



N° ENSC -



**THESE DE DOCTORAT
DE L'ÉCOLE NORMALE SUPÉRIEURE DE CACHAN**

Spécialité : Traitement du Signal

Présentée par
Alexandre Renaux

**pour obtenir le grade de
DOCTEUR DE L'ÉCOLE NORMALE SUPÉRIEURE DE CACHAN**

Sujet de la thèse :
**Contribution à l'analyse des performances d'estimation en
traitement statistique du signal**

Thèse présentée et soutenue à Cachan le 7 Juillet 2006 devant le jury composé de :

Mr. Bernard FLEURY	Aalborg University	Président
Mr. Jean-Yves TOURNERET	Inst. Nat. Polytechnique de Toulouse	Rapporteur
Mr. Luc VANDENDORPE	Université Catholique de Louvain	Rapporteur
Mr. Eric BOYER	École Normale Supérieure de Cachan	Examineur
Mr. Jean-Pierre DELMAS	Inst. Nat. des Télécommunications Evry	Examineur
Mr. Philippe FORSTER	Université Paris 10	Directeur de thèse
Mr. Pascal LARZABAL	Université Paris-Sud 11	Directeur de thèse

Laboratoire SATIE
ENS CACHAN / CNRS / UMR 8029
61, avenue du Président Wilson, 94235 CACHAN Cedex, France

à mes parents

Merci tout plein !!!

Il y a un peu plus d'une semaine, j'ai réalisé un de mes rêves, à savoir devenir Docteur. Je ne pensais pas que ces trois années auraient pu changer à ce point mes perceptions du point de vue humain comme professionnel. Avec le recul, je pense que les personnes qui ont partagé mon existence durant cette aventure sont la clé de voûte de cet accomplissement. Il m'est maintenant permis de les remercier au travers de ces quelques lignes... et je ne vais pas m'en priver !

A tout seigneur, tout honneur, je ne peux commencer ces remerciements sans exprimer ma profonde gratitude à mon cher directeur de thèse, Pascal Larzabal. Il m'a rarement été donné de côtoyer une personne aussi charismatique. Merci, Pascal, pour m'avoir accueilli dans ton équipe, pour ta patience, pour ton soutien lors de mes nombreux moments de doute, pour tes conseils et surtout pour m'avoir laissé mener mes recherches (et mes enseignements) en toute liberté. Merci aussi pour m'avoir fait voyager et fait comprendre les subtilités de ce "tout petit monde".

Pour encadrer un gaillard comme moi, il fallait bien un deuxième directeur de thèse. J'ai eu la chance que Philippe Forster soit celui-ci. Merci, Philippe, pour m'avoir apporté ta rigueur scientifique sur les points clés de ce travail, pour ta patience et pour tes nombreux conseils lors de nos discussions qu'elles aient été scientifiques ou non.

Je ne sais comment exprimer ma gratitude à ces deux personnes autrement qu'en leur promettant d'agir comme eux avec des étudiants dans ma situation, si un jour l'occasion m'en est donnée.

Mes remerciements s'adressent également aux membres du jury pour l'intérêt qu'ils ont porté à mes travaux. Je remercie tous particulièrement Jean-Yves Tourneret et Luc Vandendorpe qui ont accepté de juger ce travail ainsi que pour leurs remarques très pertinentes sur mon manuscrit. Je suis très sensible à la présence dans ce jury de Jean-Pierre Delmas, dont les articles m'ont toujours grandement intéressé durant ces trois années. Enfin, je tiens à remercier Bernard Fleury d'avoir fait le voyage jusqu'à Cachan et d'avoir présidé ce jury d'une main de maître.

Cette thèse s'est déroulée au sein de l'équipe Traitement de l'Information Multi-capteurs du laboratoire SATIE. L'ambiance qui règne dans ce groupe est tout simplement exceptionnelle. Je tiens donc à remercier toutes les personnes grâce auxquelles ce fut un vrai plaisir de venir au boulot tout les matins. Merci à Benoît pour son humour et pour nos discussions tardives sur le monde de la recherche. Merci à Jean-Pierre pour sa bonne humeur et pour nos discussions matinales. Merci à Eric Boyer pour son enthousiasme, ses conseils de "grand frère" depuis maintenant quatre ans, pour avoir décortiqué mon manuscrit et accepté de participer au jury de cette thèse. Tout spécialement, je tiens à remercier Eric Chaumette pour son soutien, son intérêt pour mes travaux et nos longues réflexions, plus ou moins intellectuelles, menées aux quatre coins du monde autour d'une bière et d'une pizza. Je remercie également Leila, Edson, Anne, Qin, et

Eric Vourc'h pour leur présence chaleureuse durant ces trois années. Enfin, je remercie Cédric Herzet, notre cher post-doc Belge, qui, bien qu'il soit dans nos murs depuis seulement trois mois, a grandement influencé ma vision de la recherche.

Plus particulièrement, je pense profondément que l'ambiance d'une équipe est essentiellement due à la présence des thésards qui sont les vrais animateurs du groupe. Comme je l'ai dit précédemment, l'ambiance qui règne dans mon groupe est exceptionnelle, et c'est pour cela que je tiens à remercier mes amis thésards grâce auxquels ces trois années ont été si agréables au labo. Merci à Marc pour ses paroles toujours justes et avisés, ses conseils amicaux, et son implication dans la vie du groupe en particulier pour motiver les foules. Merci à Fanny, ma charmante collègue, pour nos discussions, pour ses soirées, ses petits plats, sa gaieté matinale et pour avoir supporté mon humour... on va dire lourd!!! Merci à François pour ses délires et sa bonne humeur. My thanks to Angela for her enthusiasm, for her patience when I speak English and for her unbelievable "Irish" capacity to follow me in an other beer round (sláinte!). Merci à Ana pour sa gentillesse et sa bonne humeur. Et merci à Fred pour ses encouragements, son intérêt pour mes travaux et pour son humour décapant.

Plus personnellement, je voudrais remercier mes amis qui m'ont souvent fait oublier cette #!@ #?@! de figure 1.3. Merci à Tof pour sa vision paisible de la vie et nos soirées jeux video, bière, pizza. Merci à Mourroun pour ne jamais être d'accord avec moi, son écoute, et pour m'avoir suivi à New-York. Merci à Seb pour les soirées carbo et les délires. Merci à Ced, mon binôme Vosgien, pour sa sympathie et sa bonne humeur communicative. Enfin, merci à Tramb et à Anthony pour m'avoir hébergé lors de mes retraites à Grenoble et Leeds ainsi que pour nos soirées en conférences.

Pour finir, je voudrais remercier mes parents à qui je dédie cette thèse. On n'a pas souvent l'occasion de remercier ses parