i = \mathbf{x}_i \sqrt{\tau}$. Connaissant la loi conditionnelle de \mathbf{r}_i par rapport à τ , à savoir une loi Gaussienne multivariée, la règle de Bayes nous fournit une expression de la loi *a posteriori* de la *texture*, en fonction de sa "vraie" loi (inconnue), de la vraisemblance des références (connue) et de la loi du vecteur des références (inconnue). On a donc :

$$p(\tau/\mathbf{r}_i) = \frac{p(\mathbf{r}_i/\tau) p(\tau)}{p(\mathbf{r}_i)}, \quad (3.1)$$

où la constante de normalisation $p(\mathbf{r}_i)$ est calculée par intégration de la vraisemblance des données $p(\mathbf{r}_i/\tau)$ sur la loi de la *texture* :

$$p(\mathbf{r}_i) = \int_0^{+\infty} p(\mathbf{r}_i/\tau) p(\tau) d\tau. \quad (3.2)$$

Dans (3.1), seule $p(\mathbf{r}_i/\tau)$ est connue. Afin d'en déterminer une expression analytique, une densité *a priori* $g(\tau)$ est choisie pour représenter $p(\tau)$ sur les données de références. Ainsi, (3.1) se ré-écrit en remplaçant $p(\tau)$ par $g(\tau)$:

$$p(\tau/\mathbf{r}_i) = \frac{p(\mathbf{r}_i/\tau) g(\tau)}{p(\mathbf{r}_i)}, \quad (3.3)$$

et la constante de normalisation $p(\mathbf{r}_i)$ se calcule désormais en intégrant la vraisemblance des références sur le prior $g(\tau)$. On a donc :

$$p(\mathbf{r}_i) = \int_0^{+\infty} p(\mathbf{r}_i/\tau) g(\tau) d\tau. \quad (3.4)$$

La densité *a priori* que l'on attribue à la *texture* est choisie selon le degré de connaissance sur la variable aléatoire τ du modèle. Ce choix est explicité plus en détail dans le paragraphe suivant.

3.1.2 Choix du prior

La question du choix du prior est toujours sujette à beaucoup de discussions, surtout lorsqu'aucune information physique n'est disponible sur la variable. Certaines personnes restent sceptiques quant à l'application même du paradigme bayésien à un problème physique, alors que d'autres y adhèrent totalement, en se détachant quelque peu de la réalité physique du problème. En fait, le paradigme bayésien représente une toute autre façon de penser un problème de la physique classique, le but étant de "laisser parler" les données du problème afin d'en extraire les informations qu'elles contiennent.

Si les paramètres d'intérêt sont bien définis (intervalle de valeurs, distribution, ...), le choix du prior est alors imposé par ces connaissances. Il serait effectivement dommage de considérer, par exemple, un paramètre à valeurs négatives s'il est connu pour représenter une distance ou une variance. Dans ces cas, tout le monde tombe d'accord sur l'utilité du prior. Si, en revanche, aucune information, ou presque, n'est disponible sur les paramètres d'intérêt, des priors définis comme "non-informatifs" peuvent être utilisés selon la nature des paramètres. Par exemple, de tels priors existent pour tout paramètre d'échelle ou de forme. Le scepticisme de certains apparaît alors, surtout si ces méthodes permettent d'obtenir de meilleurs résultats que des méthodes déterministes : comment justifier que l'information apportée n'apporte justement aucune information ?

Un autre choix de prior possible, est un prior dit conjugué. Ce choix est plus souvent justifié pour des raisons de confort de calculs que pour des raisons physiques. Dans ce cas, la loi *a posteriori* des paramètres d'intérêt appartient à la même famille de loi que la loi du prior. Ce qui signifie que le choix du prior détermine la loi *a posteriori*. Dans pareil cas, je consens que certains puissent croire que pour voir du rouge, il suffit d'en regarder.

Dans le but d'identifier la loi *a posteriori* de la texture à partir de données de référence bruit seul, dont la statistique est inconnue, seules deux possibilités de prior s'offrent à nous : un prior non-informatif ou un prior conjugué. Cette alternative est discutée dans le paragraphe qui suit afin de justifier, à notre sens, l'emploi préférentiel d'un prior non-informatif.

3.1.2.1 Prior non-informatif

Pour un environnement radar inconnu, la seule connaissance apportée par la modélisation SIRV est la positivité de la variable. 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 ([Robe92, Robe94, Robe96, Tann96]) :

$$g(\tau) = \frac{1}{\tau}, \quad (3.5)$$

son expression garantit 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.

3.1.2.2 Prior conjugué

Une densité prior est dite *conjuguée* si la loi *a posteriori* résultante appartient à la même classe de densités. Ceci traduit l'invariance de la famille de lois par la multiplication du terme de vraisemblance. Quelques exemples de lois conjuguées, tirés de [Decl98], sont donnés à titre indicatif dans TAB. 3.1 pour la famille des lois exponentielles.

Loi <i>a posteriori</i>	$p(\tau/y)$	Vraisemblance $p(y/\tau)$	Prior conjugué	$g(\tau)$
Normale	$\mathcal{N}\left(\frac{\sigma^2\mu + \beta^2y}{\sigma^2 + \beta^2}, \frac{\sigma^2\beta^2}{\sigma^2 + \beta^2}\right)$	Normale $\mathcal{N}(\tau, \sigma^2)$	Normale	$\mathcal{N}(\mu, \beta^2)$
Gamma	$\mathcal{G}\left(\alpha + 0.5, \beta + \frac{(\mu - y)^2}{2}\right)$	Normale $\mathcal{N}(\mu, 1/\tau)$	Gamma	$\mathcal{G}(\alpha, \beta)$
Gamma	$\mathcal{G}(\alpha + y, \beta + 1)$	Poisson $\mathcal{P}(\tau)$	Gamma	$\mathcal{G}(\alpha, \beta)$
Gamma	$\mathcal{G}(\alpha + \nu, \beta + y)$	Gamma $\mathcal{G}(\nu, \tau)$	Gamma	$\mathcal{G}(\alpha, \beta)$
Inverse Gamma	$IG\left(\alpha + 1, \frac{2\beta}{2 + \beta(y - m)^2}\right)$	Normale $\mathcal{N}(m, \tau)$	Inverse Gamma	$IG(\alpha, \beta)$
Beta	$\mathcal{Be}(\alpha + y, \beta + n - y)$	Binomiale $\mathcal{B}(n, \tau)$	Beta	$\mathcal{Be}(\alpha, \beta)$
Beta	$\mathcal{Be}(\alpha + m, \beta + y)$	Neg. Binom. $\mathcal{B}(m, \tau)$	Beta	$\mathcal{Be}(\alpha, \beta)$
Dirichlet	$\mathcal{D}(\alpha_1 + y_1, \dots, \alpha_k + y_k)$	Multinom. $\mathcal{M}_k(\tau_1, \dots, \tau_k)$	Dirichlet	$\mathcal{D}(\alpha_1, \dots, \alpha_k)$
Exp.	$\mathcal{E}(\lambda - \log(1 - y))$	Beta $\mathcal{B}(\alpha, \tau)$	Exp.	$\mathcal{E}(\lambda)$

TAB. 3.1 – Priors conjugués pour quelques lois

Dans le cas où le terme de vraisemblance est une Gaussienne, et que le paramètre d'intérêt est la variance, le prior conjugué possible pour τ est une loi Inverse Gamma. La loi *a posteriori* résultante est donc une loi Inverse Gamma.

Si un tel choix est adopté, le réglage des paramètres de la loi conjuguée doit donc être effectué avec précaution afin de ne pas égarer la loi *a posteriori* qui en dépend. Cependant, une loi conjuguée devient presque non informative lorsque ses paramètres sont choisis exagérément. Par exemple, si la loi Inverse Gamma $\mathcal{IG}(\tau; a_p, b_p)$ est choisie pour $g(\tau)$ dans (3.5) alors avec a_p proche de 0 et b_p très grand on a : $\mathcal{IG}(\tau; a_p, b_p) \propto a_p/\tau$.

Le prior non-informatif de Jeffrey paraît donc bien adapté à notre degré d'ignorance sur la *texture*. En poursuivant les calculs initiés en début de ce chapitre, l'expression de la loi *a posteriori* de la texture en découle.

3.1.2.3 Expression de la loi *a posteriori* instantanée de la texture

Connaissant N_r données de référence bruit, la loi *a posteriori* résultante est obtenue en remplaçant $g(\tau)$ par (3.5) dans (3.1). Dans ce cas, la constante de normalisation $p(\mathbf{r}_i)$ est calculée par intégration du numérateur dans (3.1) sur le prior non-informatif de Jeffrey $g(\tau)$, et devient :

$$\begin{aligned} p(\mathbf{r}_i) &= \int_0^{+\infty} \frac{\tau^{-m-1}}{(2\pi)^m |\mathbf{M}|} \exp\left(-\frac{\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i}{2\tau}\right) d\tau \\ &= \frac{\Gamma(m)}{\pi^m |\mathbf{M}| (\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i)^m}. \end{aligned} \quad (3.6)$$

Ainsi, la loi *a posteriori* instantanée (3.1) devient à son tour :

$$\begin{aligned} p(\tau/\mathbf{r}_i) &= \frac{\tau^{-m-1}}{(2\pi)^m |\mathbf{M}| p(\mathbf{r}_i)} \exp\left(-\frac{\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i}{2\tau}\right) \\ &= \frac{(\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i)^m}{2^m \Gamma(m)} \tau^{-m-1} \exp\left(-\frac{\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i}{2\tau}\right). \end{aligned} \quad (3.7)$$

Cette loi représente la loi *a posteriori* instantanée de la *texture*. Elle repose sur la connaissance d'un seul vecteur de référence \mathbf{r}_i , et peut d'ailleurs être identifiée : elle correspond à une loi Inverse Gamma de paramètres m et $2/\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i$ que l'on note $\mathcal{IG}(\tau; m, 2/\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i)$. On peut déduire de ce résultat que, choisir un prior conjugué ou un prior non-informatif résulte en la même loi, aux valeurs des paramètres du prior près. Le prior (3.5) peut être vu, dans ce cas, comme un *prior conjugué non-informatif*.

Cette analyse a permis d'identifier la principale inconnue du modèle SIRV, à savoir la *texture*. Une autre inconnue à déterminer est la matrice de covariance du *speckle*, et l'estimation de cette matrice va être réalisée à partir des mêmes données de référence que celles utilisées pour la caractérisation de la loi *a posteriori* de la texture.

3.2 Estimation NSCM de la matrice de covariance

Pour un vecteur \mathbf{r} de taille m , sa matrice de covariance \mathbf{M}_r est définie par :

$$\mathbf{M}_r = \mathbb{E}(\mathbf{r} \mathbf{r}^\dagger) = \int_{\mathbb{R}^m} \mathbf{r} \mathbf{r}^\dagger p(\mathbf{r}) d\mathbf{r}, \quad (3.8)$$

ou plus simplement, en calculant chaque terme de la matrice indépendamment pour $i, j = 1, \dots, m$:

$$\mathbf{M}_{ij} = \mathbb{E}(r_i r_j^*) = \int r_i r_j^* p(r_i, r_j) dr_i dr_j. \quad (3.9)$$

Connaissant N_r vecteurs de référence bruit $\mathbf{r}_{i=1}^{N_r}$, de taille m , l'estimation Monte-Carlo de la matrice de covariance \mathbf{M}_r des données s'effectue classiquement par :

$$\widehat{\mathbf{M}}_r = \frac{1}{N_r} \sum_{i=1}^{N_r} \mathbf{r}_i \mathbf{r}_i^\dagger. \quad (3.10)$$

Pour le modèle SIRV, la matrice de covariance présente dans les expressions est la matrice de covariance \mathbf{M} du *speckle* \mathbf{x} . Sachant que cette matrice est reliée à la matrice de covariance des données observées par :

$$\mathbf{M}_r = 2\mathbb{E}(\tau) \mathbf{M}, \quad (3.11)$$

où $2\mathbb{E}(\tau)$ représente la puissance d'une observation SIRV, une estimation dite NSCM (Normalized Structured covariance Matrix) de \mathbf{M} à partir des vecteurs \mathbf{r}_i s'écrit [Gini99, Gini00b, Gini00c] :

$$\widehat{\mathbf{M}} = \frac{m}{N_r} \sum_{i=1}^{N_r} \frac{\mathbf{r}_i \mathbf{r}_i^\dagger}{\mathbf{r}_i^\dagger \mathbf{r}_i}. \quad (3.12)$$

Dans cette expression, la matrice de covariance des données est divisée par la puissance locale moyenne. Cela permet d'obtenir une estimation de la matrice de covariance du *speckle* directement à partir de données de référence SIRV.

Cette méthode, utilisée dans le cas des SIRV et des vecteurs Gaussiens composés différencie l'ALQ (Adaptive Linear Quadratic detector), pour lequel elle est utilisée du GLRT (rapport de vraisemblance généralisé obtenu sous l'hypothèse Gaussienne) pour lequel la matrice de covariance est estimée selon (3.10), avec des vecteurs Gaussiens. Les performances des deux approches sont comparées dans [Gini00b].

3.2.1 Modèle de matrice de covariance utilisé pour les simulations

Pour analyser les performances des détecteurs présentés dans ce document, la matrice de covariance des SIRV simulés est choisie d'après un modèle de densité spectrale Gaussienne dont les caractéristiques sont les suivantes :

- Taille de la matrice : $m \times m$, où $m = 10$,
- Période de répétition des impulsions : $T_r = 10^{-3}$ (le spectre s'étale donc de 0 à $1/T_r$),
- Puissance de bruit unitaire : $P_c = 1$,
- Le spectre est centré en $f_0 = 0$,

- L'étalement spectral est : $\sigma_f = \frac{\sigma_0}{2mT_r}$, où la valeur de σ_0 est fixée à 5 ou 10, selon la "force" de la corrélation voulue. Plus σ_0 est faible, plus la corrélation entre les données est forte, et inversement,
- Les coefficients d'auto-corrélation ρ_k ($k = 0, \dots, m-1$) sont donnés par la première ligne (ou la première colonne) de la matrice ainsi construite.

La figure (FIG. 3.1) représente différents vecteurs de coefficients d'auto-corrélation, selon les valeurs de σ_0 .

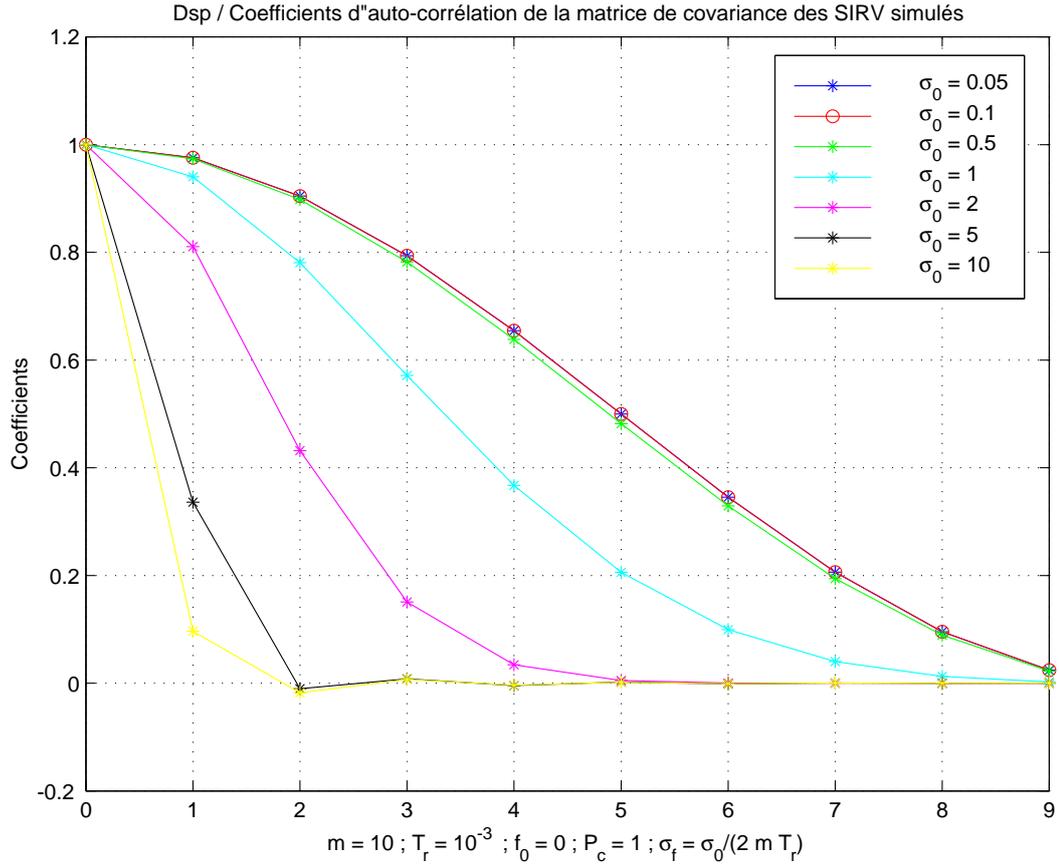


FIG. 3.1 – Coefficients d'auto-corrélation des matrices de covariance des SIRV simulés, représentés pour différentes valeurs de σ_0 (l'étalement spectral est $\sigma_f = \sigma_0 / (2mT_r)$). $T_r = 10^{-3}$; $P_c = 1$; $f_0 = 0$.

3.2.2 Propriété statistique de la matrice estimée au sens NSCM

On peut noter que l'expression de $\widehat{\mathbf{M}}$ donnée par (3.12) est statistiquement équivalente à :

$$\widehat{\mathbf{M}}_x = \frac{m}{N_r} \sum_{i=1}^{N_r} \frac{\mathbf{x}_i \mathbf{x}_i^\dagger}{\mathbf{x}_i^\dagger \mathbf{x}_i}, \quad (3.13)$$

où les vecteurs \mathbf{x}_i sont les vecteurs Gaussiens qui composent les données de référence. De cette manière, l'estimée NSCM de la matrice de covariance des données de référence ne dépend pas

de la densité de probabilité de la texture. On peut ainsi en conclure que l'estimation NSCM de \mathbf{M} selon (3.12) n'est pas un obstacle à la propriété TFAC d'un détecteur par rapport à la loi de la texture.

Pour illustrer cette propriété, la figure (FIG. 3.2) représente les coefficients d'auto-corrélation des matrices estimées au sens NSCM de SIRV K-distribué et de Student (pour $m = 10$ et $N = 1000$). A chaque SIRV correspond deux estimations : l'une à partir des données SIRV (selon (3.12)) et l'autre à partir du *speckle* du SIRV (selon (3.13)). On s'aperçoit que les estimations sont très proches ce qui justifie l'expression (3.13).

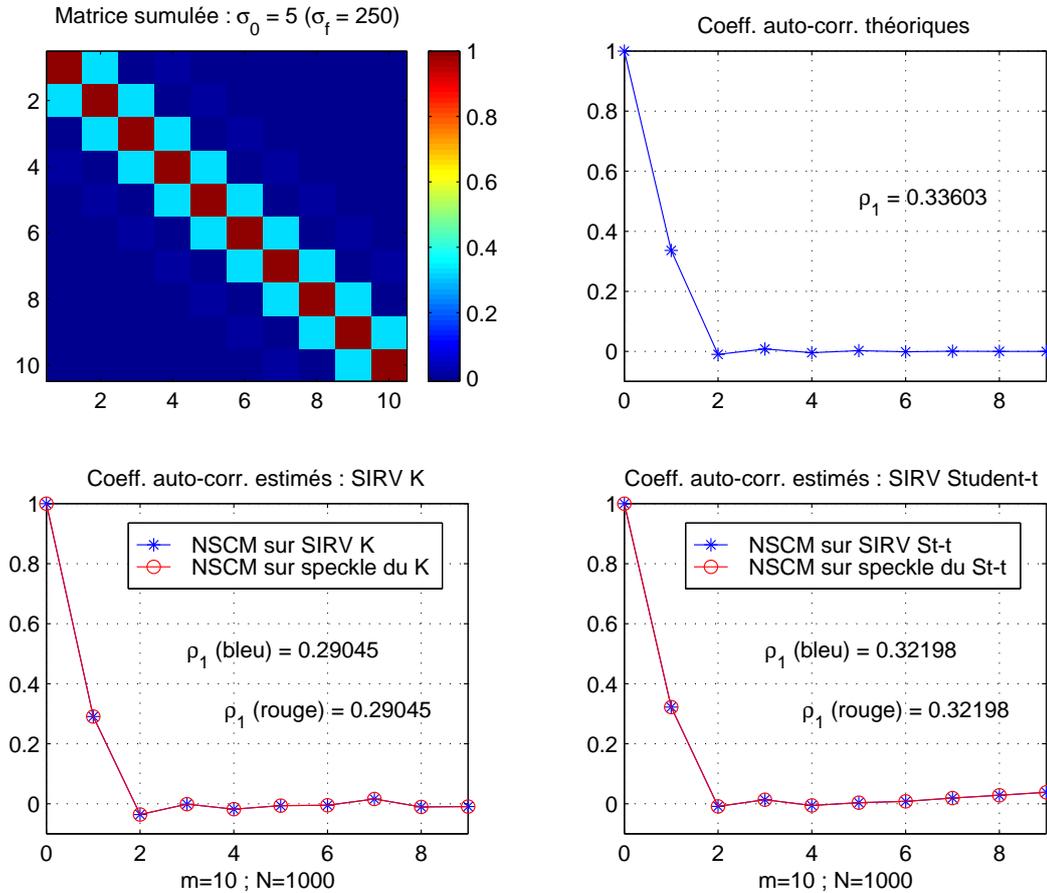


FIG. 3.2 – Indépendance des estimées NSCM par rapport à la loi de la texture du SIRV, montrée pour un SIRV K-distribué ($\nu = 0.1$) et un SIRV Student-t ($\nu = 2.24$), $m = 10$ et $N = 1000$.

Après avoir caractérisé l'environnement du radar par estimation de sa matrice de covariance et de sa loi de texture, à partir de données de référence, les traitements effectués sur les données observées deviennent adaptatifs. Toutes les quantités présentes dans l'expression du RV donné par (2.72) pour tout SIRV se déterminent à partir des données reçues, sans aucun apport extérieur nécessaire.

Dans le paragraphe suivant, deux méthodes d'estimation de la loi de la texture sont présentées. Ces deux méthodes, que sont l'approximation de Padé et une estimation MC-bayésienne,

ont l'avantage de fournir analytiquement une expression de la loi estimée. Ainsi, deux stratégies de détection peuvent être mises en place, après calcul des intégrales dans (2.72), nommées respectivement PEOD (Padé Estimated Optimum Detector) et BORD (Bayesian Optimum Radar Detector).

3.3 Le PEOD : Padé Estimated Optimum (radar) Detector

L'expression du détecteur PEOD est obtenue après estimation de la loi de la *texture* du SIRV, à partir de données de référence bruit seul. Dans un premier temps, les moments de la loi sont calculés empiriquement, en utilisant des échantillons de *texture* générés suivant la densité de probabilité *a posteriori* instantanée de τ donnée par (3.7). Une approximation de Padé de la fonction génératrice de moments permet ensuite d'obtenir une expression de l'estimée de la loi de la *texture* et de mettre en place le détecteur après formation du RV.

3.3.1 Estimation de la loi de la *texture* du SIRV par les approximants de Padé

L'étude bayésienne du modèle SIRV, menée en 3.1, a permis d'obtenir une expression de la loi *a posteriori* instantanée de la *texture*, connaissant N_r vecteurs de référence bruit. Cette loi étant identifiée comme une loi Inverse Gamma, de nouveaux échantillons de *texture* peuvent être générés, à partir desquels une estimation empirique des moments est possible. Ce "re"-tirage est nécessaire pour appliquer la méthode des approximants de Padé, basée sur une approximation de la fonction génératrice de moments (voir l'annexe C pour plus de détails).

En utilisant les N_r données de référence \mathbf{r}_i , N_r variables aléatoires $\tilde{\tau}$ sont générées de telle sorte que :

$$\tilde{\tau}_{i=1}^{N_r} \sim \mathcal{IG} \left(\tilde{\tau}; m, \frac{2}{\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i} \right), \quad (3.14)$$

en rappelant qu'une loi Inverse Gamma $\mathcal{IG}(x; a, b)$ représente la loi de $X = 1/Y$ avec Y qui suit une loi Gamma $\mathcal{G}(a, b)$.

Pour procéder à une approximation de Padé d'ordre $[L, M]$ de la fonction génératrice de moments, représentative des N_r échantillons $\tilde{\tau}$, $2M$ moments d'ordre supérieur (voir annexe C) sont calculés empiriquement comme suit :

$$\tilde{\mu}_n = \frac{1}{N_r} \sum_{i=1}^{N_r} \tilde{\tau}_i^n. \quad (3.15)$$

(L'expression littérale des moments des variables aléatoires $\tilde{\tau}$ ne peut pas s'obtenir d'après la loi des variables aléatoires, car cette dernière dépend des données.)

L'expression obtenue, notée $P^{[L/M]}(u)$, est une fraction rationnelle, dont les polynômes du numérateur et du dénominateur sont respectivement d'ordre L et M . La détermination des coefficients de ces polynômes constitue l'approximation de Padé de la fonction génératrice de moments. L'expression de la densité de probabilité des échantillons $\tilde{\tau}$ s'obtient après inversion de $P^{[L/M]}(u)$ par la méthode des résidus, qui résulte en deux séries de coefficients de taille M , α_k et λ_k , complexes, conjugués deux à deux et/ou réels.

La densité de probabilité calculée à partir de ces coefficients est définie sur \mathbb{R}_+ et admet pour expression :

$$\hat{p}(\tilde{\tau}) = \sum_{k=1}^M \lambda_k e^{-\alpha_k \tilde{\tau}}. \quad (3.16)$$

La détermination de l'ordre et les différentes étapes de l'approximation sont discutées dans l'annexe B aux paragraphes B.3.1 et B.3.5. Cependant, la valeur de M admise pour vérifier les conditions probabilistes et algébriques imposées par l'approximation de Padé n'est généralement pas très élevée. Dans la pratique, une approximation de Padé d'ordre de 4 à 8 est toujours possible.

La forme de densité de probabilité estimée après approximation de Padé (une somme d'exponentielles décroissantes) assure un certain confort dans les calculs. Remplacer $p(\tau)$ par (3.16) dans l'expression du RV donné par (2.72) permet effectivement d'obtenir une expression du RV pour tout type d'environnement du radar. L'estimation est effectuée à partir de données de référence sans nécessiter de connaissance *a priori* sur la statistique de l'environnement. Le détecteur ainsi construit est appelé PEOD pour Padé Estimated Optimum (radar) Detector et son expression est donnée au paragraphe suivant.

3.3.2 Expression du PEOD après formation du rapport de vraisemblance généralisé

Après estimation de la statistique de l'environnement du radar par une approximation de Padé et à partir de N_r données de référence bruit seul, les densités de probabilité des vecteurs observés \mathbf{y} (même taille m que \mathbf{r}_i) sous chacune des hypothèses du test de détection se calculent par intégration sur la loi estimée (3.16).

À l'hypothèse H_j , où $j = 0$ ou 1 , correspond $p(\mathbf{y}/H_j)$ donnée par :

$$\begin{aligned} p(\mathbf{y}/H_j) &= \int_0^{+\infty} \frac{\tilde{\tau}^{-m}}{(2\pi)^m |\mathbf{M}|} \exp\left(-\frac{q_j(\mathbf{y})}{2\tilde{\tau}}\right) \sum_{k=1}^M \lambda_k e^{-\alpha_k \tilde{\tau}} d\tilde{\tau} \\ &= \frac{1}{(2\pi)^m |\mathbf{M}|} \sum_{k=1}^M \lambda_k \exp\left(-\frac{q_j(\mathbf{y})}{2\tilde{\tau}} - \alpha_k \tilde{\tau}\right) d\tilde{\tau} \\ &= \frac{q_j(\mathbf{y})^{\frac{1-m}{2}}}{(2\pi)^m |\mathbf{M}|} \sum_{k=1}^M \lambda_k (\alpha_k)^{\frac{m-1}{2}} K_{1-m}\left(\sqrt{B_k^j(\mathbf{y})}\right), \end{aligned} \quad (3.17)$$

$$(3.18)$$

avec

- * $q_0(\mathbf{y}) = \mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y}$,
- * $q_1(\mathbf{y}) = q_0(\mathbf{y} - \mathbf{s})$ pour un signal cible \mathbf{s} connu,
- * $q_1(\mathbf{y}) = \mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y} - \frac{|\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{y}|^2}{\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}}$ pour un signal cible dont l'amplitude est inconnue et estimée au sens du MV selon (1.15),
- * $B_k^j(\mathbf{y}) = 2 \alpha_k q_j(\mathbf{y})$, pour $j = 0$ ou 1 ,
- * $K_\nu(x)$ est la fonction de Bessel modifiée de deuxième espèce d'ordre ν .

Le RV qui découle de ces expressions est immédiat et l'expression du PEOD devient :

$$\left(\frac{q_1(\mathbf{y})}{q_0(\mathbf{y})} \right)^{\frac{1-m}{2}} \frac{\sum_{k=1}^M \lambda_k (\alpha_k)^{\frac{m-1}{2}} K_{1-m} \left(\sqrt{B_k^1(\mathbf{y})} \right)}{\sum_{k=1}^M \lambda_k (\alpha_k)^{\frac{m-1}{2}} K_{1-m} \left(\sqrt{B_k^0(\mathbf{y})} \right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (3.19)$$

3.3.3 Avantages et inconvénients de la structure du PEOD

Pour la mise en œuvre du PEOD, la statistique de l'environnement est estimée au préalable sur données de référence. Elle est transcrite au travers des coefficients de Padé et seules les observations changent.

Cependant, le PEOD étant construit à partir d'une approximation de Padé de la loi de la *texture* du SIRV, et par conséquent à partir des moments de cette loi, quelques remarques peuvent être faites selon le degré de connaissance de la statistique de l'environnement du radar :

- Si la loi de la *texture* du SIRV est connue (les moments théoriques sont alors connus) sans toutefois permettre le calcul de la loi jointe du vecteur SIRV, alors le PEOD est "idéal". Ce cas peut se produire, soit comme un cas d'école, soit après identification de la loi de la *texture* à partir de données expérimentales de fouillis. L'erreur engendrée par l'approximation de Padé ne provient alors que de l'erreur d'arrondi dans le développement en puissance croissante à l'ordre $2M - 1$ de la fonction génératrice de moments (voir annexe B, (B.17)).
- Dans le cas le plus général où la loi de la *texture* est inconnue, les moments de cette loi sont estimés empiriquement. L'erreur engendrée par cette première estimation est difficilement quantifiable analytiquement sur l'approximation de Padé résultante ; cependant, l'étude peut être menée numériquement à travers l'évaluation des performances du PEOD. On constate alors, que si la valeur de la probabilité de fausse alarme est respectée, l'estimation des moments n'influe que très peu sur les performances du PEOD.

Pour illustrer la première remarque, considérons un cas bien précis. Supposons que la loi de la variance est une loi de Weibull de paramètres a et b . Les intégrales à calculer pour déterminer les statistiques de l'environnement sous l'hypothèse H_j ($j = 0, 1$) du test de détection s'écrivent :

$$p(\mathbf{y}/H_j) = \int_0^{+\infty} \frac{1}{(2\pi)^m |M|} \tau^{-m} \exp\left(-\frac{q_j(\mathbf{y})}{2\tau}\right) a b \tau^{b-1} e^{-a\tau^b} d\tau. \quad (3.20)$$

Cette intégrale n'admet pas de solution analytique connue. En revanche, les moments théoriques de la loi de Weibull étant :

$$\mu_n = a^{-n/b} \Gamma\left(1 + \frac{n}{b}\right), \quad (3.21)$$

une approximation de Padé à l'ordre [10/11], pour $b = 0.658$ et a fixé pour avoir une variance unité ($a = 1.8328..$), donne les coefficients réels suivants :

$$\{\alpha_k\}_{k=1..11} = \{38.596, 5.713, 2.288, 1.268, 0.825, 0.589, 0.448, \dots, 0.354, 0.287, 0.236, 0.195\} \quad (3.22)$$

$$\{\lambda_k\}_{k=1..11} = \{5.803, 1.313, 0.585, 0.261, 0.092, 0.022, 0.003, \dots, 2.520 \cdot 10^{-4}, 8.558 \cdot 10^{-6}, 9.257 \cdot 10^{-8}, 1.463 \cdot 10^{-10}\}. \quad (3.23)$$

Pour obtenir une telle approximation, les 22 premiers moments ont été utilisés. Cependant, l'ordre [10/11] n'est qu'un exemple et, dans ce cas, tous les ordres inférieurs sont également valables ; le nombre de moments utilisés est fonction de l'ordre obtenu.

Pour des valeurs de a et b différentes, et pour un ordre différent, les coefficients obtenus peuvent être complexes. Dans l'exemple ci-dessus, l'ordre [10/11] est l'ordre le plus grand pour lequel l'approximation est stable. Pour les mêmes paramètres, on peut obtenir des approximations aux ordres [9/10], [8/9], etc ... jusqu'à, raisonnablement, [1/2]. La figure (FIG. 3.3) représente les tracés des erreurs relatives entre les densités de probabilité théoriques et approximées pour différentes valeurs des paramètres a et b . Les ordres des approximations sont indiqués pour chaque estimée. On remarque nettement la convergence de l'approximation lorsque $b \rightarrow 1$, ce qui n'est pas surprenant car, lorsque $b = 1$, la loi de Weibull est une loi exponentielle. En revanche, lorsque b devient grand, l'erreur est inconsidérée ; il devient alors délicat d'estimer une loi "en $x^{b-1}e^{-ax^b}$ " avec b grand par une somme d'exponentielles, et la même conclusion est faite lorsque b est très petit.

Dans le cas où les moments théoriques ne sont pas disponibles, la qualité de l'approximation de Padé va dépendre de la qualité de l'estimation des moments, dépendante du nombre de données de référence régénérées. Pour illustrer cette remarque, les figures (FIG. 3.4), (FIG. 3.5) et (FIG. 3.6) représentent, en fonction du nombre de données, les histogrammes des échantillons de *texture* - théoriques et après retirages - et les moments de ces variables - moments théoriques, moments estimés d'après *texture* théorique et moments estimés d'après échantillons régénérés. Ces trois figures portent sur des SIRV K-distribués de paramètres $\nu = 0.1, 2$ et $\nu = 20$ et de matrice de covariance connue pour des valeurs $N_{ref} = 10000, 1000, 100$.

3.3.4 Performances de détection du PEOD

Les performances de détection du PEOD sont évaluées à la fois pour un environnement K-distribué (pour lequel l'OKD est optimal) que pour un environnement de statistique inconnue, pour lequel la loi de la *texture* est une loi de Weibull. On considère, dans les simulations, que la matrice de covariance est connue et que l'amplitude de la cible détectée est inconnue et estimée au sens du MV.

3.3.4.1 Comparaison à l'optimal

Les performances de différents détecteurs ci-dessous sont comparées :

- Performances du PEOD "idéal", (FIG. 3.7).

Le terme "idéal" signifie que les moments théoriques sont utilisés.

Dans ce cas, le PEOD atteint les performances idéales attendues pour ce détecteur, soit

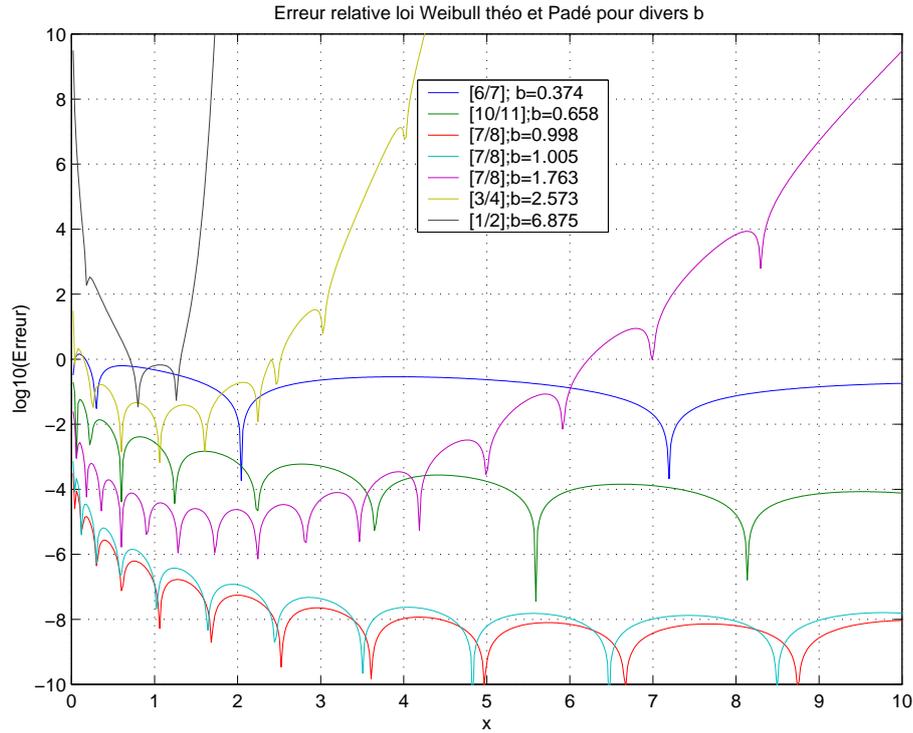


FIG. 3.3 – Erreur relative entre la loi théorique et son approximation par Padé pour différents paramètres.

une borne supérieure à celles obtenues sous des hypothèses moins favorables. L’optimalité du PEOD peut être évaluée et comparée à l’optimalité effective de l’OKD dans pareil cas. Les performances de l’OGD sont également tracées (détecteur non adapté) en vue d’une comparaison lorsque ν est grand (convergence vers le cas Gaussien).

- Performances du PEOD "non-idéal", (FIG. 3.8).

Dans ce cas, les moments de la loi de la *texture* sont estimés empiriquement à partir des retirages des échantillons de la *texture* selon (3.14).

Le PEOD est sous-optimal et cette étude permet d’évaluer l’influence de l’erreur d’estimation des moments sur ses performances, comparées à la fois à celles de l’optimal, l’OKD, et à celles du PEOD "idéal".

- Performances de l’OKD. Courbes optimales.

- Performances de l’OGD.

On constate une nette dégradation de ses performances dans un environnement non-Gaussien. En revanche, lorsque ν devient grand, la convergence des performances du PEOD et de l’OKD vers celles de l’OGD confirme bien la convergence de la K-distribution vers la Gaussienne.

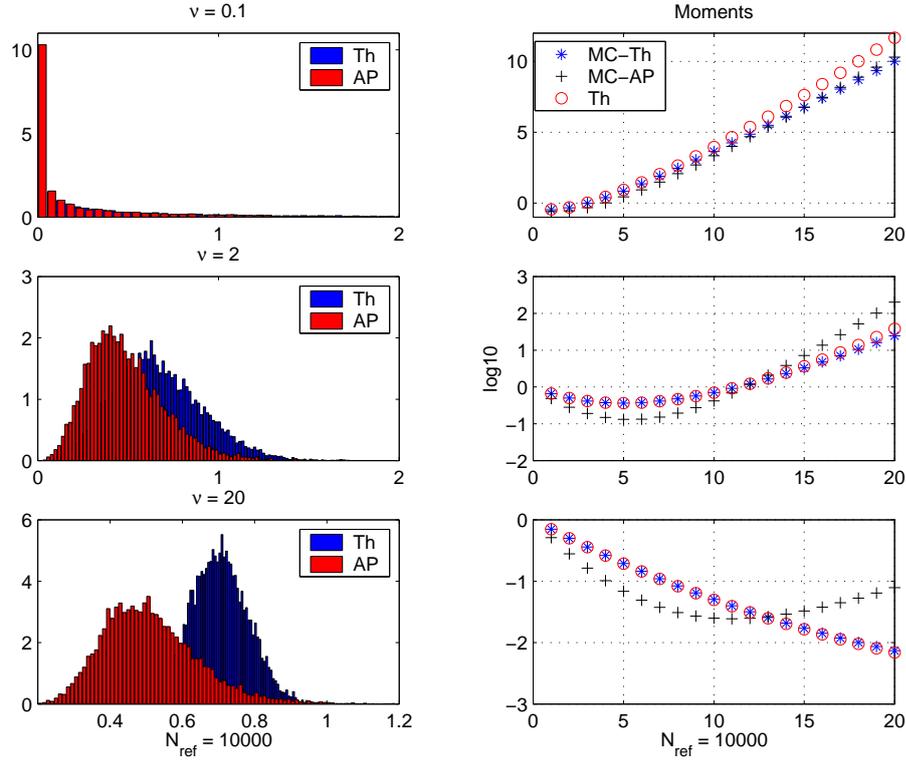


FIG. 3.4 – Histogrammes et moments des variables de *texture* : *texture* théorique (en bleu), *texture a posteriori* (en rouge) ; moments théoriques (“Th”, ronds rouges), moments empiriques sur *texture* théorique (“MC-Th”, en bleu) et moments empiriques sur *texture a posteriori* (“MC-AP”, en noir) pour un SIRV K-distribué : $\nu = 0.1, 2, 20$; $N_{ref} = 10000$.

3.3.4.2 Environnement inconnu

Les performances du PEOD “idéal” et du PEOD “non-idéal” sont évaluées et comparées dans l’environnement inconnu décrit en (3.20), pour lequel la *texture* suit une loi de Weibull. Les résultats sont montrés à la figure (FIG. 3.9). Les performances du PEOD sont comparées à celles de l’OKD, dans l’hypothèse où l’environnement serait considéré comme K-distribué, de paramètres $\nu = 0.1, 0.5, 2$, et comparées également à celles de l’OGD.

3.3.4.3 Contexte des simulations

Les différentes étapes des simulations sont décrites ci-dessous, en fonction du SIRV généré. Les analyses de performances du PEOD sont effectuées soit dans un environnement K-distribué, soit dans un environnement inconnu tel un SIRV dont la *texture* suit une loi de Weibull.

Simulation de N_r échantillons K-distribués de taille m , de paramètres ν et b_K :

- Simuler N_r vecteurs aléatoires Gaussiens complexes de taille m , de matrice de covariance \mathbf{M} (taille $m \times m$), de variance $\sigma^2 = 1$,
- Simuler N_r variables aléatoires de loi Gamma, $\mathcal{G}(\nu, \sigma^2/2\nu)$ ou $\mathcal{G}(\nu, 2/b_K^2)$, indépendantes,
- Multiplier chaque vecteur Gaussien par chaque réalisation de la loi Gamma,

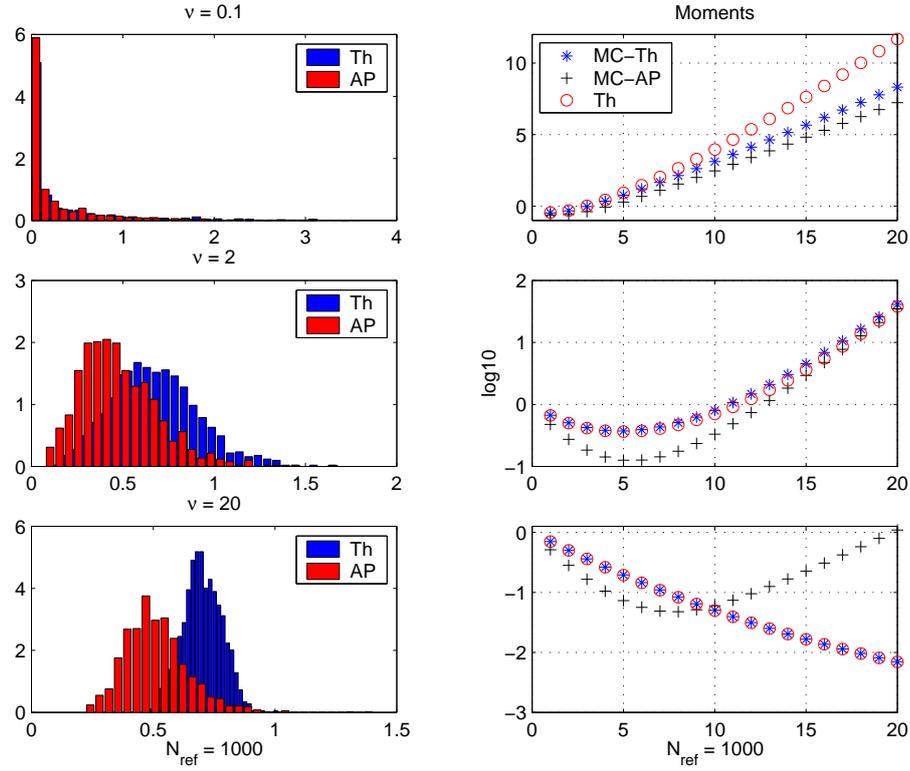


FIG. 3.5 – Histogrammes et moments des variables de *texture* : *texture* théorique (en bleu), *texture a posteriori* (en rouge); moments théoriques ("Th", ronds rouges), moments empiriques sur *texture* théorique ("MC-Th", en bleu) et moments empiriques sur *texture a posteriori* ("MC-AP", en noir) pour un SIRV K-distribué : $\nu = 0.1, 2, 20$; $N_{ref} = 1000$.

- L'enveloppe de chaque vecteur (taille m) suit une loi $K(\nu, b_K)$.

Simulation de N_r SIRV de taille m dont la loi de la texture est une loi de Weibull(a, b) :

- Simuler N_r vecteurs aléatoires Gaussiens complexes de taille m , de matrice de covariance \mathbf{M} (taille $m \times m$), de variance $\sigma^2 = 1$,
- Simuler N_r variables aléatoires de loi de Weibull, $\mathcal{W}(a, b)$, indépendantes,
- Multiplier chaque vecteur Gaussien par chaque réalisation de la loi de Weibull,
- La loi de chaque vecteur résultant n'est pas identifiée.

Calcul des coefficients de l'approximation de Padé :

- Calculer $2M$ moments théoriques de la loi Gamma,
- Calculer les deux séries α_k et λ_k de M coefficients de Padé à partir des moments théoriques.

Calcul du seuil de détection pour une $P_{fa} = 10^{-3}$:

- Prendre N_r au moins égal à 10^6 et générer autant de SIRV K-distribué,

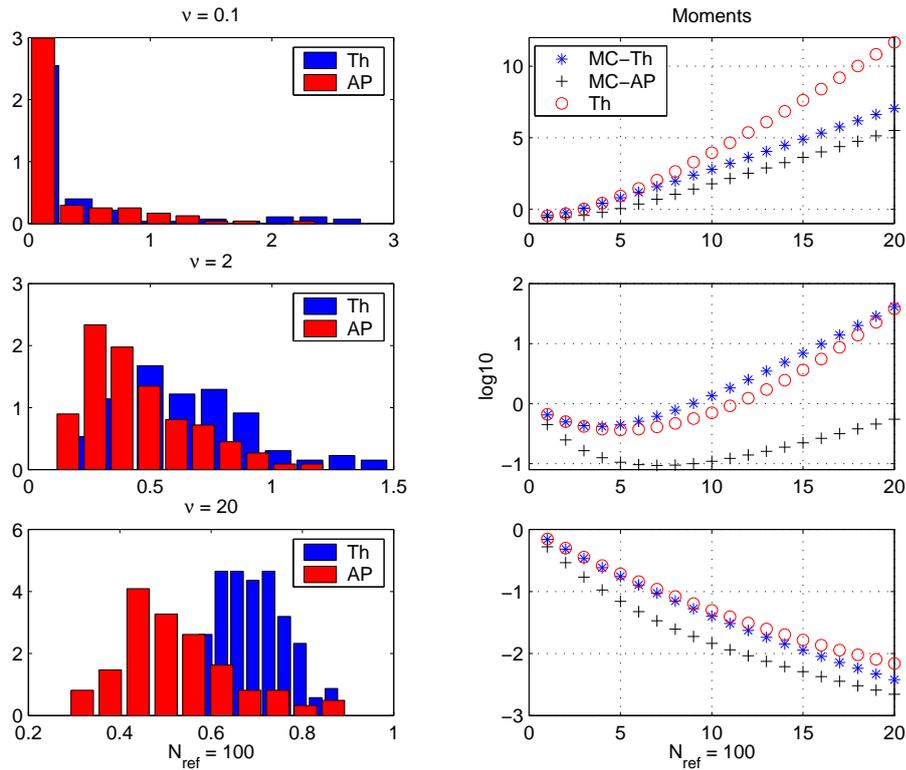


FIG. 3.6 – Histogrammes et moments des variables de *texture* : *texture* théorique (en bleu), *texture a posteriori* (en rouge) ; moments théoriques ("Th", ronds rouges), moments empiriques sur *texture* théorique ("MC-Th", en bleu) et moments empiriques sur *texture a posteriori* ("MC-AP", en noir) pour un SIRV K-distribué : $\nu = 0.1, 2, 20$; $N_{ref} = 100$.

- Par Monte Carlo, évaluer le seuil pour vérifier le taux de fausses alarmes fixé.

Remarque : La procédure de calcul du seuil de détection peut être itérée plusieurs fois, de manière à préciser le comptage (par Monte-Carlo, la convergence est asymptotique). Ainsi, plutôt que de générer 10^6 échantillons de bruit, il est possible de n'en générer que 10^4 et de procéder à 100 itérations pour le calcul du seuil.

Évaluation des performances de détection du détecteur :

- Générer N_{obs} observations K-distribuées,
- Ajouter une cible virtuelle dans les observations et faire varier le rapport signal à bruit (RSB),
- Effectuer un comptage Monte-Carlo pour évaluer P_d selon le RSB.

La procédure est identique pour l'évaluation des performances de détection de l'OKD ou de l'OGD, sans l'étape du calcul des coefficients de Padé.

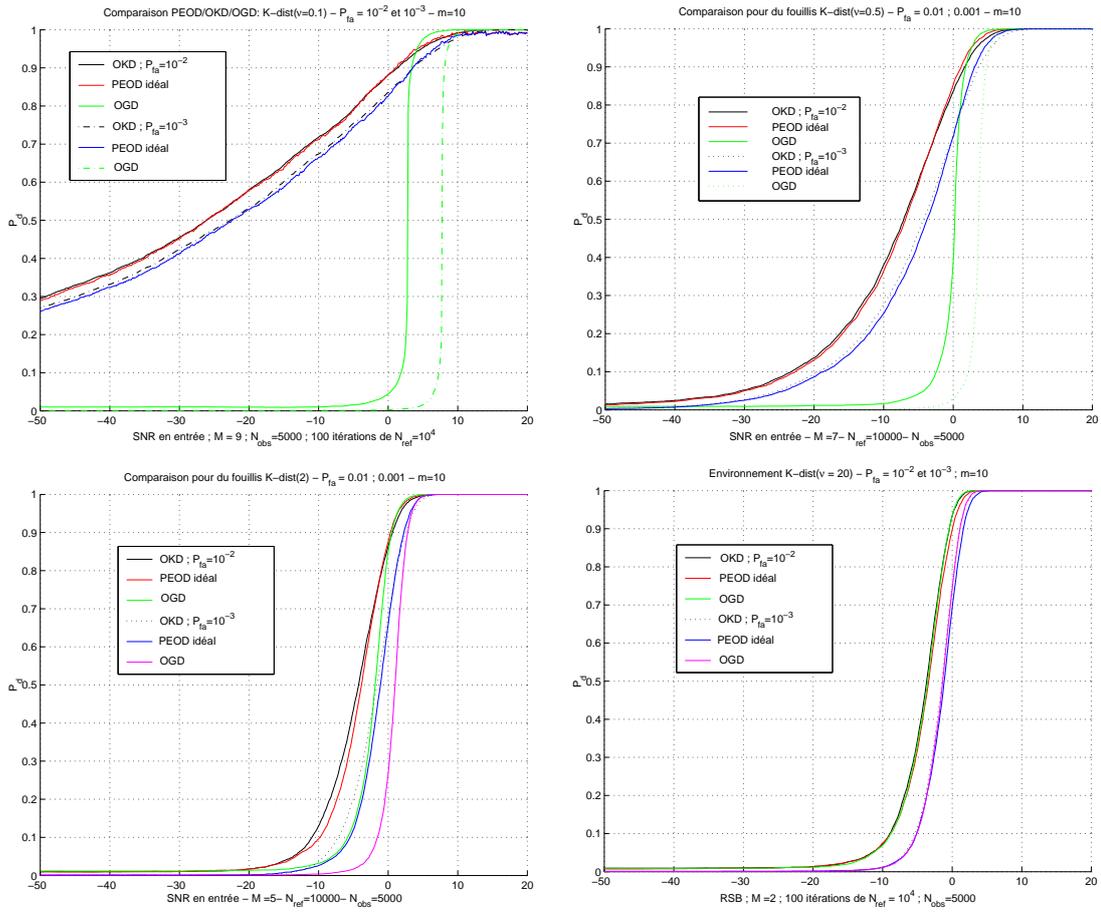


FIG. 3.7 – Performance de détection du PEOD "idéal" pour une cible d'amplitude inconnue dans un environnement K-distribué ($\nu = 0.1; 0.5; 2; 20$). Comparaison avec le détecteur optimal OKD et l'OGD. $P_{fa} = 10^{-2}, 10^{-3}$, $m = 10$, $M_{\nu=0.1} = 9$, $M_{\nu=0.5} = 7$, $M_{\nu=2} = 5$, $M_{\nu=20} = 2$.

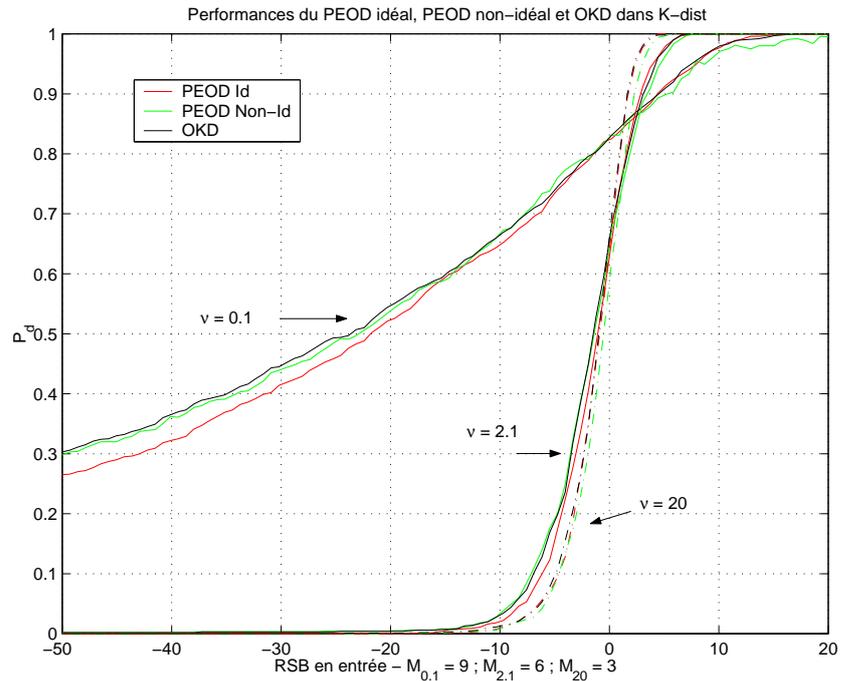


FIG. 3.8 – Performance de détection du PEOD "non-idéal" pour une cible d'amplitude inconnue dans un environnement K-distribué ($\nu = 0.1; 0.5; 2; 20$). Comparaison avec le détecteur optimal OKD et le PEOD "idéal". $P_{fa} = 10^{-3}$, $m = 10$, $M_{\nu=0.1} = 9$, $M_{\nu=2.1} = 6$, $M_{\nu=20} = 3$.

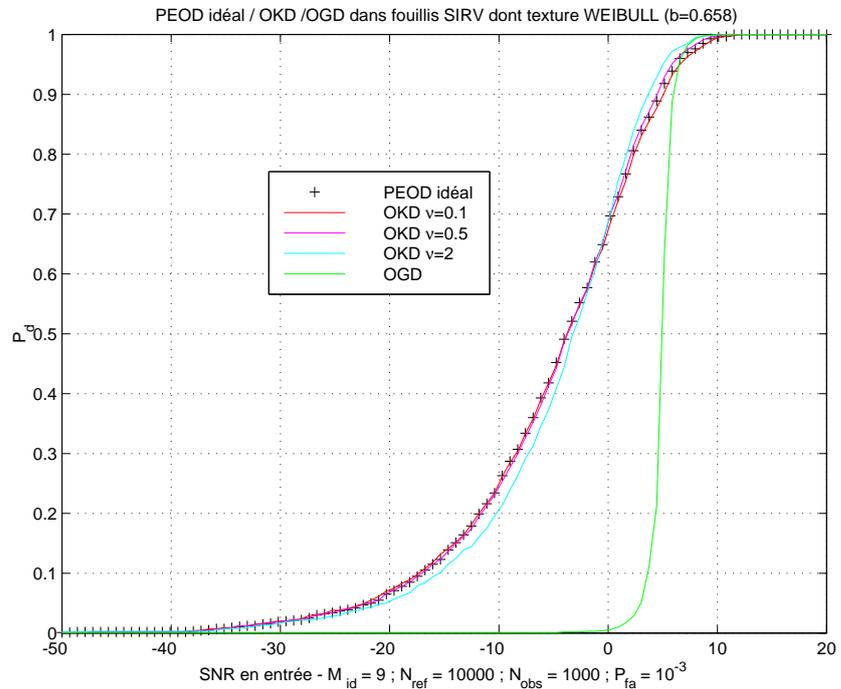


FIG. 3.9 – Performance de détection du PEOD "idéal" pour une cible d'amplitude inconnue dans un environnement inconnu : la texture suit une loi de Weibull de paramètre $b = 0.658$ et de puissance unité. Comparaison avec l'OKD pour $\nu = 0.1, 0.5, 2$ et l'OGD. $P_{fa} = 10^{-3}$, $m = 10$, $M_{id} = 9$.

3.3.5 Synthèse sur le PEOD

3.3.5.1 Le PEOD "idéal"

Le principal avantage du PEOD "idéal" est montré au travers de la figure (FIG. 3.9). Ce détecteur s'adapte à la statistique inconnue de l'environnement, ce qu'aucun autre détecteur ne réalise, sans hypothèse supplémentaire. Les courbes de la figure (FIG. 3.7) pour lesquelles le PEOD "idéal" atteint les performances de l'optimal, l'OKD, permettent de penser que le PEOD "idéal" atteint également l'optimalité de tout autre détecteur, qu'il soit connu ou non.

En revanche, dans des environnements connus pour satisfaire une statistique de Cauchy ou de Student-t, le PEOD "idéal" ne peut pas être mis en place systématiquement. En effet, les moments théoriques de telles variables aléatoires n'existent pas (Cauchy), ou ne se calculent que pour des valeurs particulières des paramètres des lois (Student-t). Dans pareil cas, il est toujours possible d'estimer empiriquement les moments à partir des données générées pour la loi de la *texture* (pour rester proche de l'idéal) ou bien après retraitage des variables aléatoires selon (3.14).

3.3.5.2 Le PEOD "non-idéal"

Pour mettre en place le PEOD "non-idéal", le calcul des moments s'effectue empiriquement à partir des variables de *texture* régénérées. Or, la loi *a posteriori* est une loi Inverse Gamma de premier paramètre m , pour lequel seuls $m - 1$ premiers moments théoriques existent (cf. Annexe A). L'ordre de l'approximation ne dépasse donc pas $(m - 1)/2$, et il peut y avoir des cas où la méthode n'est pas valide. L'utilisation du PEOD est donc recommandée vivement dans sa configuration "idéale".

La section suivante présente un autre détecteur, le BORD pour Bayesian Optimum Radar Detector, nettement plus général que le PEOD. Son expression s'obtient après une estimation MC-bayésienne de la loi *a posteriori* de la *texture* et ne dépend plus que des données de référence et des observations.

3.4 Le BORD : Bayesian Optimum Radar Detector

La stratégie de détection, appelée BORD pour Bayesian Optimum Radar Detector, s'obtient après une estimation bayésienne de la loi de la *texture* du SIRV à partir de données de référence bruit seul.

3.4.1 Estimation bayésienne de la loi de la *texture* du SIRV

D'après la règle de Bayes et une estimation Monte-Carlo d'intégrale, un estimateur bayésien de la loi de la *texture* va être déterminé.

Pour ce faire, exprimons la loi de la *texture* sous la forme d'une intégrale sur l'ensemble des données de référence \mathbf{r}_i :

$$p(\tau) = \int_{\mathbb{R}^m} p(\tau/\mathbf{r}_i) p(\mathbf{r}_i) d\mathbf{r}_i. \quad (3.24)$$

Connaissant les N_r vecteurs de référence \mathbf{r}_i , une approximation Monte-Carlo de (3.24) s'écrit :

$$\hat{p}_{N_r}(\tau) = \frac{1}{N_r} \sum_{i=1}^{N_r} p(\tau/\mathbf{r}_i). \quad (3.25)$$

Asymptotiquement et sachant que les N_r échantillons de bruit de référence \mathbf{r}_i sont statistiquement indépendants, la loi forte des grands nombres s'applique pour assurer la convergence de l'estimée vers sa valeur. Cela s'écrit :

$$\lim_{N_r \rightarrow +\infty} \hat{p}_{N_r}(\tau) \xrightarrow{p.s.} p(\tau), \quad (3.26)$$

où *p.s.* signifie *presque sûrement*.

En utilisant l'expression de $p(\tau/\mathbf{r}_i)$ donnée en (3.1), on obtient un estimateur bayésien à la loi de la *texture* du SIRV. Son expression générale est donnée par :

$$\hat{p}_{N_r}(\tau) = \frac{1}{N_r} \sum_{i=1}^{N_r} \frac{p(\mathbf{r}_i/\tau) g(\tau)}{p(\mathbf{r}_i)}, \quad (3.27)$$

où $g(\tau)$ a été désigné pour être un prior non-informatif ou prior de Jeffrey (3.5). On rappelle que ce prior admet pour expression :

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

Le terme $p(\mathbf{r}_i)$ est la constante de normalisation donnée en (3.6). Elle est calculée à partir du prior $g(\tau)$, associé à la *texture* des vecteurs \mathbf{r}_i , et admet pour expression :

$$p(\mathbf{r}_i) = \frac{\Gamma(m)}{\pi^m |\mathbf{M}| (\mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i)^m}.$$

La loi *a posteriori* instantanée de la *texture* étant donnée par :

$$p(\tau/\mathbf{r}_i) = \frac{Z_i^m}{2^m \Gamma(m)} \tau^{-m-1} \exp\left(-\frac{Z_i}{2\tau}\right), \quad (3.28)$$

où $Z_i = \mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i$, en utilisant (3.25), l'estimateur bayésien de $p(\tau)$ devient :

$$\hat{p}_{N_r}(\tau) = \frac{\tau^{-m-1}}{2^m \Gamma(m) N_r} \sum_{i=1}^{N_r} Z_i^m \exp\left(-\frac{Z_i}{2\tau}\right). \quad (3.29)$$

L'estimateur ainsi déterminé à partir de données de référence permet d'obtenir une expression du RV pour des observations du vecteur \mathbf{y} de même statistique que les références bruit. L'expression du BORD est ainsi valable quel que soit l'environnement du radar.

3.4.2 Expression du détecteur BORD

L'expression du Bayesian Optimum Radar Detector ou BORD s'obtient après formation du rapport de vraisemblance, pour lequel les expressions des lois sous les hypothèses H_j , $j = 0$ ou 1 , résultent de l'intégration de la vraisemblance des observations \mathbf{y}_{obs}/H_j sur la loi estimée $\hat{p}_{N_r}(\tau)$. On a donc :

$$\begin{aligned} p(\mathbf{y}/H_j) &= A_m^{N_r} \sum_{i=1}^{N_r} Z_i^m \int_0^{+\infty} \tau^{-2m-1} \exp\left(-\frac{W_{i,j}(\mathbf{y}_{obs})}{2\tau}\right) d\tau \\ &= 2^{2m} A_m^{N_r} \Gamma(2m) \sum_{i=1}^{N_r} \frac{Z_i^m}{(W_{i,j}(\mathbf{y}_{obs}))^{2m}}, \end{aligned} \quad (3.30)$$

avec

- * $Z_i = \mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i$,
- * $(A_m^N)^{-1} = (2\pi)^m |\mathbf{M}| N 2^m \Gamma(m)$,
- * $W_{i,j}(\mathbf{y}_{obs}) = q_j(\mathbf{y}_{obs}) + Z_i$.

Un calcul direct nous fournit l'expression du BORD, donnée par :

$$\Lambda_{N_r}(\mathbf{y}_{obs}) = \frac{\sum_{i=1}^{N_r} \left[\frac{Z_i}{(q_1(\mathbf{y}_{obs}) + Z_i)^2} \right]^m}{\sum_{i=1}^{N_r} \left[\frac{Z_i}{(q_0(\mathbf{y}_{obs}) + Z_i)^2} \right]^m} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (3.31)$$

Les deux formes quadratiques q_0 et q_1 restent les mêmes qu'au chapitre précédent, à savoir :

- * $q_0(\mathbf{y}_{obs}) = \mathbf{y}_{obs}^\dagger \mathbf{M}^{-1} \mathbf{y}_{obs}$,
- * $q_1(\mathbf{y}_{obs}) = q_0(\mathbf{y}_{obs} - \mathbf{s})$ pour un signal cible \mathbf{s} connu, et
- * $q_1(\mathbf{y}_{obs}) = \mathbf{y}_{obs}^\dagger \mathbf{M}^{-1} \mathbf{y}_{obs} - \frac{|\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{y}_{obs}|^2}{\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}}$ pour un signal cible dont l'amplitude est inconnue et estimée au sens du MV.

L'expression du BORD ne dépend que des données reçues par le radar. Cette particularité est très intéressante d'un point de vue opérationnel car cela ne nécessite aucun calcul annexe de paramètres. La matrice de covariance s'estime à partir des données de référence selon son estimée NSCM, et en remplaçant la matrice \mathbf{M} par (3.12), le BORD est alors un détecteur auto-adaptatif.

3.4.3 Version adaptative du BORD

Afin d'employer les bons termes pour les bonnes expressions, la version "correcte" du BORD adaptatif est donnée par (3.31) où désormais :

$$* \widehat{\mathbf{M}} = \frac{m}{N_r} \sum_{i=1}^{N_r} \frac{\mathbf{r}_i \mathbf{r}_i^\dagger}{\mathbf{r}_i^\dagger \mathbf{r}_i}$$

$$* Z_i \text{ devient } \widehat{Z}_i = \mathbf{r}_i^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{r}_i,$$

$$* q_0(\mathbf{y}_{obs}) \text{ devient } \widehat{q}_0(\mathbf{y}_{obs}) = \mathbf{y}_{obs}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs},$$

$$* q_1(\mathbf{y}_{obs}) \text{ devient } \widehat{q}_1(\mathbf{y}_{obs}) = \widehat{q}_0(\mathbf{y}_{obs} - \mathbf{s}) \text{ pour un signal cible } \mathbf{s} \text{ connu, et}$$

$$* q_1(\mathbf{y}_{obs}) \text{ devient } \widehat{q}_1(\mathbf{y}_{obs}) = \mathbf{y}_{obs}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs} - \frac{|\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs}|^2}{\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{p}} \text{ pour un signal cible dont l'amplitude est inconnue et estimée au sens du MV.}$$

On obtient alors :

$$\widehat{\Lambda}_{N_r}(\mathbf{y}_{obs}) = \frac{\sum_{i=1}^{N_r} \left[\frac{\widehat{Z}_i}{(\widehat{q}_1(\mathbf{y}_{obs}) + \widehat{Z}_i)^2} \right]^m}{\sum_{i=1}^{N_r} \left[\frac{\widehat{Z}_i}{(\widehat{q}_0(\mathbf{y}_{obs}) + \widehat{Z}_i)^2} \right]^m} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (3.32)$$

En examinant les structures de ses deux versions dans le cas d'une cible inconnue, on remarque que le BORD vérifie la propriété TFAC par rapport à la loi de la *texture*. Cela signifie que sa statistique est indépendante de la loi de la *texture*; ainsi le calcul du seuil de détection sera le même quelle que soit la statistique de l'environnement SIRV.

3.4.4 Propriété TFAC du BORD

Pour vérifier que le BORD admet la propriété TFAC par rapport à la loi de la *texture*, différents environnements de fouillis ont été simulés (K-distribué, Gaussien, Student-t, SIRV dont la *texture* suit une loi de Weibull) pour lesquels on a calculé les couples de valeurs (η, P_{fa}) . Les courbes de la figure (3.10) illustrent ces valeurs : les courbes de gauche ont été obtenues pour des environnements K-distribués de paramètre de forme ν variant de 0.1 à 20, et les courbes de droite ont été obtenue pour des environnements de natures différentes, à savoir un K-distribué (pour $\nu = 0.1$), un Gaussien, un Student-t et un dernier tel que la *texture* suit une loi une

Weibull. D'après ces courbes, on constate que la statistique du test est bien indépendante de la loi du SIRV, ce qui montre que le BORD admet la propriété TFAC par rapport à la loi de la *texture*.

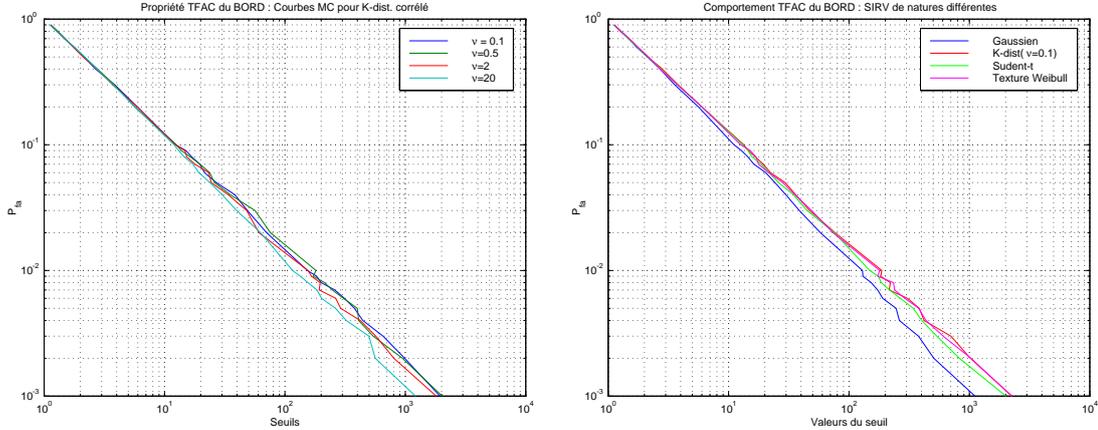


FIG. 3.10 – Propriété TFAC par rapport à la loi de la *texture* du BORD. Environnements K-distribués : $\nu = 0.1, 0.5, 2, 20$ pour les courbes de gauche. Environnements K-distribués ($\nu = 0.1$), Student-t, Gaussien et inconnu (la loi de la *texture* suit une loi de Weibull), pour les courbes de droite.

3.4.5 Performances de détection du BORD

L'évaluation des performances de détection du BORD est effectuée par Monte Carlo. Le but étant de comparer les performances du BORD aux performances de détecteurs optimaux existants, la valeur de la probabilité de fausse alarme est fixée à 10^{-3} , dans toutes les simulations. Pour des valeurs de P_{fa} plus faibles, l'optimalité relative du BORD reste vraie.

3.4.5.1 Contexte des simulations

Dans un premier temps, les performances de détection du BORD sont évaluées en environnement K-distribués corrélés. Ce choix nous paraît justifié pour diverses raisons :

- * La loi de la *texture* est une loi Gamma à partir de laquelle il est possible de générer des échantillons,
- * La K-distribution est la statistique la plus appropriée pour modéliser des données de fouillis de mer ou de terre,
- * Le détecteur optimal, l'OKD, est connu,
- * La K-distribution tend vers une Gaussienne lorsque son paramètre de forme devient grand.

De manière générale, les performances de détection du BORD sont comparées à celles de l'OKD et de l'OGD pour différentes valeurs du paramètre de la K-distribution, $\nu = 0.1, 0.5, 2, 20$. Le second paramètre de la loi K est fixée de telle sorte que la puissance σ^2 de l'enveloppe des données soit fixée. Ainsi $b = 2\sqrt{\frac{\nu}{\sigma^2}}$.

La deuxième série de courbes concerne des environnements SIRV, différents de la K-distribution. Pour l'environnement de Student-t, le détecteur optimal l'OStD est connu et ses performances

sont comparées à celles du BORD. Pour l'environnement "inconnu", pour lequel la loi de la *texture* est une loi de Weibull, les performances obtenues par le BORD ne peuvent pas être comparées aux performances optimales; en revanche, les performances du PEOD obtenues dans les mêmes conditions, sont reportées sur la figure.

La taille des vecteurs de référence et d'observation a été fixée arbitrairement à $m = 10$. Pour parvenir à une valeur empirique "acceptable" de la probabilité de fausse alarme, 100 itérations sont effectuées pour le calcul du seuil de détection avec un nombre de données de référence $N_r = 10^4$.

La matrice de covariance des données est choisie d'après le modèle décrit au paragraphe 3.2.1, pour $\sigma_0 = 10$. La figure de gauche de (FIG. 3.11) illustre une telle matrice, en indiquant (sur la droite), ses coefficients d'auto-corrélation, avec $\rho_1 = .$

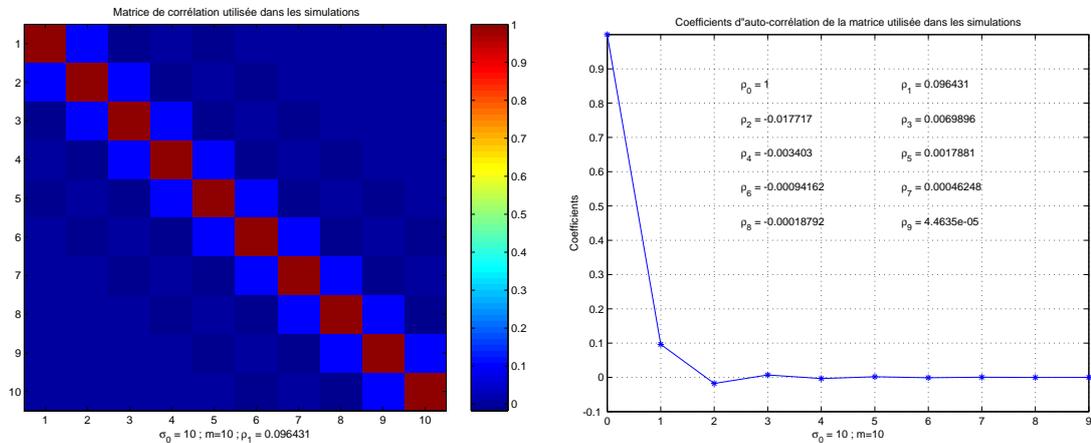


FIG. 3.11 – Caractéristiques de la matrice de covariance $\mathbf{M}(m, m)$ utilisée dans les simulations

La matrice de covariance est, soit supposée connue, soit estimée au sens NSCM. La cible introduite dans les simulations est d'amplitude inconnue estimée au sens du MV.

Toutes les courbes de performances représentent la probabilité de détection en fonction du rapport signal à bruit (RSB) en entrée (par impulsion), pour une probabilité de fausse alarme fixée.

3.4.5.2 Résultats des simulations

- * Sur la figure (FIG. 3.12), l'optimalité du BORD relative à l'optimalité de l'OKD peut être très nettement annoncée. Quelle que soit la valeur de ν , le BORD atteint l'optimal. De plus, pour une grande valeur de ν , la convergence de la K-distribution vers la Gaussienne se confirme à travers les courbes de performances. Ces résultats ont été obtenus connaissant la matrice de covariance et représentent donc les performances idéales du BORD,
- * Sur la figure (FIG. 3.13), les performances du BORD sont comparées aux performances de l'OStD dans un environnement de Student-t. Là encore, les performances du BORD sont optimales, pour une statistique différente de la K-distribution,

* Sur la figure (FIG. 3.14), les performances du BORD sont évaluées dans un environnement de statistique analytiquement inconnue : la *texture* suit une loi de Weibull, ce qui ne permet pas de construire le rapport de vraisemblance facilement.

Les courbes obtenues sont proches de l'optimal si on considère que le BORD est effectivement capable de donner les performances optimales quelle que soit la statistique de l'environnement. Aucun autre détecteur, mis à part le PEOD n'est capable de fournir de tels résultats dans des environnements de cette nature. Les performances du PEOD sont d'ailleurs reportées sur cette figure, et on constate que les deux détecteurs fournissent les mêmes performances.

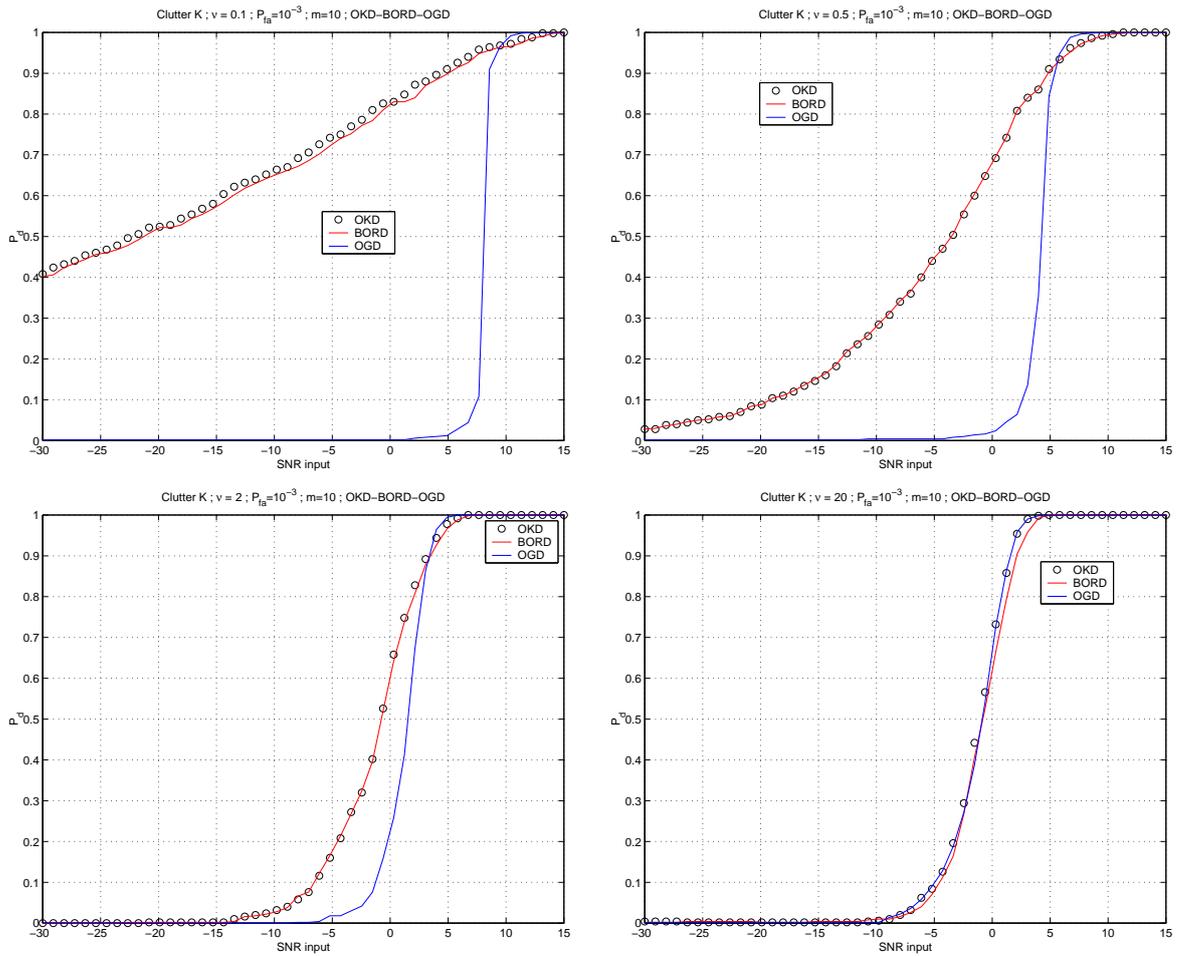


FIG. 3.12 – Performances de détection du BORD comparées aux performances de l’OGD et de l’OKD dans des environnements K-distribués ($\nu = 0.1, 0.5, 2, 20$), avec $m = 10$ et $N_{obs} = 5000$ et $P_{fa} = 10^{-3}$. La matrice de covariance est supposée connue.

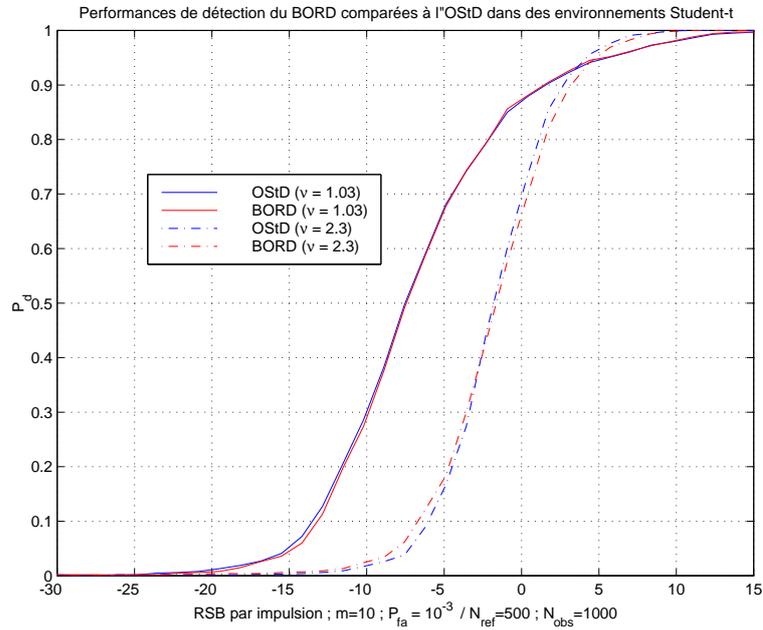


FIG. 3.13 – Performances de détection du BORD comparées aux performances de l’OSTd dans des environnements de Student-t ($\nu = 1.03, 2.3$), avec $m = 10$, $N_{obs} = 1000$, $N_{ref} = 500$ et $P_{fa} = 10^{-3}$. La matrice de covariance est supposée connue.

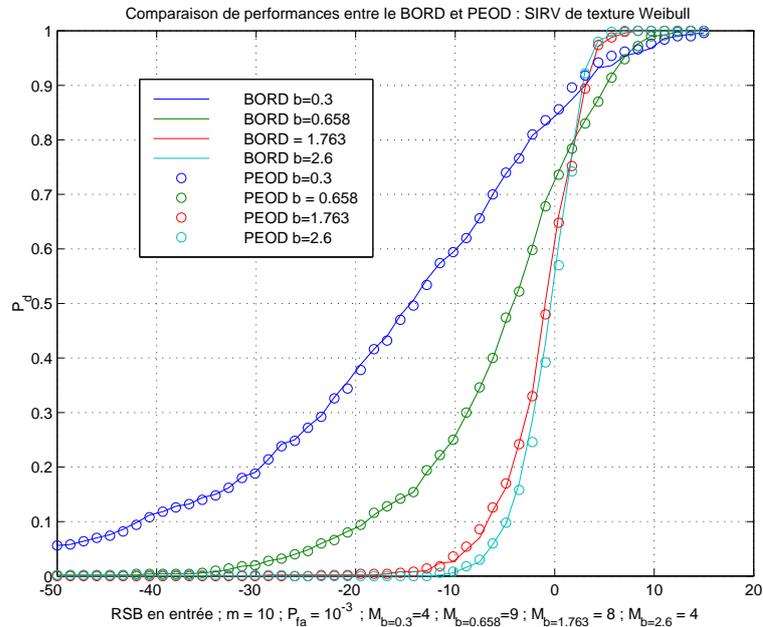


FIG. 3.14 – Performances de détection du BORD dans un environnement SIRV inconnu : la *texture* suit une loi de Weibull ($b = 0.3, 0.658, 1.763, 2.6$), avec $m = 10$, $N_{obs} = 500$, $N_{ref} = 500$, $P_{fa} = 10^{-3}$. Comparaison avec le PEOD. La matrice est supposée connue.

3.4.6 Robustesse du BORD

La robustesse du BORD s'évalue par une variation du nombre N_r de références utilisées pour "l'auto-estimation" de l'environnement. Le seuil de détection est calculé pour respecter une probabilité de fausse alarme fixée quelle que soit la valeur de N_r (seuil "idéal"). De cette manière, on constate que le BORD atteint quasiment tout le temps les performances de détection du détecteur optimal dans ce cas. Ce qui signifie que, dès que la probabilité de fausse alarme est respectée, le BORD reste proche de l'optimal quel que soit le nombre de données de référence utilisé.

La figure (FIG. 3.15) interprète la robustesse du BORD. Le seuil de détection a été calculé au préalable sur données de référence pour respecter une $P_{fa} = 10^{-3}$, et les performances du BORD sont évaluées dans un environnement K-distribué de paramètre $\nu = 0.5$ pour un nombre de données de référence variant de $N_{ref} = 10$ à $N_{ref} = 1000$. Ses performances tendent très rapidement vers les performances optimales de l'OKD. Dans un cas de mise en œuvre d'un détecteur TFAC avec le BORD, le nombre de cellules adjacentes à la cellule testée, prises comme cellules de référence, n'est pas nécessairement grand.

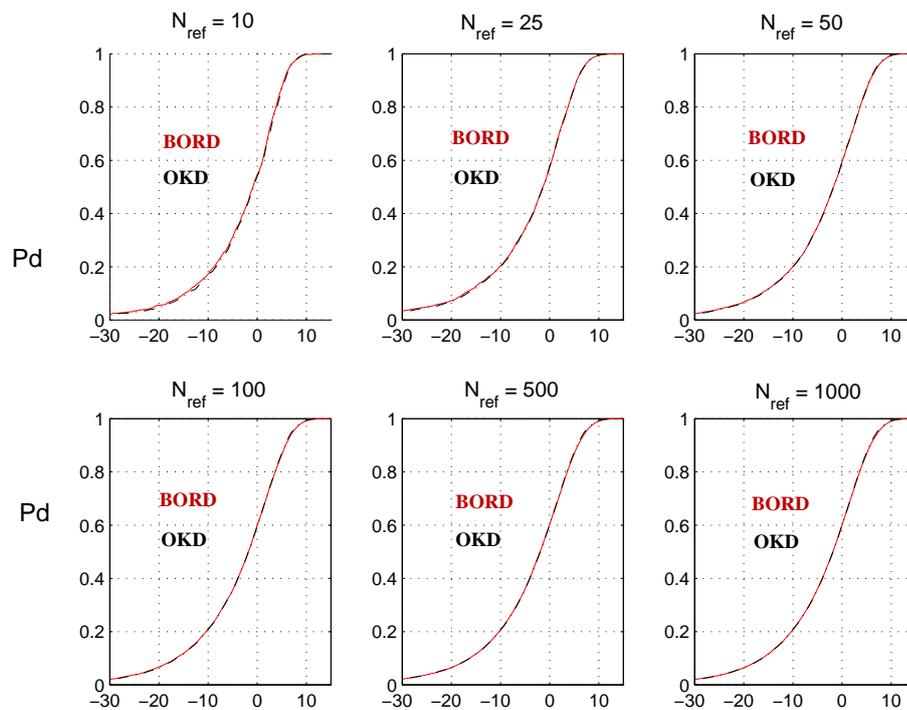


FIG. 3.15 – Robustesse du BORD (en rouge) au nombre de données de référence, évaluée en environnement K-distribué pour $\nu = 0.5$, $m = 10$. Comparaison avec l'OKD (en noir). La $P_{fa} = 10^{-3}$ est respectée. La matrice est supposée connue.

Pour un nombre fini de données de référence, il est difficile d'obtenir la statistique du BORD bien qu'il ne dépende que de formes quadratiques de vecteurs Gaussiens dont on connaît les lois. En revanche, on peut s'interroger sur sa statistique lorsque N_r devient grand, ce qui, statistiquement, revient à chercher la loi du BORD quand $N_r \rightarrow +\infty$. L'expression du BORD résultant

en partie d'une estimation Monte Carlo d'une intégrale, sa loi doit théoriquement converger vers la valeur de cette intégrale.

3.4.7 Comportement asymptotique du BORD

Sur des considérations probabilistes, et connaissant les N_r données de référence \mathbf{r}_i contenues dans chaque variable Z_i , l'expression du BORD (3.31) peut être considérée comme l'approximation Monte-Carlo de l'expression suivante :

$$\frac{\int_0^{+\infty} \frac{z^m}{(q_1(\mathbf{y}_{obs}) + z)^{2m}} p_Z(z) dz}{\int_0^{+\infty} \frac{z^m}{(q_0(\mathbf{y}_{obs}) + z)^{2m}} p_Z(z) dz}, \quad (3.33)$$

où Z est une variable aléatoire positive de loi $p_Z(z)$, dont chaque réalisation Z_i dépend des vecteurs de référence \mathbf{r}_i , tel que $Z_i = \mathbf{r}_i^\dagger \mathbf{M}^{-1} \mathbf{r}_i$.

La densité de probabilité de la forme quadratique Z a été déterminée au chapitre 2 comme étant la loi de la forme quadratique d'un SIRV complexe. Son expression est donnée par (2.8), soit :

$$p_Z(z) = \int_0^{+\infty} \frac{z^{m-1}}{2^m \Gamma(m)} \tau^{-m} \exp\left(-\frac{z}{2\tau}\right) p(\tau) d\tau.$$

La loi de Z représente la loi de la forme quadratique des données de référence sur lesquelles le prior $g(\tau) = 1/\tau$ a été choisi. La loi de Z est donc déterminée ici en intégrant non pas sur $p(\tau)$, a priori inconnue, mais sur le prior $g(\tau)$. On obtient alors :

$$p_Z(z) = \int_0^{+\infty} \frac{z^{m-1}}{2^m \Gamma(m)} \tau^{-m} \exp\left(-\frac{z}{2\tau}\right) g(\tau) d\tau = \frac{1}{Z}. \quad (3.34)$$

Le calcul de (3.33) s'effectue en intégrant le numérateur et le dénominateur par parties, ce qui donne, pour $j = 0, 1$:

$$\int_0^{+\infty} \frac{Z^m}{(q_j(\mathbf{y}_{obs}) + Z)^{2m}} \frac{1}{Z} dZ = \frac{[(m-1)!]^2}{(2m-1)!} \left(\frac{1}{q_j(\mathbf{y}_{obs})}\right)^m, \quad (3.35)$$

et (3.31) converge alors en loi vers (3.33), ce qui s'écrit :

$$\lim_{N_r \rightarrow +\infty} \Lambda_{N_r}(\mathbf{y}_{obs}) \stackrel{\mathcal{L}}{=} \left(\frac{q_0(\mathbf{y}_{obs})}{q_1(\mathbf{y}_{obs})}\right)^m. \quad (3.36)$$

L'expression du BORD Asymptotique coïncide avec l'expression de détecteurs obtenus par d'autres auteurs sous d'autres hypothèses. Par exemple, K. J. Sangston et al. dans [Sang99] 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, soit $\hat{\tau}_{j,mv} = \frac{q_j}{2m}$.

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 [Gini97a], où F.Gini en donne une expression similaire :

$$\frac{|\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{y}|^2}{(\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p})(\mathbf{y}^\dagger \mathbf{M}^{-1} \mathbf{y})} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sqrt[m]{\eta} - 1}{\sqrt[m]{\eta}}, \quad (3.37)$$

qui peut se réécrire en fonction de q_0 et de q_1 comme suit :

$$\frac{q_0(\mathbf{y}) - q_1(\mathbf{y})}{q_0(\mathbf{y})} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sqrt[m]{\eta} - 1}{\sqrt[m]{\eta}}. \quad (3.38)$$

Cette expression a également été obtenue par E.Conte, M.Lops et G.Ricci dans [Cont95] après un développement asymptotique des fonctions h_m du test optimum pour la détection cohérente en présence de K-distribution. Antérieurement, V.A. Korado dans [Kora68], et B. Picinbono et G. Vezzosi dans [Pici70] avaient déjà obtenu cette expression mais sous des hypothèses de bruit Gaussien.

3.4.8 Version adaptative du BORD Asymptotique

Dès lors que la matrice de covariance est déterminée, le BORD Asymptotique est adaptatif. En reprenant l'estimée NSCM de la matrice de covariance des données, le BORD Asymptotique admet pour expressions (dans le cas d'une cible inconnue) :

$$\left(\frac{\widehat{q}_0(\mathbf{y}_{obs})}{\widehat{q}_1(\mathbf{y}_{obs})} \right)^m \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (3.39)$$

$$\frac{|\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs}|^2}{(\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{p})(\mathbf{y}_{obs}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs})} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sqrt[m]{\eta} - 1}{\sqrt[m]{\eta}} \quad (3.40)$$

$$|\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs}|^2 - (1 - \eta^{-1/m})(\mathbf{y}_{obs}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs})(\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{p}) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \quad (3.41)$$

Avec l'estimée NSCM, la dernière expression est appelée ALQ (Adaptive Linear Quadratic) dans [Gini97a]. Son appellation provient de sa dépendance aux données par la statistique linéaire $\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs}$ et sa statistique quadratique $\mathbf{y}_{obs}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs}$ et se différencie du GLRT-LQ adaptatif simplement par l'estimation de la matrice de covariance. Les études de ces détecteurs menées dans [Cont95] et dans [Gini97a] montrent que le GLRT-LQ et l'ALQ vérifient la propriété TFAC par rapport à la loi de la *texture*. Des courbes représentant les valeurs de la probabilité de fausse alarme en fonction du seuil de détection pour différentes natures de la K-distribution montrent bien l'invariance de la statistique des détecteurs à la loi de la *texture*.

Le BORD Asymptotique adaptatif coïncide donc avec l'ALQ, qui admet la propriété TFAC par rapport à la *texture*. Dans la section suivante, cette propriété est vérifiée en équation et illustrée par des résultats de simulation.

3.4.9 Propriété TFAC du BORD Asymptotique

Le BORD Asymptotique conserve structurellement la propriété TFAC par rapport à la loi de la *texture* de la même manière que le BORD. En effet, en remplaçant q_0 et q_1 par leurs

expressions dans le cas d'une cible d'amplitude estimée selon (1.15), le BORD Asymptotique devient indépendant de la *texture*, et par conséquent, sa statistique est indépendante de la loi du SIRV considéré :

$$\frac{q_0(\mathbf{x}) - q_1(\mathbf{x})}{q_0(\mathbf{x})} \underset{H_0}{\overset{H_1}{\approx}} \frac{\sqrt[m]{\eta} - 1}{\sqrt[m]{\eta}} \quad (3.42)$$

$$\frac{|\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{x}|^2}{(\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p})(\mathbf{x}^\dagger \mathbf{M}^{-1} \mathbf{x})} \underset{H_0}{\overset{H_1}{\approx}} \frac{\sqrt[m]{\eta} - 1}{\sqrt[m]{\eta}}. \quad (3.43)$$

Cette propriété est illustrée à travers les courbes de la figure (3.16) représentant la valeur du seuil en fonction de la P_{fa} . La figure de gauche concerne un environnement K-distribué corrélé pour différentes valeurs du paramètre ν . Les courbes de la figure de droite ont été obtenues pour différentes natures de SIRV de même matrice de covariance : K-distribué de paramètre $\nu = 0.1$, Gaussien, Student-t et le dernier est de statistique inconnue avec une *texture* qui suit une loi de Weibull. Comme pour le BORD, on constate que, quelle que soit la statistique du bruit, toutes les courbes indiquent la même valeur du seuil pour une P_{fa} donnée.

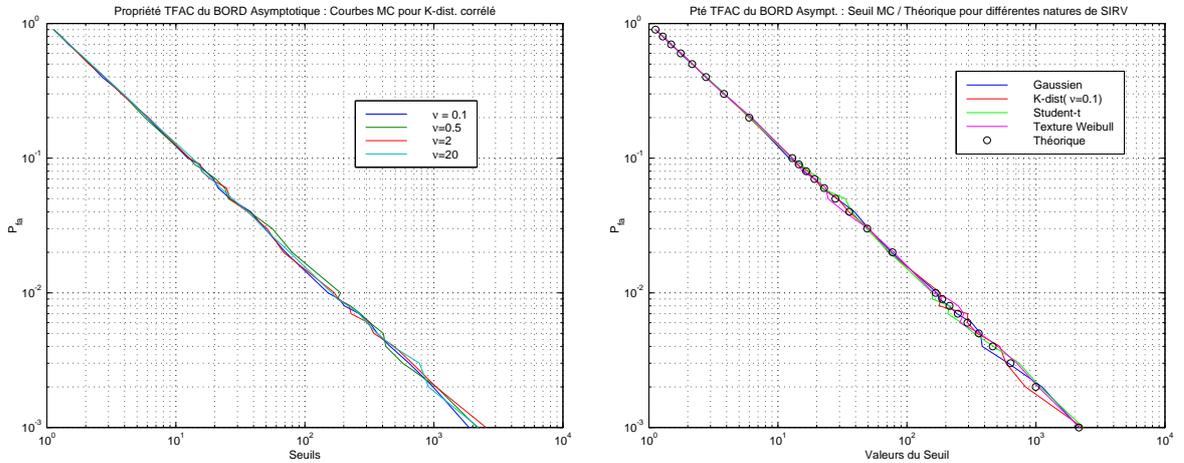


FIG. 3.16 – Propriété TFAC par rapport à la loi de la *texture* du BORD Asymptotique. Environnements K-distribués : $\nu = 0.1, 0.5, 2, 20$ (courbes de gauche) ; environnements K-distribué ($\nu = 0.1$), Gaussien, Student-t($\nu = 0.3$) et dont la texture suit une loi de Weibull (courbes de droite). Comparaison avec les valeurs théoriques (à droite).

La loi de ce détecteur se déduit donc de la loi des vecteurs Gaussiens qui le composent (et plus précisément des formes quadratiques des vecteurs Gaussiens), indépendamment de la loi de la *texture*. Contrairement au cas du BORD, la loi du BORD Asymptotique peut se déduire des lois des formes quadratiques des vecteurs Gaussiens après s'être affranchi de l'indépendance entre le numérateur et le dénominateur.

3.4.10 Loi du BORD Asymptotique

Pour déterminer la loi du BORD Asymptotique, on utilise son équivalent statistique explicite en (3.43), pour lequel les vecteurs bruit représentent des vecteurs de bruit Gaussien complexe. Indépendamment l'une de l'autre, les lois du numérateur et du dénominateur se déterminent facilement (formes quadratiques de vecteurs Gaussiens). En revanche, la loi du quotient ne peut pas se déduire directement de leurs lois respectives, puisque ces deux variables ne sont pas indépendantes l'une de l'autre.

En transformant l'écriture de (3.43), et en utilisant le théorème de Cochran, la loi du BORD Asymptotique peut être déterminée. La preuve de ce théorème est donnée dans l'annexe C (C.4.3), et son énoncé est le suivant :

Théorème 4 (Théorème de Cochran) *Soit \mathbf{z} un vecteur réel Gaussien centré, de taille m , tel que $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_m)$, où \mathbf{I}_m désigne la matrice identité de taille $m \times m$. Supposons que la forme quadratique $\mathbf{z}^T \mathbf{z}$ se décompose en k formes quadratiques $\mathbf{Q}_i = \mathbf{z}^T \mathbf{B}_i \mathbf{z}$, $1 \leq i \leq m$, avec r_i le rang de \mathbf{B}_i , matrice semi-définie positive.*

Alors chacune des trois propositions suivantes entraîne les deux autres :

- (a) *La somme des rangs r_i des matrices \mathbf{B}_i est égal à m ,*
- (b) *Chaque forme quadratique $\mathbf{Q}_i = \mathbf{z}^T \mathbf{B}_i \mathbf{z}$ suit une loi du $\chi_{r_i}^2$, loi du χ^2 à r_i degrés de liberté,*
- (c) *Les formes quadratiques \mathbf{Q}_i sont mutuellement indépendantes.*

Remarque 1 : *Ce théorème est encore valable dans le cas d'un vecteur \mathbf{x} Gaussien de matrice de covariance \mathbf{M} quelconque. Une première justification se trouve dans le fait que \mathbf{x} appartient à la classe des SIRV pour lesquels il est montré que la forme quadratique, définie par $\mathbf{x}^T \mathbf{M}^{-1} \mathbf{x}$, est invariante par transformation linéaire. Par conséquent, cette forme quadratique est statistiquement équivalente à $\mathbf{z}^T \mathbf{I}^{-1} \mathbf{z}$, où \mathbf{z} est défini comme dans le théorème de Cochran. Appliquer le théorème avec le vecteur \mathbf{z} et sa forme quadratique associée ou avec le vecteur \mathbf{x} et sa forme quadratique associée revient au même.*

Une deuxième justification peut être donnée : soit $\mathbf{w} = \mathbf{M}^{-1/2} \mathbf{x}$. Le vecteur \mathbf{w} est alors un vecteur Gaussien de matrice de covariance l'identité, puisque $\mathbb{E}(\mathbf{w} \mathbf{w}^T) = \mathbb{E}(\mathbf{M}^{-1/2} \mathbf{x} \mathbf{x}^T \mathbf{M}^{-1/2}) = \mathbf{M}^{-1/2} \mathbf{M} \mathbf{M}^{-1/2} = \mathbf{I}$. Le théorème s'applique alors avec \mathbf{w} et la forme quadratique associée $\mathbf{w}^T \mathbf{w}$.

Remarque 2 : *Dans le cas d'un vecteur Gaussien complexe \mathbf{y} , de matrice de covariance $2\mathbf{I}_m$ ou $2\mathbf{M}$ quelconque, le théorème s'applique encore. En posant $\mathbf{y} = \mathbf{a} + j\mathbf{b}$, avec $\mathbf{a}, \mathbf{b} \sim \mathcal{N}(\mathbf{0}, \mathbf{M})$ et indépendants, alors le vecteur \mathbf{y} peut être représenté dans \mathbb{R}^{2m} (au lieu de \mathbb{C}) par le vecteur $\tilde{\mathbf{y}} = [\mathbf{a} \ \mathbf{b}]^T$, vecteur à $2m$ composantes, de matrice de covariance de taille $(2m \times 2m)$ donnée par :*

$$\tilde{\mathbf{M}} = \begin{pmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{pmatrix}$$

La forme quadratique $\tilde{\mathbf{y}}^T \tilde{\mathbf{M}}^{-1} \tilde{\mathbf{y}}$ est une somme de $2m$ composantes, et se décompose en k formes quadratiques $\tilde{\mathbf{Q}}_i = \tilde{\mathbf{y}}^T \mathbf{A}_i \tilde{\mathbf{y}}$.

Les propositions (b) et (c) reste inchangées tandis que la proposition (a) devient :
 (a - C) La somme des rangs r_i des matrices \mathbf{A}_i est égal à $2m$

Notre but est de trouver la loi de (3.43), que l'on note :

$$U = \frac{|\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{x}|^2}{(\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p})(\mathbf{x}^\dagger \mathbf{M}^{-1} \mathbf{x})}, \quad (3.44)$$

où $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, 2\mathbf{M})$ et où le vecteur \mathbf{p} est un vecteur dont les m composantes sont égales (par exemple, $\mathbf{p} = \mathbf{1}_m$, un vecteur de 1).

Pour se retrouver dans les conditions d'application du théorème de Cochran, il faut modifier l'expression de U pour déterminer les matrices \mathbf{B}_i et leurs rangs associés. Nous allons vérifier (a) dans un premier temps (cas réel), puis vérifier (a - C) qui impliquera alors (b) et (c).

3.4.10.1 Cas où \mathbf{x} est réel

Dans un premier temps, considérons le vecteur \mathbf{x} réel Gaussien de matrice de covariance \mathbf{M} . En posant $V = U^{-1}$ pour (3.44) avec \mathbf{x} réel, on a :

$$V = \frac{(\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p})(\mathbf{x}^T \mathbf{M}^{-1} \mathbf{x})}{|\mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{x}|^2}. \quad (3.45)$$

Si on pose :

$$\mathbf{w} = \mathbf{M}^{-1/2} \mathbf{x}, \quad (3.46)$$

et

$$\mathbf{q} = \mathbf{M}^{-1/2} \mathbf{p}, \quad (3.47)$$

alors :

- * \mathbf{w} est un vecteur Gaussien réel centré, de matrice de covariance \mathbf{I}_m (matrice identité de taille $(m \times m)$),
- * le vecteur \mathbf{q} est un vecteur de taille $(m \times 1)$.

La variable V se réécrit alors en fonction de \mathbf{w} et \mathbf{q} comme suit :

$$V = \frac{\mathbf{w}^T \mathbf{w}}{\mathbf{w}^T \mathbf{q} (\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger \mathbf{w}} = 1 + \frac{\mathbf{w}^T (\mathbf{I}_m - \mathbf{q} (\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger) \mathbf{w}}{\mathbf{w}^T \mathbf{q} (\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger \mathbf{w}} = 1 + \frac{Q_1}{Q_2}. \quad (3.48)$$

On peut alors écrire :

$$\mathbf{w}^T \mathbf{w} = \mathbf{w}^T \mathbf{I} \mathbf{w} = \mathbf{w}^T (\mathbf{I}_m - \mathbf{q} (\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger) \mathbf{w} + \mathbf{w}^T \mathbf{q} (\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger \mathbf{w} = Q_1 + Q_2, \quad (3.49)$$

où, dans l'esprit du théorème de Cochran on a :

$$\mathbf{w}^T \mathbf{w} = Q_1 + Q_2 = \mathbf{w}^T \mathbf{B}_1 \mathbf{w} + \mathbf{w}^T \mathbf{B}_2 \mathbf{w}, \quad (3.50)$$

avec :

$$* \mathbf{B}_1 = \mathbf{I}_m - \mathbf{q} (\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger$$

$$* \mathbf{B}_2 = \mathbf{q} (\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger,$$

* et $\mathbf{B}_1 + \mathbf{B}_2 = \mathbf{I}_m$.

L'indépendance entre les deux formes quadratiques Q_1 et Q_2 ainsi que leurs lois respectives s'obtiennent par le théorème de Cochran. Il suffit de montrer que la somme des rangs des deux matrices \mathbf{B}_1 et \mathbf{B}_2 est égale à m .

Rang de la matrice \mathbf{B}_2

On a : $\mathbf{B}_2 = \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger$.

Le terme $\mathbf{q}^\dagger \mathbf{q}$ représentant le carré de la norme du vecteur \mathbf{q} , la matrice \mathbf{B}_2 est proportionnelle à la matrice $\mathbf{q} \mathbf{q}^\dagger$, qui est du même rang que \mathbf{q} (pour toute matrice \mathbf{A} , $\text{rang}(\mathbf{A}) = \text{rang}(\mathbf{A}^\dagger \mathbf{A}) = \text{rang}(\mathbf{A} \mathbf{A}^\dagger)$). On a donc $r_2 = 1$. De plus, \mathbf{B}_2 étant idempotente ($\mathbf{B}_2^2 = \mathbf{B}_2$) et symétrique ($\mathbf{B}_2^\dagger = \mathbf{B}_2$), c'est une matrice de projection orthogonale sur l'espace engendré par \mathbf{q} .

Rang de la matrice \mathbf{B}_1

On peut montrer que les matrices \mathbf{B}_1 et \mathbf{B}_2 sont orthogonales. Pour ce faire, on peut montrer, par exemple, que $\mathbf{B}_1 \mathbf{q} = \mathbf{0}$ puisque \mathbf{q} est l'espace image de \mathbf{B}_2 (noté $\text{Im}(\mathbf{B}_2)$); son noyau est noté $\text{Ker}(\mathbf{B}_2)$. On a donc :

$$\begin{aligned} \mathbf{B}_1 \mathbf{q} &= (\mathbf{I}_m - \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger) \mathbf{q} \\ &= \mathbf{q} - \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger \mathbf{q} \\ &= \mathbf{0}_m. \end{aligned} \tag{3.51}$$

Les matrices \mathbf{B}_1 et \mathbf{B}_2 étant orthogonales, alors $\dim(\text{Im}(\mathbf{B}_1)) = \dim(\text{Ker}(\mathbf{B}_2)) = m - 1$ car $\dim(\text{Im}(\mathbf{B}_2)) = 1 = \dim(\text{Im}(\mathbf{I}_m)) - \dim(\text{Ker}(\mathbf{B}_2))$. Le rang de \mathbf{B}_1 est donc $r_1 = m - 1$. La matrice \mathbf{B}_2 est une matrice de projection orthogonale sur l'espace orthogonal à \mathbf{q} (elle est idempotente et symétrique).

La proposition (a) du théorème de Cochran étant vérifiée ($r_1 + r_2 = m$), elle entraîne les résultats suivants :

- * $Q_1 = \mathbf{w}^T (\mathbf{I}_m - \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger) \mathbf{w}$ suit une loi du $\chi_{r_1}^2$, soit une loi χ_{m-1}^2 ,
- * $Q_2 = \mathbf{w}^T (\mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger) \mathbf{w}$ suit une loi du $\chi_{r_2}^2$, soit une loi χ_1^2 ,
- * Q_1 et Q_2 sont indépendantes.

Ainsi, $\frac{Q_1}{Q_2}$ est le quotient de deux variables aléatoires indépendantes.

On rappelle que si $A \sim \chi_{\nu_1}^2$ et si $B \sim \chi_{\nu_2}^2$, alors

$$C = \frac{A/\nu_1}{B/\nu_2} \sim \mathcal{F}(\nu_1, \nu_2),$$

où $\mathcal{F}(\nu_1, \nu_2)$ est une loi de Fisher de paramètres ν_1 et ν_2 dont la densité de probabilité est définie sur \mathbb{R}_+ par :

$$p_C(t) = \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)} \frac{t^{\frac{\nu_1}{2}-1}}{\left(1 + \frac{\nu_1}{\nu_2} t\right)^{\frac{\nu_1+\nu_2}{2}}}, \quad (3.52)$$

où $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ est la fonction Beta.

La loi de $D = \frac{\nu_1}{\nu_2} C$ est donc définie sur \mathbb{R}_+ par :

$$p_D(t) = \frac{t^{\frac{\nu_1}{2}-1}}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right) (1+t)^{\frac{\nu_1+\nu_2}{2}}}. \quad (3.53)$$

Ainsi, par identification, on a :

$$D = \frac{Q_1}{Q_2} = (m-1) \frac{\chi_{m-1}^2/(m-1)}{\chi_1^2/1},$$

dont l'expression est donnée par (3.53) pour $\nu_1 = m-1$ et $\nu_2 = 1$, soit encore :

$$p_{Q_1/Q_2}(t) = \frac{t^{\frac{m-3}{2}}}{B\left(\frac{m-1}{2}, \frac{1}{2}\right) (1+t)^{\frac{m}{2}}}. \quad (3.54)$$

La loi de $G = \frac{1}{1+D}$ est alors une loi Beta, de paramètres $\nu_2/2$ et $\nu_1/2$, notée $\mathcal{B}\left(\frac{\nu_2}{2}, \frac{\nu_1}{2}\right)$, définie sur $[0, 1]$ par :

$$p_G(t) = \frac{t^{\frac{\nu_2}{2}-1} (1-t)^{\frac{\nu_1}{2}-1}}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)}. \quad (3.55)$$

La densité de probabilité de $U = \frac{1}{1 + \frac{Q_1}{Q_2}}$ est donc définie sur $[0, 1]$ par :

$$p_U(u) = \frac{\sqrt{u} (1-u)^{\frac{m-3}{2}}}{B\left(\frac{m-1}{2}, \frac{1}{2}\right)}, \quad (3.56)$$

qui représente une loi Beta de paramètres $1/2$ et $(m-1)/2$, notée $\mathcal{B}\left(\frac{1}{2}, \frac{m-1}{2}\right)$.

Le test de détection (3.43), dans le cas où \mathbf{x} est réel se réécrit alors :

$$U \underset{H_0}{\overset{H_1}{\geq}} \frac{\mathfrak{V}\sqrt{\eta} - 1}{\mathfrak{V}\sqrt{\eta}} = \lambda, \quad (3.57)$$

et la probabilité de fausse alarme est donnée par :

$$\begin{aligned}
 P_{fa} &= \int_{\lambda}^1 p_U(u) du = \frac{1}{B\left(\frac{m-1}{2}, \frac{1}{2}\right)} \int_{\lambda}^1 \sqrt{u} (1-u)^{\frac{m-3}{2}} du \\
 &= \frac{2}{3 B\left(\frac{m-1}{2}, \frac{1}{2}\right)} \left[u^{3/2} {}_2F_1\left(\frac{3}{2}, \frac{3-m}{2}; \frac{5}{2}, u\right) \right]_{\lambda}^1, \tag{3.58}
 \end{aligned}$$

où ${}_2F_1(x)$ est la fonction hypergéométrique F à 2 et 1 paramètres en argument, évaluée en x . Cette fonction s'exprime avec les fonctions Gamma, dont la forme générale est donnée en annexe D.

3.4.10.2 Cas où \mathbf{x} est complexe

Dans le cas où \mathbf{x} est complexe, on transpose le problème sur \mathbb{R}^{2m} (équivalent à \mathbb{C}^m) :

- * \mathbf{w} devient le vecteur $\mathbf{z} = [\mathbf{a} \ \mathbf{b}]^T$ de taille $(2m \times 1)$, où \mathbf{a} et \mathbf{b} sont respectivement les parties réelle et imaginaire de \mathbf{w} , indépendantes, de loi $\mathcal{N}(\mathbf{0}, \mathbf{I}_m)$.
- * le vecteur \mathbf{q} devient une matrice de taille $(2m \times 2)$, de rang 2, de la forme :

$$\mathbf{q} = \begin{bmatrix} q_{1,1} & 0 \\ q_{2,1} & 0 \\ \vdots & \vdots \\ q_{m,1} & 0 \\ 0 & q_{m+1,2} \\ 0 & q_{m+2,2} \\ \vdots & \vdots \\ 0 & q_{2m,2} \end{bmatrix}. \tag{3.59}$$

La quantité $\mathbf{q}^\dagger \mathbf{q}$ devient :

$$\mathbf{q}^\dagger \mathbf{q} = \begin{pmatrix} \sum_{i=1}^m q_{i,1}^2 & 0 \\ 0 & \sum_{i=m+1}^{2m} q_{i,2}^2 \end{pmatrix} = \begin{pmatrix} \alpha_1^2 & 0 \\ 0 & \alpha_2^2 \end{pmatrix}, \tag{3.60}$$

et la matrice de projection $\mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger$, de rang 2 est une matrice bloc diagonale de la forme :

$$\mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger = \begin{pmatrix} \mathbf{E}_m/\alpha_1^2 & \mathbf{0}_m \\ \mathbf{0}_m & \mathbf{F}_m/\alpha_2^2 \end{pmatrix}, \tag{3.61}$$

où les matrices \mathbf{E}_m et \mathbf{F}_m sont des matrices de taille $(m \times m)$, de rang 1.

La forme quadratique $\mathbf{z}^T \mathbf{z}$ se décompose en somme de deux formes quadratiques $P_1 = \mathbf{z}^T \mathbf{A}_1 \mathbf{z}$ et $P_2 = \mathbf{z}^T \mathbf{A}_2 \mathbf{z}$ telles que :

$$\mathbf{A}_1 = \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger, \tag{3.62}$$

et

$$\mathbf{A}_2 = \mathbf{I}_{2m} - \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger. \quad (3.63)$$

La matrice de projection \mathbf{A}_1 est de rang 2, soit $r_1 = 2$.

Les deux matrices \mathbf{A}_1 et \mathbf{A}_2 sont orthogonales, puisque :

$$\begin{aligned} \mathbf{A}_2 \mathbf{q} &= (\mathbf{I}_{2m} - \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger) \mathbf{q} \\ &= \mathbf{q} - \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger \mathbf{q} \\ &= \mathbf{0}_m. \end{aligned} \quad (3.64)$$

Le rang de la matrice \mathbf{A}_2 est donc $r_2 = 2m - 2$.

On a bien montré que $r_1 + r_2 = 2m$, proposition (a–C) du théorème de Cochran, qui entraîne (b) et (c) :

- * $P_1 = \mathbf{z}^T (\mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger) \mathbf{z}$ suit une loi χ_2^2 ,
- * $P_2 = \mathbf{z}^T (\mathbf{I}_{2m} - \mathbf{q}(\mathbf{q}^\dagger \mathbf{q})^{-1} \mathbf{q}^\dagger) \mathbf{z}$ suit une loi χ_{2m-2}^2 ,
- * P_1 et P_2 sont indépendantes.

La loi de (3.43), dans le cas complexe, se déduit donc des lois respectives de P_1 et P_2 comme la loi de :

$$U = \frac{1}{1 + \frac{P_2}{P_1}}, \quad (3.65)$$

où la loi de $\frac{P_2}{P_1}$ est donnée par (3.53), pour $\nu_1 = 2m - 2$ et $\nu_2 = 2$, soit encore :

$$p_{P_2/P_1}(t) = \frac{t^{m-2}}{B(m-1, 1) (1+t)^m}. \quad (3.66)$$

D'après le résultat général donné par (3.55), la loi de U, dans le cas complexe, est donc une loi Beta $\mathcal{B}(1, m-1)$, définie sur $[0, 1]$ par :

$$p_u(u) = \frac{(1-u)^{m-2}}{B(1, m-1)} = (m-1)(1-u)^{m-2}. \quad (3.67)$$

Le test de détection (3.43), dans le cas où \mathbf{x} est complexe se réécrit :

$$U \underset{H_0}{\overset{H_1}{>}} \frac{\sqrt[2]{\eta} - 1}{\sqrt[2]{\eta}} = \lambda, \quad (3.68)$$

et la probabilité de fausse alarme se calcule simplement :

$$P_{fa} = \int_{\lambda}^1 (m-1)(1-u)^{m-2} du = (1-\lambda)^{m-1} \quad (3.69)$$

Le calcul du seuil de détection η pour une probabilité de fausse alarme donnée s'effectue selon :

$$\eta = P_{fa}^{\frac{m}{1-m}}. \quad (3.70)$$

Cette expression ne dépend plus que de la taille des vecteurs observés comme le montre les courbes sur la figure (FIG. 3.17). Les courbes de droite représentent les valeurs du seuil de détection en fonction de m , à P_{fa} fixée, alors que celles de gauche représentent ces mêmes valeurs en fonction de la P_{fa} à m fixé.

Les valeurs théoriques du seuil de détection sont comparées sur la figure (FIG. 3.16) pour le BORD Asymptotique en environnements divers. On constate que la loi déterminée pour ce détecteur est en adéquation avec la loi empirique, donnée par Monte-Carlo. Il en va de même pour le BORD, comme cela est montré sur la figure (FIG. 3.18).

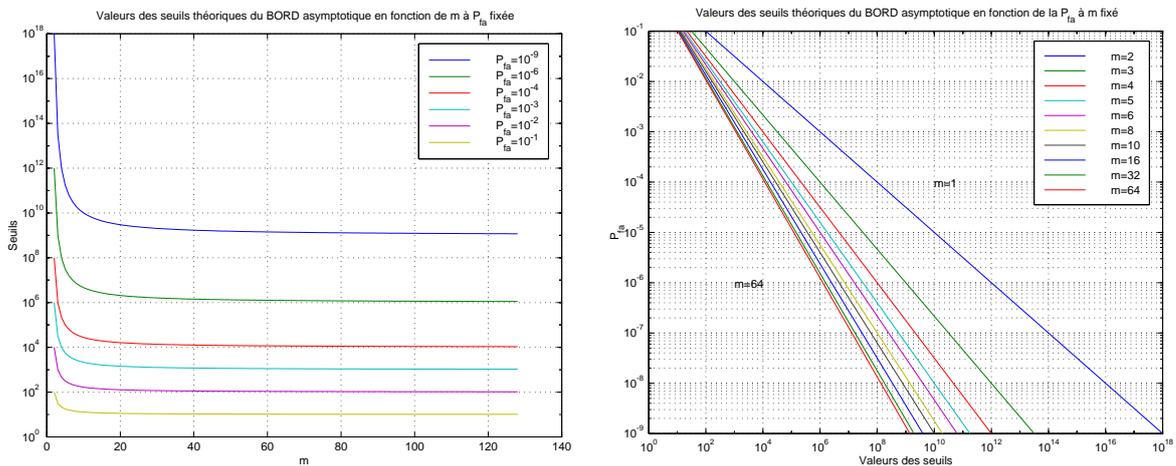


FIG. 3.17 – Courbes théoriques déterminant la valeur du seuil de détection en fonction de la P_{fa} (gauche) et en fonction de m pour une P_{fa} fixée.

3.4.11 Performances de détection du BORD Asymptotique

La simplicité de l'expression du BORD Asymptotique associée à l'expression théorique pour le calcul du seuil de détection permet d'imaginer tous les scénarios possibles pour étudier les performances de ce détecteur dans des environnements quelconques. Il est désormais facile d'analyser les performances de détection pour des P_{fa} très faibles, comme par exemple $P_{fa} = 10^{-6}$, 10^{-9} ou 10^{-12} .

Avant de présenter de telles courbes, une première étude est réalisée sur l'influence de l'estimation de la matrice sur les performances du BORD et du BORD Asymptotique, évaluées avec la valeur théorique du seuil du BORD Asymptotique. La P_{fa} est fixée à 10^{-3} afin de comparer ces résultats avec ceux obtenus par MC. La figure (FIG. 3.19) représente donc les performances de détection en environnement K-distribué de paramètre $\nu = 0.1$ selon le nombre de données de

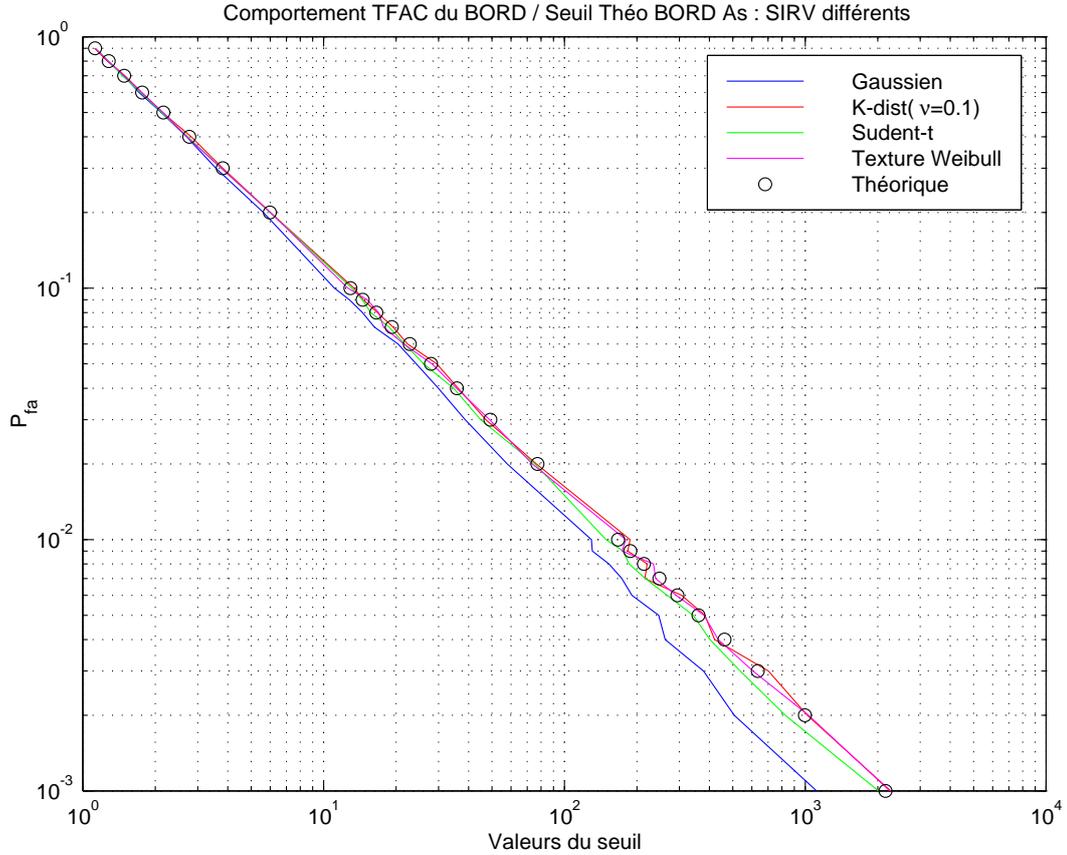


FIG. 3.18 – Adéquation entre les valeurs théoriques du seuil du BORD Asymptotique et les valeurs MC du BORD.

référence utilisé pour estimer la matrice de covariance . Le nombre d'observations est $N_{obs} = 200$. Les erreurs relatives de P_d sont également tracées ainsi que le taux de fausse alarme constaté, du fait de l'utilisation du seuil théorique.

Sur cette figure, on constate que l'erreur d'estimation de la matrice devient négligeable, dès lors que N_{ref} est supérieur ou égal à 50. Le taux de fausse alarme est dû à l'utilisation de la valeur théorique du seuil de détection, adapté au BORD Asymptotique, pour une matrice parfaitement connue et non-singulière.

La figure (FIG. 3.20) représente les performances de détection comparées du BORD et du BORD Asymptotique dans un environnement K-distribué de paramètre $\nu = 0.1, 0.5, 2, 20$. Les courbes sont tracées pour $P_{fa} = 10^{-6}$. Le seuil de détection est calculé théoriquement pour les deux détecteurs et la matrice de covariance utilisée est celle représentée sur la figure (FIG. 3.11). Les courbes en trait plein correspondent au cas où la matrice est supposée connue, et les courbes en pointillés correspondent au cas où la matrice est estimée. Sur les différentes courbes de cette figure, le nombre de données de référence utiles à l'estimation de la matrice est égal à $N_{ref} = 1000$ et le nombre de données d'observation est de $N_{obs} = 500$.

Jusqu'à présent, la corrélation des données n'était pas très forte. On s'aperçoit vite que

lorsque la corrélation augmente, l'inversion de la matrice devient un problème. La matrice n'est plus forcément de rang plein, et le calcul théorique du seuil de détection n'est plus valable. Dans ces cas-là, on peut avoir recours à la pseudo-inverse de la matrice, et ceci est étudié plus en détail dans la section suivante.

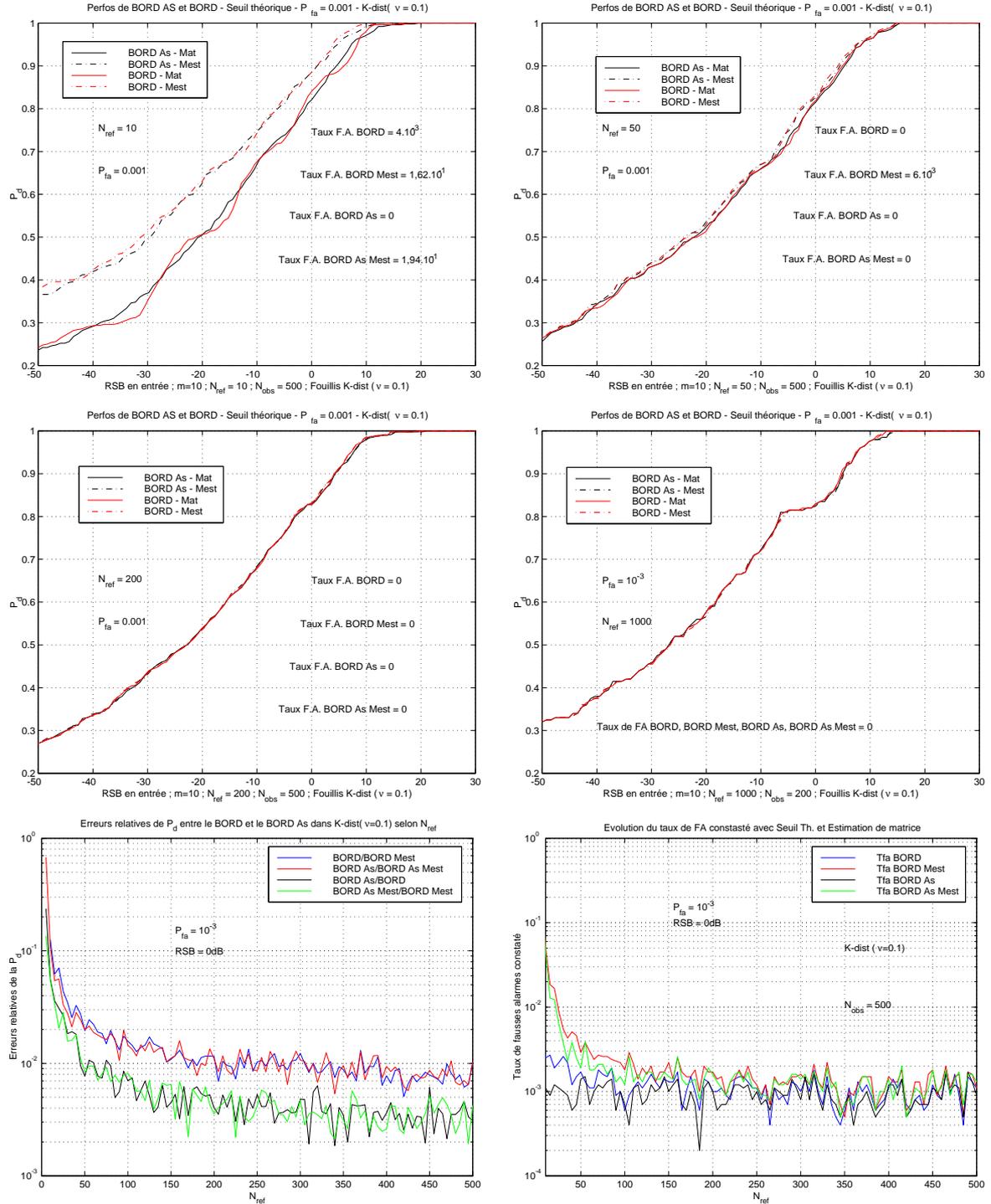


FIG. 3.19 – Performances de détection (4 premières courbes) du BORD et du BORD Asymptotique en environnement K-distribué ($\nu = 0.1$). Le seuil théorique est utilisé pour une $P_{fa} = 10^{-3}$. Le nombre de référence utile à l'estimation de la matrice varie : $N_{ref} = 10, 50, 200, 1000$. Erreur relative (courbe gauche en bas) de P_d en fonction de N_{ref} . Evolution associée du taux de fausse alarme constaté (courbe droite en bas) en fonction de N_{ref}

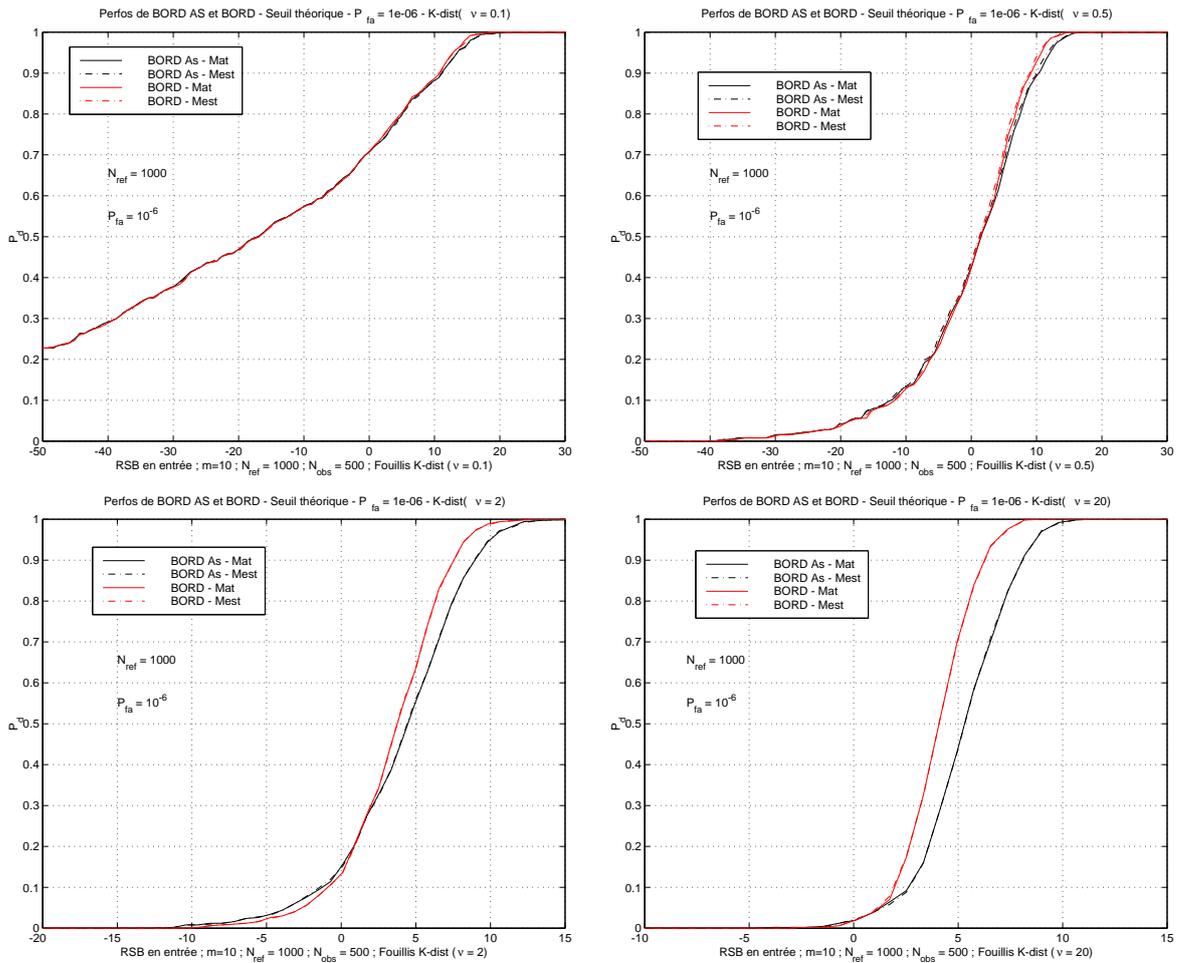


FIG. 3.20 – Performances de détection du BORD et du BORD Asymptotique en environnement K-distribué ($\nu = 0.1, 0.5, 2, 20$). Le seuil théorique est utilisé pour une $P_{fa} = 10^{-6}$. Le nombre de référence utile à l'estimation de la matrice est fixé à $N_{ref} = 1000$. Le nombre d'observations est égal à $N_{obs} = 500$.

3.4.12 Influence de la corrélation sur les performances

La corrélation des données n'influe pas spécifiquement sur l'estimation de la matrice de covariance, mais sur son inversion, dépendante de son rang. Inverser une matrice qui n'est pas de rang plein nécessite l'utilisation de la matrice pseudo-inverse. Dans ce cas-là, le théorème de Cochran ne s'applique plus, car ses hypothèses ne sont pas respectées. Les deux transformations (3.46 et 3.47), effectuées sur les vecteurs \mathbf{x} des données et le vecteur \mathbf{p} , pour se ramener à un vecteur Gaussien de matrice de covariance \mathbf{I}_m , ne sont plus du même effet si la matrice est singulière. Il devient alors difficile (voire impossible ?) de déterminer la matrice de projection qui rend les deux formes quadratiques indépendantes. Ce problème reste ouvert et quelques perspectives sont données à ce propos dans le paragraphe 3.5.3 qui synthétise les résultats obtenus dans ce chapitre.

Dans le souci d'interpréter ce phénomène sur les matrices (de taille 10×10), construites à partir de la densité spectrale Gaussienne, la figure (FIG. 3.21) représente leurs rangs, leurs conditionnements et leurs déterminants en fonction de la valeur de l'étalement spectral, $\sigma_f = 0.2, \dots, 2$. Le conditionnement d'une matrice est défini comme le rapport de sa plus grande valeur singulière avec sa plus petite. Plus ce nombre est grand et plus la matrice devient singulière. Les caractéristiques de la matrice utilisée dans les simulations de ce chapitre ne figurent pas sur les courbes. Son étalement spectral est $\sigma_f = 10$, son déterminant vaut approximativement $\det(\mathbf{M}) = 0.9076$ et son conditionnement est de 1.5467.

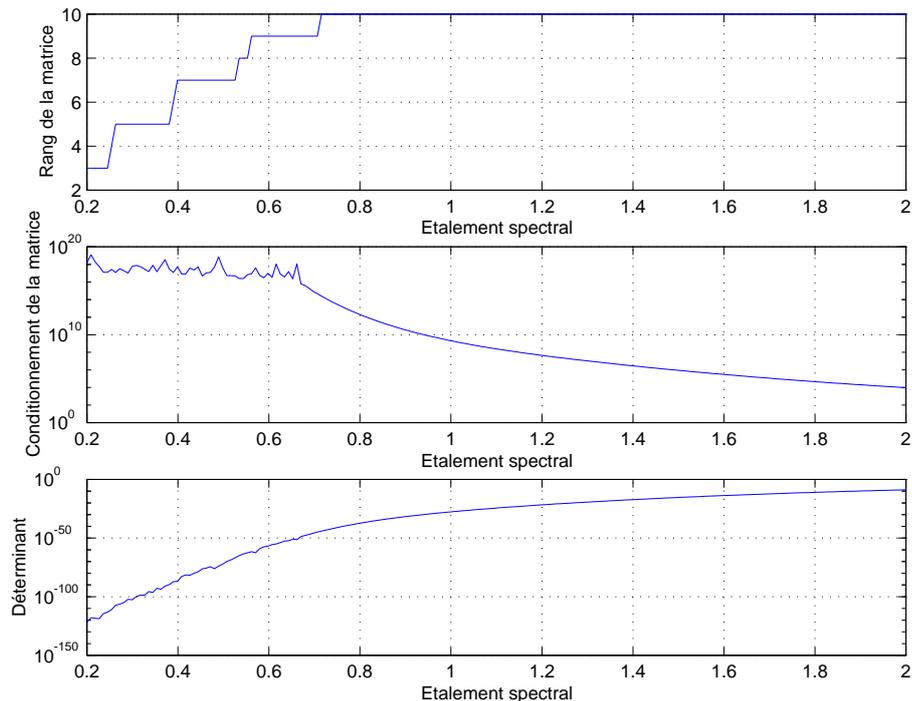


FIG. 3.21 – Rang et conditionnement de la matrice de corrélation en fonction de son étalement spectral. Les différentes matrices testées ont un étalement spectral allant de 0.2 (forte corrélation) à 2 (corrélation moyenne).

3.5 Synthèse

3.5.1 Le BORD

Les résultats de simulation obtenus avec le BORD ont montré que ce détecteur atteint les performances des détecteurs optimaux. Cela se vérifie aussi bien pour du bruit Gaussien que pour du fouillis très impulsif, comme un fouillis K-distribué de paramètre $\nu = 0.1$ (FIG. 3.12) ou un fouillis régi par une loi de Student-t (FIG. 3.13). Il en va de même lorsque la statistique globale est analytiquement inconnue, auquel cas le BORD atteint les mêmes performances que le PEOD "idéal" (FIG. 3.14). Le BORD est construit de telle sorte que la loi de la texture des données est "apprise" par la seule présence des données de référence.

Dans sa version adaptative, les données de référence sont également nécessaires à l'estimation de la matrice de covariance. Pour un vecteur de références de taille $m = 10$, la figure (FIG. 3.19) montre qu'à partir de $N_{ref} = 50$, l'erreur commise sur la P_d à $0dB$ est inférieure à 2%. La matrice de covariance ($\sigma_f = 10$) est estimée au sens NSCM et le seuil théorique du BORD Asymptotique est utilisé pour vérifier $P_{fa} = 10^{-3}$. L'utilisation de ce seuil est justifiée dans pareil cas : la figure (FIG. 3.10) a été obtenue pour une matrice de covariance de taille 10×10 , de même étalement spectral $\sigma_f = 10$. L'adéquation entre le seuil MC du BORD et le seuil théorique du BORD Asymptotique peut d'ailleurs être observée sur la figure (FIG. 3.18).

Comme cela a également été montré, le BORD vérifie structurellement la propriété TFAC par rapport à la loi de la texture. Cette propriété permet d'appliquer, à P_{fa} donnée, la même valeur du seuil de détection, quel que soit l'environnement observé. Elle reste d'ailleurs valable pour la version adaptative du BORD, puisque l'estimée au sens NSCM de la matrice de covariance est indépendante de la loi de la texture du SIRP.

Tous ces résultats, importants aux yeux des radaristes, permettent d'affirmer que le BORD est un détecteur optimal TFAC et adaptatif .

3.5.2 Le BORD Asymptotique

L'expression du BORD Asymptotique découle de la convergence en loi du BORD, lorsque le nombre de référence devient très grand. D'un point de vue pratique, cela suppose que la matrice de covariance peut être parfaitement estimée, puisque le nombre de données utilisé est bien plus grand que la taille de cette matrice. En outre, l'information apportée par les données de référence disparaît, et des différences de comportement sont alors observées par rapport au BORD. Ces différences sont d'autant plus importantes lorsque la statistique de l'environnement tend vers une Gaussienne, comme le montre la figure (FIG. 3.20). La convergence en loi est établie à partir du prior non-informatif de la texture, dont la forme est hyperbolique. Ce dernier est donc assez éloigné de la loi réelle de la texture d'un bruit Gaussien, qui, dans ce cas est une distribution de Dirac en 1.

Tout comme le BORD, le BORD Asymptotique est adaptatif, dès lors que la matrice de covariance est estimée. L'estimation au sens NSCM est également utilisée pour ce détecteur, ce qui contribue à sa propriété TFAC, vérifiée structurellement, par rapport à la loi de la texture. Cette propriété a été montrée au travers des courbes de la figure (FIG.3.16) pour plusieurs lois.

Contrairement au BORD, la loi du BORD Asymptotique se calcule assez facilement à partir de son équivalent statistique TFAC. Le BORD Asymptotique ne dépend alors plus que de formes quadratiques de vecteurs Gaussiens et le théorème de Cochran peut s'appliquer. Le seuil de détection s'obtient à partir de la valeur de la P_{fa} et de la taille m des observations. Les valeurs théoriques sont en adéquation avec les valeurs obtenues par MC, sauf, une fois encore, dans le cas d'un bruit Gaussien, où les courbes sont légèrement décalées.

Dans certains cas, l'expression du seuil théorique du BORD Asymptotique peut être utilisée pour la mise en œuvre du BORD. La figure (FIG. 3.18) en donne la preuve, mais il faut toutefois rester prudent selon les cas envisagés.

3.5.3 Validité du seuil théorique du BORD Asymptotique

Le théorème de Cochran s'applique dans le cas d'un vecteur Gaussien blanc ou d'un vecteur Gaussien de matrice de covariance \mathbf{M} quelconque, avec l'hypothèse implicite que cette dernière n'est pas singulière. Si c'est le cas, la forme quadratique du SIRV est exprimée non plus en fonction de l'inverse de la matrice mais de sa pseudo-inverse. Les deux matrices de projection, permettant d'obtenir l'indépendance entre le numérateur et le dénominateur du BORD Asymptotique, ne sont plus évidentes à déterminer et il devient difficile d'en obtenir la loi. Dans pareils cas, l'expression du seuil théorique du BORD Asymptotique n'est plus valide.

Chapitre 4

Application sur signaux expérimentaux

Ce chapitre est consacré à la validation des résultats de simulation du BORD, du BORD Asymptotique et d'autres détecteurs sur données expérimentales de fouillis. Dans un premier temps, le fonctionnement des systèmes radars actuels est rappelé. Les données expérimentales de fouillis sont ensuite décrites et analysées avant d'évaluer les performances des détecteurs dans cet environnement. L'attention est principalement portée sur le BORD et le BORD Asymptotique dont les performances sont comparées à l'OGD et à l'OKD. Ce dernier est mis en place après estimation des paramètres relatifs à une K -distribution, supposée régir la statistique de l'amplitude des données.

4.1 Principe opérationnel d'un radar

Dans un cadre opérationnel, les caractéristiques du radar déterminent la précision des mesures effectuées et les traitements à mettre en place pour la détection et l'estimation des paramètres en jeu. Il est donc important de mettre en correspondance les paramètres utilisés lors de la modélisation des signaux avec les paramètres du radar, de la cible et du fouillis.

4.1.1 Le traitement théorique optimal

L'objectif des traitements de détection mis en œuvre, est de localiser avec précision la ou les cibles détectées. Cela signifie qu'il faut être capable d'interpréter la nature physique du signal reçu par le radar et de faire le lien avec les paramètres géographiques et cinétiques de la cible. Considérons que le signal $s(t)$, émis par le radar, rencontre une cible en mouvement. Ce signal subit une transformation due à son trajet aller-retour (radar-cible-radar) et à l'effet Doppler (notés conjointement par α), et est affecté d'une phase ϕ et d'une amplitude A , dépendantes de la cible analysée. Aux distorsions qui caractérisent la cible, s'ajoutent les perturbations du fouillis ambiant, considérées comme du bruit $b(t)$. L'ensemble des paramètres de la cible est noté $\theta = (A, \phi, \alpha)$ et le signal total $x(t)$, capté en réception peut se modéliser par :

$$x(t) = s(t, \theta) + b(t) \quad (4.1)$$

$$= A e^{j\phi} s(t, \alpha) + b(t) \quad (4.2)$$

Le vecteur α est relié à la distance et la vitesse de la cible. Si on note par R_0 la distance réelle de la cible par rapport au radar, v sa vitesse radiale, ω_0 la pulsation de la porteuse du si-

gnal émis, λ_0 la longueur d'onde et c la célérité de l'onde dans le milieu, alors [Thou82, LeCh89] :

- la distance de la cible se déduit du calcul du retard $\frac{2R_0}{c}$,
- la vitesse radiale se déduit de la mesure du décalage Doppler égal à $\frac{2\omega_0 v}{c} = \frac{2v}{\lambda_0}$.

Pour déterminer ces quantités, un test de détection doit être mis en place. D'après la théorie de l'estimation au sens du MV, la statistique S à maximiser est le logarithme du rapport de vraisemblance. Dans le cas classique, où le bruit $b(t)$ est supposé Blanc et Gaussien de puissance $2\sigma^2$, cette statistique s'écrit :

$$S = \log(\Lambda(x)) = \frac{1}{2\sigma^2} \left[2 \Re \int_{-\infty}^{+\infty} x(t, \theta_0) s^*(t, \theta) dt - \int_{-\infty}^{+\infty} |s(t, \theta)|^2 dt \right], \quad (4.3)$$

avec $\theta_0 = (A_0, \phi_0, \alpha_0)$, $\theta = (A, \phi, \alpha)$ et où $\Re(\cdot)$ désigne la partie réelle.

Bien souvent, il n'est pas utile d'estimer le paramètre de déphasage. Les radaristes préfèrent déterminer avec précision les paramètres propres aux cibles tels la vitesse, l'accélération, la position angulaire, la position radiale, etc.... Les paramètres de déphasage Φ et d'atténuation A sont considérés comme étant inconnus. Ils sont donc remplacés dans (4.3) par leurs estimées au sens du MV, données par :

$$\hat{\phi}_{mv} = \arg \left[\int_{-\infty}^{+\infty} x(t, \theta_0) s^*(t, \alpha) dt \right] \quad (4.4)$$

$$\hat{A}_{mv} = \frac{\left| \int_{-\infty}^{+\infty} x(t, \theta_0) s^*(t, \alpha) dt \right|}{\int_{-\infty}^{+\infty} |s(t, \alpha)|^2 dt}. \quad (4.5)$$

Une statistique équivalente à S est donc donnée par S' :

$$S' = \log(\Lambda(x, \hat{A}_{mv}, \hat{\phi})) = \frac{1}{2\sigma^2} \frac{\left| \int_{-\infty}^{+\infty} x(t, \theta_0) s^*(t, \alpha) dt \right|^2}{\int_{-\infty}^{+\infty} |s(t, \alpha)|^2 dt}. \quad (4.6)$$

Le traitement approprié consiste donc à effectuer un traitement quadratique sur la fonction d'inter-ambiguïté entre le signal reçu et une réplique adaptée sur les bons paramètres. Le vecteur α qui rendra cette expression maximale sera le candidat désigné pour caractériser la présence de la cible et ainsi la localiser.

4.1.2 Le traitement opérationnel idéal

D'après le traitement théorique optimal, la procédure de détection efficace à mettre en œuvre est basée sur la corrélation du signal reçu avec la réplique du signal émis adapté aux effets de transformation physique qu'il a subie. Cette corrélation donne, pour les paramètres de transformation réels rencontrés, un niveau de sortie maximal, supérieur au seuil de détection. Pour les autres paramètres, ce seuil ne doit pas être dépassé, sous peine d'engendrer des fausses alarmes.

Le récepteur parfait pour les hypothèses fixées (bruit blanc Gaussien) doit alors être du type "batteries de corrélateurs adaptés à un jeu de paramètres de cible fixé". Toutes les valeurs de l'ensemble des paramètres sont balayées de manière exhaustive et les maxima sont recherchés.

Le récepteur opérationnel idéal, limité à l'ordre deux (c'est-à-dire à la position et à la vitesse) calcule le filtre adapté en sortie de *tous* les filtres Doppler et retards possibles, quadrillant ainsi l'espace observé.

Dans le cas d'un signal $\Phi_e(t)$ émis en train d'impulsions de PRF (Pulse Repetition Frequency) donnée, si on note $\Phi_r(t)$ le signal reçu, la sortie $A(\tau, f_d)$ du filtre adapté idéal à l'ordre deux, dans le filtre retard τ et le filtre Doppler f_d prend la forme suivante :

$$A(\tau, f_d) = \left| \int_{-\infty}^{+\infty} \Phi_r(t) \Phi_e^*(t - \tau) e^{-2j\pi f_d(t-\tau)} dt \right|^2. \quad (4.7)$$

Il est nécessaire de considérer les paramètres retard τ et Doppler f_d du filtre adapté les plus proches possibles des paramètres recherchés, ce qui évite d'adapter les paramètres en dehors de la zone distance-vitesse intéressante.

Ce filtre adapté idéal est capable d'estimer toutes les positions et vitesses radiales d'un cortège de cibles pour peu que le code utilisé soit performant (lobes secondaires bas), et que tous les filtres Doppler et retard soient examinés. Si la première condition est réalisable, la deuxième l'est difficilement, de par la charge de calculs trop importante que cela représente pour une zone d'observation étendue.

L'alternative à ce traitement, certes idéal mais difficilement réalisable, consiste à transcrire dans l'expression (4.7) le fait qu'un train de P impulsions élémentaires est émis. On émet alors l'hypothèse que la fréquence Doppler reste cohérente, c'est-à-dire qu'elle ne varie pas dans l'impulsion, mais d'une impulsion à l'autre. Contrairement au traitement idéal, il ne nécessite aucunement de batteries de filtres Doppler à chaque retard.

4.1.3 Le traitement opérationnel classique

La structure du récepteur idéal théorique se simplifie dans le cas de l'émission d'un train de P impulsions élémentaires $u_e(t)$ de période T_r , ce qui s'écrit :

$$\Phi_e(t) = \sum_{k=0}^{P-1} u_e(t - kT_r). \quad (4.8)$$

Dans ce cas, la corrélation avec le signal reçu $\Phi_r(t)$ a pour expression :

$$A(\tau, f_d) = \left| \sum_{k=0}^{P-1} \int_T \Phi_r(t) u_e^*(t - \tau - kT_r) e^{-2j\pi f_d(t-\tau)} dt \right|^2, \quad (4.9)$$

ce qui se réécrit comme :

$$A(\tau, f_d) = \left| \sum_{k=0}^{P-1} e^{-2j\pi f_d kT_r} \int_{-\tau/2}^{\tau/2} \Phi_r(t + \tau + kT_r) u_e^*(t) e^{-2j\pi f_d t} dt \right|^2. \quad (4.10)$$

Pour un retard donné, on remarque une somme discrète d'échantillons prélevés tous les kT_r . Ils proviennent de l'échantillonnage à l'instant $kT_r + \tau$ de l'intercorrélacion distance-Doppler de l'impulsion k , avec l'impulsion élémentaire $u_e(t)$.

Si on fait l'hypothèse, que pendant la durée τ de l'impulsion, la phase $f_d t$ due au Doppler ne fluctue pas trop, on peut alors négliger le terme de phase dans l'intégrale. Cette simplification est justifiée si la vitesse des cibles observées est telle que la variation de phase engendrée par l'effet Doppler n'est significative que d'une impulsion à l'autre, et non pas dans l'impulsion. Une autre manière d'expliquer ceci, est de dire que la durée de l'impulsion est trop brève pour permettre d'estimer cette fréquence Doppler.

Le filtre adapté devient alors :

$$A(\tau, f_d) = \left| \sum_{k=0}^{P-1} e^{-2j\pi f_d k T_r} \int_{-\tau/2}^{\tau/2} \Phi_r(t + \tau + k T_r) u_e^*(t) dt \right|^2. \quad (4.11)$$

Cette procédure se décompose alors en trois étapes, pour chaque porte-distance $c\tau/2$:

1. Acquisition au cours du temps (variation de T_r) du signal reçu, et corrélation par l'impulsion élémentaire $u_e(t)$,
2. Échantillonnage aux instants $\tau + k T_r$,
3. Transformation de Fourier Discrète des échantillons ainsi obtenus dans la case distance $c\tau/2$ pour résoudre le Doppler f_d .

Par rapport au schéma du récepteur idéal qui nécessite une batterie de corrélateurs en retard et en Doppler (plus précisément une batterie de filtres Doppler pour chaque retard), ce récepteur ne nécessite qu'un seul corrélateur par porte distance. L'analyse en distance s'effectue simplement par échantillonnage du signal en sortie du corrélateur aux différents instants τ possibles. L'analyse Doppler est réalisée en prenant, dans chaque porte distance, la transformée de Fourier Discrète des échantillons obtenus dans celle-ci tous les T_r . Cette procédure, schématisée sur la figure (FIG. 4.1), est couramment employée dans les radars actuels. Elle est d'ailleurs utilisée dans le processus d'acquisition des données expérimentales, présentées dans ce chapitre. En considérant l'expression (4.11) du filtre adapté, ces données sont issues du résultat de l'intégrale, comme fonction de τ et de k .

Remarque : *La corrélation temporelle du signal reçu par l'impulsion élémentaire est la sortie classique du filtre adapté à la vitesse $v = 0$, quelle que soit la vitesse de la cible analysée.*

Les traitements décrits ci-dessus et conduisant au *filtre adapté* découlent du caractère Gaussien du bruit. Si cette hypothèse n'est plus vérifiée (par exemple si le fouillis non-Gaussien est pris en compte), de fausses alarmes, déclenchées par des pointes de bruit, apparaissent en nombre considérable lors de la détection.

L'objet du paragraphe suivant est d'effectuer un parallèle entre les traitements existants et les détecteurs non-Gaussiens proposés dans le chapitre 3. L'étude se limite au cas des deux versions du BORD, à savoir le BORD et le BORD Asymptotique.

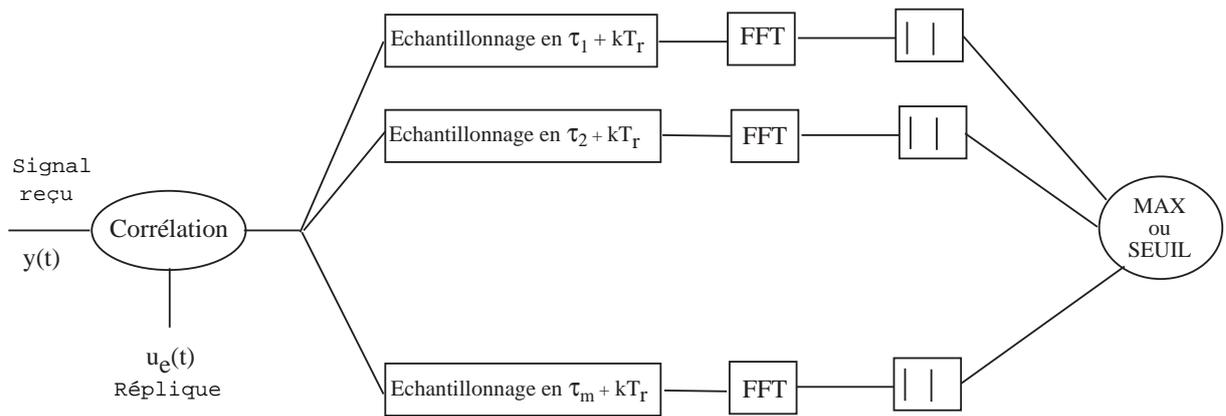


FIG. 4.1 – Récepteur optimal (OGD) en bruit blanc pour une émission d'un train de m impulsions

4.2 Mise en place des deux versions du BORD

4.2.1 Le BORD

Pour améliorer la structure des récepteurs existants en insérant un détecteur optimal tel le BORD, il est nécessaire de décomposer les différentes étapes réalisées dans ce détecteur. Dans un souci de linéarité dans la lecture, rappelons ici l'expression du BORD adaptatif :

$$\Lambda_{N_r}(\mathbf{y}_{obs}) = \frac{\sum_{i=1}^{N_r} \left[\frac{\mathbf{r}_i^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{r}_i}{(\widehat{q}_1(\mathbf{y}_{obs}) + \mathbf{r}_i^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{r}_i)^2} \right]^m}{\sum_{i=1}^{N_r} \left[\frac{\mathbf{r}_i^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{r}_i}{(\widehat{q}_0(\mathbf{y}_{obs}) + \mathbf{r}_i^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{r}_i)^2} \right]^m} \underset{H_0}{\overset{H_1}{>}} \eta, \quad (4.12)$$

où :

- * $\mathbf{r}_1, \dots, \mathbf{r}_{N_r}$ sont N_r données de référence supposées ne contenir que du fouillis,
- * \mathbf{M} est la matrice de covariance des vecteurs Gaussiens qui composent le SIRV,
- * $\widehat{\mathbf{M}}^{-1}$ est la matrice de covariance inverse estimée de \mathbf{M} , blanchissant les données,
- * $\widehat{q}_0(\mathbf{y}_{obs}) = \mathbf{y}_{obs}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs}$,
- * $\widehat{q}_1(\mathbf{y}_{obs}) = q_0(\mathbf{y}_{obs} - \mathbf{s})$ pour un signal cible \mathbf{s} connu, et
- * $\widehat{q}_1(\mathbf{y}_{obs}) = \mathbf{y}_{obs}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs} - \frac{|\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs}|^2}{\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{p}}$ si l'amplitude est estimée au sens du MV.

La première étape consiste à estimer \mathbf{M} à partir des données de référence. Ensuite, une partie du calcul de q_1 , soit l'expression $|\mathbf{p}^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{y}_{obs}|^2$, réalise intrinsèquement la dernière étape du filtre adapté classique, à savoir, la transformée de Fourier discrète et l'étape quadratique. On considère que le vecteur \mathbf{p} a déjà été "élu" à l'issue des étapes précédentes.

Le détecteur BORD peut alors être mis en place dans son intégralité, en calculant les autres quantités. Sa structure, représentée sur la figure (FIG. 4.2), est certes plus lourde que le filtre adapté, mais permet d'améliorer considérablement les performances de ce dernier.

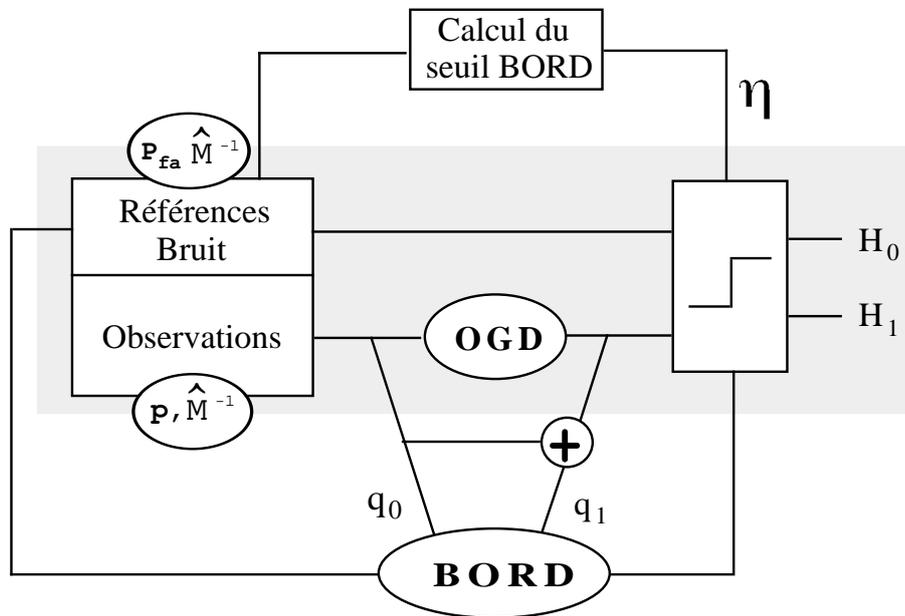


FIG. 4.2 – Mise en place du BORD en mode opérationnel, à partir du filtre adapté classique

4.2.2 Le BORD Asymptotique

L'expression du BORD Asymptotique s'exprime en fonction du filtre adapté. Pour s'en convaincre, il suffit déjà de remarquer, que le filtre adapté classique se réécrit comme la différence $q_0 - q_1$. Ensuite, en manipulant l'expression du BORD Asymptotique, qui est :

$$\left(\frac{q_0}{q_1}\right)^m \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (4.13)$$

on retrouve aisément le filtre adapté, ce qui donne :

$$\left(\frac{1}{1 - \frac{q_0 - q_1}{q_0}}\right)^m \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (4.14)$$

Il suffit donc, après avoir effectué les étapes du filtre adapté ($q_0 - q_1$), de le diviser par q_0 et d'effectuer la transformation décrite ci-dessus.

Le schéma de la figure (FIG. 4.3) résume les étapes de la mise en place du BORD Asymptotique, à partir du filtre adapté classique.

La suite de ce chapitre est consacrée à l'évaluation des performances de détection du BORD et du BORD Asymptotique en environnement de fouillis réel, comparées aux performances du détecteur classique. Les données opérationnelles de fouillis sont décrites, analysées, et des courbes de performances sont tracées pour une cible factice d'amplitude inconnue.

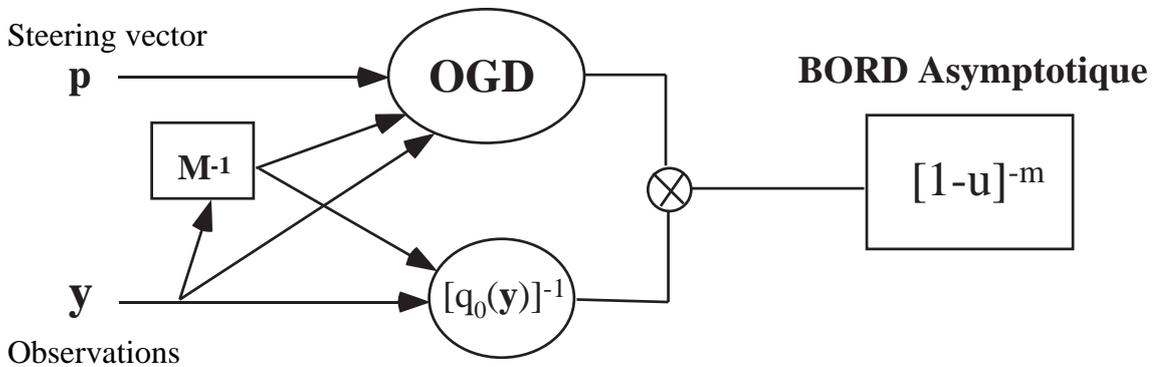


FIG. 4.3 – Mise en place du BORD Asymptotique en mode opérationnel, à partir du filtre adapté classique

4.3 Analyse des données expérimentales

Les données traitées comportent un caractère fortement non-Gaussien dans les zones les plus proches du radar. Au-delà d'une certaine distance, seul du bruit thermique est enregistré, ce qui caractérise la faible incidence du radar. L'horizon radio-électrique se trouve à $\sqrt{17h} \sim 15$ km de ce dernier [Bart88], où $h = 13$ m est la hauteur à laquelle est placé le radar. Au-delà, l'onde émise se perd à l'horizon, sans rencontrer de réflecteur, et seul le bruit thermique de l'électronique du système est enregistré.

4.3.1 Description des données

Les données représentent le signal écho enregistré par un radar opérationnel de THALES TAD, placé à une hauteur de 13 m, et illuminant une zone de terrain à site bas, 0° . Les échos du fouillis de sol (collines de la région de Colmar) sont situés dans différentes cases distance (au nombre de 868), pour différents angles de visée en azimut de l'antenne (au nombre de 70, couvrant chacun environ $1/8$ de tour d'antenne), et pour $m = 8$ récurrences. Ces récurrences représentent la fréquence de répétition de l'impulsion, ou Pulse Repetition Frequency (PRF), et, par abus de langage, on dira que chaque récurrence est une PRF.

La première case distance se situe à 2 km du radar. Chacune est de longueur 60 mètres, ce qui signifie que le radar éclaire un peu plus de 52 km de terrain. L'illustration des données sur les cartes de fouillis de la figure (FIG. 4.4), permet de constater la non-homogénéité spatiale (par case distance), ainsi qu'une forte corrélation temporelle (de PRF à PRF). La faible incidence du radar, entraînant une absence de réflecteurs au-delà de l'horizon radio-électrique est très bien illustrée sur ces cartes de fouillis, par la présence des zones de faible amplitude, significative du bruit thermique Gaussien. La figure (FIG. 4.5) illustre le rapport des puissances entre le fouillis et le bruit thermique (CNR - Clutter to Noise Ratio) par azimut, comparé à la puissance totale moyenne des données, du fouillis et du bruit thermique.

Les $m = 8$ PRF correspondent, dans notre cas, à la taille d'un vecteur de référence ou d'observation. Le nombre total d'observations (ou de références) nous est fourni par le nombre de cases distance, soit $N_{ref,obs} = 868$. Sachant que les données sont enregistrées pour 70 azimuts différents, le nombre total de données disponibles est de $70 \times 868 = 60760$ données.

La non-homogénéité spatiale du fouillis nous incite, dans un premier temps, à sélectionner les données par zones homogènes (bruit Gaussien blanc ou fouillis impulsif). Ensuite des zones transitoires, dans lesquelles se situe une rupture bruit thermique / fouillis, sont traitées.

De manière générale, le nombre de données testées est égal au produit du nombre de cases distance sélectionnées par le nombre d'azimuts pris en compte.

Dans la suite, une "zone" désigne à la fois une sélection de plusieurs données et une procédure d'évaluation de performances. On en distingue 8 (décrites au paragraphe 4.3.2), pour lesquelles plusieurs figures sont tracées : celles concernant la caractérisation du fouillis et celles concernant l'analyse des performances de détection.

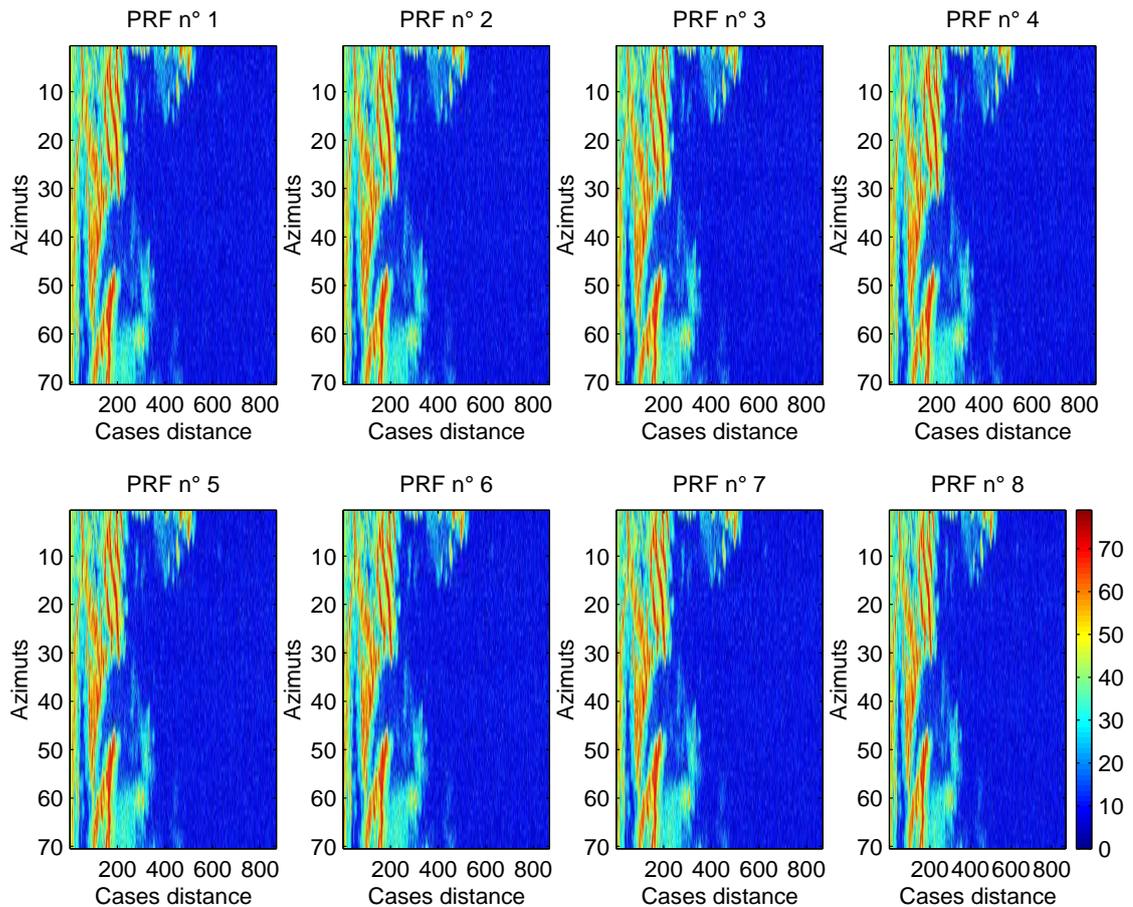


FIG. 4.4 – Cartes de fouillis de sol issu d’enregistrement réel par un radar opérationnel. Ces cartes montrent le niveau de fouillis (en dB) pour 70 angles d’azimuts, 868 cases distance et pour 8 PRF.

4.3.2 Définition des ”zones” de données

Pour évaluer les performances des détecteurs dans des situations diverses, 8 ”zones” de données, résumées dans le tableau (TAB. 4.1), sont définies :

- * Les ”zones 1, 2, 3” sélectionnent des données de référence, uniquement pour l’apprentissage du BORD. Les données d’observation sont utilisées pour évaluer les performances de tous les détecteurs, et représentent, respectivement, du bruit thermique, du fouillis et une partie transitoire bruit thermique / fouillis.
- * Les ”zones 4 et 5”, zones de bruit thermique et de fouillis, ne font intervenir que des données d’observation pour tous les détecteurs. L’estimation de la matrice de covariance s’effectue donc à partir des mêmes données pour tous les détecteurs, ainsi que le calcul du seuil de détection.
- * Les ”zones 6, 7 et 8”, représentatives de bruit thermique, de fouillis et d’une partie transitoire, utilisent des données de référence pour l’apprentissage de tous les détecteurs. Les

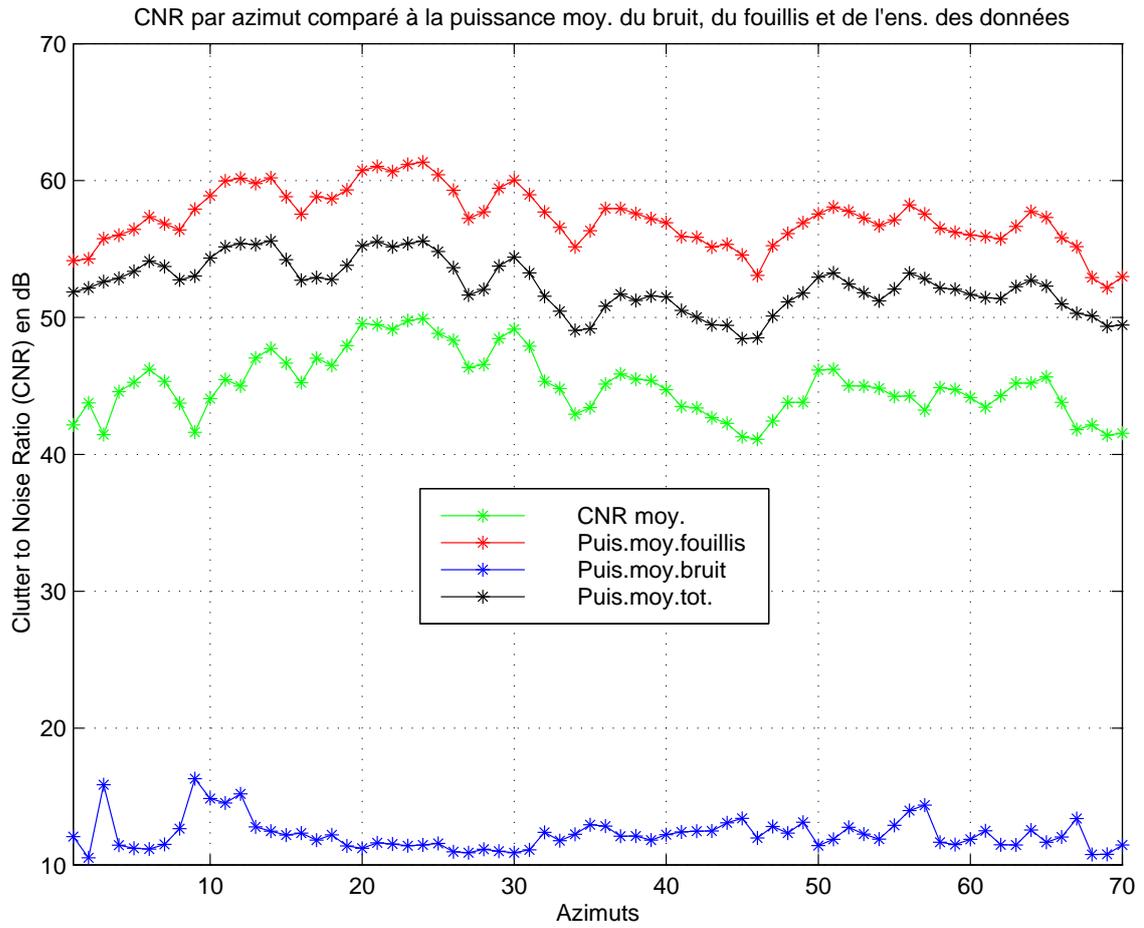


FIG. 4.5 – Evolution du CNR - Clutter to Noise Ratio - par azimuth, comparé à la puissance totale moyenne du fouillis, du bruit thermique et de l'ensemble des données.

données d'observation, de même nature que les références, permettent d'évaluer les performances des détecteurs.

"Zone"	Références	Observations	Apprentissage sur Références	Courbes de performances
Zone 1	Bruit Thermique C.d. : 478 à 642 Az. : 33 à 45 FIG. 4.18 (gauche)	Bruit Thermique C.d. : 560 à 706 Az. : 55 à 65 FIG. 4.18 (droite)	BORD	FIG. 4.20
Zone 2	Fouillis C.d. : 14 à 170 Az. : 1 à 31 FIG. 4.27 (gauche)	Fouillis C.d. : 10 à 140 Az. : 31 à 41 FIG. 4.27 (droite)	BORD	FIG. 4.29
Zone 3	Transition C.d. : 60 à 354 Az. : 42 à 57 FIG. 4.36 (gauche)	Transition C.d. : 178 à 308 Az. : 1 à 13 FIG. 4.36 (droite)	BORD	FIG. 4.38
Zone 4		Bruit Thermique C.d. : 568 à 774 Az. : 22 à 38 FIG. 4.21	non	FIG. 4.23
Zone 5		Fouillis C.d. : 22 à 212 Az. : 18 à 32 FIG. 4.30	non	FIG. 4.32
Zone 6	Bruit Thermique C.d. : 375 à 552 Az. : 22 à 34 FIG. 4.24 (gauche)	Bruit Thermique C.d. : 641 à 812 Az. : 38 à 51 FIG. 4.24 (droite)	OGD, OKD, BORD, BORD As.	FIG. 4.26
Zone 7	Fouillis C.d. : 10 à 168 Az. : 31 à 44 FIG. 4.33 (gauche)	Fouillis C.d. : 16 à 78 Az. : 2 à 27 FIG. 4.33 (droite)	OGD, OKD, BORD, BORD As.	FIG. 4.35
Zone 8	Transition C.d. : 94 à 250 Az. : 37 à 52 FIG. 4.39 (gauche)	Transition C.d. : 176 à 306 Az. : 6 à 22 FIG. 4.39 (droite)	OGD, OKD, BORD, BORD As.	FIG. 4.41

TAB. 4.1 – Définitions des "zones" utilisées pour l'évaluation des performances de l'OGD, de l'OKD, du BORD et du BORD Asymptotique

4.3.3 Caractérisation du fouillis lors de la détection

Les ruptures fouillis non-Gaussien / bruit thermique se situent entre la 200^{ème} et la 500^{ème} case distance. Pour préciser leurs localisations, la figure (FIG. 4.6) indique, pour chaque azimuth, le numéro de la case distance à laquelle la rupture définitive est observée. La courbe est approximative, mais permet une sélection des zones de fouillis homogènes ou hétérogènes.

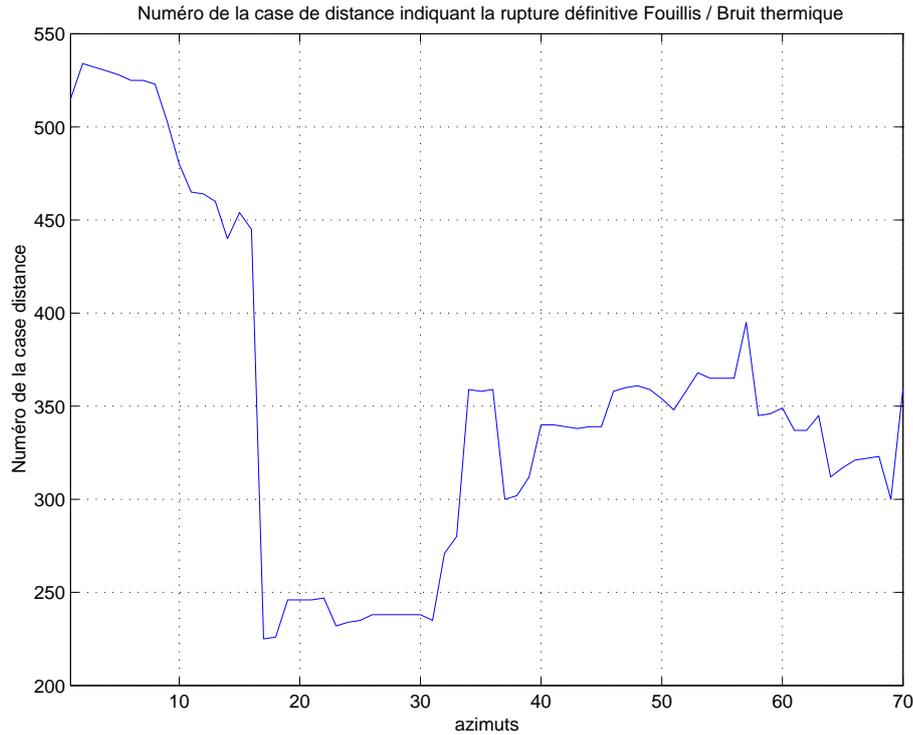


FIG. 4.6 – Case distance où la rupture définitive fouillis non-Gaussien / bruit thermique a lieu, pour chaque azimuth. Moyenne sur les 8 PRF.

Pour illustrer le caractère composé du processus de fouillis, sa puissance instantanée est tracée pour chaque zone de même nature explorée et testée pour la détection (FIG. 4.9, FIG. 4.12, FIG. 4.15). La variation de ces puissances représente la variation des échantillons de texture au cours du temps, pour un azimuth donné, dans toutes les cases distance et pour les 8 PRF. On constate ainsi que les puissances sont de même niveau d'une PRF à l'autre (superposition des courbes), alors qu'elles varient d'une case distance à l'autre.

Sur les figures (FIG. 4.10, FIG. 4.13, FIG. 4.16), on constate que les phases entre les voies I et Q des données sont uniformément réparties entre $-\pi$ et $+\pi$, ce qui confirme les suppositions théoriques.

Pour chaque zone explorée, une comparaison à la K-distribution est effectuée (FIG. 4.11, FIG. 4.14, FIG. 4.17). Pour ce faire, on estime les paramètres ν et b d'une K-distribution supposée en utilisant la méthode des moments. Les histogrammes empiriques et théoriques sont comparés et les valeurs estimées de ν sont reportées pour chacun d'entre eux. Cette estimation est réalisée à partir de l'amplitude des données de chaque zone, pour chacune des $m = 8$ PRF. Les paramètres ainsi estimés sont utilisés, par la suite, pour la mise en place de l'OKD.

Sur ces mêmes figures, les résultats des approximations de Padé de la loi de l'enveloppe des données sont reportés. Les approximations de Padé sont calculées à partir des moments empiriques de l'enveloppe des données pour chaque PRF, et l'ordre des approximations est indiqué.

La dernière étape de caractérisation du fouillis consiste en l'estimation de la matrice de covariance des données. Selon le détecteur utilisé, cette étape est réalisée, soit à partir de données de référence (cas du BORD), soit sur les données d'observation sans cible (cas de l'OGD, de l'OKD et du BORD Asymptotique). Pour l'OGD, on utilise l'estimation empirique classique (SCM), alors que pour les trois autres détecteurs, l'estimée NSCM (donnée en (3.12)) est calculée. Les matrices estimées sont représentées pour certaines zones sur les figures (FIG. 4.7, FIG. 4.8).

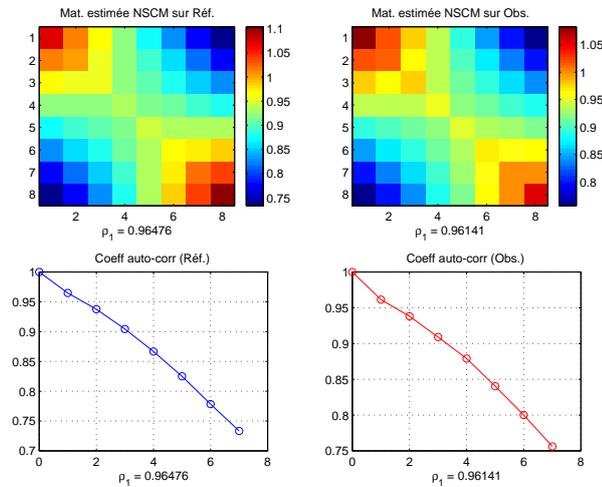


FIG. 4.7 – Zone de fouillis : Estimations NSCM des matrices de covariance des références et des observations de la "zone 2" et coefficients d'auto-corrélation correspondants.

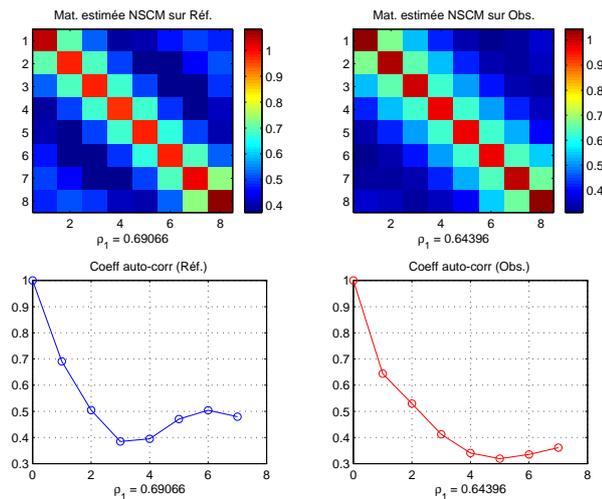


FIG. 4.8 – Zone transitoire bruit thermique / fouillis : Estimations NSCM des matrices de covariance des références et des observations de la "zone 8" et coefficients d'auto-corrélation correspondants.

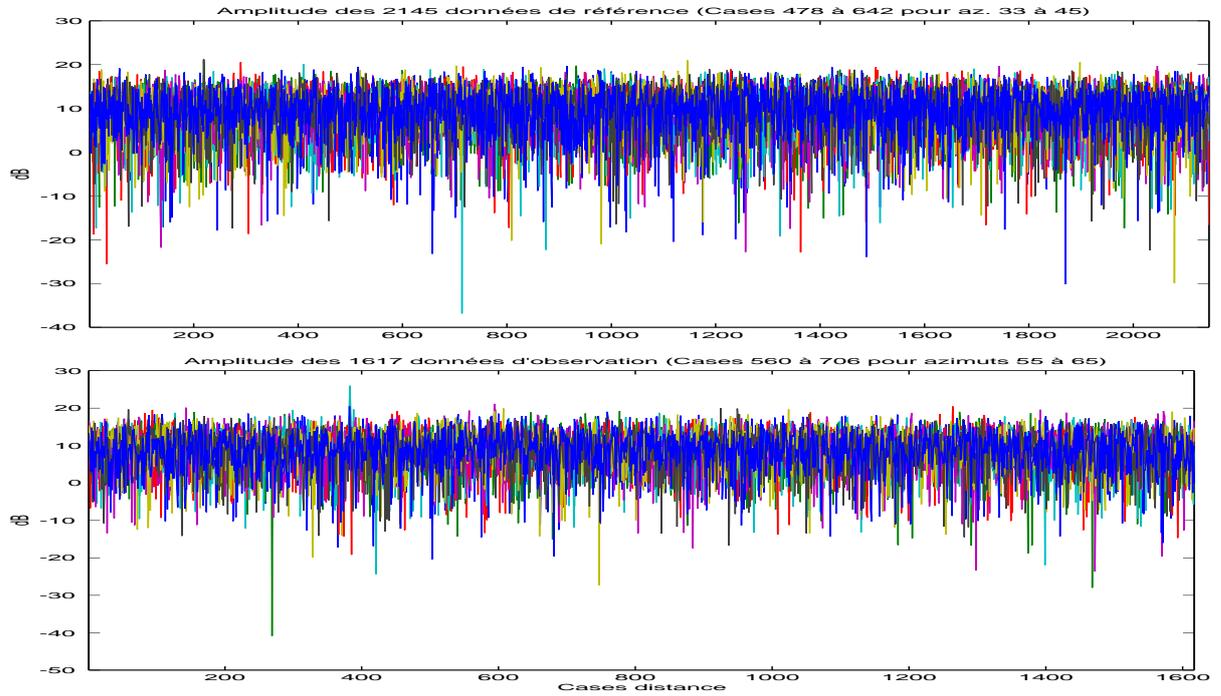


FIG. 4.9 – Zone de bruit thermique ("Zone 1") : Amplitude des données de référence (en haut) et d'observation (en bas).

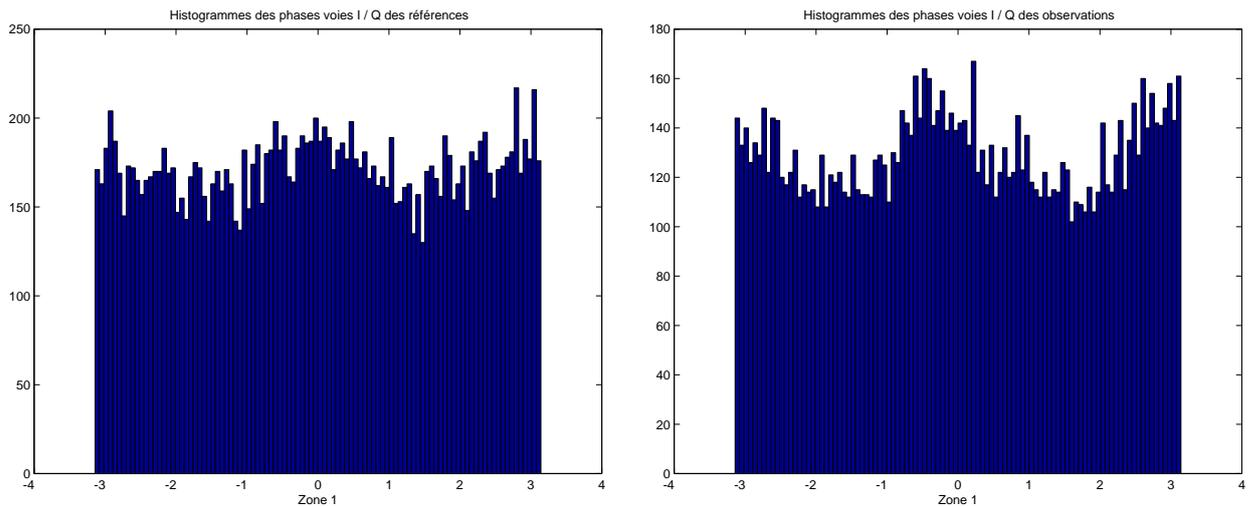


FIG. 4.10 – Zone de bruit thermique ("Zone 1") : Histogrammes des 8×2145 phases des données de référence (à gauche) et des 8×1617 données d'observation (à droite).

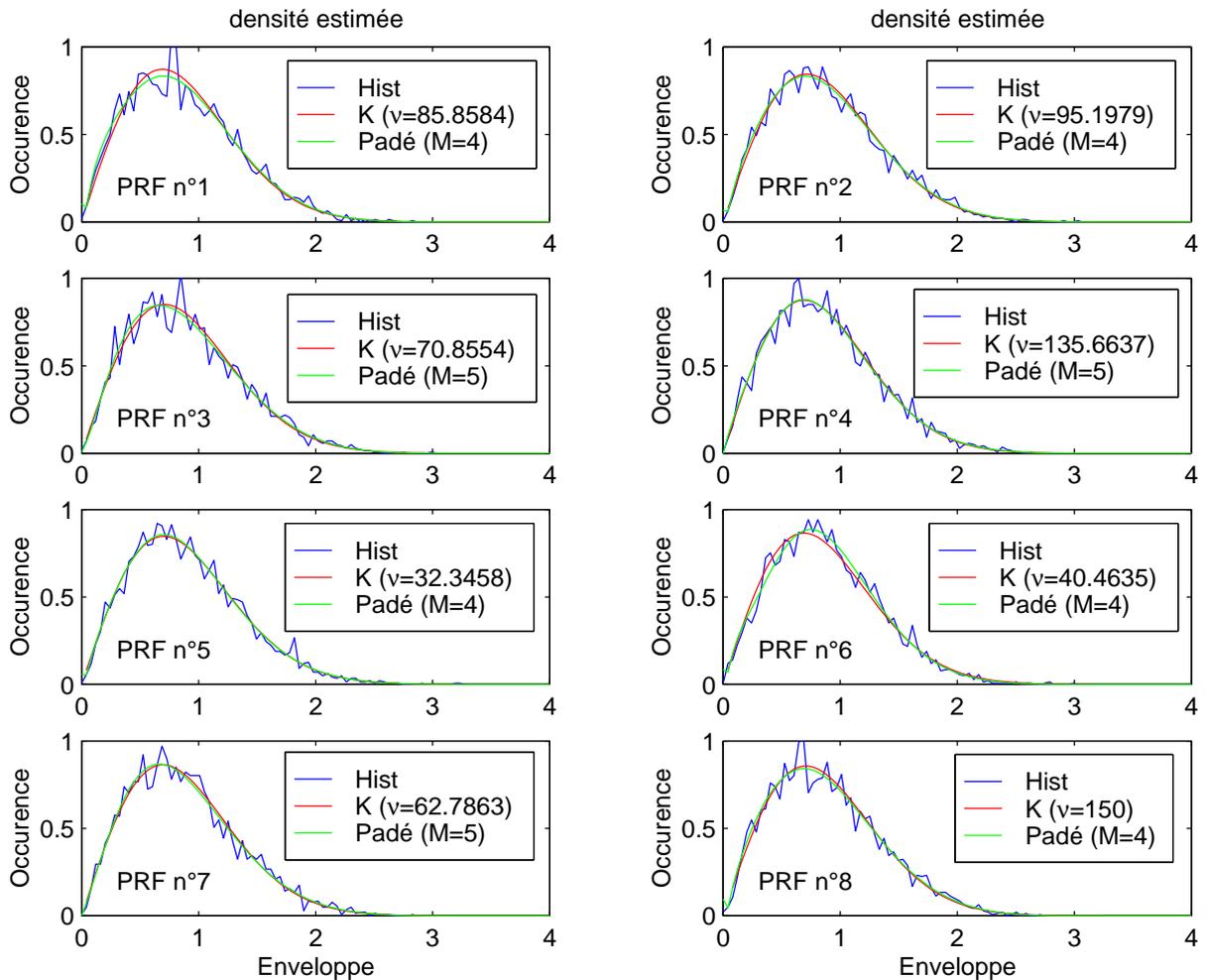


FIG. 4.11 – Zone de bruit thermique ("Zone 4") : Histogrammes empiriques (en bleu), lois $K(\nu, b)$ estimées (en rouge) et approximations de Padé (en vert) de l'enveloppe des observations de bruit thermique pour chaque PRF.

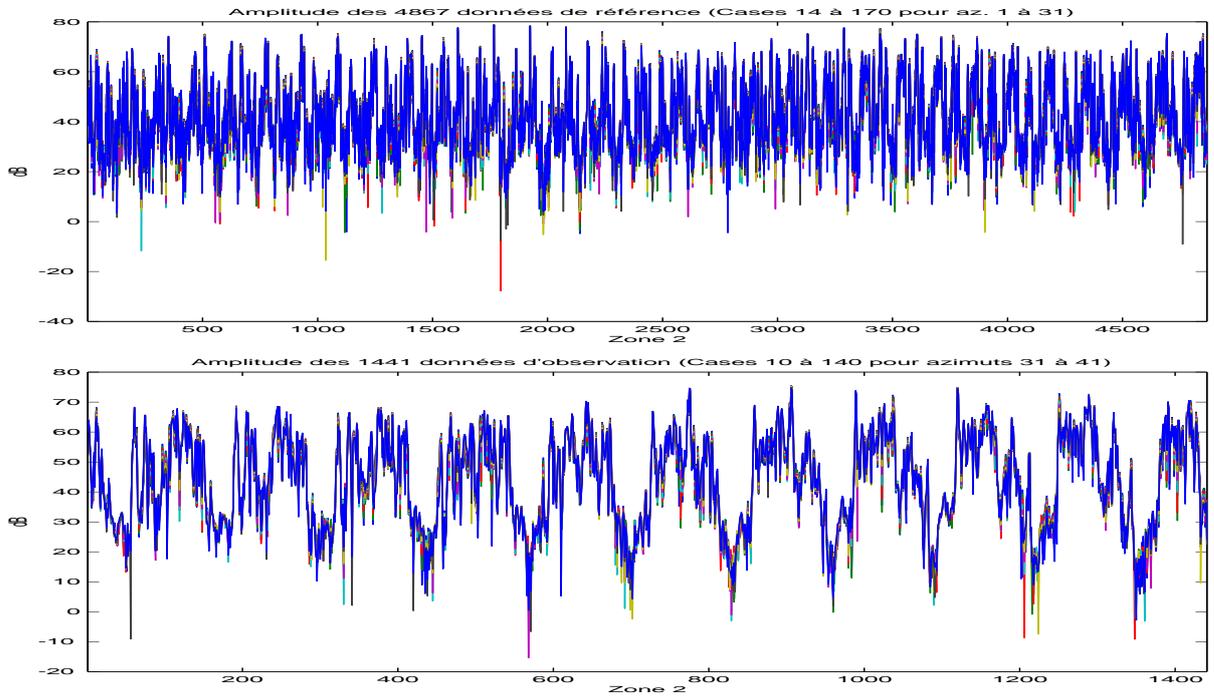


FIG. 4.12 – Zone de fouillis ("Zone 2") : Amplitude des données de référence (en haut) et d'observation (en bas).

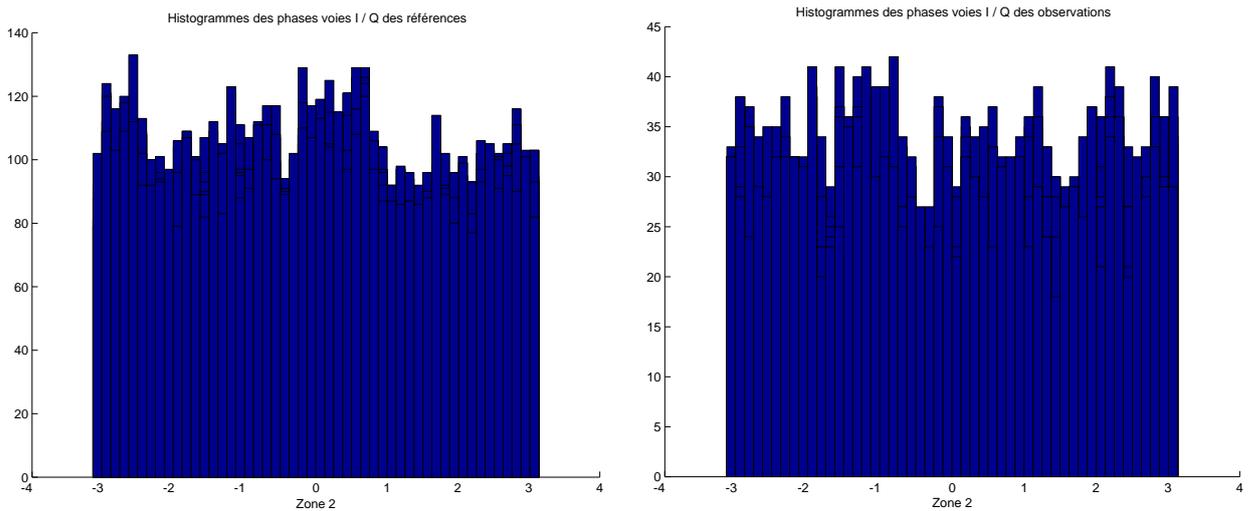


FIG. 4.13 – Zone de fouillis ("Zone 2") : Histogrammes des 8×4867 phases des données de référence (à gauche) et des 8×1441 données d'observation (à droite).

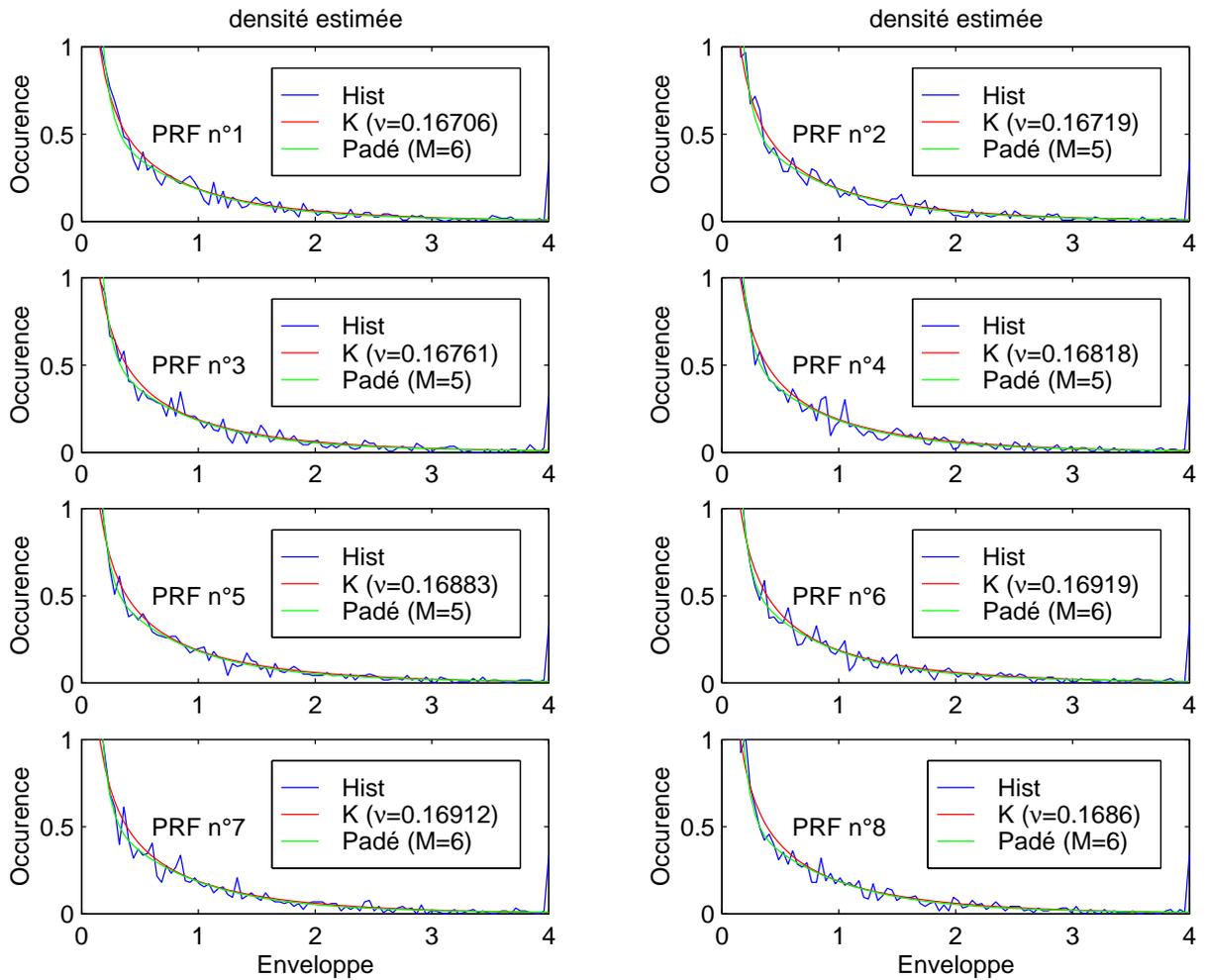


FIG. 4.14 – Zone de fouillis ("Zone 5") : Histogrammes empiriques (en bleu), lois $K(\nu, b)$ estimées (en rouge) et approximations de Padé (en vert) de l'enveloppe des observations de fouillis pour chaque PRF.

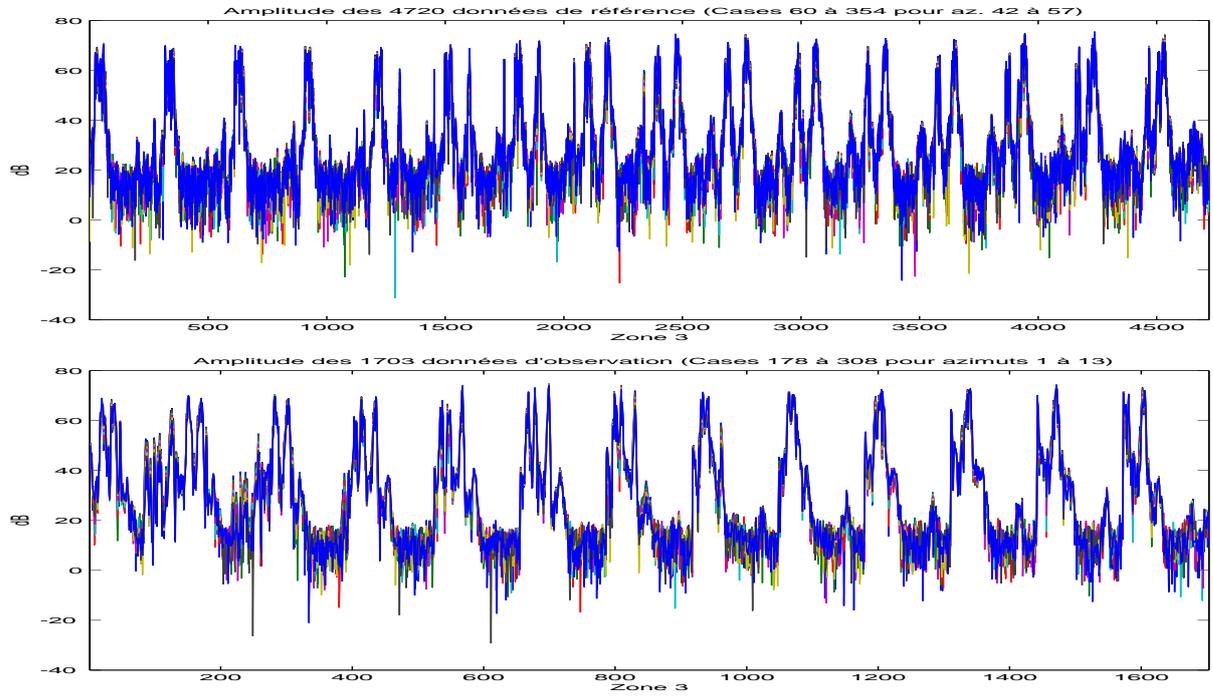


FIG. 4.15 – Zone transitoire bruit thermique / fouillis ("Zone 3") : Amplitude des données de référence (en haut) et d'observation (en bas).

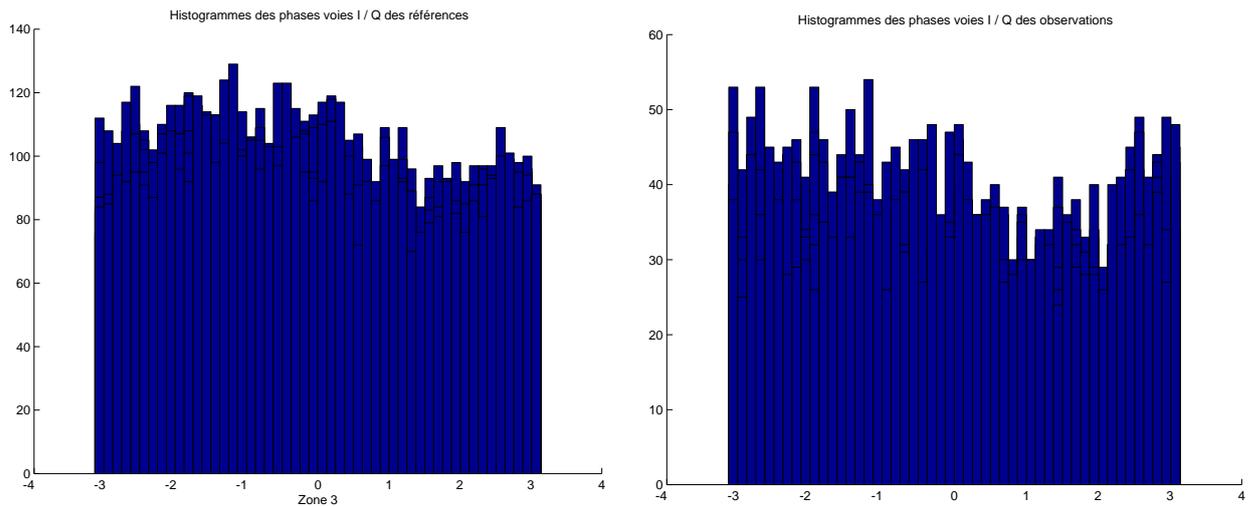


FIG. 4.16 – Zone transitoire bruit thermique / fouillis ("Zone 3") : Histogrammes des 8×4720 phases des données de référence (à gauche) et des 8×1703 données d'observation (à droite).

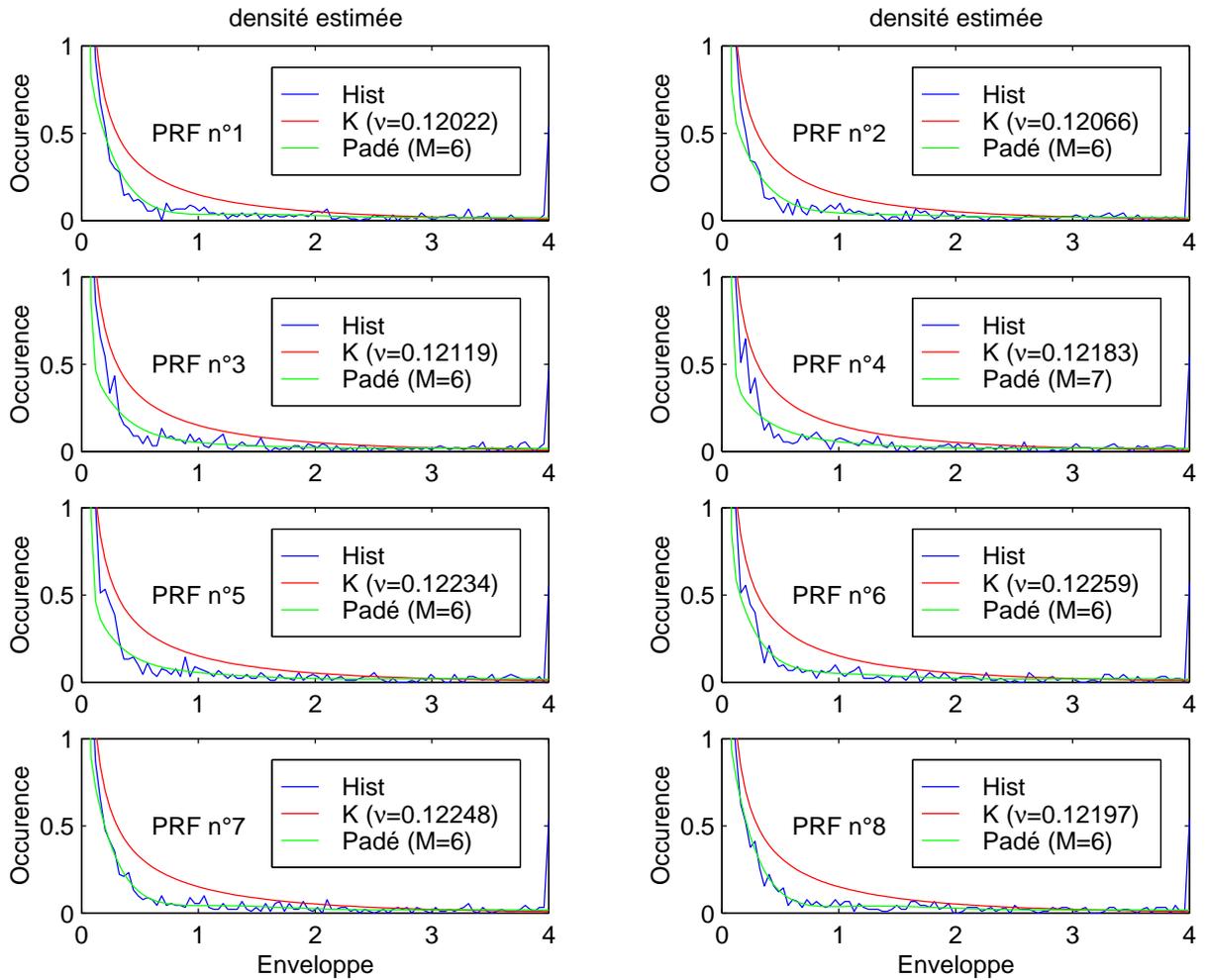


FIG. 4.17 – Zone transitoire bruit thermique / fouillis ("Zone 8") : Histogrammes empiriques (en bleu), lois $K(\nu, b)$ estimées (en rouge) et approximations de Padé (en vert) de l'enveloppe des observations pour une zone transitoire bruit thermique / fouillis et pour chaque PRF.

4.4 Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique

La matrice de covariance est estimée, soit classiquement (SCM), pour l'OGD, soit d'après l'estimation NSCM pour l'OKD, le BORD et le BORD Asymptotique. Selon le rang de la matrice obtenue, soit l'inverse, soit la pseudo-inverse est calculée, pour la mise en place des détecteurs. Les problèmes engendrés par le mauvais conditionnement de la matrice ne sont qu'évoqués dans les analyses de performances et n'ont pas constitué d'études plus poussées.

Pour la mise en place de l'OKD, les paramètres ν et b de la K-distribution, supposée régir la loi de l'amplitude des données, sont estimés pour chaque PRF. Les valeurs moyennes de ces paramètres sont retenues et alimentent ainsi l'OKD.

Le seuil de détection est ensuite déterminé, pour chaque détecteur, de manière à respecter la P_{fa} , fixée à 10^{-2} .

Ces étapes sont réalisées à partir de données, considérées comme étant les références. Elles correspondent, soit à des données annexes, soit aux observations, dans lesquelles sont évaluées les performances. La distinction est faite pour chaque "zone" étudiée.

Pour illustrer le procédé de calcul du seuil de détection, le logarithme du rapport de vraisemblance est tracé pour chaque détecteur, relativement à la "zone" testée. Ces courbes sont représentées sur les figures (FIG. 4.19, FIG. 4.28, FIG. 4.37, FIG. 4.22, FIG. 4.31, FIG. 4.25, FIG. 4.34, FIG. 4.40).

La valeur théorique du BORD Asymptotique, donnée par (3.70), est reportée dans les cas du BORD et du BORD Asymptotique, et la valeur théorique du seuil de détection, donnée par (1.18) dans le cas Gaussien, est reportée pour l'OGD.

Les performances des détecteurs sont ensuite évaluées pour chaque "zone" définie, à partir des données d'observation sélectionnées. Les premiers tests se rapportent aux "zones 1, 4, 6", zones de bruit thermique. Suivent les tests en environnement de fouillis des "zones 2, 5, 7", puis ceux pour les zones transitoires bruit thermique / fouillis, "zones 3, 8".

Après toute évaluation de performances, le Taux de Fausse Alarme (TFA) constaté est donné. Cela permet de vérifier si la P_{fa} initiale est respectée ou non, dans les données d'observation en absence de cible. Si les observations ont été utilisées comme références pour le calcul du seuil, ce TFA est alors égal à la P_{fa} . Sinon, il valide plus ou moins la valeur du seuil de détection.

4.4.1 Zones de bruit thermique Gaussien

4.4.1.1 "Zone 1"

On rappelle que la "zone 1" sélectionne des références pour seul l'apprentissage du BORD. La matrice de covariance et les paramètres de la supposée loi K sont estimés directement à partir des données d'observation, dans lesquelles les performances de détection sont ensuite évaluées. Les données de référence, sélectionnées pour l'apprentissage du BORD, sont désignées par le rectangle noir, sur la carte de gauche de la figure (FIG. 4.18). Ces données comprennent 165 cases distance (de la 478 à la 642) et 13 azimuts (du 33 au 45). Au total, $N_{ref} = 2145$. Sur la carte de droite, $N_{obs} = 1617$ données d'observation sont sélectionnées, pour 147 cases distance (de la 560 à 706) et 11 azimuts (du 55 au 65).

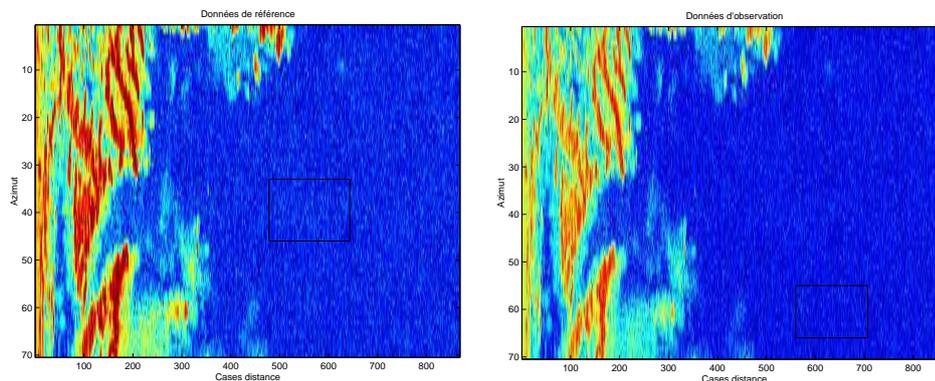


FIG. 4.18 – Zone 1 : Bruit thermique Gaussien pour les références (gauche) et les observations (droite). Au total, $N_{ref} = 2145$ et $N_{obs} = 1617$.

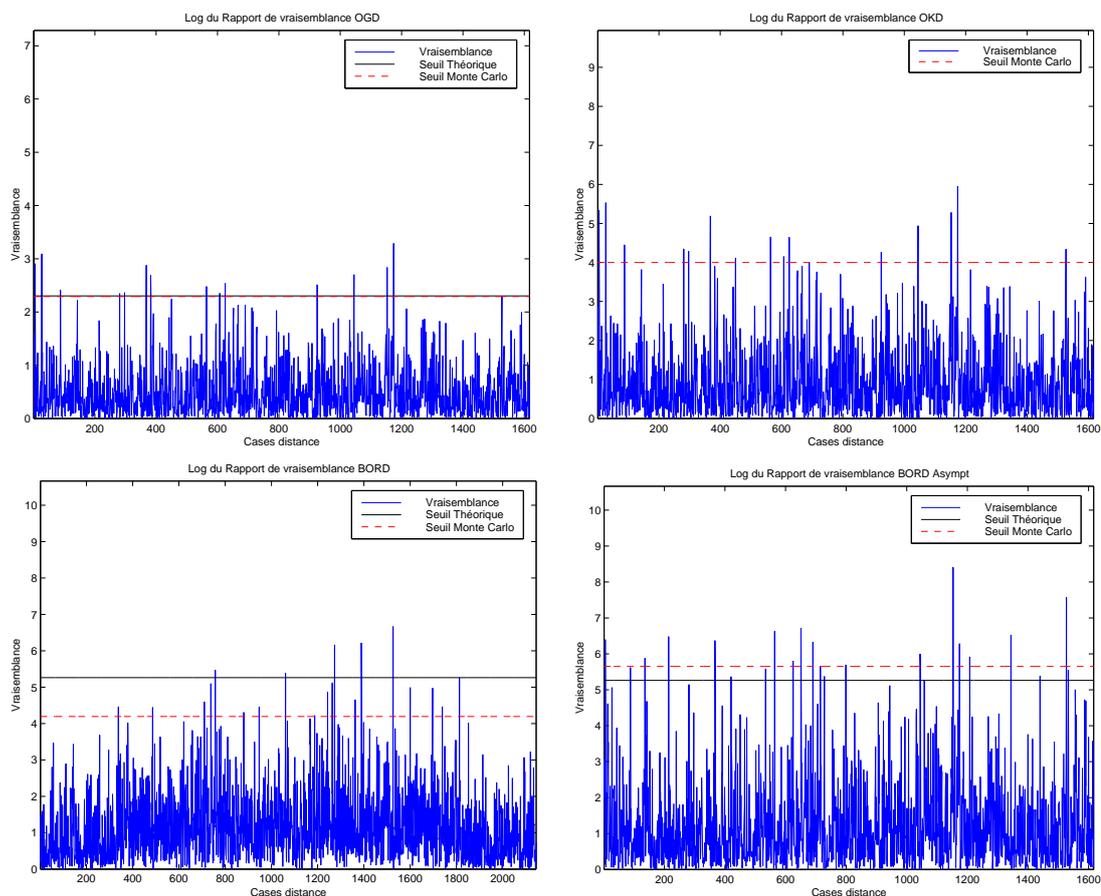


FIG. 4.19 – Zone 1 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).

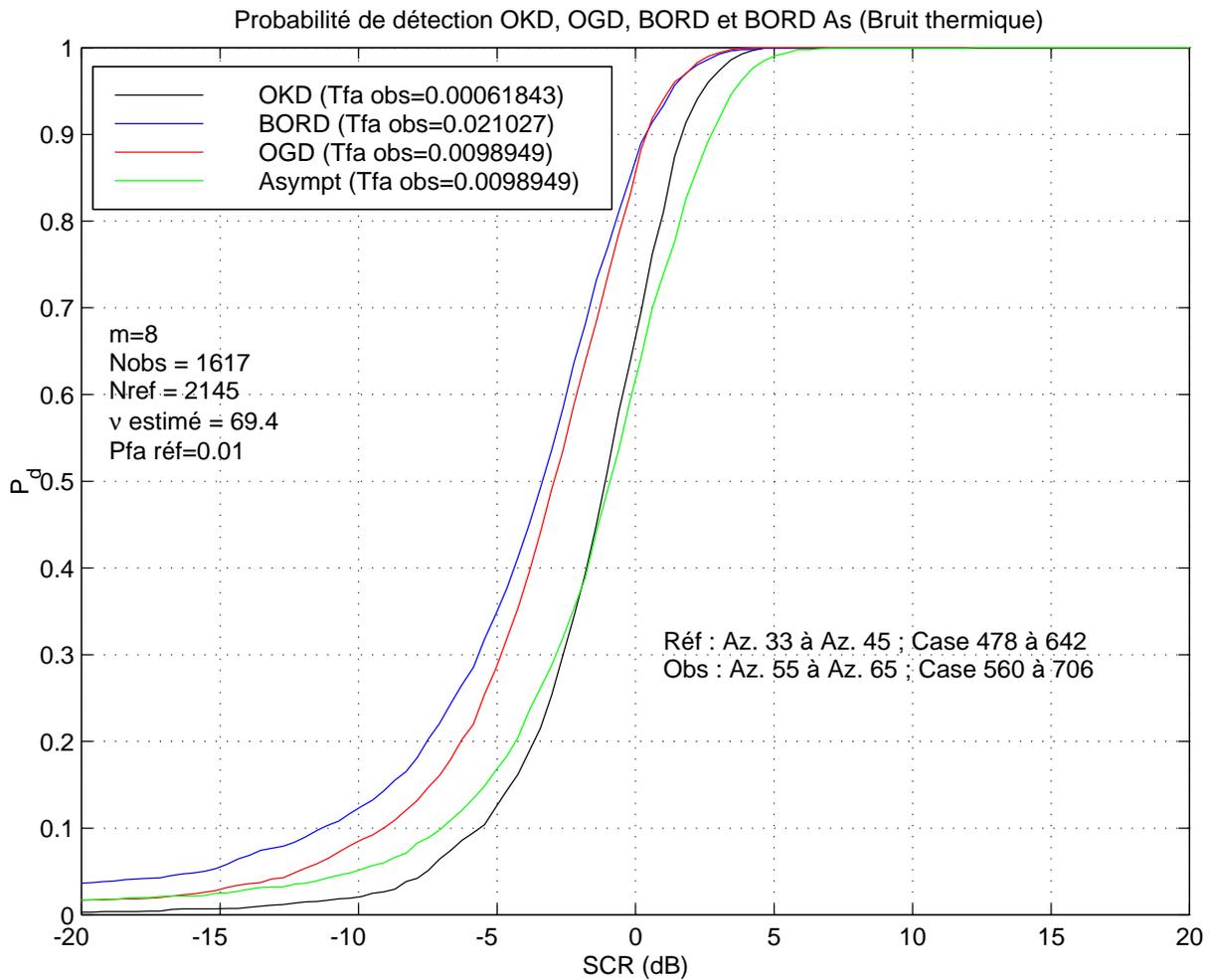


FIG. 4.20 – Zone 1 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du bruit thermique Gaussien

Analyses des performances de la "zone 1"

En environnement Gaussien, les détecteurs doivent fournir les mêmes performances de détection que l'OGD, dès lors que la probabilité de fausse alarme est respectée. Cela est le cas pour l'OGD et le BORD Asymptotique, dont les mêmes données ont été utilisées pour le calcul du seuil de détection et l'évaluation de leurs performances. Cette procédure est identique pour l'OKD, qui présente cependant un TFA plus faible que la P_{fa} demandée. Cela peut s'expliquer à cause des fortes valeurs estimées dans les différentes PRF pour le paramètre ν . Ces valeurs sont moyennées et la différence provient de ce moyennage. La courbe représentant les performances de l'OKD (en noir) devrait se trouver plus à gauche, et se rapprocher ainsi de celle de l'OGD.

En utilisant des données de référence uniquement pour le BORD, son TFA constaté sur les observations, est, de fait, un peu différent de la P_{fa} souhaitée initialement. Il est à peu près son double, soit $T_{FA} = 2.10^{-2}$, et les performances de détection, indiquées par le BORD ne sont pas celles attendues (FIG. 4.20). Leur courbe représentative (en bleu) devrait être plus en retrait sur la droite, si $P_{fa} = 10^{-2}$ était respectée. Ses performances seraient alors plus proches de celles de l'OGD.

Les performances du BORD Asymptotique décrochent par rapport à celles de l'OGD. En comparaison des performances du BORD, elles leur sont tout à fait parallèles, ce qui ne peut que confirmer son comportement asymptotique. Cet écart a déjà été constaté lors des simulations effectuées en environnement Gaussien (K-distribué de paramètre $\nu = 20$), présentées sur une des courbes de la figure (FIG. 3.20), du chapitre 3. Le BORD Asymptotique est mal adapté en environnement Gaussien, et cela se confirme sur les données expérimentales.

4.4.1.2 "Zone 4"

La "zone 4" comporte une procédure idéale, mais peu réaliste. On suppose que les données d'observation sont également utilisées comme références. Cela signifie que la P_{fa} est respectée pour tous les détecteurs (excepté peut-être pour l'OKD, en fonction des valeurs estimées des paramètres ν et b) et que les matrices de covariance estimées (SCM pour l'OGD et NSCM pour les autres) représentent vraiment la corrélation des données testées.

Les données sélectionnées sont représentées par le rectangle noir de la figure (FIG. 4.21), et comportent $N_{ref,obs} = 3296$ données, dont 206 cases distance (de la 568 à la 774) et 17 azimuts (du 22 au 38).

Sur la figure (FIG. 4.22), on constate que les valeurs théoriques des seuils de détection (pour l'OGD, le BORD et le BORD Asymptotique) sont représentatives de la P_{fa} demandée.

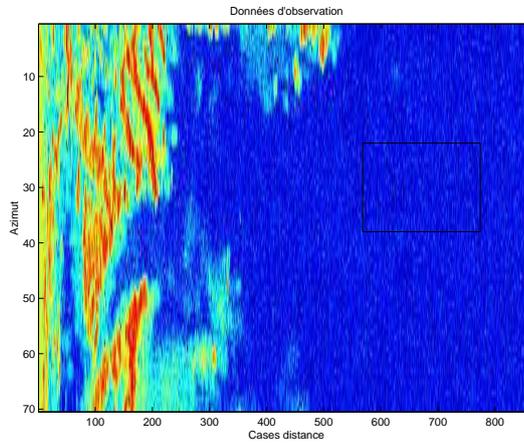


FIG. 4.21 – Zone 4 : Bruit thermique Gaussien des observations : $N_{obs} = 3296$.

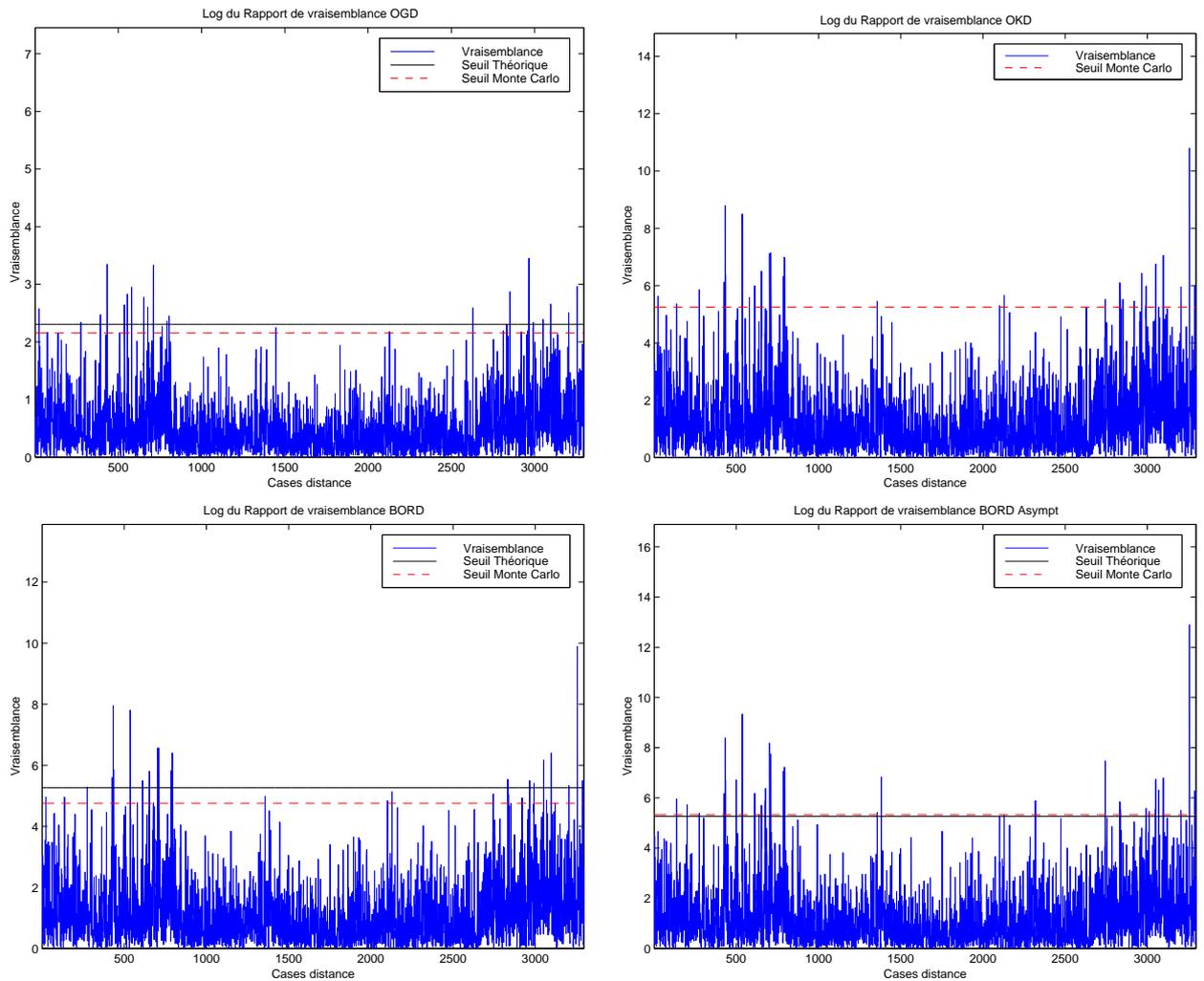


FIG. 4.22 – Zone 4 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l’OGD), ou au seuil théorique du BORD Asymptotique.

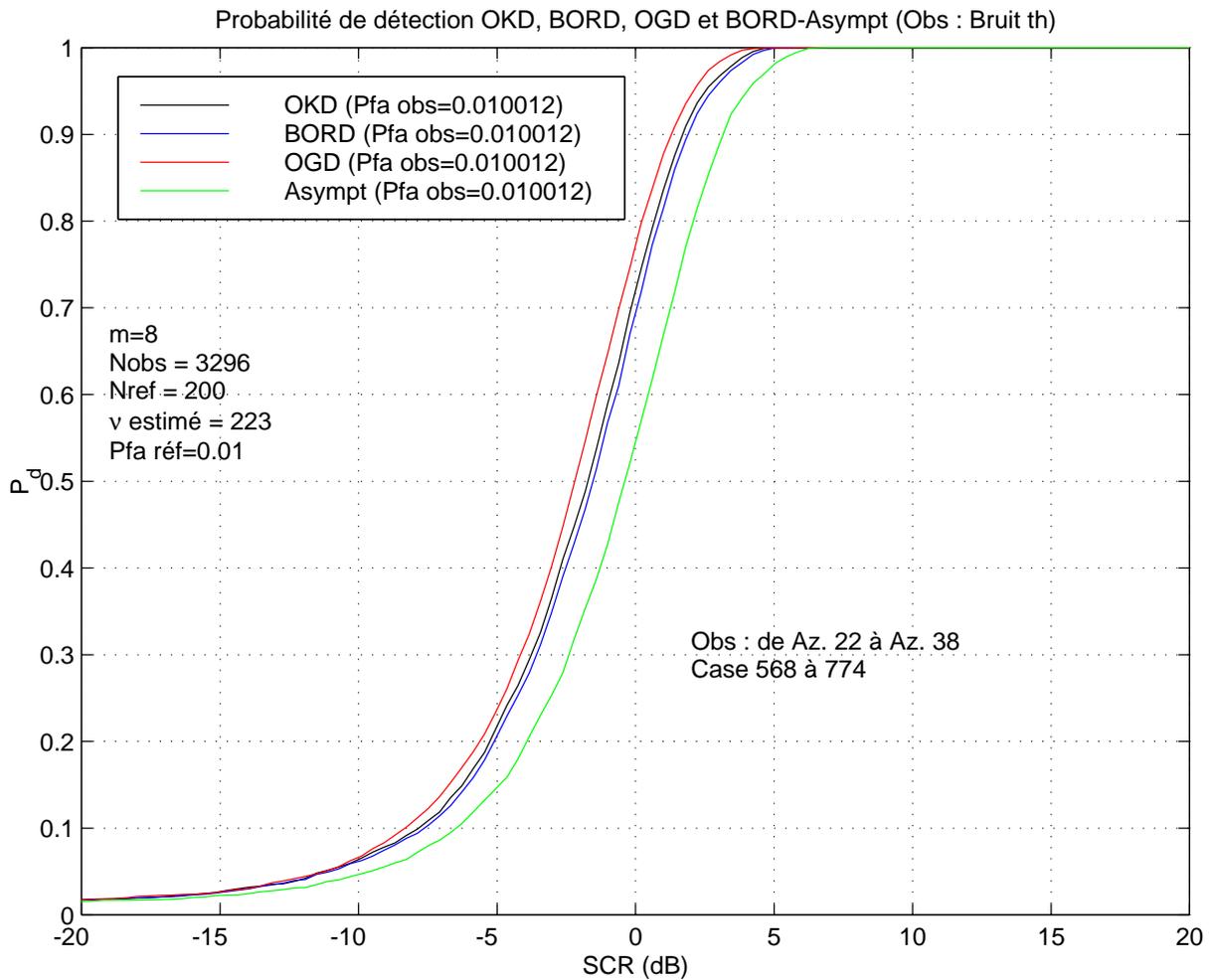


FIG. 4.23 – Zone 4 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du bruit thermique Gaussien

Analyses des performances de la "zone 4"

Les performances des détecteurs sont présentées sur la figure (FIG. 4.23). La procédure d'évaluation étant idéale, les performances indiquées pour cette "zone 4" sont donc "plus optimales" que celles de la "zone 1".

Les TFA constatés restent cohérents avec la P_{fa} demandée, pour tous les détecteurs, même pour l'OKD, dont le paramètre ν retenu est très élevé. L'OGD présente de meilleures performances que le BORD et l'OKD, malgré leur optimalité en simulations. Cependant, l'écart reste faible : à P_d constante (entre 0.1 et 1), l'OGD gagne un peu moins de 0.5 dB en RSB, et pour un RSB fixé, l'OGD gagne entre 3% et 7% en détectabilité.

En revanche, le BORD Asymptotique reste égal à lui-même en bruit Gaussien. Il perd à peu près 2 dB de RSB pour une P_d donnée, et de 10% à 25% en détectabilité pour un RSB fixé.

4.4.1.3 "Zone 6"

Dans la "zone 6", les caractéristiques du bruit thermique (matrice de covariance et paramètres de la loi K supposée) sont évaluées à partir des mêmes données de référence, pour tous les détecteurs (rectangle noir sur la carte de gauche de FIG. 4.24). Cette sélection comporte au total $N_{ref} = 2314$.

Les données d'observation, différentes des références pour tous les détecteurs, comportent $N_{obs} = 2408$ échantillons. Les performances des détecteurs sont évaluées dans cette zone, représentée sur la carte de droite de la figure (FIG. 4.24).

Les valeurs des seuils de détection, obtenues en sortie de chaque détecteurs sous l'hypothèse H_0 , pour respecter la $P_{fa} = 10^{-2}$ fixée initialement, sont tracées sur les courbes de la figure (FIG. 4.25) et comparées aux seuils théoriques. Comme dans la "zone 4", il existe une bonne adéquation entre la théorie et l'empirique.

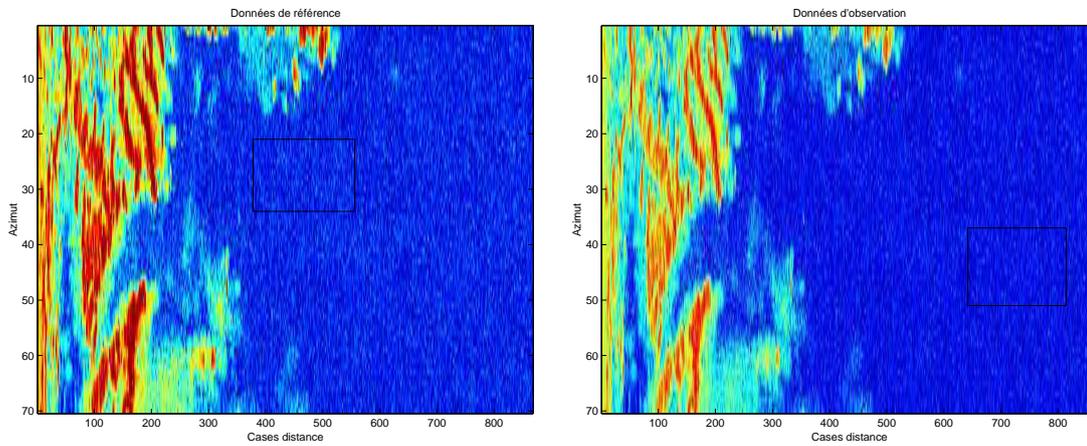


FIG. 4.24 – Zone 6 : Bruit thermique Gaussien des références (gauche) et des observations (droite) : $N_{obs} = 2408$, $N_{ref} = 2314$.

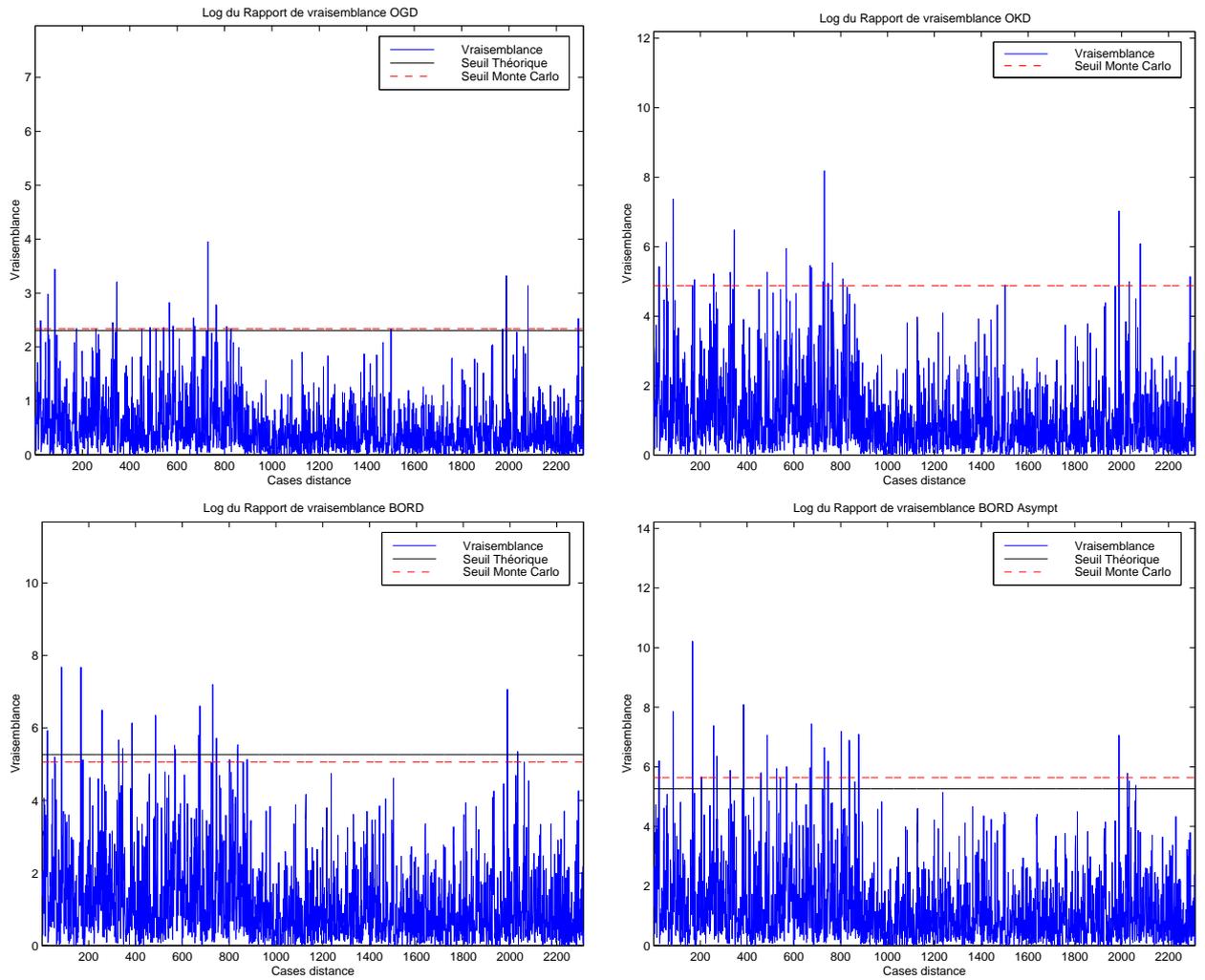


FIG. 4.25 – Zone 6 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien ou BORD Asymptotique.

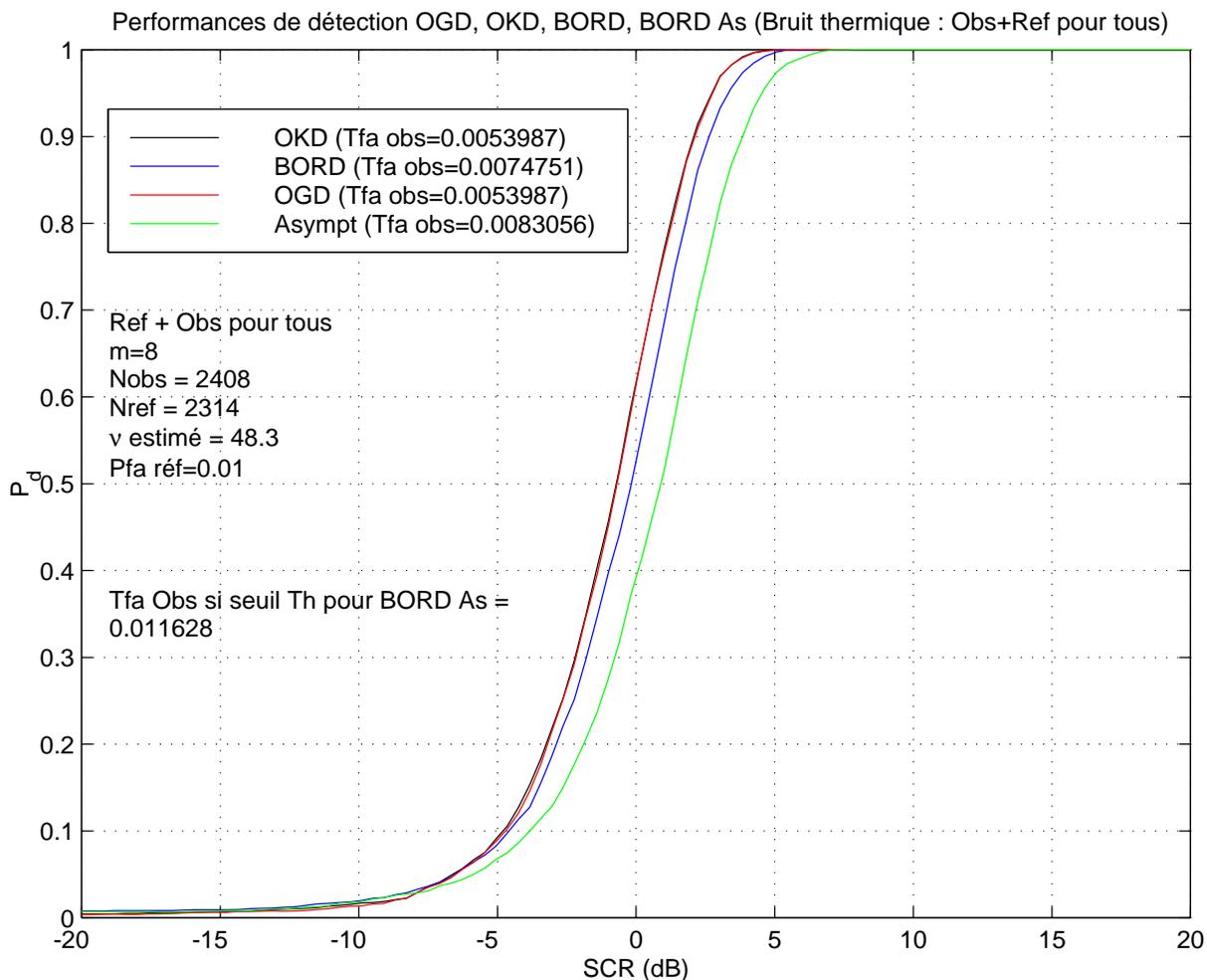


FIG. 4.26 – Zone 6 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du bruit thermique Gaussien

Analyses des performances de la "zone 6"

Pour les quatre détecteurs, la P_{fa} n'est pas exactement respectée. Cependant, les TFA constatés pour l'OGD et l'OKD sont les mêmes, ce qui permet d'avoir deux éléments de comparaison proches. De manière générale, les TFA sont plus faibles que la P_{fa} , ce qui signifie que toutes les courbes devraient se trouver plus à gauche qu'elles ne le sont. On constate, une fois encore, que le BORD approche les performances de l'OGD, avec toujours un écart faible, comparable à celui constaté pour la "zone 4". Le BORD Asymptotique présente également un écart du même ordre de grandeur que précédemment.

Sur cette figure, un TFA supplémentaire est indiqué. Il s'agit du TFA que l'on aurait constaté (comparable à la valeur de P_{fa}) si le seuil théorique du BORD Asymptotique avait été utilisé, au lieu du seuil obtenu par MC. Cet exemple est trop singulier pour devenir général, mais la validité de ce seuil théorique trouve toutefois une justification pratique.

Après avoir éprouvé les performances des détecteurs en environnement "classique" de bruit Gaussien, les trois prochaines "zones" testées se rapportent à des zones de fouillis fortement non-Gaussien.

4.4.2 Zones de fouillis

Dans les zones de fouillis, les mêmes comparaisons sont effectuées. Les caractéristiques du fouillis sont données sur les figures (FIG. 4.12, 4.13, FIG. 4.14), pour la "zone 2" et la "zone 5". Les paramètres estimés de la K-distribution semblent indiquer une bonne adéquation entre la statistique des données et cette loi. Pour la "zone 2", les estimations des matrices de covariance sont reportées sur la figure (FIG. 4.7).

4.4.2.1 "Zone 2"

Des données de référence (sur la carte gauche de FIG. 4.27) sont utilisées uniquement pour l'apprentissage du BORD. Elles comportent $N_{ref} = 4836$ échantillons, prélevés de 156 cases distance (de la 14 à la 170) dans 31 azimuts (du 1 au 31).

Les performances sont évaluées avec les données d'observation indiquées sur la carte droite de FIG. 4.27. On observe $N_{obs} = 1300$ données, provenant de 130 cases distance (de la 10 à la 140) pour 10 azimuts différents (du 31 au 41).

Comme dans la "zone 1", l'OGD, l'OKD et le BORD Asymptotique sont mis en place uniquement à partir des observations. Les vraisemblances des données sont reportées sur la figure (FIG. 4.28). Dans pareille situation, les seuils théoriques de l'OGD et du BORD Asymptotique ne satisfont pas la P_{fa} demandée. La corrélation des données de référence est trop forte pour que les matrices estimées ne soient pas singulières, ou mal conditionnées. Les expressions de ces seuils deviennent caduques et les seuils calculés par MC sont conservés.

Les performances des quatre détecteurs sont présentées sur la figure (FIG. 4.29).

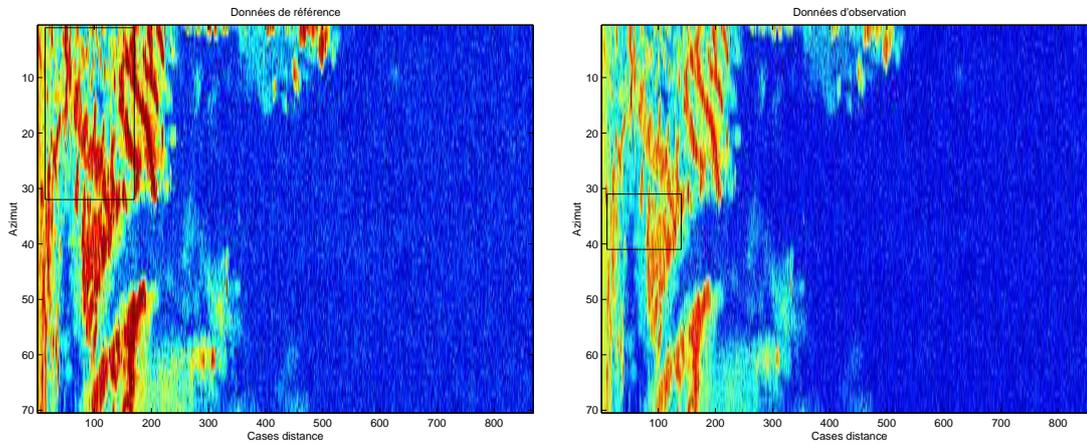


FIG. 4.27 – Zone 2 : Fouillis pour les références (gauche) et les observations (droite). Au total, $N_{ref} = 4836$ et $N_{obs} = 1300$.

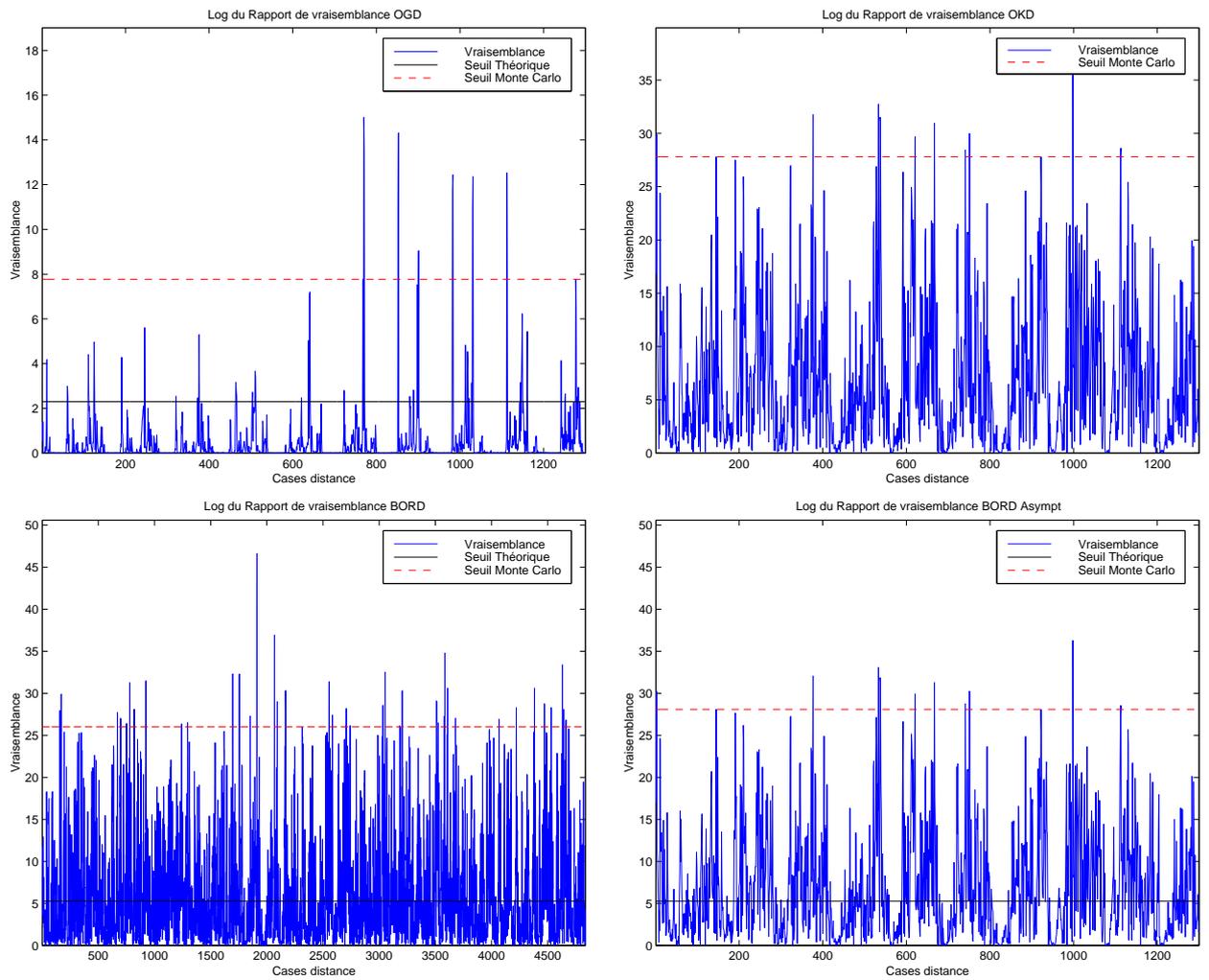


FIG. 4.28 – Zone 2 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).

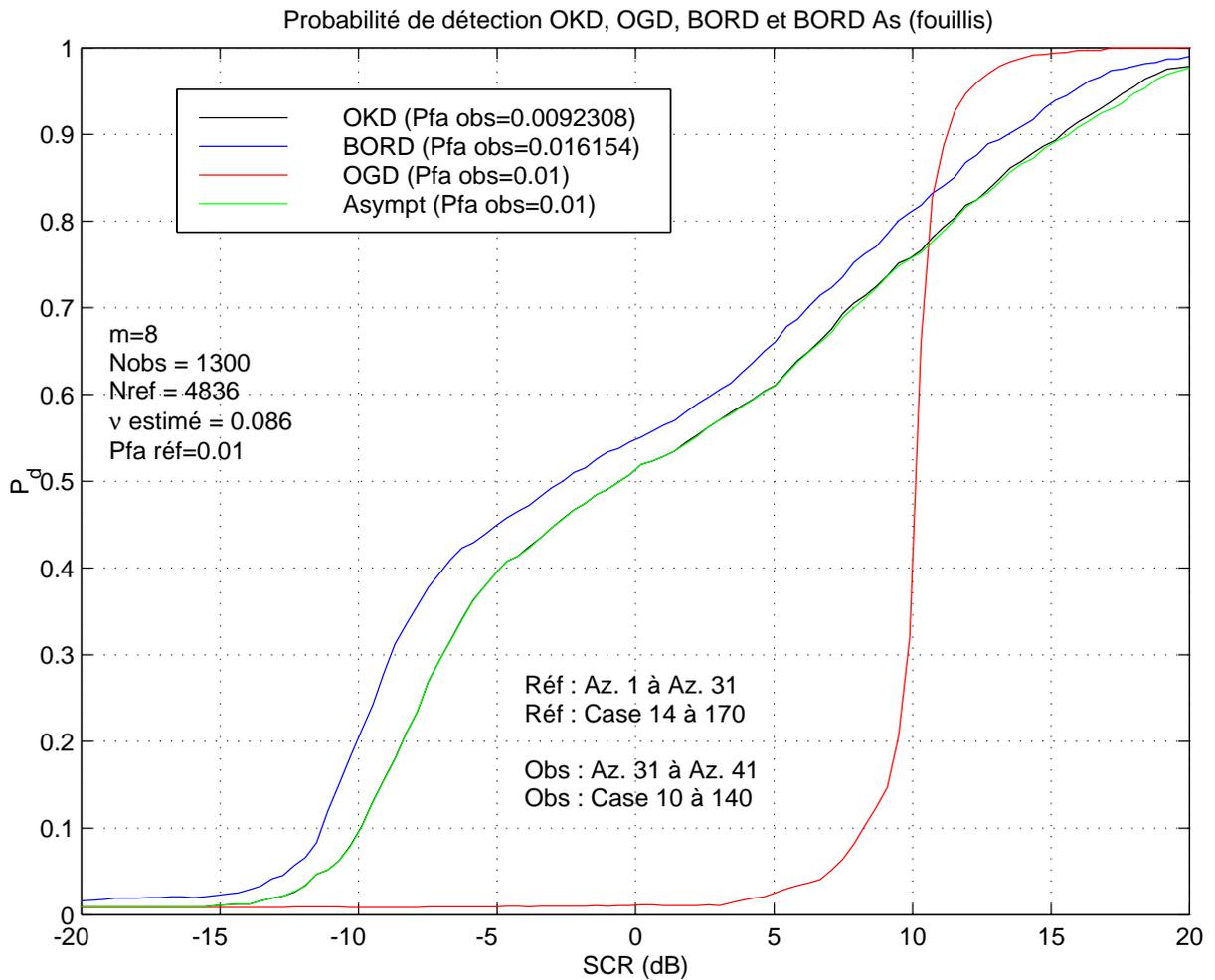


FIG. 4.29 – Zone 2 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis

Analyses des performances de la "zone 2"

D'après les résultats de simulation du chapitre précédent, l'OKD, le BORD et le BORD Asymptotique doivent se comporter en détecteurs optimaux, à condition que les paramètres estimés de la loi K soient réalistes pour l'OKD. Comme attendu, les performances de l'OGD sont largement dépassées par celles des autres détecteurs, à faible RSB.

Toutes les courbes de performances sont comparables, puisque les TFA constatés sont en cohérence avec la P_{fa} désirée. Le TFA du BORD est cependant légèrement supérieur, et sa courbe de performances devrait se trouver légèrement plus à droite. Comme pour la "zone 1", cette différence entre le TFA et la P_{fa} est due aux données d'observation, autres que les références. La matrice de covariance estimée sur les références ne correspond pas parfaitement à la corrélation des données d'observation, et cela suffit à créer quelques fausses alarmes.

La forme particulière des courbes de performances du BORD, de l'OKD et du BORD Asymptotique, entre -10 dB et 0 dB, est due à la présence de bruit thermique dans les premières cases distance des données d'observation. La baisse d'amplitude du bruit augmente le pouvoir de détection du radar, malgré le faible RSB. L'OGD, quant à lui, est insensible à ce changement, puisque sa détectabilité n'est effective qu'à partir de $+5$ dB. En revanche, pour de forts RSB (au-delà de $+10$ dB), l'OGD indique de meilleures performances. Cette particularité a été constatée lors des simulations en environnement fortement non-Gaussien et s'explique difficilement. L'intérêt est cependant plus particulièrement porté sur le gain apporté par les autres détecteurs à faible RSB.

Contrairement aux cas du bruit Gaussien, le BORD Asymptotique se comporte parallèlement au BORD. Ses performances sont confondues avec celles de l'OKD, pour lequel une très faible valeur de ν a été estimée. Ceci confirme d'ailleurs le caractère fortement non-Gaussien des zones de fouillis.

Entre les performances du BORD et du BORD Asymptotique (et donc de l'OKD), il est difficile de savoir lesquelles sont les plus optimales. Ces détecteurs n'ont pas été mis en place identiquement, et la réponse est apportée par les tests effectués dans les "zones" 5 et 7.

4.4.2.2 "Zone 5"

De même que dans la "zone 4", seul un jeu de données est sélectionné. Ces données servent de références et d'observation à tous les détecteurs et représentent $N_{ref,obs} = 2660$. Elles proviennent de 190 cases distance (de la 22 à la 212) prélevées dans 14 azimuts différents (du 18 au 32). Elles sont indiquées par le rectangle noir de la figure (FIG. 4.30).

Une fois encore, les expressions des seuils théoriques tombent en désuétude pour garantir la P_{fa} . Cela peut être constaté sur la figure (FIG. 4.31), représentant les logarithmes des rapports de vraisemblance des détecteurs.

La différence de performances, observée dans la "zone 2", entre le BORD et le BORD Asymptotique, devrait disparaître, ou du moins s'amoinrir. L'estimation de la matrice de covariance correspond réellement aux données testées et le TFA devrait lui aussi s'accorder avec la P_{fa} .

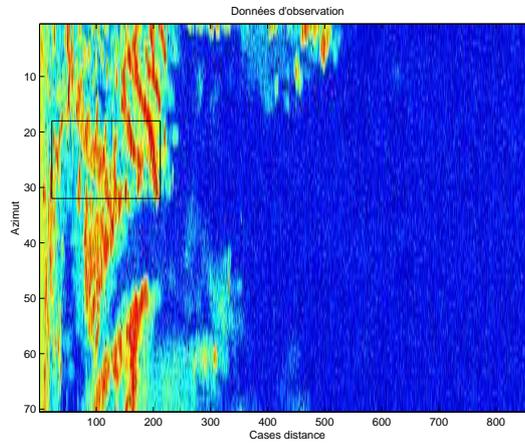


FIG. 4.30 – Zone 5 : Fouillis sélectionné pour les observations. Au total $N_{obs} = 2660$.

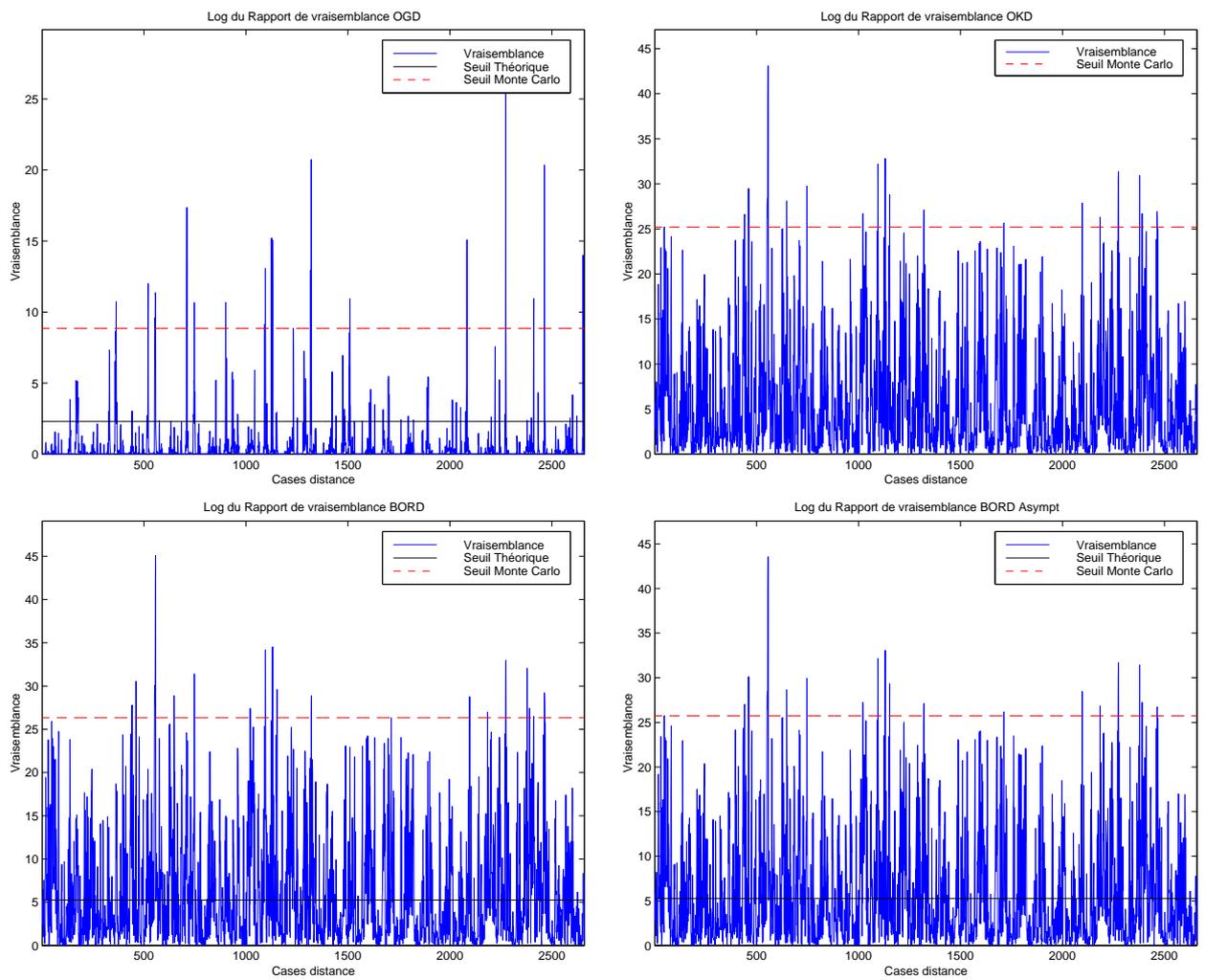


FIG. 4.31 – Zone 5 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).

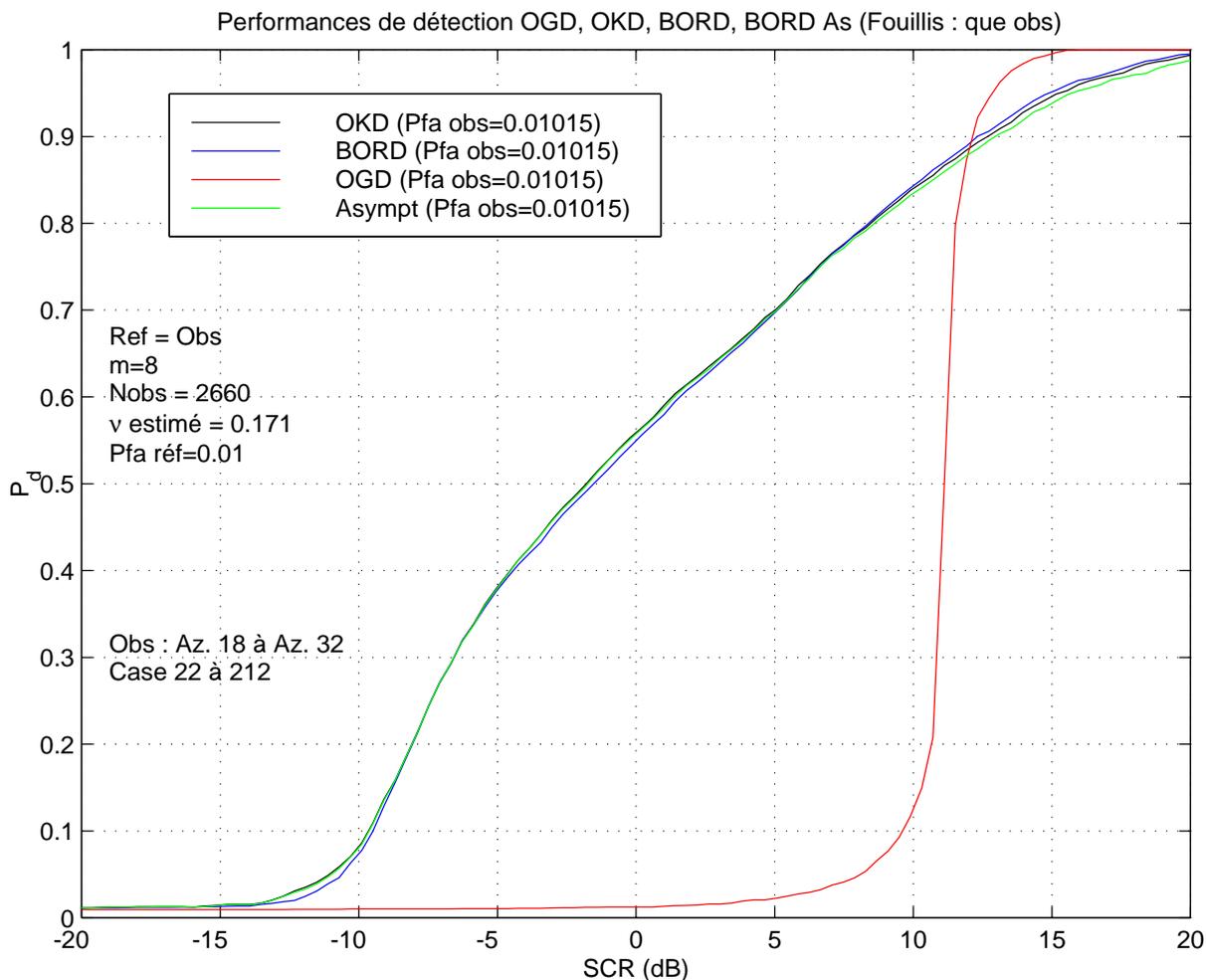


FIG. 4.32 – Zone 5 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis

Analyses des performances de la "zone 5"

Les courbes de détection de la figure (FIG. 4.31) sont tout à fait satisfaisantes. Pour chaque détecteur, la P_{fa} est conservée, ce qui est logique de par l'emploi d'un seul jeu de données pour l'apprentissage et la détection. Cependant, dans la "zone 4", la même procédure était utilisée, et le TFA de l'OKD ne s'accordait pas avec la P_{fa} . On peut donc supposer, que la valeur estimée retenue ($\nu = 0.171$) pour les paramètres de la loi K coïncide parfaitement avec la statistique des données.

Les performances du BORD, de l'OKD et du BORD Asymptotique sont donc tout à fait comparables, et sont même confondues. Ceci nous permet d'annoncer, que le BORD et le BORD Asymptotique restent optimaux en situations réelles de fouillis.

Comme précédemment, la forme légèrement bombée des courbes de performances (non constatée sur simulations), entre -10 dB et 0 dB, est due à la faible amplitude des données dans les premières cases distance.

L'OGD reste toujours insensible, mais, une fois encore, indique de meilleures performances à partir de $+12$ dB. Le gain apporté se situe, à P_d constante, entre 0 et 5 dB, alors qu'à RSB donné, l'OGD gagne de 0% à 7% en détectabilité.

Cela n'est rien comparé aux quelques 19 dB gagné en RSB à $P_d = 0.2$ par les autres détecteurs. A RSB = 5 dB, le BORD, le BORD Asymptotique et l'OKD gagnent 68% en détectabilité par rapport à l'OGD. La petite faiblesse de ces détecteurs, à fort RSB, n'est donc pas si importante.

4.4.2.3 "Zone 7"

La "zone 7" est constituée de références, utilisées pour l'apprentissage de tous les détecteurs, et de données d'observation utiles à l'évaluation des performances de ces mêmes détecteurs. Les $N_{ref} = 2054$ références se composent de 158 cases distance (de la 10 à la 168), provenant de 13 azimuts (du 31 au 44), et sont représentées par le rectangle noir de la carte gauche de la figure (FIG. 4.33).

Les observations comportent $N_{obs} = 1550$ échantillons de 62 cases distance (de la 16 à la 78), et représentant 25 azimuts (du 2 au 27). La carte droite de la figure (FIG. 4.33) en donne un aperçu.

Les seuils théoriques tracés sur les courbes de la figure (FIG. 4.34) sont une fois de plus inutilisables, si on désire conserver la P_{fa} initiale.

Les performances des détecteurs sont analysées d'après la figure (FIG. 4.35). La situation est plus réaliste que la précédente, et l'optimalité annoncée du BORD et du BORD Asymptotique pour la "zone 3", peut être plus justement décrite.

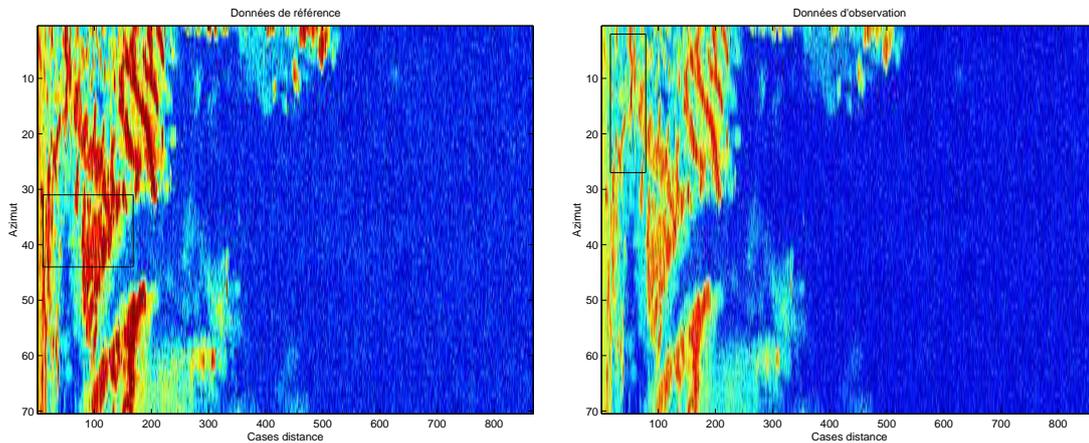


FIG. 4.33 – Zone 7 : Fouillis pour les références (gauche) et les observations (droite). Au total, $N_{ref} = 2054$ et $N_{obs} = 1550$.

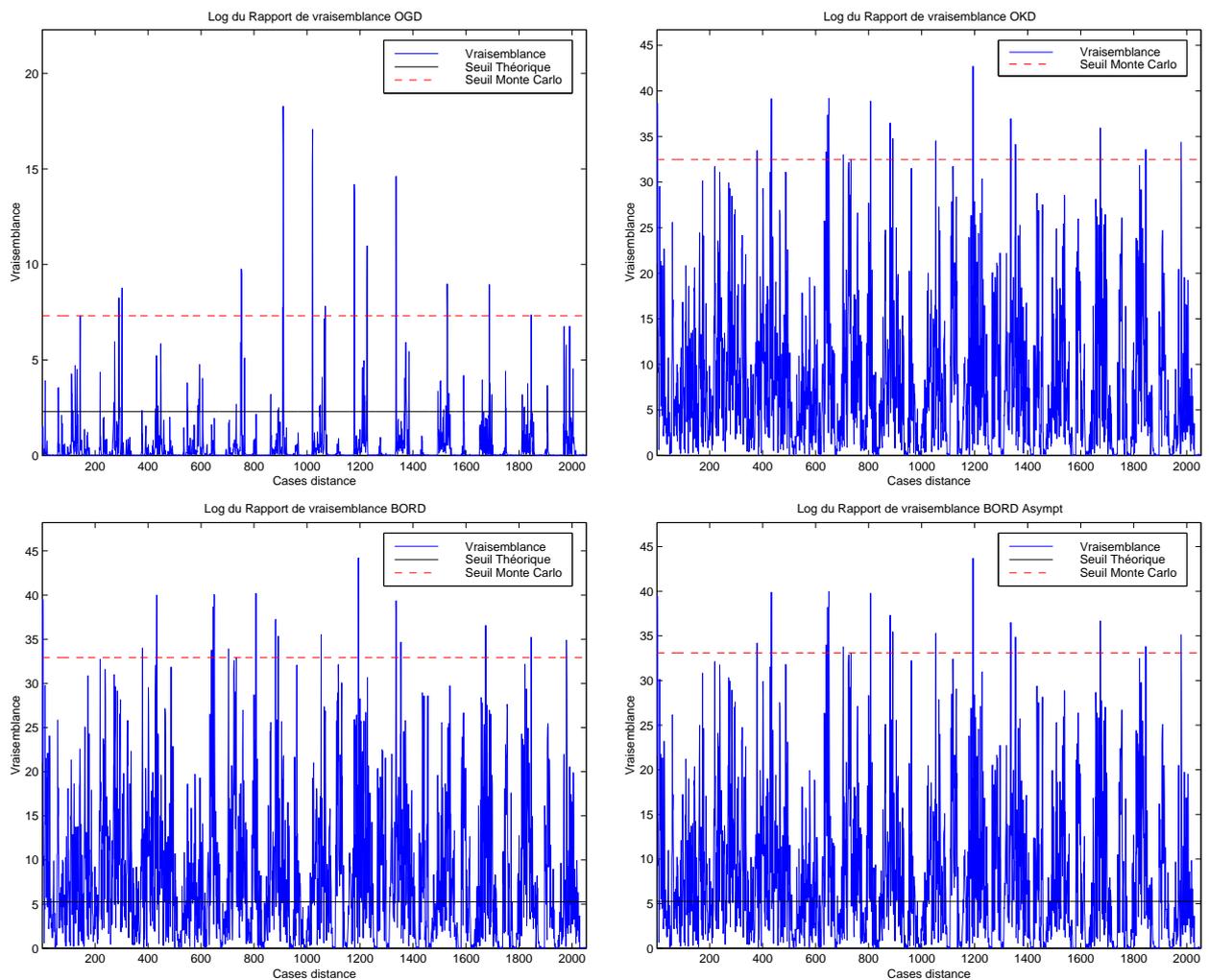


FIG. 4.34 – Zone 7 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).

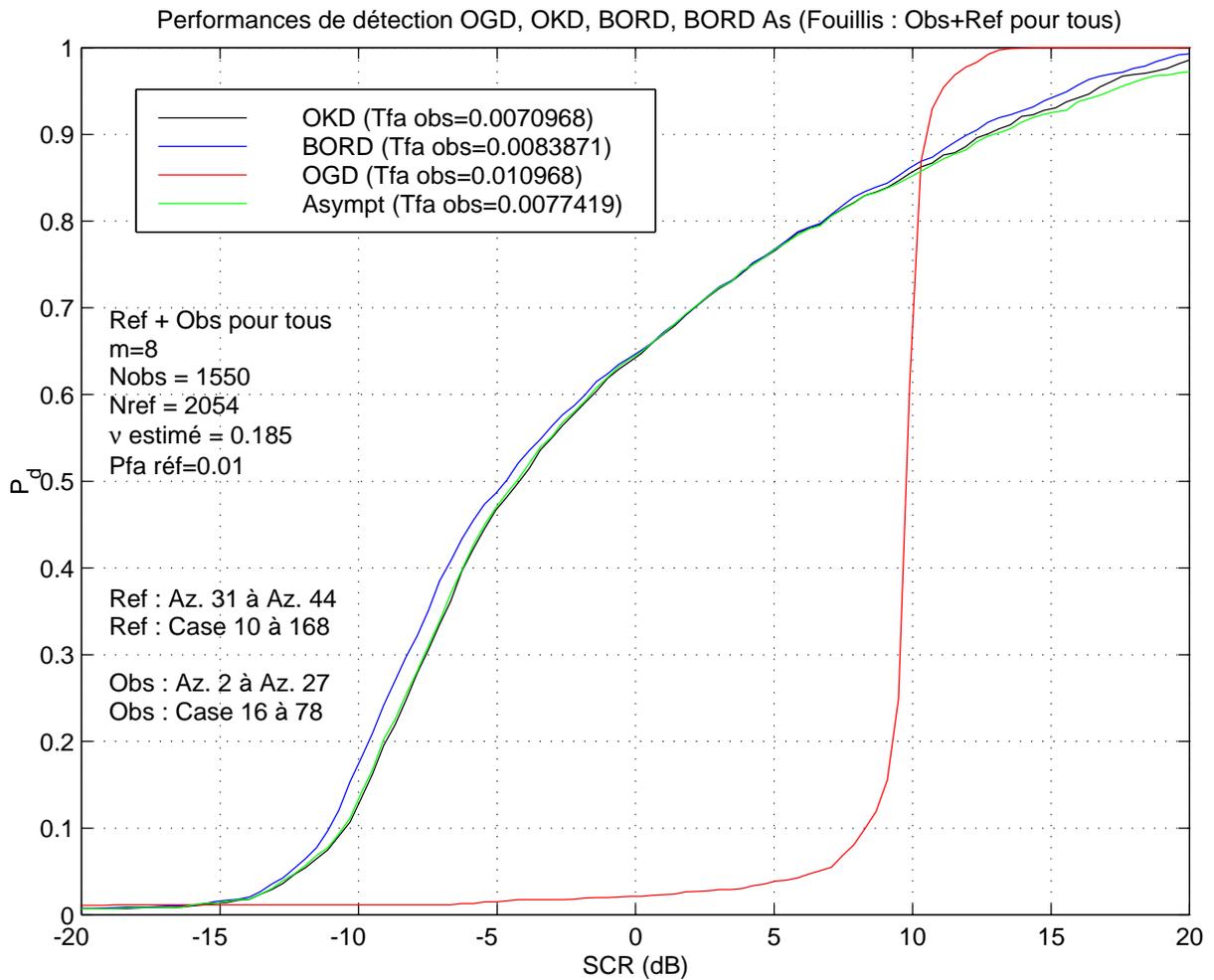


FIG. 4.35 – Zone 7 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis

Analyses des performances de la "zone 7"

Les courbes de performances sont similaires à celles obtenues dans la "zone 5", à ceci près que les TFA constatés pour le BORD, le BORD Asymptotique et l'OKD sont plus faibles qu'espéré. Leurs valeurs sont toutefois proches, et les performances de ces trois détecteurs peuvent donc se comparer. L'OGD, bizarrement, conserve un TFA égal à la P_{fa} , sans toutefois égaler les performances des autres détecteurs.

Une fois encore l'estimation des paramètres de la K-distribution semblent satisfaire à l'environnement observé. Les performances de l'OKD se confondent avec celles du BORD Asymptotique, tandis que le BORD indique des performances légèrement meilleures. Cela est dû à son TFA, un peu plus élevé que celui de ses voisins, et sa courbe de performances devrait être légèrement plus à droite pour satisfaire leur TFA. Il n'y a cependant pas de comparaison possible avec les performances de détection de l'OGD, qui s'affirme, cette fois-ci, à partir de +10 dB.

Les "zones 3 et 8" comportent une transition entre le bruit thermique Gaussien et le fouillis. Les courbes de performances attendues devraient donc avoir une forme bombée prononcée, significative du changement d'amplitude brutal dans les données.

4.4.3 Zones transitoires bruit thermique / fouillis

Les "zones 3 et 8", définies à partir de données présentant une rupture bruit thermique / fouillis, procèdent à l'apprentissage des détecteurs sur données de référence, différentes des observations. La "zone 3" utilisent ces données uniquement pour le BORD, alors que pour la "zone 8", les caractéristiques des données sont apprises sur références pour tous les détecteurs. La "zone 3" est similaire aux "zones 1 et 6" et la "zone 8", aux "zones 2 et 4".

4.4.3.1 "Zone 3"

Les données de références nécessaires à l'apprentissage du BORD sont désignées par le rectangle noir sur la carte gauche de la figure (FIG. 4.36). La sélection représente $N_{ref} = 4410$ échantillons, provenant de 294 cases distance (de la 60 à la 354) et de 15 azimuts (du 42 au 57). Les données d'observation, prises comme références pour l'OGD, l'OKD et le BORD Asymptotique, sont représentées sur la carte droite de la même figure. On observe $N_{obs} = 1560$ données, comprenant 130 cases distance (de la 178 à la 308) et 12 azimuts (du 1 au 13).

Les valeurs des seuils théoriques, données sur la figure (FIG. 4.37) et comparées aux valeurs empiriques, ne respectent pas du tout la P_{fa} initiale. Une fois encore, le mauvais conditionnement de la matrice devrait être pris en compte pour adapter le seuil théorique du BORD Asymptotique dans pareilles situations.

Le seuil théorique de l'OGD est normalement inadapté, puisque les données ne sont pas Gaussiennes.

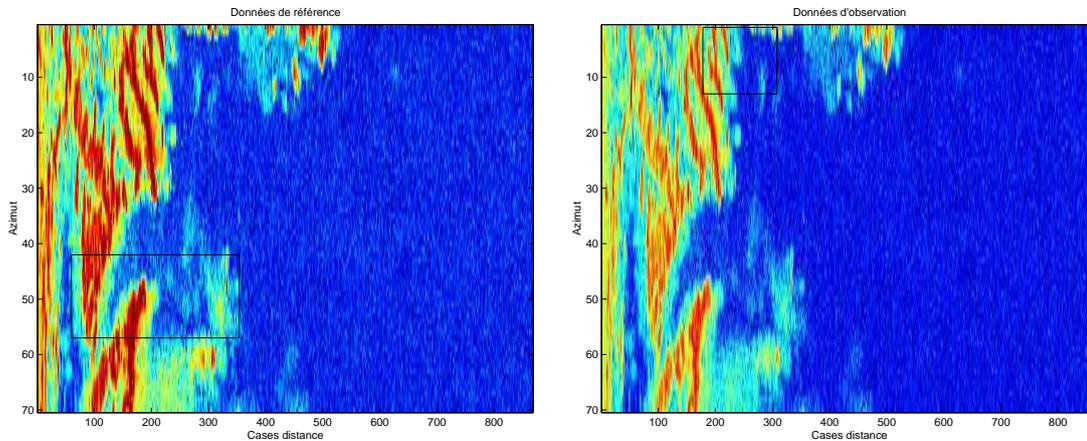


FIG. 4.36 – Zone 3 : Références (à gauche) et observations (à droite) pour un zone transitoire bruit thermique / fouillis. Au total, $N_{ref} = 4410$ et $N_{obs} = 1560$.

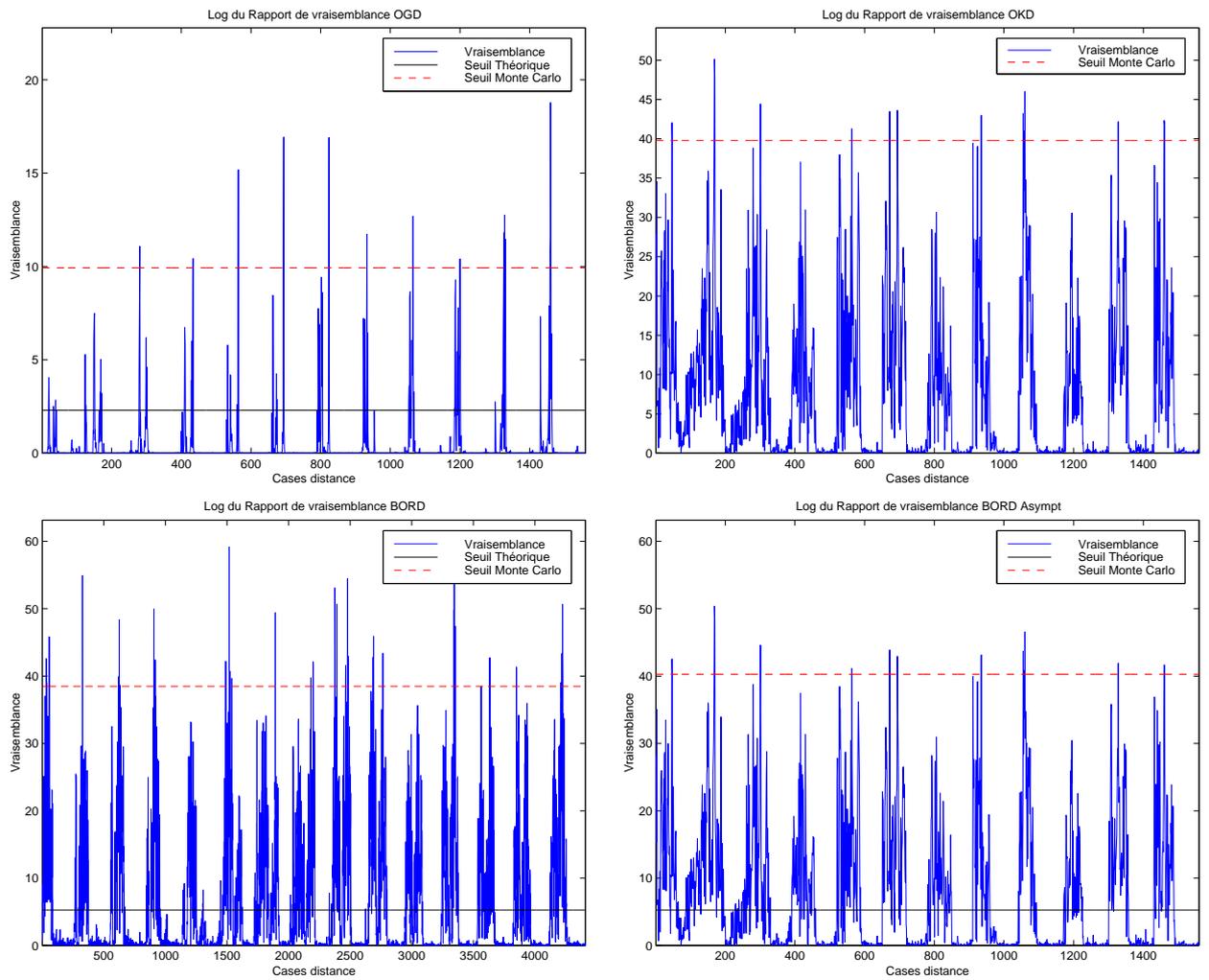


FIG. 4.37 – Zone 3 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).

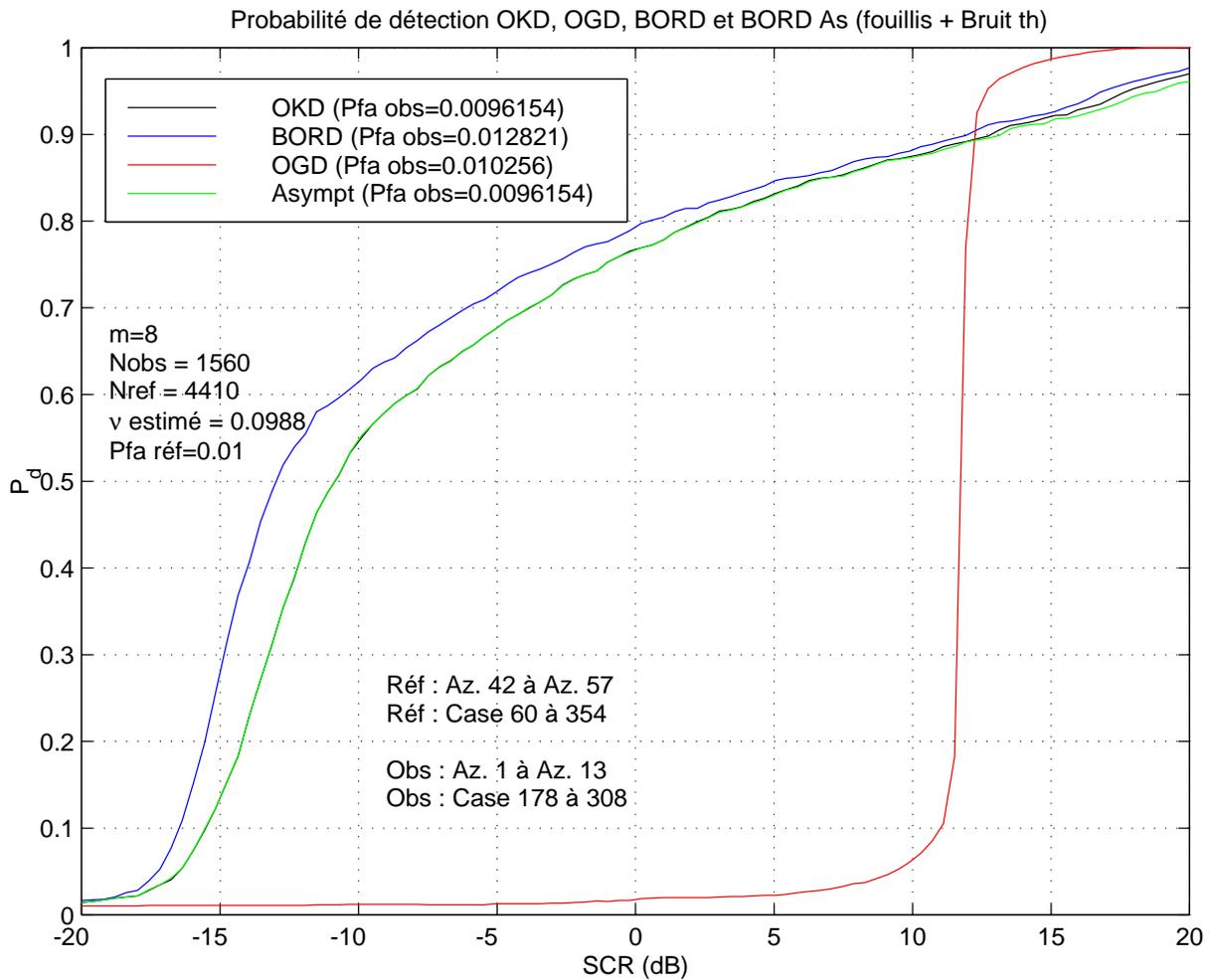


FIG. 4.38 – Zone 3 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis

Analyses des performances de la "zone 3"

Comme pour les "zones" cousines, les "zone 1" et "zone 6", les performances du BORD sont plus optimistes que celles de l'OKD et du BORD Asymptotique. La matrice de covariance estimée pour le BORD représente la corrélation des références et non celle des observations. La différence entre ces détecteurs ne peut provenir que de là, puisque les TFA constatés sont tous cohérents avec la P_{fa} .

Les courbes de l'OKD et du BORD Asymptotique sont confondues. Cela laisse à penser qu'elles indiquent les performances optimales, ou tout du moins les meilleures possibles. La procédure utilisée dans cette "zone" est effectivement idéale pour ces deux détecteurs, puisque les observations sont également les références. La matrice de covariance estimée correspond bien à la matrice observée. Les paramètres estimés de la K-distribution supposée, semblent convenir parfaitement à la statistique des données, et nous indiquent un caractère largement non-Gaussien ($\nu = 0.0988$). Les composantes de fouillis semblent inhiber le caractère Gaussien des autres composantes.

La forme bombée des courbes est bien confirmée. La rupture entre le bruit thermique et le fouillis engendre une baisse de l'amplitude des données, et les performances des détecteurs en est accrue, dans les faibles RSB. Cela n'est cependant pas vrai pour l'OGD, insensible à de telles valeurs.

Comme précédemment, les performances de l'OGD dépassent celles des autres détecteurs à partir de +12/13 dB, avec un gain maximal de 6/7% en terme de détectabilité. De même, à $P_d = 0.95$, l'OGD est capable de détecter pour un $RSB = 12.5$ dB, alors que 18 dB sont nécessaires aux autres détecteurs.

Cette particularité a déjà été soulignée dans les "zones" de fouillis. Cependant, cet avantage que présente l'OGD à forts RSB, n'est rien à côté du gain apporté par le BORD Asymptotique à faibles RSB. Par exemple, à -10 dB, ce dernier détecte la cible dans 55% des cas. L'OGD, non. Pour obtenir $P_d = 0.55$, l'OGD nécessite approximativement 22 dB supplémentaires.

La "zone 8" procède également pour tous les détecteurs. L'apprentissage des données est effectué au préalable pour tous les détecteurs, et les performances sont évaluées à partir d'observations de même nature que les références.

4.4.3.2 "Zone 8"

Cette "zone" est la dernière testée. Les références utilisées pour l'estimation des matrices de covariance et des paramètres de la loi K sont représentées par le rectangle noir sur la carte de gauche de la figure (FIG. 4.39). On totalise $N_{ref} = 2340$ données, dont 156 cases distance (de la 94 à la 250) et 15 azimuts (du 37 au 52). Les transitions apparaissent surtout pour les premiers azimuts mais ne représentent pas la rupture définitive donnée en (FIG. 4.6).

Les $N_{obs} = 2080$ observations prennent en compte 130 cases distance (de la 176 à la 306) et 16 azimuts (du 6 au 22). Cette sélection comporte des ruptures définitives, entre le 17^{ème} et le 22^{ème} azimut, au-delà de la 230^{ème} case distance. Elles sont représentées sur la carte droite de la même figure.

Les transitions entre le bruit thermique et le fouillis, dans les références, sont très bien illustrées à travers les courbes de la figure (FIG. 4.40). Les rapports de vraisemblance sont très chahutés et chaque partie, comprenant un "pic de forte amplitude" suivi d'une zone homogène de faible amplitude, correspond aux données prélevées dans un azimut. L'OKD, le BORD et le BORD Asymptotique illustrent très bien la nature des données, alors que l'OGD ne prend en compte que les fortes perturbations. Au-delà de l'échantillon 1600, la forme de la vraisemblance de l'OGD ne laisse pas entrevoir la zone, de nouveau homogène, de fouillis. Cela est très net dans les cas des trois autres détecteurs.

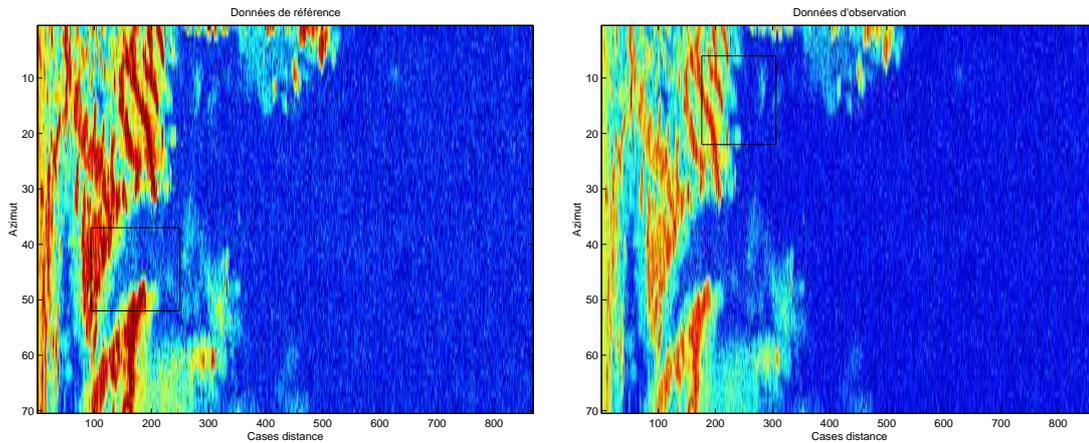


FIG. 4.39 – Zone 8 : Zone transitoire bruit thermique / fouillis pour les références (gauche) et les observations (droite). Au total, $N_{ref} = 2340$ et $N_{obs} = 2080$.

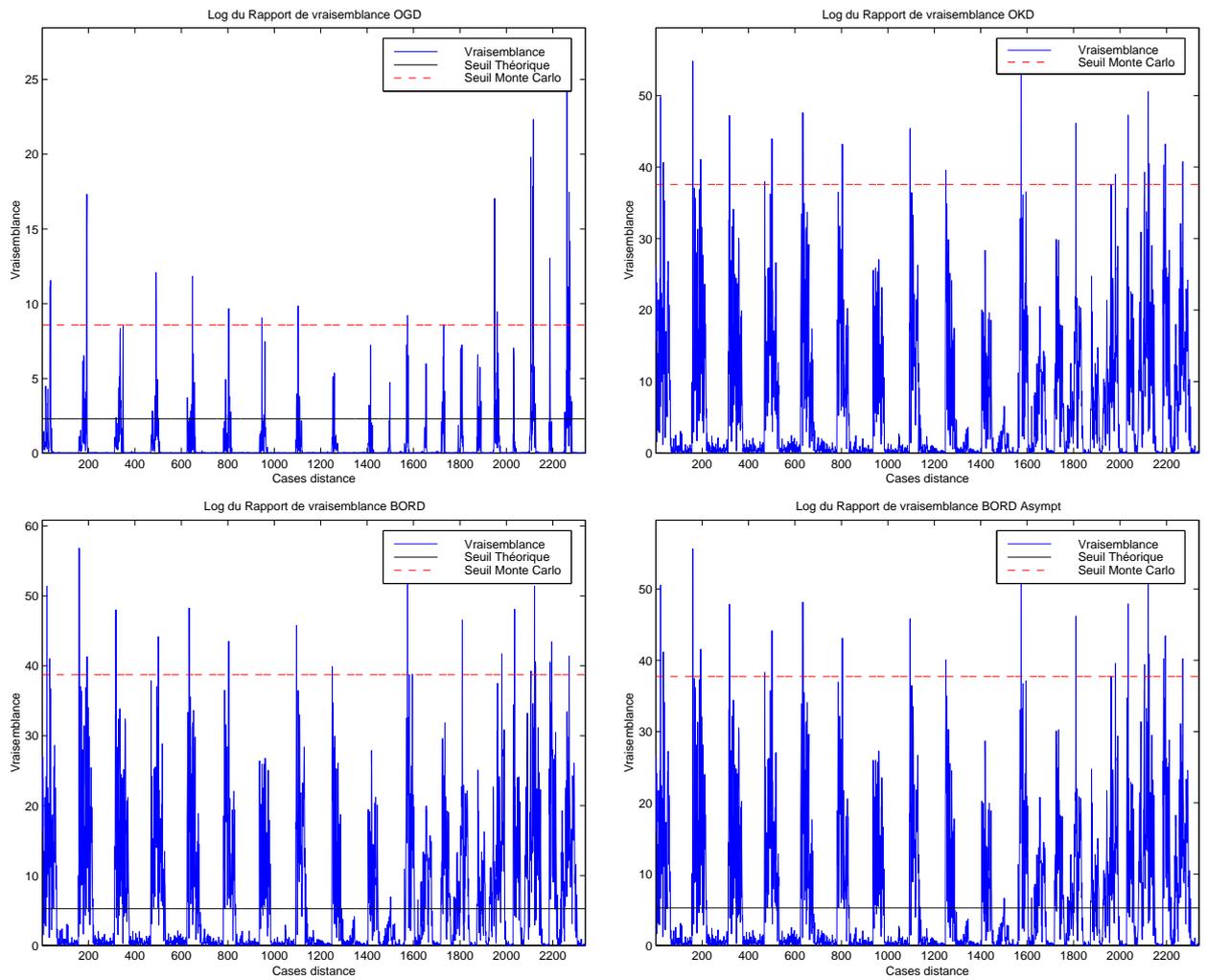


FIG. 4.40 – Zone 8 : Vraisemblance des données utilisées pour le calcul du seuil de détection, comparé au seuil théorique Gaussien (pour l'OGD) et au seuil théorique du BORD Asymptotique (pour le BORD et le BORD Asymptotique).

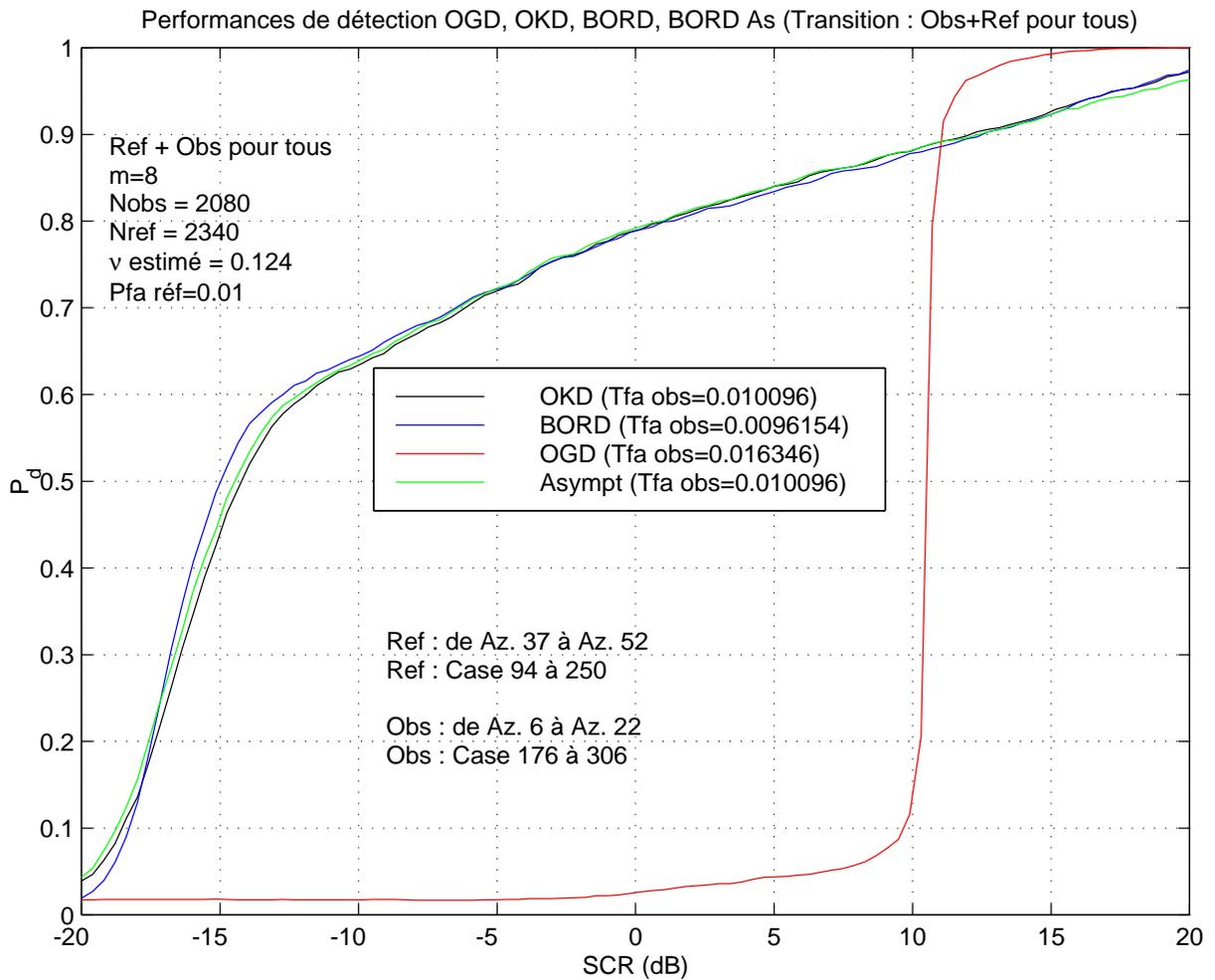


FIG. 4.41 – Zone 8 : Performances de détection de l'OGD, de l'OKD, du BORD et du BORD Asymptotique dans du fouillis

Analyses des performances de la "zone 8"

Les courbes de performances peuvent être comparées telles qu'elles sont présentées, puisque les TFA respectent la P_{fa} initiale. Une petite hausse est cependant effective pour l'OGD, et sa courbe de performances devrait se trouver plus à droite.

Les trois autres détecteurs affichent approximativement les mêmes performances. Le BORD est cependant légèrement supérieur à l'OKD et au BORD Asymptotique (une fois de plus confondus) entre -8 dB et -17 dB. A P_d constante, il gagne au maximum 0.2 dB et pour un RSB donné, sa P_d est améliorée d'environ 3% . Autant dire que, dans un tel environnement, ces différences sont minimales.

D'après les courbes obtenues dans la "zone 3", les performances du BORD restent comparables à celles-ci. Bien que la statistique des données est différente d'une "zone" à l'autre, les procédures d'apprentissage et d'évaluation sont identiques pour ce détecteur. Les performances de l'OKD et du BORD Asymptotique se sont donc rapprochées, dans la "zone 8", de celles du BORD.

La forme bombée des courbes se retrouve encore, caractéristique des transitions rencontrées dans les données. Le comportement de l'OGD à forts RSB suscite les mêmes remarques que précédemment, et il est toujours intéressant de noter, qu'avec $P_d = 0.5$, le BORD, l'OKD et le BORD Asymptotique sont capables de détecter une cible, pour un RSB 25 fois plus faible (environ -15 dB) que ne l'aurait fait l'OGD (à environ $+10$ dB).

4.5 Synthèse des résultats

En comparant les résultats de simulation avec les résultats obtenus sur données expérimentales, on constate que les détecteurs se comportent de la même manière. Cela valide l'utilisation du modèle des SIRP pour des données de fouillis radar non-Gaussiens, ainsi que les détecteurs construits sur ce modèle.

Pour le bruit thermique Gaussien des zones 1, 4 et 6, le détecteur classique apporte les meilleures performances. Le BORD se comporte en détecteur sous-optimal, et nécessite moins de 1 dB supplémentaire pour atteindre les performances de l'OGD. L'écart provient très certainement de la différence entre la densité prior utilisée pour la texture et sa distribution réelle de Dirac en 1. Ce phénomène, qui n'a pas été constaté sur simulations pour le BORD, est bien plus important pour le BORD Asymptotique. La convergence en loi a été établie d'après la densité prior de la texture des références. Cet *a priori* a donc servi pour l'estimation de la loi de la texture et pour le calcul de l'expression du BORD Asymptotique. Ce détecteur coïncide pourtant avec le GLRT-LQ, obtenu sous l'hypothèse d'un bruit Gaussien, mais pour lequel la matrice de covariance est estimée classiquement. Tel qu'il est mis en place, il est donc plus proche de l'ALQ, obtenu après estimation au sens du MV de la texture. En présence de bruit Gaussien, le GLRT-LQ tend asymptotiquement vers l'OGD et admet de meilleures performances que l'ALQ. C'est un des seuls cas où l'ALQ est dépassé par le GLRT-LQ, ce qui confirme le comportement du BORD Asymptotique en présence de bruit Gaussien.

Les performances de l'OKD sont difficiles à commenter. Les estimations des paramètres ν et b permettent de considérer les données comme étant régies par une K-distribution, mais l'hypothèse est difficilement vérifiable. Les performances de l'OKD sont toutefois proches de celles de l'OGD et du BORD, avec de fortes valeurs estimées pour ν . Cela confirme bien que les données de faible amplitude sont Gaussiennes.

Dans les zones de fouillis 2, 5 et 7, le BORD, l'OKD et le BORD Asymptotique se comportent en détecteurs optimaux. Bien qu'il est difficile de déterminer lequel des trois apporte les performances les plus vraisemblables, la zone 5 est la situation la plus idéale pour effectuer une comparaison. Les trois détecteurs fournissent les mêmes performances malgré leurs structures tout à fait différentes, et sont certainement les meilleures possibles.

Les résultats obtenus dans les zones transitoires 3 et 8 sont analogues à ceux des zones de fouillis seul. La statistique non-Gaussienne des données met l'OGD à défaut, mais permet de démontrer, une fois de plus, le fort pouvoir d'adaptation des deux versions du BORD en environnements difficiles. Les courbes de la zone 8 sont plus significatives que celles de la zone 3, puisque tous les détecteurs ont été soumis au même régime : apprentissage sur données de référence puis analyse sur données d'observation.

A l'issue de tous les résultats présentés, on peut émettre les remarques suivantes :

- * L'OGD est toujours optimal en environnement Gaussien et présente de meilleures performances que l'OKD, le BORD et le BORD Asymptotique pour de très forts RSB, en présence de fouillis non-Gaussien,
- * Le BORD est un détecteur qui s'adapte parfaitement aux situations de fouillis non-Gaussien, même non-homogène, et présente un caractère quasi-optimal en environnement

Gaussien. Quel que soit l'environnement rencontré, ses performances sont quasi-optimales, voire même optimales, dès lors que la probabilité de fausse alarme est respectée,

- * Le BORD Asymptotique est particulièrement bien adapté aux situations fortement non-Gaussiennes même non-homogènes. En revanche, ses performances se dégradent en environnement Gaussien. Lorsque la corrélation des données n'est pas trop forte (matrice de covariance non singulière), l'expression théorique de son seuil de détection est valide et ne dépend que de la taille des observations et de la P_{fa} souhaitée,
- * L'OKD, mis en place à partir des paramètres estimés de la loi K par la méthode des moments, présente des performances identiques à celles du BORD Asymptotique. Cela se vérifie notamment en présence de fouillis non-Gaussien même non-homogène. Ces résultats permettent de conclure, que les données expérimentales analysées dans ce chapitre sont en adéquation avec une K-distribution. Les paramètres restent à déterminer selon les zones étudiées.

Conclusion

Conclusion générale

Cette thèse, présentant une nouvelle génération de détecteurs optimaux, s'est articulée autour de la démarche suivante.

Dans le premier chapitre, nous avons rappelé les bases de la théorie de la détection et présenté les différentes techniques de détection utilisées en environnement de fouillis non-Gaussiens. Dans un souci de conserver la structure du filtre adapté, simple à mettre en œuvre, les diverses procédures TFAC présentées visent à en améliorer les performances, en adaptant le seuil de détection à la puissance locale du fouillis. Ces techniques ne sont ni optimales, ni robustes aux environnements de fouillis non-homogènes. Seuls des détecteurs adaptés à la statistique de l'environnement peuvent vérifier ces propriétés, ce qui nécessite de repenser les structures des récepteurs radar actuels.

Les modèles couramment cités et utilisés sont les SIRP, ou Spherically Invariant Random Process, qui sont particulièrement bien adaptés aux environnements de fouillis impulsionnels. Les détecteurs construits d'après ces modèles sont optimaux pour la nature du fouillis désigné, mais en nécessitent une connaissance *a priori*. La restriction imposée par cet *a priori* empêche toutefois une mise en œuvre unique du détecteur optimal, et notre travail a pour objectif de s'en affranchir. Cela requiert une analyse plus fine des modèles SIRP afin d'élaborer des détecteurs optimaux, capables de s'adapter à la statistique de l'environnement rencontré.

Dans le chapitre 2, les modèles SIRP de processus non-Gaussiens sont détaillés. Pour le besoin de l'étude, nous avons étendu certains résultats théoriques au cas des processus complexes, et notamment le théorème de représentation de K.Yao qui nous indique que ces processus sont naturellement des processus Gaussiens, dont la variance, appelée *texture* en radar, est une variable aléatoire positive. Ainsi, la loi jointe des vecteurs issus de ces processus s'obtient en intégrant la loi jointe du vecteur Gaussien sur la loi de la *texture*. Elle seule détermine donc la nature statistique du SIRP.

Le travail s'est alors orienté sur la recherche de méthodes permettant d'estimer la loi de la *texture* à partir de données de fouillis complexes, et sans aucune connaissance *a priori* sur leur statistique. Le paradigme bayésien nous est apparu comme le concept le plus approprié à ce problème et a contribué à l'originalité de ce document.

Le chapitre 3 constitue la contribution majeure de ma thèse. Après une analyse bayésienne du modèle SIRP, réalisée sur les données complexes du radar, deux méthodes d'estimation de la loi de la *texture* sont proposées. Deux stratégies optimales de détection en découlent et s'adaptent à la statistique de l'environnement, sans émettre d'hypothèse supplémentaire.

Le premier détecteur proposé, le PEOD, repose sur une approximation de Padé de la fonction

génératrice de moments de la *texture*. Lorsque les moments sont connus, elle s'effectue directement à partir de leurs valeurs et le PEOD est "idéal". Son expression permet une mise en œuvre unique quel que soit l'environnement rencontré et ses performances de détection atteignent celles des détecteurs optimaux.

Pour appliquer cette méthode dans un cadre opérationnel, les moments de la *texture* doivent être déterminés. Des échantillons de *texture* sont donc générés selon sa loi *a posteriori* instantanée et une estimation empirique des moments est réalisée.

Même si le PEOD ne constitue pas fondamentalement une révolution en détection radar, il faut avant tout en souligner la démarche. L'analyse bayésienne permet de reconstruire des échantillons de *texture* représentatifs de la nature des données de référence, et l'approximation de Padé possède l'avantage indéniable de fournir une estimation de qualité, notamment lorsque les moments sont connus, ou, bien estimés.

La deuxième méthode proposée repose sur l'estimation MC-bayésienne de la loi de la *texture*, obtenue en moyennant la densité *a posteriori* instantanée sur toutes les cellules de référence. Injectée dans l'expression du rapport de vraisemblance sous chacune des hypothèses, cette expression permet de construire le détecteur BORD, qui possède plusieurs propriétés majeures.

En premier lieu, sa structure peut être mise en place après le traitement classique opéré par le filtre adapté. Le BORD peut alors être considéré comme un post-traitement à effectuer afin d'optimiser la détection. Ceci évite de procéder à un changement radical dans la conception des récepteurs radar actuels.

On constate également que la statistique du test de détection est indépendante de la loi de la texture des données. Cela confère au BORD la propriété TFAC et la valeur du seuil de détection ne dépend pas cette loi. Cela signifie, qu'avec une estimation NSCM de la matrice de covariance des références, le BORD est un détecteur adaptatif et TFAC par rapport à la loi de la *texture*. L'évaluation de ses performances face à différents types de fouillis (Gaussiens et non-Gaussiens) montre qu'il atteint les performances des détecteurs optimaux, quels qu'ils soient. Dans les situations où le détecteur optimal n'est pas connu sous forme simple, ses performances se confondent avec celles du "PEOD idéal", et les deux approches se rejoignent alors.

D'un point de vue mathématique, la formulation complexe du BORD rend difficile l'obtention de sa loi. Une étude asymptotique de ce détecteur (convergence en loi) a toutefois permis d'en établir une forme plus simple : le BORD Asymptotique. Son expression coïncide avec un autre détecteur connu dans la littérature sous le nom de ALQ, établi après estimation de la variable de *texture* au sens du MV.

Dans les mêmes conditions que le BORD, ce détecteur est adaptatif et vérifie structurellement la propriété TFAC par rapport à la loi de la *texture*. Cela a été prouvé analytiquement et confirmé à l'issue des résultats des simulations.

La loi du BORD Asymptotique a également été obtenue. Ce résultat enrichit considérablement le travail, car il permet de garantir le taux de fausses alarmes voulu, indépendamment du nombre de données traitées. Le seuil de détection peut alors être fixé théoriquement. Les simulations, reprises avec ce seuil théorique, montrent que ce dernier reste valable pour le BORD et ceci, quelle que soit la statistique du fouillis.

La boucle est pour ainsi dire "bouclée", puisque l'on dispose de structures optimales de dé-

tection, adaptatives, TFAC par rapport à la statistique du fouillis et dont les performances sont maîtrisées. Ce constat mérite toutefois d'être nuancé : tous ces résultats ont été obtenus dans le cas où la matrice de covariance des données est supposée connue et non-singulière, ce qui signifie que la corrélation temporelle n'est pas trop forte. Dans le cas où la matrice est singulière, son inversion devient un problème, qui peut se résoudre soit en considérant sa pseudo-inverse, soit en utilisant des techniques de régularisation de matrice comme la technique du *diagonal loading*. Le théorème de Cochran ne s'applique plus dans les mêmes conditions et l'arrivée récente de ce résultat ne nous a pas permis d'approfondir le sujet dans le cadre de la thèse. Quelques réflexions à ce propos seront données en perspectives.

Pour leur originalité et leurs résultats, les travaux relatifs au BORD et au BORD Asymptotique ont été intégrés dans un projet d'étude amont (PEA) avec THALES TAD, qui vise à améliorer les détecteurs radar actuels, voire à en développer de nouveaux.

Le dernier chapitre est consacré à la validation des résultats de simulation du BORD et du BORD Asymptotique sur données expérimentales. Celles-ci proviennent de mesures de fouillis effectuées par un radar de THALES TAD dans la région de Colmar. Ne disposant d'aucune information quant à la statistique de ces données, il nous a paru logique de comparer les performances du BORD et du BORD Asymptotique à celles d'un détecteur OKD optimal. Dans cette optique, les paramètres de forme de la K-distribution ont été estimés au sens du MV à partir des données expérimentales. On constate que les performances du BORD, du BORD Asymptotique, de l'OGD et de l'OKD sont analogues aux résultats des simulations effectuées sur signaux synthétiques. Dans les zones de fouillis non-Gaussien, les performances de l'OGD sont très nettement dépassées par celles des trois autres. Dans les zones de bruit thermique Gaussien, l'OGD prouve bien son optimalité, atteinte partiellement par les trois autres détecteurs. De manière générale, le BORD, le BORD Asymptotique et l'OKD fournissent toujours les mêmes performances, ce qui laisse à penser que les données ont un caractère K-distribué. Les paramètres de forme estimés et les résultats des approximations de Padé, réalisées sur l'enveloppe des données, confirment d'ailleurs cette adéquation entre les données et une loi supposée K-distribuée.

Les données expérimentales présentant une corrélation temporelle très forte (notamment dans les zones de présence de fouillis), la valeur théorique du seuil du BORD Asymptotique ne correspond plus au seuil estimé par Monte-Carlo. En revanche, dans les zones de bruit thermique, les deux valeurs des seuils concordent. Comme cela n'est pas un résultat généralisable à toutes les zones, les valeurs MC des seuils de détection ont été conservées pour tous les détecteurs, et dans toutes les zones traitées. Du fait d'un nombre limité de données, cela ne permet pas d'atteindre de très faibles valeurs de la P_{fa} .

Cette étude a cependant permis de valider la pertinence du modèle SIRP, tout comme les détecteurs construits à partir de ce modèle.

D'après les résultats présentés dans la thèse, un constat d'ordre général s'impose concernant l'apport de l'analyse bayésienne menée sur le modèle SIRP. La variable de *texture* a pu être isolée en tant que composante principale du modèle, et est alors devenue l'élément incontournable pour la caractérisation de la statistique de l'environnement.

Jusqu'à présent, la majorité des études menées d'après le modèle SIRP s'attachaient à identifier les paramètres de ce modèle, à partir d'une ou plusieurs lois fixées *a priori* pour la statistique du fouillis. Cela a certes l'avantage d'en identifier la nature, pour permettre la mise en place du détecteur optimal, mais le choix final du modèle n'est validé que dans ce cas particulier. Chaque

changement de situation impose de renouveler cette étude préliminaire.

D'autres études se sont approchées de l'analyse bayésienne. La *texture* a parfois été estimée au sens du *maximum a posteriori*, mais là encore en supposant connue la densité *a priori* de la *texture*. Dans toutes ces études, on ressent nettement le besoin de rester attaché aux quantités physiques du modèle, et c'est sûrement pour cette raison que l'approche totalement bayésienne n'avait pas encore été utilisée.

Perspectives

Dans la formulation de l'analyse bayésienne, une densité *a priori* est requise pour accéder à la densité *a posteriori*. Notre choix d'un prior "non-informatif" reflète le degré d'ignorance sur les paramètres du modèle et confère une certaine objectivité à l'analyse.

Un autre choix pourrait s'orienter vers un prior conjugué, en considérant, par exemple, une loi conjuguée de mélange. Toute densité de probabilité pouvant s'estimer à partir d'un mélange de lois, la densité prior représenterait un mélange de lois conjuguées, et plus précisément un mélange de lois Inverse Gamma. Cette étude pourrait faire l'objet de nouvelles recherches, s'inscrivant dans la démarche de trouver un prior représentatif des données traitées.

L'estimateur MC-bayésien de la loi de la *texture* est parfois appelé l'estimateur de la méthode des noyaux. Dans ce cas, le noyau est la densité *a posteriori* instantanée de la *texture*, et son expression dépend, une fois de plus, du choix effectué pour le prior.

On pourrait donc envisager d'effectuer une comparaison *a posteriori* entre les résultats provenant de choix différents de priors afin de comprendre l'évolution de la loi de la *texture* selon l'origine des données. Cette étude pourrait aboutir à une sélection *a posteriori* adaptative pour la loi du prior, ce qui la rendrait encore plus attractive.

Revenons également à l'expression du BORD telle que nous l'avons obtenue, et plus précisément à son expression Asymptotique.

La loi du BORD Asymptotique a été établie sous des conditions idéales. La matrice de covariance des vecteurs Gaussiens est implicitement non-singulière et supposée connue.

Lorsqu'elle est inconnue et/ou singulière, deux problèmes peuvent être mis en évidence, tant dans la complexité apparente du problème que dans l'application du théorème de Cochran :

- Dans le cas où la matrice de covariance est inconnue, son estimation dépend alors des données et admet, de ce fait, une densité de probabilité : par exemple, la loi de Wishart est associée à l'estimation Monte-Carlo de la matrice de covariance de vecteurs Gaussiens. La détermination de la loi de la version adaptative du BORD Asymptotique nécessiterait, en premier lieu, d'identifier la loi de la matrice estimée au sens NSCM. Ensuite, il faudrait rechercher les matrices de projection orthogonales qui rendraient indépendants le numérateur et le dénominateur du BORD Asymptotique. L'expression résultante permettrait de quantifier les pertes TFAC dues à l'estimation de la matrice de covariance, en fonction du nombre de données utilisé.

- Dans l'hypothèse où un grand nombre de données est disponible pour l'estimation de la

matrice de covariance, on peut supposer qu'elle est parfaitement estimée et donc connue. Si, en revanche, elle s'avère singulière, son inversion devient un problème à part entière. La solution la plus répandue consiste à déterminer une pseudo-inverse de la matrice, auquel cas l'expression du BORD Asymptotique se doit d'être redéfinie. D'autres méthodes permettent de résoudre numériquement ce problème, comme les méthodes appelées *voltage domain methods* qui opèrent directement sur les données radar sans calculer la matrice de covariance. Ces méthodes mériteraient d'être étudiées plus en détail, bien qu'elles ne paraissent pas fournir de résultats analytiques. Quelle que soit l'option choisie, il n'est plus du tout certain qu'une décomposition sur des sous-espaces orthogonaux soit réalisable en vue d'appliquer le théorème de Cochran. Dans l'hypothèse contraire, on pourrait envisager que la loi du BORD Asymptotique dépende non seulement de la taille des vecteurs d'observation, mais également du rang de leur matrice de covariance.

Enfin, le dernier point qu'il faut évoquer concerne la validation des détecteurs sur les données expérimentales.

Les résultats ont clairement montré que les modèles, et leurs détecteurs associés, pouvaient être utilisés de manière pertinente dans un cadre opérationnel. Cela avait déjà été prouvé pour les modèles SIRP en général, et pour certains détecteurs, tels l'OKD en particulier. Cependant, le mode applicatif du filtre adapté (ou de l'OGD) consiste non pas à fixer un seuil de détection sur l'ensemble des données (seuil fixe), mais à utiliser une procédure TFAC par fenêtres glissantes (seuil adaptatif). Les performances de ce détecteur en sont améliorées, et si l'on suit l'évolution de l'OGD, les performances du BORD et du BORD Asymptotique en auraient peut-être été accrues. Les résultats qui seront obtenus à l'issue du PEA avec THALES éclaireront sûrement ce point laissé en suspens.

Les données expérimentales fournies par THALES TAD n'ont été qu'un exemple d'application pour le BORD parmi tant d'autres. On pourrait envisager d'en étudier le comportement face à du fouillis de mer ou encore face à des zones du littoral présentant des transitions de fouillis terre/mer. Ces zones de transition sont souvent problématiques, car les différents détecteurs existants sont très peu robustes dans ces situations. D'après les résultats que nous avons obtenu dans les zones transitoires entre le bruit thermique Gaussien et le fouillis, tout laisse à penser que le BORD s'adapterait à ces zones de transition du littoral, tout du moins sur l'ensemble des données traitées.

De nombreux autres cas pourraient être envisagés, selon les caractéristiques du radar, selon les situations rencontrées pour le fouillis ou encore selon le type de cible à détecter. L'utilisation d'un radar polarisé nécessiterait, par exemple, de prendre en compte la polarisation des données dans le modèle comme dans la procédure de détection. Que se passerait-il dans le cas de fouillis bistatique? Dans le cas de brouilleurs additionnés au fouillis? Dans le cas de cibles affectées d'un Doppler (par exemple les pales d'un hélicoptère) ou encore dans le cas de cibles décrites comme des processus non-Gaussiens? Toutes ces études permettraient de comprendre plus en détail le comportement du BORD et d'en établir une sorte de fiche technique selon les situations rencontrées.

Toutes ces perspectives théoriques et applicatives montrent que tout l'avenir du BORD est devant lui. A bon entendeur ...

Annexe A

Densités de probabilité d'une variable aléatoire et propriétés

A.1 Généralités

A.1.1 Probabilité d'occurrence, loi et variables aléatoires

En théorie des probabilités, l'ensemble des événements probables est appelé univers et la probabilité associée à tout sous-ensemble de l'univers est égale au nombre des événements probables de ce sous-ensemble divisé par le nombre des événements contenus dans l'univers. Ce comptage d'événements inclus dans tout ensemble E est appelé cardinal de l'ensemble E , se notant $Card(E)$, toujours positif. Soit un univers Ω et deux sous-ensembles $A, B \subset \Omega$. La probabilité d'occurrence du sous-ensemble A se note $\mathbb{P}(A)$ et a pour valeur $\mathbb{P}(A) = \frac{Card(A)}{Card(\Omega)}$. On comprendra bien que la probabilité associée à tout l'univers est égale à 1.

De plus, on a les propriétés évidentes suivantes :

- $\mathbb{P}(\Omega) = 1$,
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$,
- $\mathbb{P}(\emptyset) = 0$,
- Si $A \cap B = \emptyset$ alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$,

et on dira que \mathbb{P} est une probabilité sur Ω si :

$$\mathbb{P}(\Omega) = 1,$$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \text{ si } A \cap B = \emptyset,$$

Si $(A_n, n \in \mathbb{N})$ est une suite croissante dans l'ensemble probabilisé Ω avec $A = \cup_n A_n$, alors $\mathbb{P}(A) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n)$.

Il existe des variables aléatoires dites discrètes (de lois de probabilités discrètes comme Bernoulli, binomiale, Poisson, géométrique, ...) et des v.a. de lois de probabilité à densités continues, dont plusieurs exemples seront donnés dans cette section.

Une v.a. réelle, par définition, est une application mesurable de Ω , univers dénombrable, dans un sous-ensemble E de \mathbb{R} . Une v.a. permet en fait de transporter la mesure de probabilité \mathbb{P} sur E . On dira que la mesure de probabilité $P_X(E) = \mathbb{P}(X \in E)$ est la loi de X .

A.1.2 Fonction de répartition

On définit la fonction de répartition F_X d'une variable aléatoire X comme l'application de \mathbb{R} dans $[0, 1]$ définie à partir de la loi de X comme suit :

$$F_X(x) = P_X(]-\infty; x]) = \mathbb{P}(X \leq x). \quad (\text{A.1})$$

La fonction de répartition vérifie les propriétés suivantes :

- F_X est une fonction croissante, continue à droite,
- $\lim_{x \rightarrow -\infty} F_X(x) = 0$,
- $\lim_{x \rightarrow +\infty} F_X(x) = 1$,
- $F_X(x^-) = \lim_{y \rightarrow x^-} F_X(y) = \mathbb{P}(X < x)$.

D'après ces propriétés, $\mathbb{P}(X = x) = \mathbb{P}(X \leq x) - \mathbb{P}(X < x) = F_X(x) - F_X(x^-) = P_X(x)$. De plus,

- La fonction de répartition caractérise la loi,
- Si une fonction $F : \mathbb{R} \rightarrow \mathbb{R}$ est croissante et continue à droite alors il existe une unique mesure notée dF telle que $dF(]a, b]) = F(b) - F(a)$,
- Cas particulier (le plus usuel) où F_X est dérivable au sens de Lebesgue :

$$F_X(b) - F_X(a) = \int_a^b f(x) dx, \text{ ce qui signifie que } dF_X(x) = f(x), \text{ soit encore, que}$$

la dérivée en x de la fonction de répartition de X est la densité de X si $F_X \in L^1$.

A.1.3 Fonction caractéristique

Par définition, la fonction caractéristique d'une variable aléatoire X est la transformée de Fourier de la loi P_X de X , soit :

$$\Phi_X(u) = \mathbb{E}(e^{iuX}) = \int_{\mathbb{R}} e^{iuX} P_X(x) dx \quad (\text{A.2})$$

La fonction caractéristique vérifie les propriétés suivantes :

- $\Phi_X(0) = 1$,
- $|\Phi_X(u)| \leq 1$,
- $\Phi_X(-u) = \overline{\Phi_X(u)}$,
- Φ_X est continue.

La fonction caractéristique caractérise également la loi et on a la relation d'équivalence suivante ((d) signifie en distribution) :

$$X \stackrel{(d)}{=} Y \iff \Phi_X = \Phi_Y. \quad (\text{A.3})$$

A.1.4 Fonction caractéristique cohérente radiale (f.c.c.r.)

La fonction caractéristique cohérente radiale de l'enveloppe d'un processus aléatoire complexe est la fonction caractéristique bidimensionnelle de ce processus, exprimée sous forme radiale.

Soit un processus aléatoire complexe $u(t) = a(t) + jb(t)$. Ce processus est caractérisé soit par la densité de probabilité jointe de a et de b , $p(a, b)$, soit dualement, par sa fonction caractéristique bidimensionnelle $C(u_1, u_2)$, transformée de Fourier bidimensionnelle de $p(a, b)$:

$$C(u_1, u_2) = \int \int_{\mathbb{R}^2} p(a, b) e^{i(u_1 a + u_2 b)} da db \quad (\text{A.4})$$

Cette expression peut être transformée en une fonction unidimensionnelle de la variable $\rho = \sqrt{u_1^2 + u_2^2}$. Pour ce faire on effectue le changement de variables $(a, b) \mapsto (r, \phi)$ tel que :

$$\begin{aligned} r &= \sqrt{a^2 + b^2} \\ \phi &= \arctan\left(\frac{b}{a}\right) \end{aligned}$$

avec l'hypothèse que ϕ est indépendante de r et uniformément distribuée sur $[0, 2\pi]$, soit $p_\phi(\phi) = 1/2\pi$.

On obtient alors $p(a, b) da db = p_r(r) p_\phi(\phi) dr d\phi$ et :

$$\begin{aligned} C(\rho) &= \int_0^{2\pi} p_\phi(\phi) \left[\int_0^{+\infty} p_r(r) e^{i\rho r \cos(\phi)} dr \right] d\phi \\ &= \int_0^{+\infty} p_r(r) \left[\frac{1}{2\pi} \int_0^{2\pi} e^{i\rho r \cos(\phi)} d\phi \right] dr \\ &= \int_0^{+\infty} p_r(r) J_0(\rho r) dr \end{aligned} \quad (\text{A.5})$$

où $J_0(x)$ est la fonction de Bessel J ordinaire d'ordre 0.

En inversant (A.5), la densité de probabilité de l'enveloppe du processus s'exprime en fonction de $C(\rho)$ comme :

$$p_r(r) = \int_0^{+\infty} r \rho C(\rho) J_0(\rho r) d\rho. \quad (\text{A.6})$$

Remarque : Pour un signal d'enveloppe constante A et de phase uniforme, sa densité de probabilité est une distribution de Dirac en A , soit $p_r(r) = \delta(r - A)$, et sa f.c.c.r. est donnée par $C(\rho) = J_0(\rho A)$.

A.1.5 Propriété d'indépendance de variables aléatoires

Propriété 3 Pour que X_1, \dots, X_d soient indépendantes, il faut et il suffit que

$$\Phi_X(u) = \prod_{j=1}^d \Phi_{X_j}(u_j), \quad (\text{A.7})$$

où $X = (X_1, \dots, X_d)$.

A.1.6 Propriétés sur la somme de n variables aléatoires X_1, \dots, X_n

Propriété 4 Si $S = X_1 + \dots + X_n$, alors $\forall u \in \mathbb{R}$

$$\Phi_S(u) = \prod_{i=1}^n \Phi_{X_i}(u) \quad (\text{A.8})$$

Propriété 5 Si X_1, \dots, X_n sont des variables aléatoires indépendantes, alors la loi de $S = X_1 + \dots + X_n$ est le produit de convolution des lois des X_i , soit

$$P_S = P_{X_1} * \dots * P_{X_n} \quad (\text{A.9})$$

A.1.6.1 Relations de dérivation

Si $\mathbb{E}(|X|^n) < +\infty$, alors, $\Phi_X(u)$ est de classe $\mathcal{C}^n(\mathbb{R})$ et on a :

$$\Phi_X^{(n)}(u) = (i)^n \mathbb{E}(X^n e^{iuX}). \quad (\text{A.10})$$

On note par \widehat{f} , la transformée de Fourier de f .

Si $f \in L^1$ et $\widehat{f} \in L^1$, alors

$$\mathbb{E}(f(X)) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \widehat{f}(-u) \Phi_X(u) du. \quad (\text{A.11})$$

A.1.7 Moments d'une variable aléatoire

On appelle les moments d'ordre k ($k \in \mathbb{N}$), notés μ_k , d'une variable aléatoire le résultat du calcul de l'espérance mathématique $\mathbb{E}(\cdot)$ de la v.a. élevée à la puissance k . On peut donc écrire :

$$\mu_k = \mathbb{E}(X^k) = \int_E X^k P_X(E). \quad (\text{A.12})$$

Dans les sous-sections suivantes, plusieurs densités de probabilité sont rappelées ainsi que les fonctions caractéristiques, les fonctions de répartition et l'expression des moments d'ordre supérieur relativement à la variable aléatoire représentée.

A.2 Loi uniforme sur $[a, b] \subset \mathbb{R}$, $\mathcal{U}([a, b])$

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 :

$$p(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x), \quad (\text{A.13})$$

où $\mathbf{1}_{[a,b]}(x) = 1$ sur $[a, b]$ et 0 sinon.

A.2.1 Fonction de répartition de la loi uniforme

$$F(x) = \int_{-\infty}^x \frac{1}{b-a} \mathbf{1}_{[a,b]}(t) dt = \begin{cases} 0 & \text{si } x < a, \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases} \quad (\text{A.14})$$

A.2.2 Fonction caractéristique de la loi uniforme

$$\Phi(u) = \frac{e^{iub} - e^{iua}}{iu(b-a)} = \frac{\sin[u(b-a)/2]}{u(b-a)/2} e^{iu(b+a)/2}. \quad (\text{A.15})$$

A.2.3 Moments de la loi uniforme

$$\mu_n = \frac{b^{n+1} - a^{n+1}}{(n+1)(b-a)}. \quad (\text{A.16})$$

A.3 Lois Gaussienne ou normale $\mathcal{N}(m, \sigma^2)$ définie sur \mathbb{R}

Une v.a. x , de moyenne m et de variance σ^2 suit une loi Gaussienne (ou normale) $\mathcal{N}(m, \sigma^2)$ quand sa fonction densité de probabilité s'écrit, $\forall x \in \mathbb{R}$:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad (\text{A.17})$$

Quand $m = 0$ et $\sigma^2 = 1$, $\mathcal{N}(0, 1)$ est appelée loi normale standard. Une Gaussienne est entièrement déterminée par m et σ^2 , et si X suit une loi normale standard $\mathcal{N}(0, 1)$, alors $Y = \sigma X + m$ suit la loi normale $\mathcal{N}(m, \sigma^2)$.

A.3.1 Fonction de répartition de la Gaussienne

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) dt = 1 - \mathcal{Q}(x). \quad (\text{A.18})$$

La fonction $\mathcal{Q}(x)$, classiquement appelée *fonction de répartition complémentaire*, ne se calcule pas sous forme analytique. Étant utilisée fréquemment pour évaluer des erreurs de tests, des tables de valeurs existent, que l'on peut trouver, par exemple dans [Abra64].

A.3.2 Fonction caractéristique de la Gaussienne

$$\Phi(u) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) e^{iux} dx = \exp\left(ium - \frac{\sigma^2 u^2}{2}\right). \quad (\text{A.19})$$

A.3.3 Moments d'une gaussienne

Par définition, $\mu_n = \mathbb{E}(X^n)$, pour toute variable aléatoire X . Si la moyenne $m = 0$, on a, $\forall p \in \mathbb{N}$:

$$\mu_{2p+1} = 0 \quad (\text{A.20})$$

$$\mu_{2p} = \frac{(2p)!}{2^p p!} \sigma^{2p} \quad (\text{A.21})$$

Quand $m \neq 0$, on utilise la relation suivante (formule du binôme) :

$$\mathbb{E}[(X+m)^n] = \sum_{k=0}^n C_n^k \mathbb{E}(X^k) m^{n-k}, \quad (\text{A.22})$$

où $\mathbb{E}(X^k)$ est donné par (A.20) et (A.21) selon la parité de k .

A.4 Loi log-normale définie sur \mathbb{R}_+^*

Si X suit une loi normale $\mathcal{N}(m, \sigma^2)$, alors la variable $Y = e^X$ (définie alors sur \mathbb{R}_+^*) suit une loi log-normal :

$$p(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln(y) - \mu)^2}{2\sigma^2}\right) \quad (\text{A.23})$$

A.4.1 Fonction de répartition de la loi log-normale

$$F(y) = \int_{-\infty}^{\log(y)} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx, \quad (\text{A.24})$$

De même que pour la Gaussienne, cette fonction ne se calcule pas sous forme analytique. Les mêmes tables de valeurs doivent être utilisées [Abra64].

A.4.2 Moments de la loi log-normale

$$\mu_n = \exp\left[n\left(m + n\frac{\sigma^2}{2}\right)\right] \quad (\text{A.25})$$

A.5 Loi de Cauchy

La loi de Cauchy est une loi à un paramètre a définie sur \mathbb{R} par :

$$p(x) = \frac{a}{\pi(x^2 + a^2)} \quad (\text{A.26})$$

A.5.1 Fonction de répartition de la loi de Cauchy

$$F(x) = \frac{1}{\pi} \left(\arctan\left(\frac{x}{a}\right) + \frac{\pi}{2} \right) \quad (\text{A.27})$$

A.5.2 Fonction caractéristique de la loi de Cauchy

$$\Phi(u) = e^{-a|u|} \quad (\text{A.28})$$

Les moments de la loi de Cauchy n'existent pas ; ils sont infinis de par la divergence de l'intégrale.

A.5.3 Loi déduite des deux lois Gaussiennes $\mathcal{N}(0, \sigma_1^2)$ et $\mathcal{N}(0, \sigma_2^2)$

Si deux variables X et Y sont indépendantes et telles que $X \sim \mathcal{N}(0, \sigma_1^2)$ et $Y \sim \mathcal{N}(0, \sigma_2^2)$, alors $U = \frac{X}{Y}$ suit une loi de Cauchy de paramètre $a = \frac{\sigma_1}{\sigma_2}$.

A.6 Loi exponentielle $\mathcal{E}(b)$ définie sur \mathbb{R}_+

La loi exponentielle de paramètre b , notée $\mathcal{E}(b)$ admet pour densité de probabilité :

$$p(x) = b e^{-bx}. \quad (\text{A.29})$$

A.6.1 Fonction de répartition de la loi exponentielle

$$F(y) = \int_0^y b e^{-bt} dt = 1 - e^{-bx}. \quad (\text{A.30})$$

A.6.2 Fonction caractéristique de la loi exponentielle

$$\Phi(u) = \int_0^{+\infty} b e^{-bx} e^{iux} dx = \frac{b}{b - iu}. \quad (\text{A.31})$$

A.6.3 Moments de la loi exponentielle

$$\mu_n = \frac{n!}{b^n}. \quad (\text{A.32})$$

A.7 Loi de Laplace (ou exp. bilatérale) $\mathcal{L}(b)$ définie sur \mathbb{R}

$$p(x) = \frac{b}{2} e^{-b|x|}. \quad (\text{A.33})$$

A.7.1 Fonction de répartition de la loi de Laplace

$$F(y) = \int_{-\infty}^x b e^{-b|t|} dt = \begin{cases} \frac{e^{bx}}{2} & \text{si } x \leq 0 \\ 1 - \frac{e^{-bx}}{2} & \text{si } x \geq 0 \end{cases} \quad (\text{A.34})$$

A.7.2 Fonction caractéristique de la loi de Laplace

$$\Phi(u) = \int_{-\infty}^{+\infty} b e^{-b|x|} e^{iux} dx = \frac{2b^2}{b^2 + u^2}. \quad (\text{A.35})$$

A.7.3 Moments de la loi de Laplace

$$\mu_n = \begin{cases} 0 & \text{si } n \text{ est impair} \\ \frac{2n!}{b^n} & \text{si } n \text{ est pair.} \end{cases} \quad (\text{A.36})$$

A.8 Loi de Rayleigh de paramètre b , définie sur \mathbb{R}_+

$$p(x) = \frac{x}{b^2} \exp\left(-\frac{x^2}{2b^2}\right) \quad (\text{A.37})$$

A.8.1 Fonction de répartition de la loi de Rayleigh

$$F(x) = \int_0^x \frac{t}{b^2} \exp\left(-\frac{t^2}{2b^2}\right) dt = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\frac{x^2}{2b^2}} & \text{si } x \geq 0 \end{cases} \quad (\text{A.38})$$

A.8.2 Fonction caractéristique de la loi de Rayleigh

$$\Phi(u) = \int_0^{+\infty} \frac{x}{b^2} e^{-\frac{x^2}{2b^2}} e^{iux} dx = 1 + \sqrt{\pi} e^{-\frac{u^2 b^2}{2}} \frac{i u b}{\sqrt{2}} \operatorname{Erfc} \left(-\frac{i u b}{\sqrt{2}} \right), \quad (\text{A.39})$$

où

$$\operatorname{Erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{+\infty} e^{-t^2} dt = 1 - \operatorname{Erf}(z),$$

et

$$\operatorname{Erfc}(-z) = 1 + \operatorname{Erf}(z) = 2 - \operatorname{Erfc}(z).$$

A.8.3 Moments de la loi de Rayleigh

$$\mu_n = \left(\sqrt{2} b \right)^n \Gamma \left(\frac{n}{2} + 1 \right) \quad (\text{A.40})$$

A.8.4 Loi de Rayleigh déduite de la loi normale $\mathcal{N}(0, \sigma^2)$

Si $X \sim \mathcal{N}(0, \sigma^2)$ et $Y \sim \mathcal{N}(0, \sigma^2)$, alors $Z = \sqrt{X^2 + Y^2}$ suit une loi de Rayleigh de paramètre σ^2 .

A.8.5 Lien entre la loi de Rayleigh et la loi exponentielle

Si X suit une loi de Rayleigh de paramètre b , alors $Y = X^2$ suit une loi exponentielle de paramètre $\frac{1}{2b^2}$.

A.9 Loi de Rice

La loi de Rice est une loi à deux paramètres σ et α définie sur \mathbb{R}_+ ayant pour expression :

$$p(x) = \frac{x}{\sigma^2} \exp \left(-\frac{x^2 + \alpha^2}{2\sigma^2} \right) I_0 \left(\frac{\alpha x}{\sigma^2} \right), \quad (\text{A.41})$$

où $I_\nu(u)$ est la fonction de Bessel modifiée de première espèce dont on peut donner des représentations intégrales :

$$I_\nu(u) = \frac{(u/2)^\nu}{\sqrt{\pi}\Gamma\nu + 1/2} \int_0^\pi e^{\pm u \cos(\theta)} (\sin(\theta))^{2\nu} d\theta \quad (\text{A.42})$$

$$= \frac{(u/2)^\nu}{\sqrt{\pi}\Gamma\nu + 1/2} \int_{-1}^1 (1-t^2)^{\nu-1/2} e^{\pm ut} dt, \quad \text{avec } \Re(\nu) > -\frac{1}{2} \quad (\text{A.43})$$

A.9.1 Lien entre la loi Gaussienne et la loi de Rice

La loi de Rice s'obtient comme la loi de $R = \sqrt{X^2 + Y^2}$ où $X \sim \mathcal{N}(m_1, \sigma^2)$ et $Y \sim \mathcal{N}(m_2, \sigma^2)$ sont indépendantes. La loi de Rice obtenue a pour paramètres σ et $\alpha^2 = m_1^2 + m_2^2$.

A.10 Loi gamma $\mathcal{G}(a, b)$ définie sur \mathbb{R}_+

La loi Gamma est une loi à deux paramètres a et b et a pour expression :

$$p(x) = \frac{1}{b^a \Gamma(a)} x^{a-1} e^{-\frac{x}{b}} \quad (\text{A.44})$$

où $\Gamma(\cdot)$ est la fonction Gamma donnée $\forall z > 0$ par : $\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt$.

A.10.1 Fonction de répartition de la loi Gamma

$$F(x) = \int_0^x \frac{1}{b^a \Gamma(a)} t^{a-1} e^{-\frac{t}{b}} dt \quad (\text{A.45})$$

Cette fonction n'a pas d'expression analytique mais est appelée, dans les tables, fonction Gamma incomplète (GammaInc).

A.10.2 Fonction caractéristique de la loi Gamma

$$\Phi(u) = \int_0^{+\infty} \frac{1}{b^a \Gamma(a)} t^{a-1} e^{-\frac{t}{b}} e^{iut} dt = (1 - i u b)^{-a}. \quad (\text{A.46})$$

A.10.3 Moments de la loi Gamma

$$\mu_n = \frac{b^n \Gamma(a + n)}{\Gamma(a)}. \quad (\text{A.47})$$

A.11 Loi Inverse Gamma $\mathcal{IG}(a, b)$ définie sur \mathbb{R}_+

La loi inverse Gamma de paramètre a et b est la loi de la variable $1/X$ où X suit une loi Gamma(a, b). Elle est définie sur \mathbb{R}_+ et son expression est la suivante :

$$p(x) = \frac{1}{b^a \Gamma(a)} x^{-a-1} e^{-\frac{1}{xb}}, \quad (\text{A.48})$$

A.11.1 Moments de la loi Inverse Gamma

Les moments de la loi Inverse Gamma sont donnés par :

$$\mu_n = \frac{\Gamma(a - n)}{b^n \Gamma(a)}, \quad (\text{A.49})$$

et n'existent que si $(a - n) \notin \mathbb{Z}$, soit $a \notin \mathbb{N}^*$.

A.12 La loi Beta

La loi Beta est une loi à deux paramètres $r > 0$ et $s > 0$, définie sur $[0, 1]$ par :

$$p(x) = \frac{1}{B(r, s)} x^{r-1} (1-x)^{s-1}, \quad (\text{A.50})$$

où $B(r, s)$ est la fonction Beta, reliée à la fonction Gamma par :

$$B(r, s) = \int_0^1 t^{r-1} (1-t)^{s-1} dt = \frac{\Gamma(r) \Gamma(s)}{\Gamma(r+s)}. \quad (\text{A.51})$$

A.12.1 Moments de la loi Beta

Les moments de la loi Beta s'exprime en fonction des fonctions Beta comme :

$$\mu_n = \frac{B(n+r, s)}{B(r, s)}. \quad (\text{A.52})$$

A.13 La K-distribution

La K-distribution, loi à deux paramètres b et ν , est définie sur \mathbb{R}_+ et doit son nom à la fonction de Bessel modifiée de deuxième espèce, $K_\nu(x)$.

$$p(x) = \frac{b^{\nu+1}}{2^{\nu-1} \Gamma(\nu)} x^\nu K_{\nu-1}(bx) \quad (\text{A.53})$$

A.13.1 Fonction de répartition de la loi K

$$F(x) = \int_0^x \frac{b^{\nu+1}}{2^{\nu-1} \Gamma(\nu)} t^\nu K_{\nu-1}(bt) dt = 1 - \frac{(bx)^\nu}{\Gamma(\nu) 2^{\nu-1}} K_\nu(bx). \quad (\text{A.54})$$

A.13.2 Fonction caractéristique cohérente radiale de la loi K

$$C(\rho) = \left(\frac{1}{1 + \frac{\rho^2}{b^2}} \right)^\nu \quad (\text{A.55})$$

A.13.3 Moments de la loi K

$$\mu_n = \frac{\Gamma(n/2 + 1) \Gamma(n/2 + \nu)}{\Gamma(\nu)} \left(\frac{2}{b} \right)^n. \quad (\text{A.56})$$

A.14 Loi de Weibull

La loi de Weibull est une loi à deux paramètres a et b , définie sur \mathbb{R}_+ :

$$p(x) = a b x^{b-1} \exp(-ax^b) \quad (\text{A.57})$$

A.14.1 Fonction de répartition de la loi de Weibull

$$F(x) = \int_0^x a b t^{b-1} \exp(-a t^b) dt = 1 - e^{-x} \quad (\text{A.58})$$

A.14.2 Moments de la loi de Weibull

$$\mu_n = a^{-n/b} \Gamma(1 + n/b). \quad (\text{A.59})$$

A.15 Loi du χ^2 centrée

La loi du χ^2 centrée est une loi à un paramètre ν ($\in \mathbb{N} - \{0\}$, noté \mathbb{N}^*) déterminant le nombre de *degrés de liberté* de la loi. On dit qu'une variable aléatoire X suit une loi χ_ν^2 (loi du χ^2 à ν degrés), si $\forall x > 0$, sa densité de probabilité est donnée par :

$$p(x) = \frac{1}{2^{\nu/2} \Gamma(\frac{\nu}{2})} x^{\nu/2-1} \exp\left(-\frac{x}{2}\right) \quad (\text{A.60})$$

La loi du χ_ν^2 coïncide avec la loi Gamma de paramètres $\nu/2$ et 2, soit $\mathcal{G}\left(\frac{\nu}{2}, 2\right)$. Pour $\nu > 2$, cette loi admet un mode en $x = \nu - 2$. Lorsque $\nu = 2$, la loi χ_ν^2 est une loi exponentielle de paramètre 2 et pour $1 \leq \nu < 2$, la loi tend vers $+\infty$ en 0.

A.15.1 Fonction de répartition de la loi χ_ν^2

La fonction de répartition de la loi χ_ν^2 admet deux expressions, selon la parité de ν . Elle se définit sur \mathbb{R}_+ par :

$$F(x) = \int_0^x p(t) dt = 1 - \mathcal{Q}_{\chi_\nu^2}(x), \quad (\text{A.61})$$

où, pour $\nu \geq 2$ pair, on a [Abra64] :

$$\mathcal{Q}_{\chi_\nu^2}(x) = \exp\left(-\frac{x}{2}\right) \sum_{k=0}^{\nu/2-1} \frac{x^k}{2^k k!}, \quad (\text{A.62})$$

et pour ν impair, on a :

$$\mathcal{Q}_{\chi_\nu^2}(x) = \begin{cases} 2 \mathcal{Q}(\sqrt{x}) & \nu = 1 \\ 2 \mathcal{Q}(\sqrt{x}) + \frac{\exp(-x/2)}{\sqrt{\pi}} \sum_{k=1}^{(\nu-1)/2} \frac{(k-1)! (2x)^{k-1/2}}{(2k-1)!} & \nu \geq 3 \end{cases} \quad (\text{A.63})$$

La fonction \mathcal{Q} est la fonction de répartition complémentaire de la Gaussienne, dont l'expression est donnée par (A.18).

A.15.2 Fonction caractéristique de la loi χ_ν^2

$$\Phi(u) = \int_0^{+\infty} p(t) e^{iut} dt = \frac{1}{(1 - 2iu)^{\nu/2}}. \quad (\text{A.64})$$

A.15.3 Moments de la loi χ_ν^2

$$\mu_n = 2^n \frac{\Gamma(n + \frac{\nu}{2})}{\Gamma(\frac{\nu}{2})}. \quad (\text{A.65})$$

A.15.4 Loi déduite la Gaussienne

La loi du χ_ν^2 est en fait la loi résultante de la somme de ν variables aléatoires Gaussiennes standard i.i.d. au carré, soit encore la loi de $\sum_{k=1}^{\nu} x_k^2$, où $x_i \sim \mathcal{N}(0, 1)$.

A.15.5 Somme de K variables aléatoires $\chi_{\nu_i}^2$

Soient K variables aléatoires indépendantes X_i telles que $X_i \sim \chi_{\nu_i}^2$. Alors $S = \sum_{i=1}^K X_i \sim \chi_r^2$, où $r = \sum_{i=1}^K \nu_i$.

A.16 Loi du χ^2 décentrée

La loi du χ^2 décentrée est une loi à deux paramètres ν , le nombre de *degrés de liberté*, et λ , la paramètre de *décentralité*, notée $\chi_\nu'^2(\lambda)$.

Cette loi est la loi de $x = \sum_{k=1}^{\nu} x_k^2$, où $x_i \sim \mathcal{N}(\mu_i, 1)$ et ainsi le paramètre λ est défini d'après les moyennes μ_i comme $\lambda = \sum_{k=1}^{\nu} \mu_k^2$. La densité de probabilité correspondante est définie sur \mathbb{R}_+ par une intégrale ou une série par :

$$p(x) = \frac{1}{2} \left(\frac{x}{\lambda}\right)^{(\nu-2)/4} \exp\left(\frac{x+\lambda}{2}\right) I_{\nu/2-1}(\sqrt{\lambda x}) \quad (\text{A.66})$$

où $I_a(u)$ est la fonction de Bessel de première espèce d'ordre a . Cette fonction est définie soit sous forme intégrale, soit sous forme d'une série [Abra64].

Lorsque $\lambda = 0$, la loi $\chi_\nu'^2(\lambda)$ se réduit à la loi du χ_ν^2 centrée.

A.16.1 Moments de la loi $\chi_\nu'^2$

$$\mu_n = 2^n \exp\left(-\frac{\lambda}{2}\right) \sum_{k=0}^{+\infty} \left(\frac{\lambda}{4}\right)^k \frac{\Gamma(k + n + \frac{\nu}{2})}{k! \Gamma(k + \frac{\nu}{2})}. \quad (\text{A.67})$$

A.17 Loi de Fisher F centrée

La loi de Fisher, ou loi F centrée est la loi de :

$$x = \frac{x_1/\nu_1}{x_2/\nu_2}, \quad (\text{A.68})$$

où $x_1 \sim \chi_{\nu_1}^2$ et $x_2 \sim \chi_{\nu_2}^2$ sont indépendantes. La densité de probabilité résultante, notée F_{ν_1, ν_2} est définie sur \mathbb{R}_+ par :

$$\frac{\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2}}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)} \frac{x^{\nu_1/2-1}}{\left(1 + \frac{\nu_1}{\nu_2} x\right)^{\frac{\nu_1+\nu_2}{2}}}, \quad (\text{A.69})$$

où $B(a, b)$ est la fonction Beta, reliée à la fonction Gamma par (A.51).

Remarque : La loi Beta $\mathcal{B}(r, s)$, définie sur $[0, 1]$ avec $r > 0$ et $s > 0$, est la loi de $\frac{1}{1 + \frac{s}{r} Y}$, où $Y \sim F_{2s, 2r}$.

A.18 Loi de Fisher F' décentrée

La loi de Fisher décentrée est la loi de :

$$x = \frac{x_1/\nu_1}{x_2/\nu_2}, \quad (\text{A.70})$$

où $x_1 \sim \chi_{\nu_1}^2(\lambda)$ et $x_2 \sim \chi_{\nu_2}^2(\lambda)$ sont indépendantes. La densité de probabilité résultante, notée $F'_{\nu_1, \nu_2}(\lambda)$ est définie sur \mathbb{R}_+ par :

$$F'_{\nu_1, \nu_2}(\lambda) = \exp\left(-\frac{\lambda}{2}\right) \sum_{k=0}^{+\infty} \frac{\left(\frac{\lambda}{2}\right)^k}{k!} \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2+k}}{B\left(\frac{\nu_1+2k}{2}, \frac{\nu_2}{2}\right)} x^{\nu_1/2+k+1} \left(1 + \frac{\nu_1}{\nu_2} x\right)^{-(\nu_1+\nu_2)/2-k}. \quad (\text{A.71})$$

Lorsque $\lambda = 0$, la loi F' décentrée est réduite à la loi F centrée.

A.19 Loi de Student's -t

Soit $X \sim \mathcal{N}(0, 1)$ et $Y \sim \chi_r^2$ deux variables aléatoires indépendantes. La loi de

$$T = \frac{X}{\sqrt{Y/r}}$$

est une loi de Student's-t à r degrés de liberté, dont la densité de probabilité, définie sur \mathbb{R} est donnée par :

$$p(t) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi r} \Gamma\left(\frac{r}{2}\right)} \left(1 + \frac{t^2}{r}\right)^{-(r+1)/2} \quad (\text{A.72})$$

A.20 Convergences vers la loi normale

A.20.1 Théorème central limite (TCL) pour des v.a. i.i.d.

Soit X_1, \dots, X_n une suite de v.a. i.i.d. réelles. On suppose que $\mathbb{E}(X^2) < +\infty$ et soit $m = \mathbb{E}(X)$, $\sigma^2 = \text{Var}(X)$. Alors :

$$\frac{X_1 + \dots + X_n - nm}{\sqrt{n}} \xrightarrow{(d), n \rightarrow +\infty} \mathcal{N}(0, \sigma^2) \quad (\text{A.73})$$

A.20.2 TCL pour des v.a. indépendantes non identiquement distribuées

Soit X_1, \dots, X_n une suite de v.a. réelles indépendantes. Si $\forall j \in \mathbb{N}$, $m_j = \mathbb{E}(X_j)$ et $\text{Var}(X_j) = \sigma_j^2$, alors

$$\frac{(X_1 - m_1) + \dots + (X_n - m_n)}{\sqrt{\sigma_1^2 + \dots + \sigma_n^2}} \xrightarrow{(d), n \rightarrow +\infty} \mathcal{N}(0, 1) \quad (\text{A.74})$$

si et seulement si $\forall 1 \leq j \leq n$, on a : $\frac{\sigma_j^2}{\sigma_1^2 + \dots + \sigma_n^2} \xrightarrow{n \rightarrow +\infty} 0$.

Remarque 1 : Le TCL permet de montrer le caractère gaussien de toute somme de n v.a. indépendantes quand n devient grand.

Remarque 2 : On peut, par une approche probabiliste et grâce au TCL, retrouver la formule de Stirling (qui donne un équivalent asymptotique de $n!$, soit $n! \sim n^n e^{-n} \sqrt{2\pi n}$).

A.20.3 Convergence de la K-distribution vers la Gaussienne

On peut montrer par les fonctions caractéristiques radiales (f.c.r.) que la f.c.r. de l'enveloppe de deux v.a. K-distribuées tend vers la f.c.r. de la Gaussienne quand $\nu \rightarrow +\infty$ (paramètre de forme de la K-distribution).

La f.c.r. de l'enveloppe d'une v.a. K-distribuée, de paramètres ν et b s'écrit :

$$C(\rho) = \left(\frac{1}{1 + \frac{\rho^2}{b^2}} \right)^\nu \quad (\text{A.75})$$

où les deux paramètres sont reliés par la puissance σ^2 par $b = 2\sqrt{\frac{\nu}{\sigma^2}}$. Ainsi, (A.75) peut s'exprimer en fonction de ν et σ^2 , ce qui donne :

$$C(\rho) = \left(1 + \frac{\sigma^2 \rho^2}{4\nu} \right)^{-\nu} = \exp \left[-\nu \left(\ln \left(1 + \frac{\sigma^2 \rho^2}{4\nu} \right) \right) \right]$$

Lorsque $\nu \rightarrow +\infty$, $\sigma^2 \rho^2 / (4\nu) \rightarrow 0$ et donc

$$\ln \left(1 + \frac{\sigma^2 \rho^2}{4\nu} \right) = \frac{\sigma^2 \rho^2}{4\nu} - \frac{\sigma^4 \rho^4}{32\nu^2} + \mathcal{O} \left(\frac{\sigma^4 \rho^4}{32\nu^2} \right)$$

D'où

$$C(\rho) \xrightarrow{\nu \rightarrow \infty} \exp\left(-\frac{\sigma^2 \rho^2}{4}\right), \quad (\text{A.76})$$

qui coïncide avec la fonction caractéristique de la Gaussienne $\mathcal{N}(0, \frac{\sigma^2}{2})$.

Annexe B

Méthodes d'estimation de densités de probabilité

Estimer une densité de probabilité est un problème délicat à résoudre si on recherche une expression analytique de la loi sous la forme d'une fonction continue. La tâche devient encore plus complexe lorsque la densité à estimer est multivariée (loi jointe d'un vecteur par exemple). Dans cette section nous allons présenter trois méthodes d'estimation de densités de probabilité : la première, la méthode des histogrammes est une méthode empirique puisqu'il ne s'agit que d'un comptage de valeurs, à partir d'échantillons issus d'une même loi et le résultat de l'estimation se présente sous la forme d'un histogramme de valeurs (d'où son nom). La deuxième méthode généralise la première en apportant une pondération au comptage des valeurs sous forme d'une fonction continue appelée *noyau*. Le résultat est en quelque sorte le lissage de l'histogramme obtenu par la première méthode. Ces deux méthodes dépendent des échantillons de la loi. La troisième méthode, les approximations de Padé, se base sur l'approximation de la fonction génératrice des moments (transformée de Laplace de la loi) sous forme de fraction rationnelle. Cette fraction rationnelle est ensuite développée en éléments simples et on retrouve la densité de probabilité par transformée de Laplace inverse. Cette méthode permet d'obtenir une expression analytique simple de la loi estimée, indépendante des échantillons traités (contrairement aux deux premières méthodes).

B.1 Estimation empirique par la méthode des histogrammes

L'estimation \hat{f} d'une densité de probabilité f par un histogramme s'effectue en comptant le nombre d'échantillons X_i égaux aux valeurs considérées, ramené au nombre total des échantillons (N). Le résultat obtenu n'est pas une fonction continue et on peut l'exprimer comme suit :

$$\hat{f}(x) = \frac{1}{N} \sum_{i=1}^N \delta(x - X_i), \quad (\text{B.1})$$

où $\delta(y) = 1$ si $y = 0$ et 0 sinon. Cette méthode, dépendante des données, est très simple et permet d'obtenir rapidement une idée de la forme de la loi. En revanche pour une estimation plus précise, il n'est pas très judicieux de l'utiliser.

Pour introduire assez rapidement la méthode des noyaux, on peut généraliser, dans un premier temps, la méthode décrite ci-dessus, en ne comptant pas le nombre d'échantillons égaux à certaines valeurs, mais en comptant le nombre d'échantillons *tombant* dans un intervalle donné

de longueur $2h$ ($h \geq 0$). L'estimateur obtenu est appelé l'estimateur *naïf* ou *rectangulaire* et s'exprime comme suit :

$$\hat{f}(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{h} w\left(\frac{x - X_i}{h}\right), \quad (\text{B.2})$$

où $w(\cdot)$ est la fonction *poids* définie par :

$$w(x) = \begin{cases} \frac{1}{2} & \text{si } |x| < 1 \\ 0 & \text{sinon} \end{cases} \quad (\text{B.3})$$

Cet estimateur est loin d'être satisfaisant puisque \hat{f} n'est pas une fonction continue mais a des sauts aux points $X_i \pm h$ et admet des dérivées nulles partout ailleurs. On peut donc généraliser la méthode en remplaçant la fonction poids $w(\cdot)$ par une fonction noyau.

B.2 Généralisation de la méthode des histogrammes : la méthode des noyaux

Soit K une fonction noyau, supposée symétrique et qui vérifie les propriétés suivantes :

$$\begin{aligned} \int K(t) dt &= 1 \\ \int t K(t) dt &= 0 \\ \int t^2 K(t) dt &= k_2 \neq 0 \end{aligned} \quad (\text{B.4})$$

On suppose de plus que la densité inconnue f est de classe \mathcal{C}^∞ . L'estimateur noyau de \hat{f} admet donc pour expression :

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - X_i}{h}\right) \quad (\text{B.5})$$

où h est la largeur de la fenêtre. Dans la suite, nous donnons quelques exemples de noyaux avec leur efficacité, notée $eff(K)$ sans en donner le détail du calcul que l'on peut trouver dans ([Silv86]).

B.2.1 Noyau triangulaire

Le noyau triangulaire K_t est défini comme suit, et représenté sur la figure B.1 :

$$K_t(t) = \begin{cases} 1 - |t| & \text{si } |t| < 1 \\ 0 & \text{sinon} \end{cases} \quad (\text{B.6})$$

avec

$$eff(K_t) = \left(\frac{243}{250}\right)^{1/2} \approx 0.9859$$

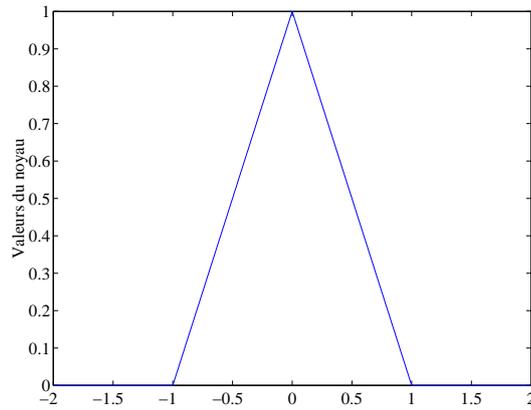


FIG. B.1 – Noyau Triangulaire

B.2.2 Noyau bipoids

Le noyau bipoids K_b est défini comme suit, et représenté sur la figure B.2 :

$$K_b(t) = \begin{cases} \frac{15}{16} (1 - t^2)^2 & \text{si } |t| < 1 \\ 0 & \text{sinon} \end{cases} \quad (\text{B.7})$$

avec

$$eff(K_b) = \left(\frac{3087}{3215} \right)^{1/2} \approx 0.9939$$

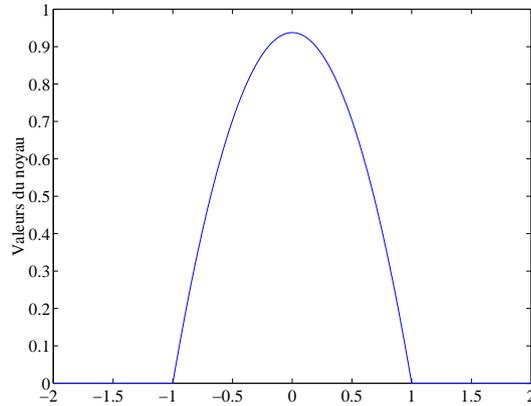


FIG. B.2 – Noyau Bipoids

B.2.3 Noyau gaussien

Le noyau gaussien K_g est défini comme suit, et représenté sur la figure B.3 :

$$K_g(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \quad t \in \mathbb{R} \quad (\text{B.8})$$

avec

$$eff(K_g) = \left(\frac{36\pi}{125} \right)^{1/2} \approx 0.9512$$

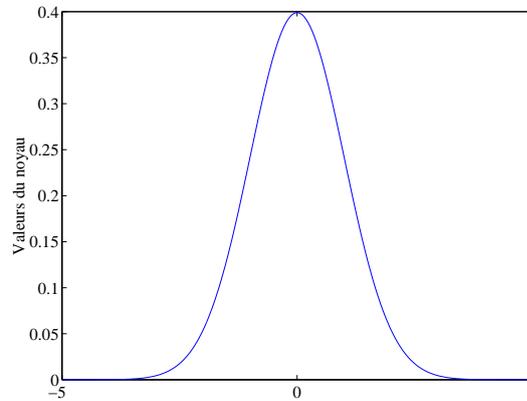


FIG. B.3 – Noyau Gaussien

B.2.4 Noyau d'Epanechnikov

Le noyau d'Epanechnikov $K_e(t)$ (du nom d'Epanechnikov 1969) est défini comme suit, et représenté sur la figure B.4 :

$$K_e(t) = \begin{cases} \frac{3}{4\sqrt{5}} \left(1 - \frac{t^2}{5} \right) & \text{si } -\sqrt{5} < t < \sqrt{5} \\ 0 & \text{sinon} \end{cases} \quad (\text{B.9})$$

avec

$$eff(K_e) = 1$$

On notera que son efficacité est égale à 1, du fait que le calcul de l'efficacité, telle que nous l'avons donnée ici, est relative au noyau d'Epanechnikov.

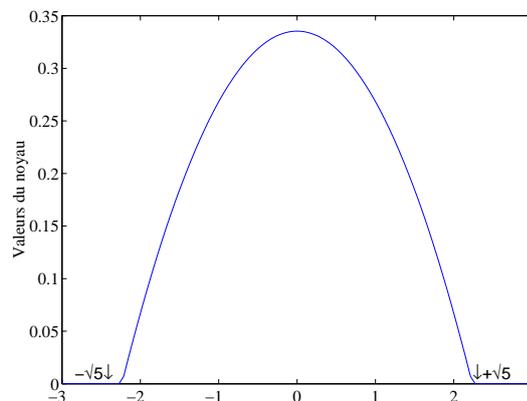


FIG. B.4 – Noyau d'Epanechnikov

B.2.5 Choix du paramètre h

La valeur idéale pour h , du point de vue minimisation de l'erreur quadratique moyenne intégrée (MISE) approximée (B.10)

$$\frac{1}{4}h^4 k_2^2 \int f''(x)^2 dx + \frac{1}{Nh} \int K(t)^2 dt \quad (\text{B.10})$$

a été trouvée par PARZEN, 1962 comme égale à h_{opt} :

$$h_{opt} = (N k_2^2)^{-1/5} \left\{ \int K(t)^2 dt \right\}^{1/5} \left\{ \int f''(x)^2 dx \right\}^{-1/5} \quad (\text{B.11})$$

Expliquons-nous un peu : par définition, l'erreur quadratique moyenne intégrée (MISE) est définie comme suit :

$$\begin{aligned} \text{MISE}(\tilde{f}) &= \int \mathbb{E} [\tilde{f}(x) - f(x)]^2 dx \\ &= \int [\mathbb{E}(\tilde{f}(x)) - f(x)]^2 dx + \int \text{var}(\tilde{f}(x)) dx \end{aligned} \quad (\text{B.12})$$

où $\mathbb{E}(\cdot)$ désigne l'espérance mathématique.

Le premier terme du second membre de (B.12) $\int [\mathbb{E}(\tilde{f}(x)) - f(x)]^2 dx$, n'est rien d'autre que l'intégrale du carré du biais $\mathbb{E}(\tilde{f}(x)) - f(x)$, avec :

$$\mathbb{E}(\tilde{f}(x)) = \int h^{-1} K\left(\frac{x-y}{h}\right) f(y) dy$$

si on considère \tilde{f} comme étant l'estimée par le noyau K , et donc, en effectuant le changement de variable $y = x - ht$, et sachant que $\int K(t) dt = 1$:

$$\begin{aligned} \text{biais}_h(x) &= \int h^{-1} K\left(\frac{x-y}{h}\right) f(y) dy - f(x) \\ &= \int K(t) [f(x - ht) - f(x)] dt \end{aligned}$$

Un développement en série de Taylor donne :

$$f(x - ht) = f(x) - ht f'(x) + 1/2 h^2 t^2 f''(x) + \dots$$

Avec les hypothèses (B.4) émises sur K , nous obtenons :

$$\begin{aligned} \text{biais}_h(x) &= -h f'(x) \int t K(t) dt + 1/2 h^2 f''(x) \int t^2 K(t)^2 dt + \dots \\ &= 1/2 h^2 f''(x) k_2 + \text{termes d'ordre supérieur en } h \end{aligned}$$

Ce qui nous procure une approximation du carré intégré du biais, nécessaire à (B.10) :

$$\int \text{biais}_h(x)^2 dx \approx 1/4 h^4 k_2^2 \int f''(x)^2 dx$$

Revenons maintenant au deuxième terme du second membre de (B.12), $\int \text{var}(\tilde{f}(x)) dx$.

Pour le calcul de la variance, on a :

$$\begin{aligned}
 \text{var}(\tilde{f}(x)) &= \frac{1}{N} \text{var}[w(X_i, t)] = \frac{1}{N} \left[\int w(x, t)^2 f(x) dx - \left(\int w(x, t) f(x) dx \right)^2 \right] \\
 &= \int \frac{1}{N h^2} K \left(\frac{x-y}{h} \right)^2 f(y) dy - \left[\frac{1}{N h} \int K \left(\frac{x-y}{h} \right) f(y) dy \right]^2 \\
 &= \frac{1}{N h^2} \int K \left(\frac{x-y}{h} \right)^2 f(y) dy - \frac{1}{N h} [f(x) + \text{biais}_h(x)]^2 \\
 &\approx \frac{1}{N h} \int f(x - h t) K(t)^2 dt - \frac{1}{N} [f(x) + \mathcal{O}(h^2)]^2 \\
 &\approx \frac{1}{N h} \int [f(x) - h t f'(x) + \dots] K(t)^2 dt + \mathcal{O}(n^{-1}) \\
 &= \frac{1}{N h} f(x) \int K(t)^2 dt + \mathcal{O}(n^{-1}) \\
 \text{var}(\tilde{f}(x)) &\approx \frac{1}{N h} f(x) \int K(t)^2 dt
 \end{aligned}$$

En remarquant que f est une densité, en intégrant cette dernière expression, on obtient une simple approximation de l'intégrale de la variance, nécessaire à (B.10) :

$$\int \text{var}(\tilde{f}(x)) \approx \frac{1}{N h} \int K(t)^2 dt$$

Comme le montre (B.11), h_{opt} dépend lui-même de la densité f inconnue; l'idée est donc de prendre en référence la famille des lois normales pour attribuer une valeur à f et d'en déduire un h_{opt} de référence. Ceci n'est pas gênant du tout pour notre approche puisque nous effectuons notre estimation à l'aide d'un noyau ADAPTATIF : f est une première fois estimée (\tilde{f}) avec h_{opt} , puis est ré-estimée avec cette fois-ci un h_{opt2} adapté à \tilde{f} :

$$h_{opt2} = h_{opt} \times \lambda$$

où

$$\begin{aligned}
 \lambda &= \left(\frac{\tilde{f}}{g} \right)^{-1/2} \\
 \log g &= \frac{1}{N} \sum \log(\tilde{f})
 \end{aligned}$$

En prenant donc pour référence une loi normale d'écart-type σ , et en posant Φ une densité de loi normale standard, on peut affecter une valeur à $\int f''(x)^2 dx$

$$\begin{aligned}
 \int f''(x)^2 dx &= \sigma^{-5} \int \Phi''(x)^2 dx \\
 &= \frac{3}{8} \pi^{-1/2} \sigma^{-5} \\
 &\approx 0.212 \sigma^{-5}
 \end{aligned} \tag{B.13}$$

Pour déterminer h_{opt} , substituons (B.13) dans (B.11) pour obtenir (B.14) :

$$\begin{aligned} h_{opt} &= (4\pi)^{-1/10} \left(\frac{3}{8} \pi^{-1/2} \right)^{-1/5} \sigma N^{-5} \\ &= \left(\frac{4}{3} \right)^{1/5} \sigma N^{-5} \\ &= 1.06 \sigma N^{-5} \end{aligned} \tag{B.14}$$

B.3 Les approximants de Padé

Cette méthode nécessite de connaître ou d'estimer, jusqu'à un ordre donné, les différents moments des variables étudiées. Connaissant les N données réelles x_i , les différents moments μ_k peuvent être estimés classiquement comme :

$$\mu_k = \frac{1}{N} \sum_{i=1}^N x_i^k$$

Soit $\Phi(u)$ la *fonction génératrice de moments*, elle peut s'écrire sous la forme d'une série infinie dont les coefficients du développement dépendent des moments μ_k :

$$\Phi(u) = \sum_{n=0}^{\infty} c_n u^n = \sum_{n=0}^{\infty} \mu_n \frac{(-u)^n}{n!} \tag{B.15}$$

Classiquement, la fonction génératrice d'un processus Z est la transformée de Laplace de sa densité de probabilité $p(z)$ (unilatérale si $p(z)$ est à valeur dans \mathbb{R}_+) et s'écrit :

$$\Phi_Z(u) = \mathbb{E}(e^{-uZ}) = \int_{-\infty}^{+\infty} p(z) e^{-uz} dz$$

avec $\mathbb{E}(\cdot)$ désignant l'espérance mathématique. Si on note par $\mu_n = \mathbb{E}(Z^n) = \int_0^{+\infty} z^n p(z) dz$, les moments d'ordre n du processus Z , l'expression de $\Phi_Z(u)$ en dépend, en développant en série e^{-uz} ; on obtient alors (B.15).

La théorie des approximants de Padé consiste à construire une approximation de $\Phi(u)$, notée $\mathbf{P}^{[L/M]}(\mathbf{u})$, sous forme de fraction rationnelle avec un dénominateur de degré M et un numérateur de degré L , telle que son développement en série de puissance croissante coïncide avec celui de $\Phi(u)$ jusqu'à l'ordre $L + M$.

B.3.1 L'approximation de Padé à 1 point

L'approximation de Padé à 1 point de la fonction génératrice de moments $\Phi(u)$, d'ordre $[L/M]$ et notée $P^{[L/M]}(u)$ est, par définition, la fonction rationnelle suivante :

$$P^{[L/M]}(u) = \frac{\sum_{n=0}^L a_n u^n}{\sum_{n=0}^M b_n u^n} \tag{B.16}$$

où les coefficients $\{a_n\}$ et $\{b_n\}$ sont définis pour qu'on ait l'égalité :

$$\frac{\sum_{n=0}^L a_n u^n}{\sum_{n=0}^M b_n u^n} = \sum_{n=0}^{L+M} c_n u^n + \mathcal{O}(u^{L+M+1}) \quad (\text{B.17})$$

La notation $\mathcal{O}(u^{L+M+1})$ signifie simplement que l'on néglige les termes de puissance d'ordre strictement supérieur à $L + M$. Pour déterminer les deux jeux de coefficients $\{a_n\}$ et $\{b_n\}$, il suffit d'égaliser les coefficients

$$\sum_{n=0}^M b_n u^n \sum_{n=0}^{L+M} c_n u^n = \sum_{n=0}^L a_n u^n + \mathcal{O}(u^{L+M+1}) \quad (\text{B.18})$$

La condition d'égalité des moments détermine dans un premier temps les coefficients $\{b_n\}$, puis $\{a_n\}$ par multiplication. En imposant

$$b_0 = 1$$

on obtient les équations matricielles suivante pour $\{b_n\}$:

$$\begin{aligned} \sum_{n=0}^M b_n c_{L-n+j} &= 0 & 1 \leq j \leq M \\ \sum_{n=1}^M b_n c_{L-n+j} &= -c_{L+j} & 1 \leq j \leq M \end{aligned} \quad (\text{B.19})$$

qui peuvent être mises sous la forme plus lisible suivante :

$$\begin{pmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_L \\ \vdots & \vdots & \vdots & \vdots \\ c_{L-M+k} & c_{L-M+k+1} & \cdots & c_{L+k-1} \\ \vdots & \vdots & \vdots & \vdots \\ c_L & c_{L+1} & \cdots & c_{L+M-1} \end{pmatrix} \begin{pmatrix} b_M \\ \vdots \\ b_k \\ \vdots \\ b_1 \end{pmatrix} = - \begin{pmatrix} c_{L+1} \\ \vdots \\ c_{L+k+1} \\ \vdots \\ c_{L+M} \end{pmatrix} \quad (\text{B.20})$$

Les coefficients du numérateur $\{a_n\}$ sont alors déterminés par (B.18), ce qui revient à effectuer un produit de convolution entre les coefficients $\{b_n\}$ et $\{c_n\}$:

$$a_j = c_j + \sum_{i=1}^{\min(M,j)} b_i c_{j-i} \quad 0 \leq j \leq L \quad (\text{B.21})$$

La détermination des coefficients $\{a_n\}$ et $\{b_n\}$ constitue l'approximation de Padé de la fonction génératrice de moments.

Cette approximation est réalisée au voisinage de $u = 0$, ce qui permet de juger l'approximation de la queue de distribution de la densité de probabilité ; pour des voisinages autres que 0 (cette méthode n'étant pas valable puisque dans ce cas la série est divergente et ne peut donc pas être approximée par une fraction rationnelle tendant vers 0 à l'infini), l'APPROXIMATION DE PADÉ À DEUX POINTS existe et fait intervenir les *coefficients de Markov* d_n .

B.3.2 L'approximation de Padé à 2 points

Le principe est le même sauf que l'approximation s'effectue à la fois en $u = 0$ et en $u = \infty$. Il nous faut donc connaître le développement *asymptotique* de la fonction caractéristique $\Psi(u)$ (autour de $u = \infty$, développement en série en puissance de $1/u$), et, toujours, son développement en série autour de $u = 0$ pour obtenir :

$$\Psi(u) = \sum_{n=0}^{\infty} c_n u^n \text{ pour } u \rightarrow 0 \quad (\text{B.22})$$

$$\Psi(u) = \sum_{n=0}^{\infty} d_n u^{-n} \text{ pour } u \rightarrow +\infty \quad (\text{B.23})$$

L'approximation de Padé à deux points de la fonction génératrice $\Phi(u)$, d'ordre $[L/M]$, notée $\mathbf{P}_{(J,K)}^{[L/M]}(\mathbf{u})$ est définie par la fraction rationnelle construite pour avoir l'égalité des J premiers termes de (B.22) et les K premiers termes de (B.23). Ainsi, pour $L = M - 1$, nous obtenons un jeu d'équations pour les $\{a_n\}$ et les $\{b_n\}$:

$$P_{J,K}^{[M-1/M]}(u) = \frac{\sum_{n=0}^L a_n u^n}{1 + \sum_{n=1}^M b_n u^n} = \begin{cases} \sum_{n=0}^J c_n u^n & u \rightarrow 0 \\ \sum_{n=1}^K d_n u^{-n} & u \rightarrow +\infty \end{cases} \quad (\text{B.24})$$

qui se transforme en $J + K + 1 = 2M$ équations matricielles à $2M$ inconnues données par :

$$a_j = \sum_{i=0}^j b_i c_{j-i}, \quad 0 \leq j \leq J \quad (\text{B.25})$$

et $\begin{cases} a_j = 0 & j \geq M \\ b_i = 0 & i \geq M + 1 \end{cases}$

$$a_{M-l-1} = \sum_{i=0}^l d_{i+1} b_{M-l+i}, \quad 0 \leq l \leq K - 1 \quad (\text{B.26})$$

et $\begin{cases} a_j = 0 & j \geq M \\ b_i = 0 & i \geq M + 1 \end{cases}$

B.3.3 Expressions des fonctions densité de probabilité et de répartition

Une fois l'ordre de l'approximation déterminé (cf. paragraphe B.3.5), il est possible de calculer très simplement la densité de probabilité et la fonction de répartition par la méthode des résidus, dès lors que la fonction génératrice de moments peut se mettre sous forme d'une fraction rationnelle du type (B.16) :

$$P^{[L/M]}(u) = \frac{\sum_{n=0}^L a_n u^n}{\sum_{n=0}^M b_n u^n} = \sum_{k=1}^M \frac{\lambda_k}{u - \alpha_k} \quad (\text{B.27})$$

Supposons que la fraction rationnelle $P^{[L/M]}(u)$ n'ait que M pôles simples α_k distincts et à partie réelle négative (pour assurer la convergence de la densité de probabilité); la densité de probabilité $p(x)$ et la fonction de répartition $F(x)$ s'écrivent alors (calculées par la méthode des résidus) :

$$p(x) = \sum_{k=1}^M \lambda_k e^{\alpha_k x} \quad (\text{B.28})$$

$$F(x) = 1 + \sum_{k=1}^M \frac{\lambda_k}{\alpha_k} e^{\alpha_k x} \quad (\text{B.29})$$

B.3.4 Détermination de $p(x)$ et de $F(x)$ par la méthode des résidus

Nous avons vu que la fonction génératrice de moments $\Phi(u)$ était la transformée de Laplace de la densité de probabilité $p(x)$; il est possible d'exprimer la densité $p(x)$ à partir de $\Phi(u)$, par transformée de Laplace inverse et on a :

$$p(x) = \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \Phi(u) e^{ux} du = \sum_k \text{res}[\Phi(u) e^{ux}, \alpha_k] \quad (\text{B.30})$$

où

- * le contour d'intégration est défini à droite de toutes les singularités de $\Phi(u)$, situées dans le demi-plan gauche,
- * α_k représente les pôles à parties réelles négatives de $\Phi(u)$ (dans notre cas les pôles sont simples),
- * $\text{res}[f(z), z_0]$ représente le résidu de la fonction f au pôle z_0 . Si z_0 est un pôle d'ordre m , on a

$$\text{res}[f(z), z_0] = \frac{1}{\Gamma(m)} \lim_{z \rightarrow z_0} \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)] \quad (\text{B.31})$$

Les M pôles de $\Phi(u)$ sont distincts (conjugués deux à deux) et simples; on va donc avoir, pour chaque k :

$$\begin{aligned} \text{res}[\Phi(u) e^{ux}, \alpha_k] &= \lim_{u \rightarrow \alpha_k} (u - \alpha_k) e^{ux} \sum_{i=1}^M \frac{\lambda_i}{u - \alpha_i} = \lim_{u \rightarrow \alpha_k} e^{ux} \left[\lambda_k + (u - \alpha_k) \sum_{i \neq k} \frac{\lambda_i}{u - \alpha_i} \right] \\ &= \lambda_k e^{\alpha_k x} \end{aligned}$$

Ainsi, on obtient bien (B.28).

Pour la fonction de répartition, calculons, $\forall x \geq 0$

$$F(x) = \int_{-\infty}^x p(u) du = 1 - \int_x^{+\infty} p(u) du$$

à l'aide de l'expression de $p(x)$ de (B.28). On obtient :

$$\begin{aligned} F(x) &= 1 - \int_x^{+\infty} \sum_{k=1}^M \lambda_k e^{\alpha_k u} du = 1 - \sum_{k=1}^M \lambda_k \int_x^{+\infty} e^{\alpha_k u} du \\ &= 1 - \sum_{k=1}^M \lambda_k \left[\frac{e^{\alpha_k u}}{\alpha_k} \right]_{u=x}^{u=+\infty} = 1 + \sum_{k=1}^M \lambda_k \frac{e^{\alpha_k x}}{\alpha_k} \end{aligned}$$

D'où l'expression (B.29).

B.3.5 Détermination de l'ordre de l'approximation

Il n'existe pas de méthode exacte pour déterminer l'ordre de l'approximation de Padé lorsque ses coefficients sont utilisés pour estimer une densité de probabilité. La condition à respecter dans ce cas-là est la négativité des parties réelles des pôles α_k et si, à une valeur de M fixée, un ou plusieurs pôles sont à parties réelles négatives, alors la valeur de M ne peut être acceptée. Il est nécessaire de changer l'ordre et de calculer de nouveaux coefficients. Pour éviter ce problème, H.Amindavar propose, dans [Amin91], une procédure de stabilisation des pôles. Cette méthode consiste à trouver la plus petite perturbation possible à apporter aux coefficients afin de rendre stable l'approximation, c'est-à-dire que toutes les singularités de la fonction génératrice de moments se trouvent dans le demi-plan gauche. Cependant, perturber les coefficients perturbe les moments (les moments peuvent s'exprimer en fonction des coefficients) et la perturbation doit donc être choisie avec soins.

La solution proposée ici (et présentée dans [Jay00b]) pour trouver une approximation stable consiste, tout d'abord, à poser $L = M - 1$. Cette égalité assure la convergence vers 0 de la fonction génératrice de moments lorsque $u \rightarrow +\infty$ et constitue une approximation dite *sub-diagonale* [Amin91]. Ensuite, une valeur supérieure K de M (et donc également de L) est choisie à partir de laquelle les coefficients sont calculés. Si tous les pôles sont à partie réelle négative, alors $M = K$, sinon, on pose $M = K - 1$, de nouveaux coefficients sont calculés, et ainsi de suite.

La valeur supérieure K est déterminée d'après le rang de la matrice de Hankel donnée en (B.20). Cette matrice n'est pas toujours de rang plein, sauf quand il est égal à M . Ainsi, plusieurs valeurs de M sont testées et on recherche la plus grande valeur K qui coïncide avec M parmi les valeurs des rangs de la matrice de Hankel coïncidant avec M . Cela signifie que lorsque $M \geq K$, le rang de la matrice n'est pas plein.

Cette méthode permet de trouver plusieurs ordres pour lesquelles l'approximation est stable.

B.3.6 Qualité de l'estimation sur quelques exemples

Basée sur le développement en série autour de 0 de la fonction génératrice de moments de la densité à estimer, l'approximation de Padé est particulièrement bien adaptée pour estimer les queues de densités. Pour illustrer cette particularité, quelques études sur l'approximation de la fonction de répartition sont présentées aux paragraphes suivants, :

- * Qualité de l'approximation de Padé d'ordre $[M - 1/M]$ calculée d'après les moments théoriques, pour la K-distribution (avec $\nu = 0.1$ et $\nu = 1.67$), pour la loi Gamma (avec $a = 0.15$ et $a = 2.24$) et pour une loi de Rayleigh de paramètre $b = 0.76$,
- * Comparaisons des approximations de Padé en fonction de tous les ordres M déterminés pour la K-distribution de paramètre $\nu = 0.1$,
- * Qualité de l'approximation de Padé en fonction du nombre N d'échantillons disponibles pour le calcul des moments empiriques, pour la K-distribution (avec $\nu = 0.1$). La qualité de l'approximation est exprimée par l'erreur quadratique moyenne intégrée (MISE en anglais, cf. (B.12) pour son expression) pour $N = 10$ à $N = 5000$.

B.3.6.1 A partir des moments théoriques, pour un ordre M

Le cas idéal pour effectuer une approximation de Padé à une densité de probabilité ou à une fonction de répartition se présente lorsque les moments théoriques de la loi sont disponibles. Les lois de référence sont ici la K-distribution, la loi Gamma et la loi de Rayleigh. Pour chacune de ces lois, les 30 premiers moments théoriques sont calculés, à partir desquels les coefficients de Padé sont déterminés selon la procédure explicitée au paragraphe B.3.1. L'ordre de l'approximation choisi est le plus grand des ordres possibles déterminés d'après les conditions données en B.3.5.

Les figures (FIG. B.5) et (FIG. B.6) représentent les approximations de Padé de la fonction de répartition de la loi K de paramètres respectifs $\nu = 0.1$ et $\nu = 1.67$. Les erreurs relatives correspondantes sont indiquées par les courbes inférieures à celles sus-citées.

Les figures (FIG. B.7) et (FIG. B.8) se rapportent à deux lois Gamma de paramètres respectifs $a = 0.15$ et $a = 2.24$, et la figure (FIG. B.9) se rapporte à une loi de Rayleigh de paramètre $b = 0.76$.

Ces trois illustrations permettent de juger de la qualité de l'approximation de Padé sur les queues de distribution, lorsque les moments théoriques sont disponibles. Ainsi, on constate que la valeur du seuil de détection peut être estimée parfaitement jusqu'à une valeur de P_{fa} très faible (allant jusqu'à 10^{-10} voire encore plus faible), dans le cadre d'une procédure de détection d'enveloppe dans les trois environnements cités.

B.3.7 Comparaison des approximations de Padé en fonction des ordres déterminés

Pour effectuer une comparaison des approximations de Padé obtenues pour chacun des ordres possibles, les cas de la K-distribution de paramètre $\nu = 0.1$ et de la loi Gamma de paramètre $a = 2.68$ ont été étudiés. Les figures (FIG. B.10) et (FIG. B.11) représentent les différentes fonctions de répartition obtenues pour chaque ordre et les erreurs commises relativement à l'expression théorique (courbes du bas).

B.3.8 Qualité de l'approximation de Padé en fonction du nombre d'échantillons disponibles pour l'estimation des moments

Les erreurs commises sur l'estimation empirique des moments d'ordre supérieur est toujours un sujet délicat à traiter. Nous nous proposons donc ici de quantifier l'erreur quadratique moyenne intégrée (MISE en anglais) en fonction du nombre d'échantillons K-distribués ($\nu = 0.1$) disponibles pour l'estimation des moments. Les ordres des approximations varient pour chaque N , et l'ordre retenu est toujours le plus grand possible. Plus N est grand et plus la MISE est faible, et on constate qu'à partir de $N = 500$, l'erreur commise sur l'ensemble de l'approximation de la fonction de répartition est de l'ordre de 10^{-4} . Ces résultats sont représentés sur la figure (FIG. B.12).

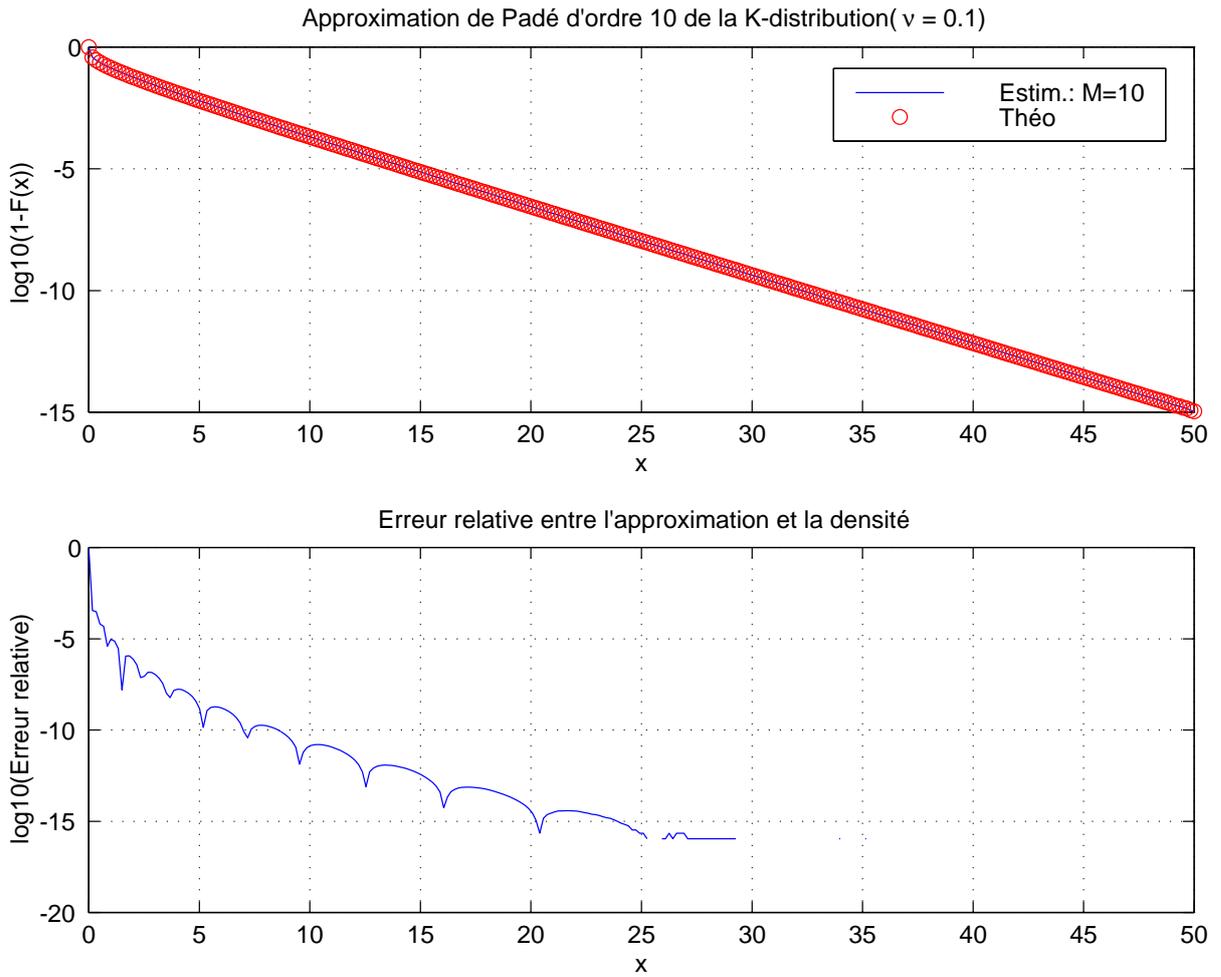


FIG. B.5 – Approximation de Padé d’ordre $M = 10$ pour la K-distribution de paramètre $\nu = 0.1$. Erreur relative correspondante sur la courbe inférieure.

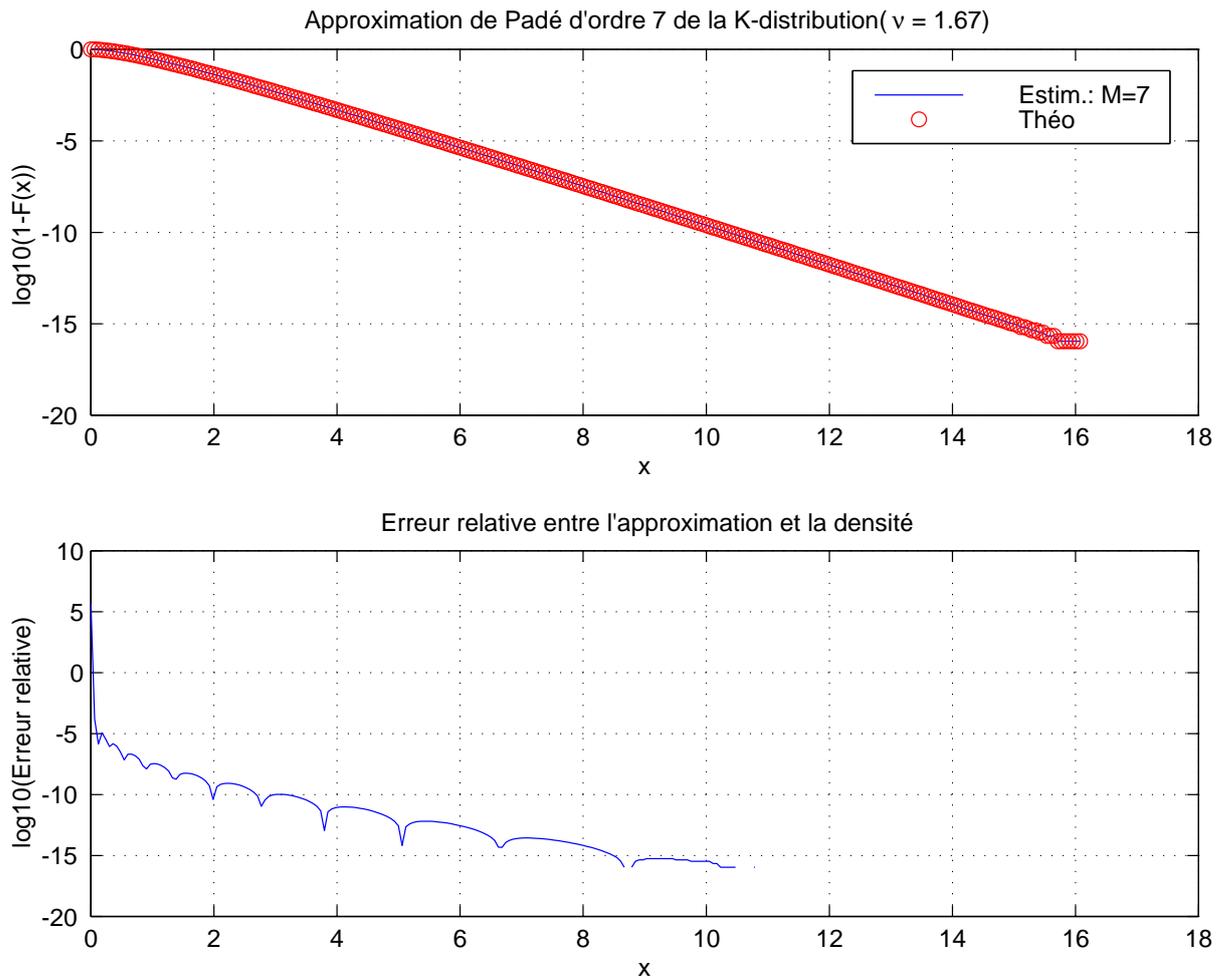


FIG. B.6 – Approximation de Padé d'ordre $M = 7$ pour la K-distribution de paramètre $\nu = 1.67$. Erreur relative correspondante sur la courbe inférieure.

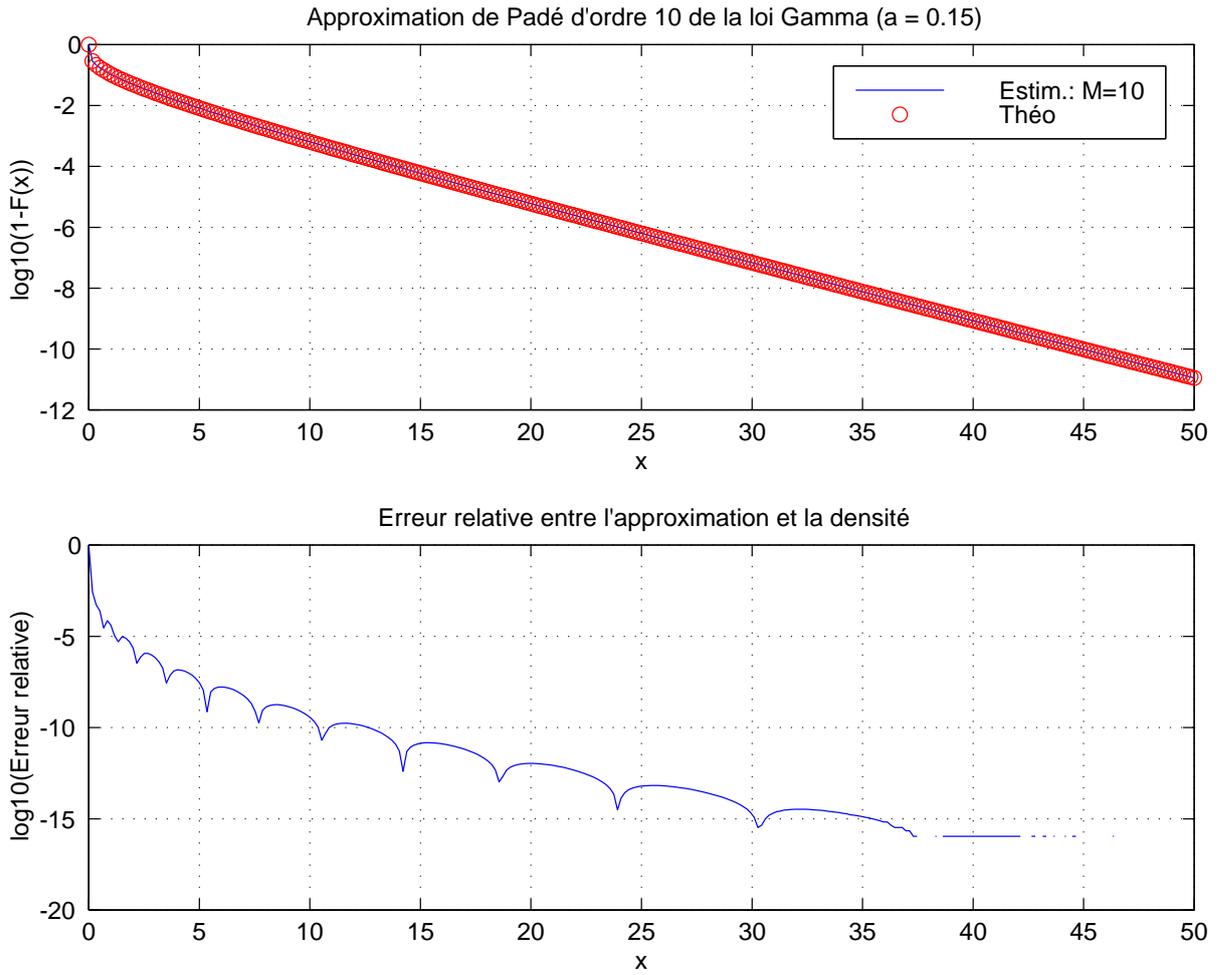


FIG. B.7 – Approximation de Padé d’ordre $M = 10$ pour la loi Gamma de paramètre $a = 0.15$. Erreur relative correspondante sur la courbe inférieure.

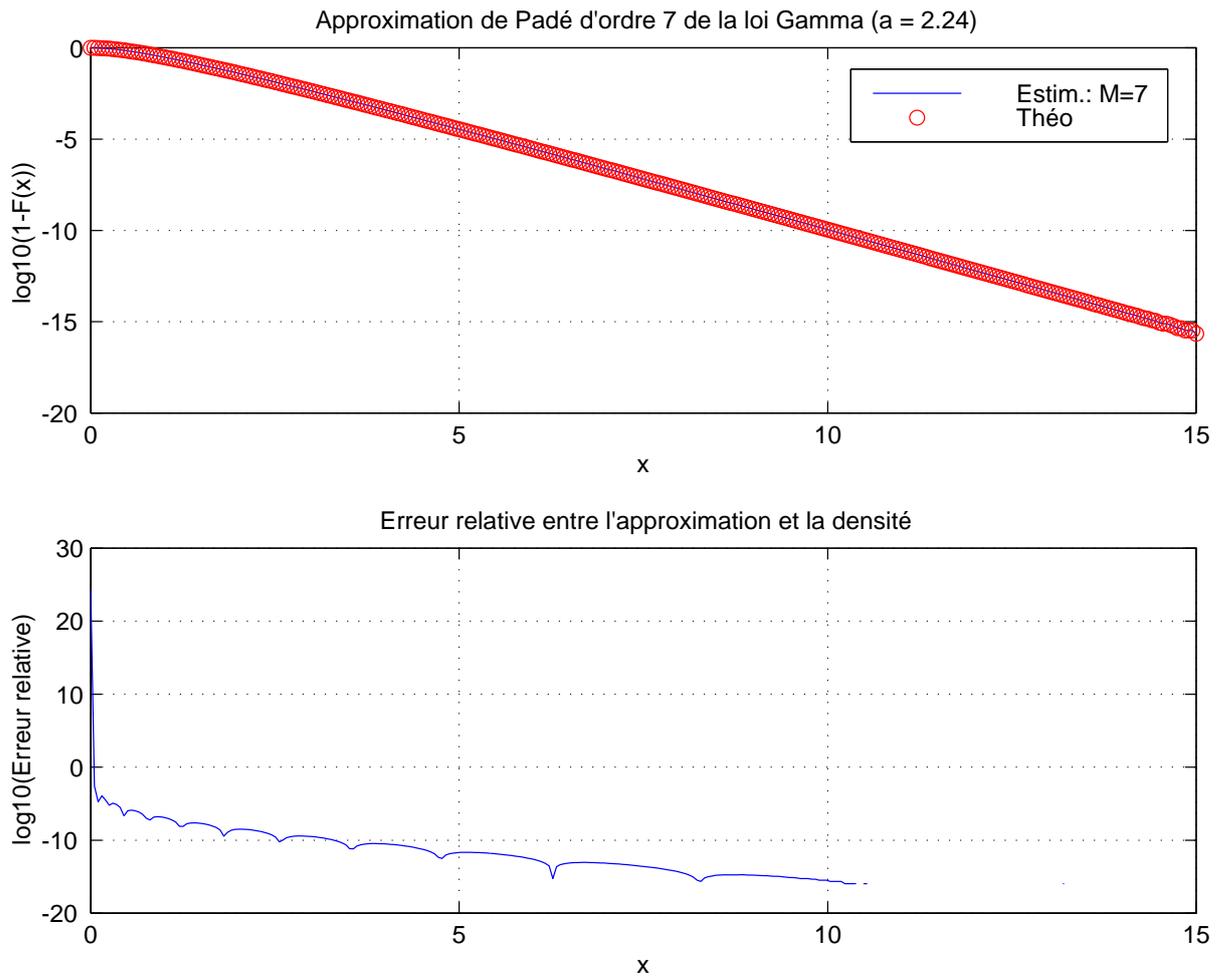


FIG. B.8 – Approximation de Padé d'ordre $M = 7$ pour la loi Gamma de paramètre $a = 2.24$. Erreur relative correspondante sur la courbe inférieure.

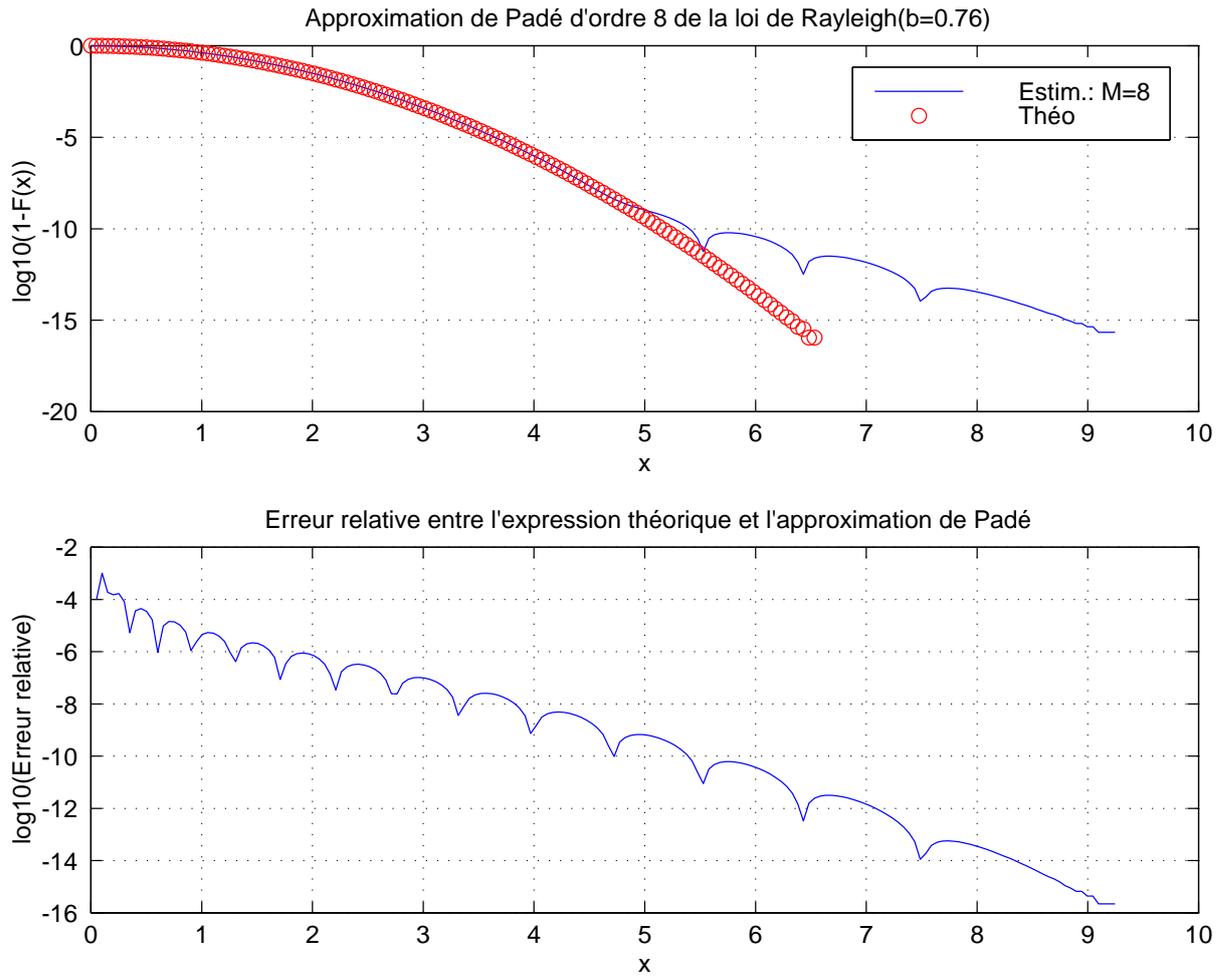


FIG. B.9 – Approximation de Padé d’ordre $M = 8$ pour la loi de Rayleigh de paramètre $b = 0.76$ et erreur relative correspondante sur la courbe inférieure.

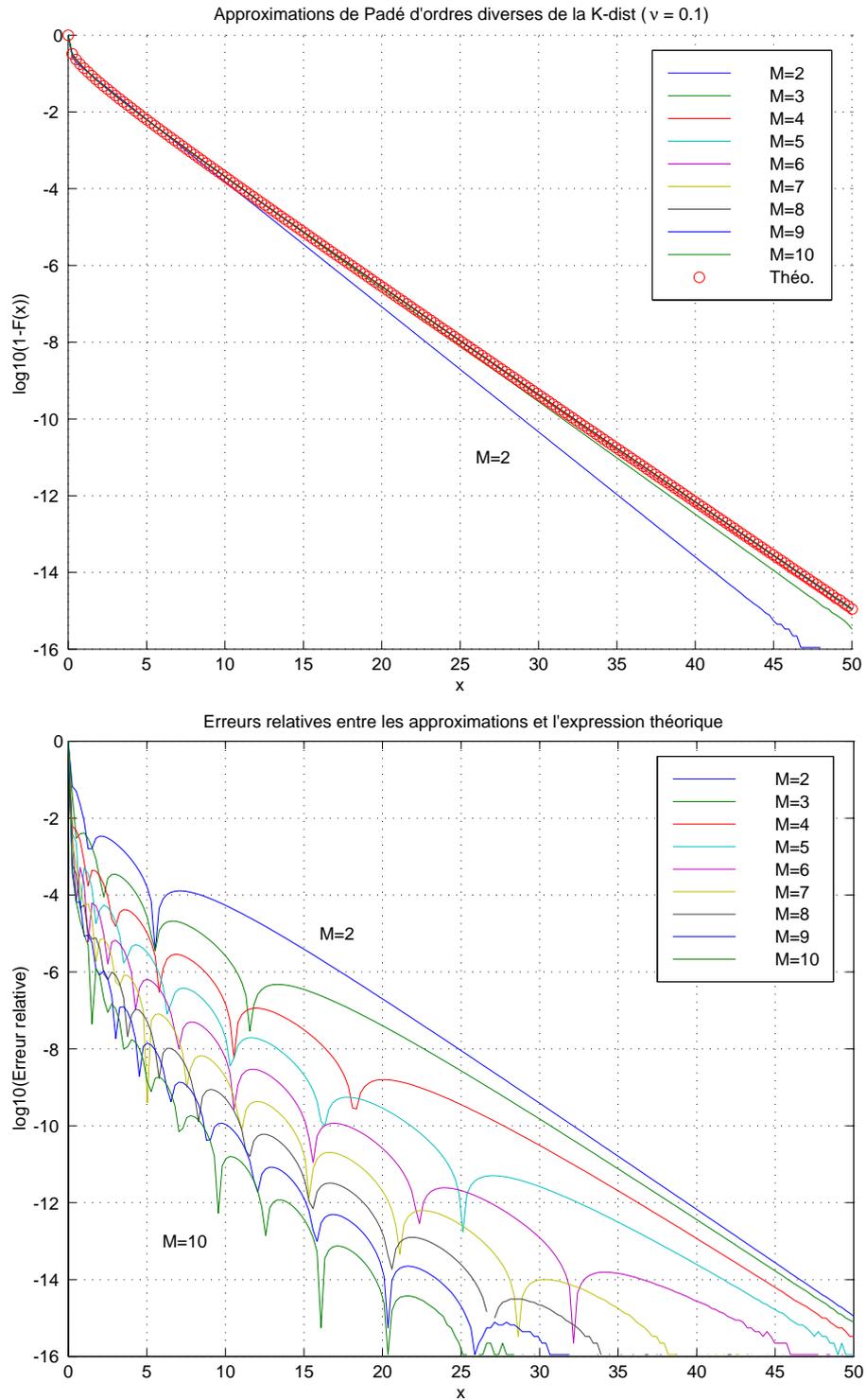


FIG. B.10 – Approximations de Padé d'ordres $M = 2$ à $M = 10$ pour la K-distribution de paramètre $\nu = 0.1$ (en haut) et erreurs relatives associées (en bas).

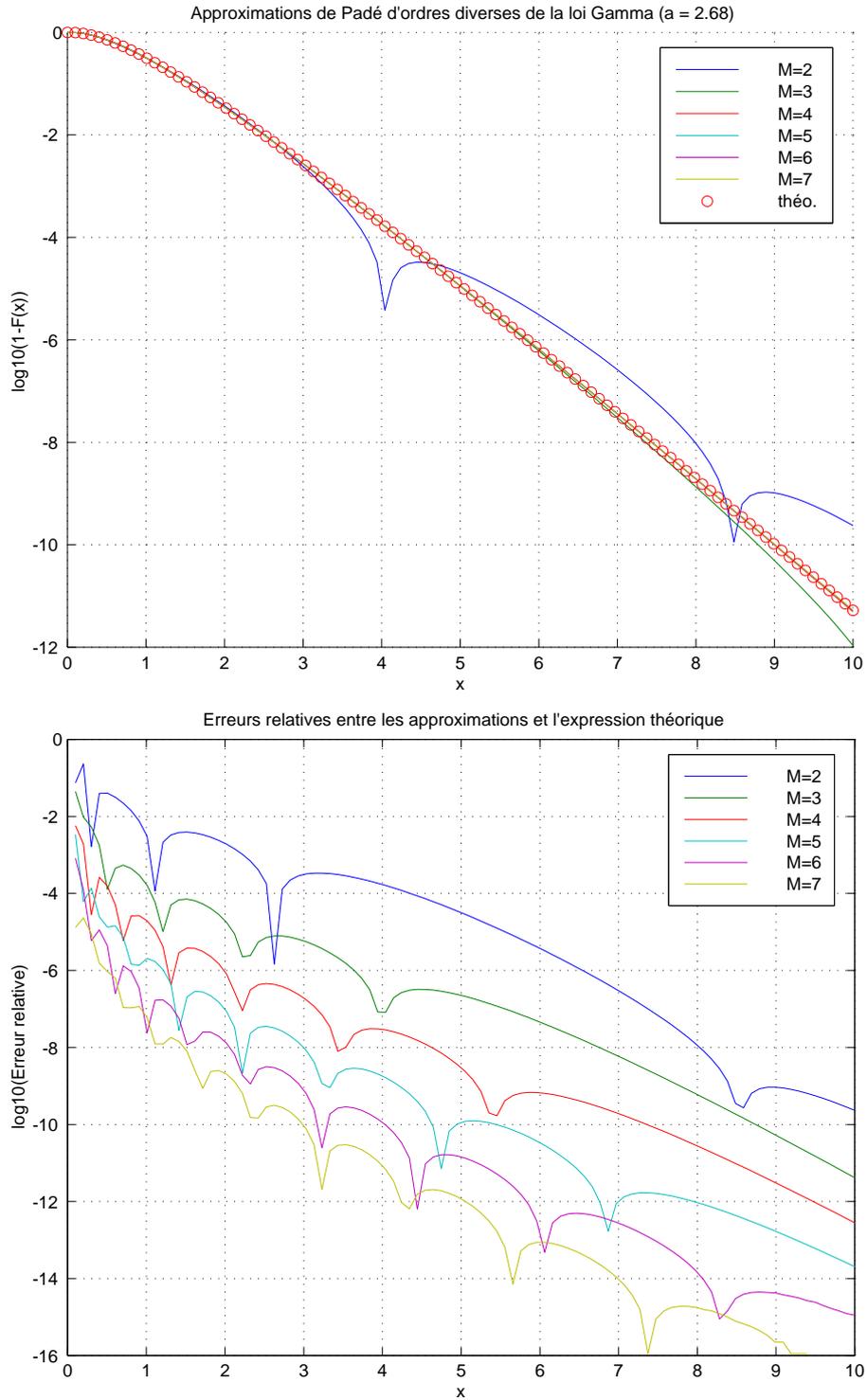


FIG. B.11 – Approximations de Padé d'ordres $M = 2$ à $M = 7$ pour la loi Gamma de paramètre $a = 2.68$ (en haut) et erreurs relatives associées (en bas).

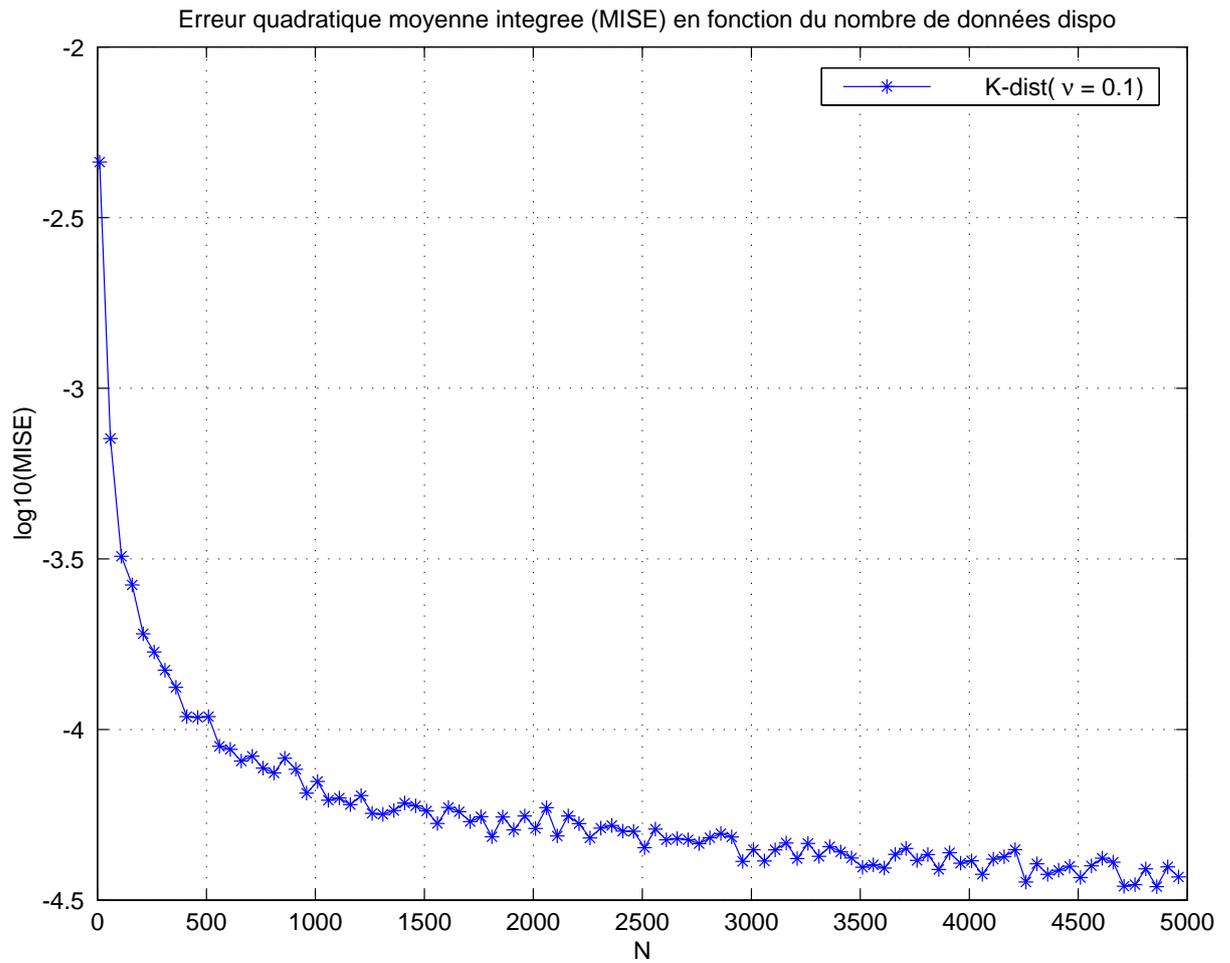


FIG. B.12 – Erreur quadratique moyenne intégrée calculée entre la fonction de répartition théorique et son approximation de Padé en fonction du nombre $N = 10$ à 5000 d'échantillons K-distribués de paramètre $\nu = 0.1$.

Annexe C

Les vecteurs Gaussiens

En probabilité et statistique la loi Gaussienne joue un rôle très important puisque nous la retrouvons très souvent. Elle est très utilisée, par exemple, pour la modélisation de bruits, et beaucoup de résultats tournent autour de la Gaussienne ou y aboutissent. Pour généraliser les résultats donnés à la section 2.1. pour un loi unidimensionnelle, nous allons traiter, dans cette section le cas des vecteurs Gaussiens en rappelant quelques résultats généraux de probabilité ainsi que des théorèmes ou propriétés sur les vecteurs Gaussiens.

C.1 Rappels de probabilité

C.1.1 Convergence presque sûre (p.s.)

On dit que $X_n \xrightarrow{p.s.} X$ s'il existe Ω , avec $\mathbb{P}(\Omega) = 1$, tel que

$$\forall w \in \Omega, \lim_{n \rightarrow +\infty} X_n(w) = X(w)$$

.

C.1.2 Convergence en probabilité (IP)

On dit que $X_n \xrightarrow{IP} X$ si $\forall \epsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$.

Proposition 1

$$\left. \begin{array}{l} \text{Si } X_n \longrightarrow X \text{ en IP ou p.s.} \\ \text{et si } X_n \longrightarrow Y \text{ en IP ou p.s.} \end{array} \right\} \text{ Alors } X = Y \text{ presque partout (p.p.)} \quad (\text{C.1})$$

Proposition 2

$$\left. \begin{array}{l} \text{Si } X_n \xrightarrow{p.s.} X \\ \text{et si } Y_n \xrightarrow{p.s.} Y \end{array} \right\} \text{ Alors } \left\{ \begin{array}{l} X_n + Y_n \xrightarrow{p.s.} X + Y \\ X_n Y_n \xrightarrow{p.s.} XY \end{array} \right. \quad (\text{C.2})$$

Proposition 3

$$\left. \begin{array}{l} \text{Si } X_n \longrightarrow X \text{ en IP et p.s.} \\ \text{et si } Y_n \longrightarrow Y \text{ en IP et p.s.} \end{array} \right\} \text{ Alors } (X_n, Y_n) \longrightarrow (X, Y) \text{ en IP et p.s.} \quad (\text{C.3})$$

Proposition 4 Si $X_n \xrightarrow{p.s.} X$, alors $X_n \xrightarrow{IP} X$.

Note : Nous ne pouvons en revanche rien dire quant à la convergence en loi.

C.1.3 Convergence en loi (\mathcal{L})

$$\forall f \in \mathcal{C}_K : (X_n \xrightarrow{\mathcal{L}} X) \iff \mathbb{E}(f(X_n)) \xrightarrow{n \rightarrow +\infty} \mathbb{E}(f(X)).$$

Proposition 5

$$X_n \xrightarrow{\mathbb{P}} X \implies X_n \xrightarrow{\mathcal{L}} X$$

Proposition 6

$$X_n \xrightarrow{\mathcal{L}} X = \mathbf{Cte}, p.s. \implies X_n \xrightarrow{\mathbb{P}} X$$

C.2 Définitions, propriétés

Les notions et notations de convergence étant introduites, on peut énoncer les différentes définitions, propriétés et théorèmes concernant les vecteurs Gaussiens.

Rappel : La masse de Dirac en m , δ_m est une loi de Gauss, de moyenne m et de variance nulle.

Définition 5 *Définition d'un vecteur Gaussien* Soit $\mathbf{X} = (X_1, \dots, X_d)^T$ un vecteur aléatoire à valeurs dans \mathbb{R}^d . Le vecteur \mathbf{X} est Gaussien si toute combinaison linéaire de ses coordonnées est une v.a. Gaussienne, soit encore :

$$(\mathbf{X} \text{ Gaussien}) \iff (\forall \lambda \in \mathbb{R}^d, \lambda_1 X_1 + \dots, \lambda_d X_d = \langle \lambda, \mathbf{X} \rangle \text{ Gaussien}).$$

C.2.1 Caractéristiques d'un vecteur Gaussien

- Vecteur moyenne : $\mathbf{m} = (m_1, \dots, m_d)^T$, avec $m_i = \mathbb{E}(X_i)$
- Matrice de covariance ou de dispersion $\mathbf{D} = (D_{ij}, i, j = 1, \dots, d)$, symétrique et positive avec $D_{ij} = \mathbb{E}(X_i X_j) - m_i m_j = \text{Cov}(X_i, X_j)$.
- Notation : $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{D})$

Proposition 7 Soit $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{D})$ à valeurs dans \mathbb{R}^d .

$\forall \lambda \in \mathbb{R}^d$, $\langle \lambda, \mathbf{X} \rangle$ suit une loi de Gauss de moyenne $\langle \lambda, \mathbf{m} \rangle$ et de variance $\sigma^2 = \langle \lambda, \mathbf{D} \lambda \rangle$.

C.2.2 Fonction caractéristique

La fonction caractéristique d'un vecteur Gaussien de moyenne \mathbf{m} et de matrice de dispersion \mathbf{D} est donnée par :

$$\Phi_X(u) = \exp \left(i \langle \lambda, \mathbf{m} \rangle - \frac{1}{2} \langle \lambda, \mathbf{D} \lambda \rangle \right) \quad (\text{C.4})$$

Proposition 8 Si $\mathbf{m} \in \mathbb{R}^d$ est un vecteur et $\mathbf{D} = (D_{ij})$ une matrice symétrique positive alors il existe un vecteur Gaussien de moyenne \mathbf{m} et de matrice de dispersion \mathbf{D} .

C.2.3 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 ($\mathbf{x} \in \mathbb{R}^d$) :

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} \sqrt{\det(\mathbf{D})}} \exp\left(-\frac{\langle \mathbf{x} - \mathbf{m}, \mathbf{D}^{-1}(\mathbf{x} - \mathbf{m}) \rangle}{2}\right) \quad (\text{C.5})$$

Proposition 9 \mathbf{X} est un vecteur Gaussien et X_1, \dots, X_d indépendantes

\iff Matrice de dispersion diagonale,

$\iff \mathbb{E}(X_i X_j) = \mathbb{E}(X_i) \mathbb{E}(X_j)$.

$\iff \text{Cov}(X_i, X_j) = 0$.

C.3 Les vecteurs Gaussiens complexes circulaires

Soit un vecteur complexe $\mathbf{Z} = \mathbf{X} + i\mathbf{Y}$ où les vecteurs \mathbf{X} et \mathbf{Y} sont des vecteurs Gaussiens de taille N à valeurs réelles de matrices de dispersion respectivement égales à $\mathbf{M}_x = \mathbb{E}(\mathbf{X}\mathbf{X}^T)$ et $\mathbf{M}_y = \mathbb{E}(\mathbf{Y}\mathbf{Y}^T)$. On note par $\mathbf{M}_{xy} = \mathbb{E}(\mathbf{X}\mathbf{Y}^T)$ et $\mathbf{M}_{yx} = \mathbb{E}(\mathbf{Y}\mathbf{X}^T)$ les matrices d'intercorrélation des vecteurs \mathbf{X} et \mathbf{Y} .

C.3.1 Définition

Définition 6 Le vecteur \mathbf{Z} , de taille $2N$ est Gaussien complexe circulaire si et seulement si les vecteurs \mathbf{X} et \mathbf{Y} sont conjointement Gaussiens et si, de plus :

$$\begin{cases} \mathbf{M}_x &= \mathbf{M}_y \\ \mathbf{M}_{xy} &= -\mathbf{M}_{yx} \end{cases} \quad (\text{C.6})$$

Il vient de ces relations que la matrice de dispersion de \mathbf{Z} s'expriment en fonction de $\mathbf{M}_x (= \mathbf{M}_y)$ et de $\mathbf{M}_{xy} (= -\mathbf{M}_{yx})$ comme suit :

$$\mathbf{M}_z = 2(\mathbf{M}_x + i\mathbf{M}_{xy}) \quad (\text{C.7})$$

Les deux relations de (C.6) expriment la circularité d'ordre 2 et permettent la détermination de \mathbf{M}_x et de \mathbf{M}_{xy} à partir de \mathbf{M}_z . Sans cette condition supplémentaire, il serait impossible de les retrouver. Ces relations expriment également la décorrélation entre \mathbf{Z} et \mathbf{Z}^* car $\mathbb{E}(\mathbf{Z}\mathbf{Z}^{*\dagger}) = \mathbf{0}$.

C.3.2 Loi d'un vecteur Gaussien complexe circulaire

La loi de \mathbf{Z} est la forme privilégiée de la loi conjointe de \mathbf{X} et \mathbf{Y} , lorsqu'on les exprime en fonction de \mathbf{Z} et de son conjugué \mathbf{Z}^* . Si on note par $\mathbf{T} = [\mathbf{X}\mathbf{Y}]^T$, on a par définition :

$$\mathbf{M}_t = \begin{pmatrix} \mathbf{M}_x & \mathbf{M}_{xy} \\ \mathbf{M}_{yx} & \mathbf{M}_y \end{pmatrix}$$

La loi de \mathbf{T} est parfaitement déterminée par les quatre matrices ci-dessus et s'écrit de manière similaire à (C.5) pour \mathbf{T} de taille $2N$:

$$f(\mathbf{T}) = \frac{1}{(2\pi)^N \sqrt{\det(\mathbf{M}_t)}} \exp\left(-\frac{\langle \mathbf{T}, \mathbf{M}_t^{-1}\mathbf{T} \rangle}{2}\right) \quad (\text{C.8})$$

On va chercher à déterminer les expressions de $\mathbf{T}^\dagger \mathbf{M}_t^{-1} \mathbf{T}$ et de $\det(\mathbf{M}_t)$ en fonction de $\mathbf{Z}, \mathbf{Z}^*, \mathbf{M}_z$ et on a :

$$\mathbf{T}^\dagger \mathbf{M}_t^{-1} \mathbf{T} = 2 \mathbf{Z}^\dagger \mathbf{M}_z^{-1} \mathbf{Z} \quad (\text{C.9})$$

$$\det(\mathbf{M}_t) = \left(\frac{\det(\mathbf{M}_z)}{2^N} \right)^2, \quad (\text{C.10})$$

où \dagger désigne le transposé conjugué.

La loi du vecteur Gaussien complexe circulaire \mathbf{Z} est donc :

$$f(\mathbf{Z}) = \frac{1}{\pi^N \det(\mathbf{M}_z)} \exp(-\mathbf{Z}^\dagger \mathbf{M}_z^{-1} \mathbf{Z}). \quad (\text{C.11})$$

C.4 Formes quadratiques de vecteurs Gaussiens

Quelques théorèmes et résultats utiles sur l'indépendance et les lois de formes quadratiques de vecteurs Gaussiens réels sont reportés dans cette section. Dans un premier temps les théorèmes principaux sont énoncés, puis appliqués aux formes quadratiques particulières présentes dans les lois jointes des SIRV.

C.4.1 Indépendance de formes quadratiques

Théorème 5 (Indépendance de formes quadratiques) Soient x_1, \dots, x_p, p variables aléatoires $\mathcal{N}(0, 1)$, i.i.d. et soit le vecteur $\mathbf{X} = (x_1, \dots, x_p)$. Soient les formes quadratiques Q_1 et Q_2 définies par :

$$Q_1 = \mathbf{X}^T \mathbf{B} \mathbf{X}, \quad Q_2 = \mathbf{X}^T \mathbf{C} \mathbf{X},$$

où \mathbf{B} et \mathbf{C} sont des matrices ($p \times p$) de rang inférieur ou égal à p .

Alors Q_1 et Q_2 sont indépendantes si et seulement si

$$\mathbf{B} \mathbf{C} = \mathbf{0}.$$

C.4.2 Lois de formes quadratiques

Théorème 6 Soit \mathbf{X} un vecteur de p composantes x_1, \dots, x_p , i.i.d. $\mathcal{N}(0, 1)$ et $Q = \mathbf{X}^T \mathbf{B} \mathbf{X}$, où \mathbf{B} est une matrice ($p \times p$) de rang $r \leq p$.

La loi de Q :

(a) admet pour n^{eme} cumulant $\kappa_n = 2^{n-1} (n-1)! \text{tr}(\mathbf{B}^n)$,

(b) est une loi du χ_r^2 si et seulement si \mathbf{B} est idempotente ($\mathbf{B}^2 = \mathbf{B}$).

Les deux théorèmes qui suivent couvrent des cas plus généraux.

Théorème 7 Soit $\mathbf{X} \sim \mathcal{N}_p(\mathbf{0}, \sigma^2 \mathbf{I}_p)$ et $Q = \mathbf{X}^T \mathbf{B} \mathbf{X}$, où \mathbf{B} est symétrique de rang r .

Alors $\frac{Q}{\sigma^2} \sim \chi_r^2$ si et seulement si \mathbf{B} est idempotente.

La forme de \mathbf{B} n'est pas forcément évidente à trouver. On peut toujours vérifier si les matrices de projection $\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ et $\mathbf{I} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ sont idempotentes.

Théorème 8 Soit $\mathbf{X} \sim \mathcal{N}_p(\mathbf{0}, \Sigma)$ où Σ est définie positive. Soit $Q = \mathbf{X}^T \mathbf{B} \mathbf{X}$, où \mathbf{B} est symétrique de rang r .

Alors $Q \sim \chi_r^2$ si et seulement si $\mathbf{B} \Sigma \mathbf{B} = \mathbf{B}$.

C.4.3 Le théorème de Cochran

Ce théorème permet de décomposer des sommes de carrés de variables aléatoires en plusieurs formes quadratiques, d'identifier leurs lois et d'établir leur indépendance.

Théorème 9 (Théorème de Cochran) Soit \mathbf{x} un vecteur réel Gaussien centré, de taille m , tel que $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_m)$, où \mathbf{I}_m désigne la matrice identité ($m \times m$). Supposons que la forme quadratique $\mathbf{x}^T \mathbf{x}$ se décompose en k formes quadratiques $\mathbf{Q}_i = \mathbf{x}^T \mathbf{B}_i \mathbf{x}$, $1 \leq i \leq k$, avec r_i le rang de \mathbf{B}_i , matrice semi-définie positive.

Alors, chacune des propositions suivantes entraîne les deux autres :

- (a) La somme des rangs r_i des matrices \mathbf{B}_i est égal à m ,
- (b) Chaque forme quadratique $\mathbf{Q}_i = \mathbf{x}^T \mathbf{B}_i \mathbf{x}$ suit une loi du $\chi_{r_i}^2$, loi du χ^2 à r_i degrés de liberté,
- (c) Les formes quadratiques \mathbf{Q}_i sont mutuellement indépendantes.

Preuve :

On peut écrire :

$$\mathbf{x}^T \mathbf{x} = \mathbf{x}^T \mathbf{I}_m \mathbf{x} = \sum_{i=1}^k \mathbf{x}^T \mathbf{B}_i \mathbf{x},$$

soit $\mathbf{I}_m = \sum_{i=1}^k \mathbf{B}_i$.

(a) \implies (b)

Choisissons une forme quadratique Q_i arbitrairement, soit $Q_1 = \mathbf{x}^T \mathbf{B}_1 \mathbf{x}$. En effectuant une transformation orthogonale $\mathbf{x} = \mathbf{P} \mathbf{y}$, qui diagonalise \mathbf{B}_1 , on obtient alors, en notant $\mathbf{B}_1 = \mathbf{B}_1 + (\mathbf{I}_m - \mathbf{B}_1)$:

$$\begin{aligned} \mathbf{x}^T \mathbf{B}_1 \mathbf{x} + \mathbf{x}^T (\mathbf{I}_m - \mathbf{B}_1) \mathbf{x} &= \mathbf{x}^T \mathbf{I}_m \mathbf{x} \\ \mathbf{y}^T \mathbf{P}^T \mathbf{B}_1 \mathbf{P} \mathbf{y} + \mathbf{y}^T \mathbf{P}^T (\mathbf{I}_m - \mathbf{B}_1) \mathbf{P} \mathbf{y} &= \mathbf{y}^T \mathbf{P}^T \mathbf{I}_m \mathbf{P} \mathbf{y} \\ &= \mathbf{y}^T \mathbf{I}_m \mathbf{y} \end{aligned}$$

Comme $\mathbf{P}^T \mathbf{B}_1 \mathbf{P}$ et $\mathbf{P}^T \mathbf{I}_m \mathbf{P}$ sont diagonales, alors $\mathbf{P}^T (\mathbf{I}_m - \mathbf{B}_1) \mathbf{P}$ l'est également. De plus, $r(\mathbf{B}_1) = r_1$, ce qui entraîne $r(\mathbf{P}^T \mathbf{B}_1 \mathbf{P}) = r_1$ et, d'après (a), $r(\mathbf{P}^T (\mathbf{I}_m - \mathbf{B}_1) \mathbf{P}) = m - r_1$. En appliquant le théorème (. TH. C.4.2), $Q_1 \sim \chi_{r_1}^2$ et \mathbf{B}_1 est idempotente. Les résultats sont les mêmes pour les autres \mathbf{B}_i , et ainsi (a) implique (b).

(b) \implies (c)

Rappelons :

$$\mathbf{I} = \mathbf{B}_1 + \mathbf{B}_2 + \cdots + \mathbf{B}_k \quad (\text{C.12})$$

La proposition (b) implique que chaque \mathbf{B}_i est idempotente (de rang r_i). En choisissant arbitrairement une de ces matrices, soit \mathbf{B}_j , alors il existe une matrice orthogonale \mathbf{C} telle que

$$\mathbf{C}^T \mathbf{B}_j \mathbf{C} = \begin{bmatrix} \mathbf{I}_{r_j} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

En multipliant (C.12) à gauche par \mathbf{C}^T et à droite par \mathbf{C} , on obtient alors :

$$\mathbf{C}^T \mathbf{I} \mathbf{C} = \mathbf{I} = \sum_{i=1, i \neq j}^k \mathbf{C}^T \mathbf{B}_i \mathbf{C} + \begin{bmatrix} \mathbf{I}_{r_j} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Désormais, chaque $\mathbf{C}^T \mathbf{B}_i \mathbf{C}$ est idempotente et ne peut pas avoir d'éléments diagonaux négatifs. Ainsi, les r_j premiers éléments diagonaux de $\mathbf{C}^T \mathbf{B}_i \mathbf{C}$ doivent être nuls et les sous-matrices composées des éléments $(r_j + 1, \dots, p \times 1, \dots, r_j)$ et des éléments $(1, \dots, r_j \times r_j + 1, \dots, p)$ doivent être les matrices $\mathbf{0}$. De la sorte :

$$\mathbf{C}^T \mathbf{B}_i \mathbf{C} \mathbf{C}^T \mathbf{B}_j \mathbf{C} = \mathbf{0}, \quad i = 1, \dots, k, \quad i \neq j,$$

soit encore $\mathbf{C}^T \mathbf{B}_i \mathbf{B}_j \mathbf{C} = \mathbf{0}$ qui n'est vrai que si $\mathbf{B}_i \mathbf{B}_j = \mathbf{0}$. Ceci reste vrai pour toutes les matrices \mathbf{B}_i . Ainsi (b) implique (c).

(b) \implies (a)

Si (b) est vérifiée alors \mathbf{B}_i admet r_i valeurs propres égales à 1 et $p - r_i$ valeurs propres nulles. Comme $\mathbf{I} = \sum_{i=1}^k \mathbf{B}_i$, prenant la trace de cette expression, on obtient $p = \sum_{i=1}^k r_i$. Ainsi (b) implique (c).

(c) \implies (b)

Si (c) est vérifiée, en élevant $\mathbf{I} = \sum_{i=1}^k \mathbf{B}_i$ à une puissance $s \in \mathbb{N}^*$ quelconque, on a $\sum_{i=1}^k \mathbf{B}_i^s = \mathbf{I}$. Calculant la trace de cette expression, on obtient

$$\text{tr} \left(\sum_{i=1}^k \mathbf{B}_i^s \right) = p, \quad \forall s \in \mathbb{N}^*,$$

égalité vraie si et seulement si toutes les valeurs propres de \mathbf{B}_i sont égales à 1, c'est-à-dire, si chaque $\mathbf{B}_i \sim \chi^2$. Ainsi (c) implique (b).

Une version plus générale du théorème de Cochran existe. Nous l'énonçons sans le démontrer :

Théorème 10 (Version générale du théorème de Cochran) *Étant donné $\mathbf{X} \sim \mathcal{N}_p(\mathbf{0}, \sigma^2 \mathbf{I})$, on suppose que $\mathbf{X}^T \mathbf{X}$ se décompose en k formes quadratiques, $Q_i = \mathbf{X}^T \mathbf{B}_i \mathbf{X}$ avec $r(\mathbf{B}_i) = r_i$.*

Alors Q_1, \dots, Q_k sont mutuellement indépendantes et $Q_i / \sigma^2 \sim \chi_{r_i}^2$ si et seulement si $\sum_{i=1}^k r_i = p$.

C.4.4 Formes quadratiques particulières

C.4.4.1 Vecteurs Gaussiens réels

On définit la forme quadratique particulière, associée au vecteur Gaussien réel $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{D})$, de taille N , par :

$$Q = \mathbf{x}^T \mathbf{D}^{-1} \mathbf{x}. \quad (\text{C.13})$$

Cette forme quadratique est une variable aléatoire positive, dont la densité de probabilité est déterminée en fonction du vecteur moyenne \mathbf{m} et la matrice \mathbf{D} , symétrique définie positive, par une loi $\chi_N^2(\lambda)$, où $\lambda = \mathbf{m}^T \mathbf{D}^{-1} \mathbf{m}$, dont l'expression est donnée par :

$$p_Q(x) = \frac{1}{2} \left(\frac{x}{\lambda} \right)^{(N-2)/4} \exp\left(-\frac{x+\lambda}{2}\right) I_{N/2-1}(\sqrt{\lambda x}) \quad (\text{C.14})$$

Preuve :

Pour prouver ce résultat, posons $\mathbf{w} = \mathbf{D}^{-1/2} \mathbf{x}$. Ce vecteur, de taille N est un vecteur Gaussien de vecteur moyenne

$$\mu = \mathbb{E}(\mathbf{D}^{-1/2} \mathbf{x}) = \mathbf{D}^{-1/2} \mathbf{m}$$

et de matrice de covariance

$$\mathbf{W} = \mathbb{E}(\mathbf{w} \mathbf{w}^T) = \mathbb{E}(\mathbf{D}^{-1/2} \mathbf{x} \mathbf{x}^T \mathbf{D}^{-1/2}) = \mathbf{D}^{-1/2} \mathbf{D} \mathbf{D}^{-1/2} = \mathbf{I}.$$

La forme quadratique Q s'écrit alors :

$$Q = \mathbf{w}^T \mathbf{w},$$

et représente une somme de N carrés de variables aléatoires indépendantes $\mathcal{N}(\mu_i, 1)$.

D'après le résultat de l'annexe A (A.66), la loi de Q est une loi du $\chi_N^2(\lambda)$, où $\lambda = \sum_{i=1}^N \mu_i^2 = \mathbf{m}^T \mathbf{D}^{-1} \mathbf{m}$.

Si $\mathbf{m} = \mathbf{0}$, la loi de Q devient une loi χ_N^2 , d'expression :

$$p_Q(x) = \frac{1}{2^{N/2} \Gamma(N/2)} x^{N/2-1} \exp\left(-\frac{x}{2}\right) \quad (\text{C.15})$$

C.4.4.2 Vecteurs Gaussiens complexes

Pour déterminer la loi de $R = \mathbf{z}^\dagger \mathbf{D}^{-1} \mathbf{z}$, où $\mathbf{z} \sim \mathcal{CN}(\mathbf{m}, 2\mathbf{D})$ est un vecteur Gaussien complexe circulaire de taille N , on se rapporte au vecteur \mathbf{v} , réel, de taille $2N$, défini par les vecteurs réels indépendants, partie réelle et partie imaginaire de \mathbf{z} . Si on note $\mathbf{z} = \mathbf{x} + j \mathbf{y}$, alors $\mathbf{v} = [\mathbf{x} \ \mathbf{y}]^T$ est un vecteur Gaussien réel de taille $2N$ et de matrice de covariance \mathbf{M} définie à partir de \mathbf{D} comme suit :

$$\mathbf{M} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}$$

La forme quadratique R est alors la somme de $2N$ carrés de variables aléatoires indépendantes Gaussiennes $\mathcal{N}(m_i, 1)$ dont la loi est une loi $\chi_{2N}^2(\lambda)$, où $\lambda = \mathbf{m}^T \mathbf{D}^{-1} \mathbf{m}$, d'expression :

$$p_R(x) = \frac{1}{2} \left(\frac{x}{\lambda} \right)^{(N-1)/2} \exp\left(-\frac{x+\lambda}{2}\right) I_{N-1}(\sqrt{\lambda x}). \quad (\text{C.16})$$

Lorsque $\mathbf{m} = \mathbf{0}$, la loi de R devient une loi χ_{2N}^2 , soit encore une loi Gamma $\mathcal{G}(N, 2)$ d'expression :

$$p_R(x) = \frac{1}{2^N \Gamma(N)} x^{N-1} \exp\left(-\frac{x}{2}\right) \quad (\text{C.17})$$

C.4.4.3 SIRV réels

Soit un SIRV $\mathbf{Y} = \sqrt{\tau} \mathbf{X}$, où τ est une variable aléatoire positive de loi $p(\tau)$ et où $\mathbf{X} \sim \mathcal{N}_m(\mathbf{0}, \mathbf{M})$.

Soit

$$Q = \mathbf{Y}^T \mathbf{M}^{-1} \mathbf{Y} = \tau \mathbf{X}^T \mathbf{M}^{-1} \mathbf{X}.$$

D'après (C.15), la densité de probabilité de Q conditionnellement à τ est la loi de τZ , où $Z \sim \chi_m^2$, soit une loi Gamma $\mathcal{G}(m, 2\tau)$:

$$p_{Q/\tau}(u) = \frac{\tau^{-\frac{m}{2}}}{2^{\frac{m}{2}} \Gamma(\frac{m}{2})} u^{\frac{m}{2}-1} \exp\left(-\frac{u}{2\tau}\right). \quad (\text{C.18})$$

La loi de Q , indépendante du SIRV considéré, s'obtient en intégrant (C.18) sur $p(\tau)$, soit :

$$p_Q(u) = \int_0^{+\infty} \frac{\tau^{-\frac{m}{2}}}{2^{\frac{m}{2}} \Gamma(\frac{m}{2})} u^{\frac{m}{2}-1} \exp\left(-\frac{u}{2\tau}\right) p(\tau) d\tau. \quad (\text{C.19})$$

C.4.4.4 SIRV complexes

La forme quadratique Q' d'un SIRV complexe $\mathbf{Y} = \sqrt{\tau} \mathbf{X}$, où τ est une variable aléatoire positive de loi $p(\tau)$, $\mathbf{X} \sim \mathcal{CN}_m(\mathbf{0}, \mathbf{M}_y = 2\mathbf{M})$ (\mathbf{M} est la matrice de covariance commune des vecteurs partie réelle et partie imaginaire mutuellement indépendants), définie par

$$Q' = \mathbf{Y}^\dagger \mathbf{M}_y^{-1} \mathbf{Y} = 2\tau \mathbf{X}^\dagger \mathbf{M}^{-1} \mathbf{X}$$

admet la densité de probabilité suivante :

$$p_{Q'}(u) = \int_0^{+\infty} \frac{\tau^{-m}}{2^m \Gamma(m)} u^{m-1} \exp\left(-\frac{u}{2\tau}\right) p(\tau) d\tau, \quad (\text{C.20})$$

indépendante du SIRV considéré. La loi de $U = \tau \mathbf{X}^\dagger \mathbf{M}^{-1} \mathbf{X}$ est donc la loi de $Q'/2$, soit :

$$p_U(u) = \int_0^{+\infty} \frac{\tau^{-m}}{\Gamma(m)} u^{m-1} \exp\left(-\frac{u}{\tau}\right) p(\tau) d\tau, \quad (\text{C.21})$$

Publications

Revue à comité de lecture

- [Jay02a] E. Jay, J.P. Ovarlez, D. Declercq, P. Duvaut, *BORD : Bayesian Optimum Radar Detector*, Signal Processing - à paraître,
- [Jay00b] E. Jay, J.P. Ovarlez, D. Declercq, P. Duvaut, *New Methods of Radar Performance Analysis*, Signal Processing - Vol. 80-12, pp.2527-2540, Décembre 2000.

Actes de congrès

- [Jay02b] E. Jay, J.P. Ovarlez, D. Declercq, P. Duvaut, *BORD : Bayesian Optimum Radar Detector*, Proc.-IEEE-ICASSP02, Vol. II, pp.1289-1292, Orlando, FL, USA,
- [Jay01] E. Jay, J.P. Ovarlez, D. Declercq, P. Duvaut, *PEOD : Padé Estimated Optimum radar Detector*, IEEE-RADAR Conference, pp.270-274, Atlanta, USA, Mai 2001,
- [Jay00a] E. Jay, J.P. Ovarlez, "Evaluation of radar performances in low grazing angle environment", Symposium OTAN organisé par Sensors & Electronics Technology Panel : "Low grazing angle clutter : its characterization, measurement and applications", Paper No.38, pp.1-10, Laurel, MD-USA, Avril 2000,
- [Ovar99] J.P. Ovarlez, E. Jay, *New Methods of Radar Performance Analysis*, Proc.-IEEE-ICASSP99, Phoenix, Arizona-USA, SPTM 2-6, Vol.3, pp.1181-1184, March 1999,
- [Jay99a] E. Jay, J.P. Ovarlez, "Nouvelles méthodes d'analyse de performances de détection radar", 5^{ème} Colloque International sur les Systèmes Radars, Brest, France, Mai 1999,
- [Jay99b] E. Jay, J.P. Ovarlez, "Estimation paramétrique des performances de détection radar en environnement inconnu", 17^{ème} Colloque GRETSI, pp.997-1000, Vannes, France, Septembre 1999.

Rapport

- [Jay98] E. Jay, *Détection dans du fouillis non-Gaussien*, Rapport de stage de DEA, Juillet 1998.

Bibliographie

- [Abra64] M. ABRAMOWITZ AND I.A. STEGUN, "*Handbook of Mathematical Functions*", National Bureau of Standard, AMS 55, June 1964
- [Alte79] 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
- [Amin91] H. AMINDAVAR, "*Applications of Padé approximations in signal analysis*", Ph.D. dissertation, Dept. of Electrical Engineering, University of Washington, Seattle, June 1991
- [Amin94a] H. AMINDAVAR, J. A. RITCEY, "*Padé Approximations for Detectability in K-Clutter and Noise*", IEEE Trans.-AES, Vol.30, No.2, pp.425-434, April 1994
- [Amin94b] H. AMINDAVAR, J. A. RITCEY, "*Padé Approximations of Probability Density Functions*", IEEE Trans.-AES, Vol.30, No.2, pp. 416-424, April 1994
- [Anas95] V. ANASTASSOPOULOS AND G. A. LAMPROPOULOS, "*Optimal CFAR detection in Weibull clutter*", IEEE Trans.-AES, Vol.31, No.1, pp.52-64, January 1995
- [Anas99] V. ANASTASSOPOULOS, G. A. LAMPROPOULOS, A. DROSOPOULOS AND M. REY, "*High resolution radar clutter statistics*", IEEE Trans.-AES, Vol.35, No.1, pp.43-60, January 1999
- [Azza95] T. AZZARELLI, "*General class of non-Gaussian coherent clutter models*", IEE Proc.-Radar, sonar navig., Vol.142, No.2, pp.61-70, April 1995
- [Bark99] M. BARKAT AND F. SOLTANI, "*Cell-averaging CFAR detection in compound clutter with spatially correlated texture and speckle*", IEE Proc.-Radar, sonar navig., Vol.146, No.6, pp.279-284, December 1999
- [Barn96] 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
- [Bart88] D.K. BARTON, "*Modern radar analysis*", Artech House, Inc., 1988
- [Bill93] J.B. BILLINGSLEY, "*Ground Clutter Measurements for Surface-Sited Radar*", Technical Report 780, MIT, February 1993
- [Bill02] J.B. BILLINGSLEY, "*Low-angle radar land clutter. Measurements and empirical models*", SciTech Publishing, February 2002
- [Bird95] J.S. BIRD, "*Calculating the performance of linear and square-law detectors*", IEEE Trans.-AES, Vol.31, No.1, pp.39-51, January 1995
- [Blac00] D. BLACKNELL, "*Target detection in correlated SAR clutter*", IEE Proc-Radar, sonar navig., Vol.147, No.1, pp.9-16, February 2000
- [Camp47] G.A. CAMPBELL AND R.M. FOSTER, "*Fourier Integrals for Practical Applications*", Van Nostrand Reinhold, Princeton, New Jersey, 1947

- [Cont87] E. CONTE, 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
- [Cont91] 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
- [Cont94] E. CONTE, G. RICCI, "Performance prediction in compound-Gaussian clutter", IEEE Trans.-AES, Vol.30, No.2, pp.611-616, April 1994
- [Cont95] 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
- [Cont98] E. CONTE, G. RICCI, "Sensitivity study of GLRT detection in compound-Gaussian clutter", IEEE Trans.-AES, Vol.34, No.1, pp.308-316, January 1998
- [Cont99] 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
- [Cont00] 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
- [Darr81] J. DARRICAU, "Physique et théorie du radar", 2ème Edition, Ed. Sodipe, Paris, 1981
- [DeAn95] R. DE ANGELIS, A. FARINA AND F. ZIRILLI, "Use of perturbation methods in the problem of radar detection against a K-clutter background", IEE Proc.-Radar, sonar navig., Vol.142, No.5, pp.243-251, October 1995
- [Decl98] D. DECLERCQ, "Apport des polynômes d'Hermite à la modélisation non Gaussienne et tests statistiques associés", Thèse de l'Université de Cergy-Pontoise, 1998
- [Dolb90] M. DOLBEAULT, "Analyse complexe", Ed. Masson, Paris 1990
- [Drum97] D.M. DRUMHELLER AND M.L. GLASSER "Detection of Chi-square fluctuating target in arbitrary clutter", IEEE Trans.-AES, Vol.33, No.3, pp.784-794, July 1997
- [Drum99] D.M. DRUMHELLER "Padé approximations to matched filter amplitude probability functions", IEEE Trans.-AES, Vol.35, No.3, pp.1033-1045, July 1999
- [Duva91] P. DUVAUT, "Traitement du signal - Concepts et applications", Ed. Hermès, Paris, 1991
- [Efr91] J. EFRON, P.F. SWASZEK AND D.W. TUFTS, "Insight into detection of deterministic and Gaussian signals in correlated plus impulsive noise environments", IEEE Trans.-SP, Vol.39, No.3, pp.603-611, March 1991
- [Fant01] R.L. FANTE "Central limit theorem : use with caution", IEEE Trans.-AES, Vol.37, No.2, pp.739-740, April 2001
- [Fari85] A. FARINA, A. RUSSO AND F.A. STUDER "Advanced models of target and disturbances and related processors", Proc.of Intern. Radar Conference CCIR, pp.151-158, Washington D.C., May 1985
- [Fari86] A. FARINA, A. RUSSO AND F. SCANNAPIECO "Radar detection of target signals in non-Gaussian clutter : theory and Applications", Proc.CCIR, pp.92-99, 1986
- [Fari87] 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
- [Gand88] P.P. GANDHI, S.A. KASSAM "Analysis of CFAR processors in nonhomogeneous background", IEEE Trans.-AES, Vol.24, No.4, pp.427-445, July 1998

-
- [Gand89] P.P. GANDHI, S.A. KASSAM "An adaptive order statistic constant false alarm rate detector", Proc. of the IEEE Intern.Conf. on Syst. Eng., New-York, pp.85-88, 1989
- [Gile73] J. GILEWICZ "Detection of the Best Padé Approximant and the Determination of Fourier Coefficients of the Insufficiently Sampled Functions", In P.R.Graves-Morris (Ed.), New-York : Academic Press, 1973
- [Gini97a] 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
- [Gini97b] 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
- [Gini98] 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
- [Gini99a] 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
- [Gini99] 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
- [Gini00a] 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
- [Gini00b] F. GINI "Performance analysis of two structured covariance matrix estimators in compound-Gaussian clutter", Signal Processing, Vol.80, pp.365-371, 2000
- [Gini00c] 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
- [Gini01a] 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
- [Gini01b] 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
- [Gold76] J. GOLDMAN "Detection in the Presence of Spherically Symmetric Random Vectors", IEEE Trans.-IT, Vol.22, No.1, pp.52-59, January 1976
- [Gord95] S.D. GORDON AND J.A. RITCEY "Calculating the K -distribution by saddlepoint integration", IEE Proc.-Radar, sonar navig., Vol.142, No.4, pp.162-166, August 1995
- [Grec01a] M. GRECO, F. GINI, A. FARINA AND J.B. BILLINGSLEY "Validation of windblown radar ground clutter spectral shape", IEEE Trans.-AES, Vol.37, No.2, pp.538-548, April 2001
- [Grec01b] M. GRECO, F. GINI AND M. DIANI "Robust CFAR detection of random signals in compound-Gaussian clutter plus thermal noise", IEE Proc.-Radar, sonar navig., Vol.148, No.4, pp.227-232, August 2001
- [Guan00] J. GUAN, Y. HEE AND Y.N. PENG "Distributed CFAR detector based on local test statistic", Signal Processing, Vol.80, pp.373-379, 2000
- [Hels84] C.W. HELSTROM AND J.A. RITCEY "Evaluating radar detection probabilities by steepest descent integration", IEEE Trans.-AES, Vol.20, No.5, pp.624-633, September 1984

- [Hou87] X.Y. HOU, N. MORINAGA AND T. NAMEKAWA "Direct evaluation of radar detection probabilities", IEEE Trans.-AES, Vol.23, No.4, pp.418-424, July 1987
- [Hou89] X.Y. HOU AND N. MORINAGA "Detection performance in K -distributed and correlated Rayleigh clutters", IEEE Trans.-AES, Vol.25, No.5, pp.634-641, September 1989
- [Jake76] E. JAKEMAN AND P.N. PUSEY "A model for non-Rayleigh sea echo", IEEE-AP, Vol.24, pp.806-814, 1976
- [Jake80] E. JAKEMAN "On the statistics of K -distributed noise", J.Phys.A : Math Gen., Vol.13, pp.31-48, 1980
- [Jao84] J.K. JAO "Amplitude distribution of composite terrain radar clutter and the K -distribution", IEEE Trans.-AP, Vol.32, No.10, pp.1049-1062, October 1984
- [Kapo99] R. KAPOOR, A. BANERJEE, G.A. TSIHRINTZIS AND N. NANDHAKUMAR "UWB radar detection of targets in foliage using Alpha-stable clutter models", IEEE Trans.-AES, Vol.35, No.3, pp.819-834, July 1999
- [Kay98] S.M. KAY "Fundamentals of statistical signal processing - Detection theory - Vol.2", Prentice Hall PTR, 1998
- [Kell86] E.J. KELLY "An adaptive detection algorithm", IEEE Trans.-AES, Vol.23, November 1986, pp.115-127
- [Kim95] C.J. KIM AND H.S. LEE "Detection analysis of a generalized order statistics CFAR detector for a correlated Rayleigh target", Signal Processing, Vol.47, pp.227-233, 1995
- [Kora68] 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
- [LeCh89] F. LE CHEVALIER "Principes de traitement des signaux radar et sonar", Ed.Masson, Paris, Milan, Barcelona et Mexico, 1989
- [Li89] G. LI AND K.-B. YU "Modelling and simulation of coherent Weibull clutter", IEE Proc.Pt.-F, Vol.136, No.1, pp.2-12, February 1989
- [Lomb96] P. LOMBARDO AND A. FARINA "Coherent radar detection against K -distributed clutter with partially correlated texture", Signal Processing, Vol.48, No.1, pp.1-16, January 1996
- [Mari95] L.J. MARIER "Correlated K -distributed clutter generation for radar detection and track", IEEE Trans.-AES, Vol.31, No.2, pp.568-580, April 1995
- [Meye73] D.P. MEYER AND H.A. MAYER "Radar Target Detection - Handbook of Theory and Practice", Academic Press, New-York and London, 1973
- [deMi97] G. DE MIGUEL AND J.R. CASAR, "CFAR detection for Weibull and other log-log-linear tail clutter distributions", IEE Proc.-Radar, sonar navig., Vol.144, No.2, pp.64-70, April 1997
- [Nagl95] D.T. NAGLE AND J. SANIIE "Performance analysis of linearly combined order statistic CFAR detectors", IEEE Trans.-AES, Vol.31, No.2, pp.522-533, April 1995
- [Ovar95] J.P. OVARLEZ, "Analyse des Données de l'Expérimentation Site-Bas 1995", Rapport ONERA RT 30/5272 SY, septembre 1996
- [Ovar96] J.P. OVARLEZ, "Détection en Environnement Non Gaussien", Rapport ONERA RT 6/5275 SN, Octobre 1996
- [Ovar98] J.P. OVARLEZ, "Modélisation Affinée du Fouillis de Sol", Rapport ONERA RT 12/4401 DEMR/Y, Février 1998

-
- [Pent92] F. ALUFFI-PENTINI, A. FARINA, 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
- [Pici70] B. PICINBONO, G. VEZZOSI, "Détection d'un signal certain dans un bruit non stationnaire et non Gaussien", Annales des Télécomm., Vol.25, pp.433-439, 1970
- [Puls00] N.B. PULSONE AND M.A. ZATMAN, "A computationally efficient two-step implementation of the GLRT", IEEE Trans.-SP, Vol.48, No.3, pp.609-616, March 2000
- [Ragh95] 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
- [Rang91] M. RANGASWAMY, D.D. WEINER, A. ÖZTÜRK, "Simulation of correlated non-Gaussian interference for radar signal detection", Proc.Conference on signals, systems and computers, Pacific Grove, CA, USA, 4-6 Nov.1991, pp.148-152
- [Rang93] M. RANGASWAMY, D.D. WEINER, A. ÖZTÜRK, "Non-Gaussian Random Vector Identification Using Spherically Invariant Random Processes", IEEE Trans.-AES, Vol.29, No.1, pp.111-123, January 1993
- [Rang95] M. RANGASWAMY, D.D. WEINER, A. ÖZTÜRK, "Computer Generation of Correlated Non-Gaussian Radar Clutter", IEEE Trans.-AES, Vol.31, No.1, pp.106-116, January 1995
- [Rifk94] R. RIFKIN, "Analysis of CFAR performance in Weibull clutter", IEEE Trans.-AES, Vol.30, No.2, pp.315-328, April 1994
- [Robe92] C.P. ROBERT, "L'analyse Statistique Bayésienne", Economica, Paris, 1992
- [Robe94] C.P. ROBERT, "The Bayesian Choice : a Decision-Theoretic Motivation", Springer-Verlag, New York, 1994
- [Robe96] C.P. ROBERT, "Méthodes de Monte Carlo par chaînes de Markov", Economica, Paris, 1996
- [Sadl94] B.M. SADLER, G.B. GIANNAKIS AND K-S. LII, "Estimation and detection in non-Gaussian noise using higher order statistics", IEEE Trans.-SP, Vol.42, No.10, pp.2729-2741, October 1994
- [Sang94] K.J. SANGSTON, K.R. GERLACH, "Coherent detection of radar targets in a non-Gaussian background", IEEE Trans.-AES, Vol.30, No.2, pp. 330-340, April 1994
- [Sang99] K.J. SANGSTON, F. GINI, M. V. GRECO, A. FARINA, "Structures for radar detection in compound Gaussian clutter", IEEE Trans.-AES, Vol.35, No.2, pp.445-458, April 1999
- [Scha71] L. SCHARF, D.W. LYTLE, "Signal detection in Gaussian noise of unknown level : an invariance application", IEEE Trans.-IT, Vol.17, pp.404-411, July 1971
- [Schl76] D.C. SCHLEHER, "Radar detection in Weibull clutter", IEEE Trans.-AES, Vol.12, No.6, pp.736-743, November 1976
- [Shei98] A. SHEIKHI, M.M. NAYEBI AND M.R. AREF, "Adaptive detection algorithm for radar signals in autoregressive interference", IEE Proc. Radar, Sonar navig., Vol.145, No.5, pp.309-314, October 1998
- [Shni95] D.A. SHNIDMAN, "Radar detection and their calculation", IEEE Trans.-AES, Vol.31, No.3, pp.928-950, July 1995
- [Shni99] D.A. SHNIDMAN, "Generalized radar clutter model", IEEE Trans.-AES, Vol.35, No.3, pp.857-865, July 1999

- [Silv86] B.W. SILVERMAN *"Density Estimation for Statistics and Data Analysis"*, Chapman & Hall, London, 1986
- [Smit00] M.E. SMITH AND P.K. VARSHNEY, *"Intelligent CFAR processor based on data variability"*, IEEE Trans.-AES, Vol.36, No.3, pp. 837-847, July 2000
- [Solt97] F. SOLTANI AND M. BARKAT, *"CFAR binary integration detection in nonhomogeneous partially correlated clutter"*, IEE Proc. Radar, Sonar navig., Vol.144, No.5, pp.293-300, October 1997
- [Ste197] D.W.J. STEIN AND G.M. DILLARD *"Applying hidden Markov models to radar detection in clutter"*, Proc. IEE-Radar 97, Publication No. 449, October 1997
- [Swer54] P. SWERLING *"Probability of Detection for Fluctuating Targets"*, The RAND Corporation, Research Memorandum RM-1217, March 17, 1954
- [Sza01] W.J. SZAJNOWSKI AND J.B. WYNNE *"Simulation of dependent samples of symmetric Alpha-stable clutter"*, IEEE-SP letters, Vol.8, No. 8, pp.151-152, May 2001
- [Tann96] M.A. TANNER *"Tools for Statistical Inference : Methods for the Exploration of Posterior Distributions and Likelihood Functions"*, Springer-Verlag, 3rd ed., 1996
- [Thou82] L. THOUREL, *"Initiation aux techniques modernes des radars"*, Cepadues ed., Toulouse, 1982
- [Tsih97] G.A. TSIHRINTZIS, C.L.. NIKIAS, *"Data-adaptive algorithms for signal detection in sub-Gaussian impulsive interference"*, IEEE Trans.-SP, Vol.45, No.7, pp.1873-1878, July 1997
- [Trun70] 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
- [VanT71] H.L. VAN TREES, *"Detection, Estimation and Modulation Theory"*, Part I, II and III, John Wiley and Sons, New York 1971
- [Visw92] R. VISWANATHAN AND A. EFTEKHARI, *"A selection and estimation test for multiple target detection"*, IEEE Trans.-AES, Vol.28, pp. 505-519, April 1992
- [Ward81] K.D. WARD, *"Compound representation of high resolution sea clutter"*, Electronics letters, Vol.17, No.16, pp.561-563, 6th August 1981
- [Watt85] 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
- [Watt87] S. WATTS, *"Radar detection prediction in K-distributed sea clutter and thermal noise"*, IEEE Trans.-AES, Vol.23, No.1, pp.40-45, January 1987
- [Watt96] S. WATTS, *"Cell-averaging gain in spatially correlated K-distributed clutter"*, IEE Proc.-Radar, sonar navig., Vol.143, No.5, pp.321-327, October 1996
- [Yao73] K. YAO *"A Representation Theorem and its Applications to Spherically Invariant Random Processes"*, IEEE Trans.-IT, Vol.19, No.5, pp.600-608 Sept. 1973

Résumé

Les échos radar provenant des diverses réflexions du signal émis sur les éléments de l'environnement (le fouillis) ont longtemps été modélisés par des vecteurs Gaussiens. La procédure optimale de détection se résumait alors en la mise en œuvre du filtre adapté classique. Avec l'évolution technologique des systèmes radar, la nature réelle du fouillis s'est révélée ne plus être Gaussienne. Bien que l'optimalité du filtre adapté soit mise en défaut dans pareils cas, des techniques TFAC (Taux de Fausses Alarmes Constant) ont été proposées pour ce détecteur, dans le but d'adapter la valeur du seuil de détection aux multiples variations locales du fouillis. Malgré leur diversité, ces techniques se sont avérées n'être ni robustes ni optimales dans ces situations. A partir de la modélisation du fouillis par des processus complexes non-Gaussiens, tels les SIRP (Spherically Invariant Random Process), des structures optimales de détection cohérente ont pu être déterminées. Ces modèles englobent de nombreuses lois non-Gaussiennes, comme la K-distribution ou la loi de Weibull, et sont reconnus dans la littérature pour modéliser de manière pertinente de nombreuses situations expérimentales. Dans le but d'identifier la loi de leur composante caractéristique qu'est la *texture*, sans *a priori* statistique sur le modèle, nous proposons, dans cette thèse, d'aborder le problème par une approche bayésienne. Deux nouvelles méthodes d'estimation de la loi de la *texture* en découlent : la première est une méthode paramétrique, basée sur une approximation de Padé de la fonction génératrice de moments, et la seconde résulte d'une estimation Monte Carlo. Ces estimations sont réalisées sur des données de fouillis de référence et donnent lieu à deux nouvelles stratégies de détection optimales, respectivement nommées PEOD (Padé Estimated Optimum Detector) et BORD (Bayesian Optimum Radar Detector). L'expression asymptotique du BORD (convergence en loi), appelée le "BORD Asymptotique", est établie ainsi que sa loi. Ce dernier résultat permet d'accéder aux performances théoriques optimales du BORD Asymptotique qui s'appliquent également au BORD dans le cas où la matrice de corrélation des données est non singulière. Les performances de détection du BORD et du BORD Asymptotique sont évaluées sur des données expérimentales de fouillis de sol. Les résultats obtenus valident aussi bien la pertinence du modèle SIRP pour le fouillis que l'optimalité et la capacité d'adaptation du BORD à tout type d'environnement.

Mots-clés: Détection, Non-Gaussien, Processus SIRP, Méthodes bayésiennes, Approximation de Padé

Abstract

For a long time, radar echoes coming from the various returns of the transmitted signal on many objects of the environment (clutter) have been exclusively modelled by Gaussian vectors. The related optimal detection procedure was then performed by the classical matched filter. Then, the technological improvement of radar systems showed that the true nature of the clutter could not be considered as Gaussian any more. Though the optimality of the matched filter is no more valid in such cases, CFAR techniques (Constant False Alarm Rate) were proposed for this detector in order to adapt the value of the detection threshold to the multiple local variations of the clutter. In spite of their diversity, none of these techniques turned to be either robust or optimal in these situations. With the modelling of the clutter by non-Gaussian complex processes, such as SIRP (Spherically Invariant Random Process), optimal structures of coherent detection have been found. These models describe many non-Gaussian laws, like K-distribution or Weibull laws, and are acknowledged in the literature to model many experimental situations in a relevant way. To identify the law of their characteristic component (namely the *texture*) without statistical *a priori* on the model, we propose, in this thesis, to tackle the problem by a Bayesian approach. Two new estimation methods of the *texture* law emerge from this proposition : the first one is a parametric method, based on a Padé approximation of the moment generating function, and the second one results from a Monte Carlo estimation. These estimations are carried out on reference clutter data and lead to two new optimal detection strategies, respectively named PEOD (Padé Estimated Optimum Detector) and BORD (Bayesian Optimum Detector Radar). The asymptotic expression of the BORD (convergence in law), called the "Asymptotic BORD", is established together with its law. This last result gives access to the optimal theoretical performances of the Asymptotic BORD, and may also applied to the BORD if the data correlation matrix is non-singular. The detection performances of BORD and those of Asymptotic BORD are evaluated on experimental ground clutter data. We obtained results that validate both the relevance of SIRP model for the clutter, the optimality of the BORD and its adaptability to any type of environment.

Keywords: Detection, Non-Gaussian, SIRP Processes, Bayesian methods, Padé approximation

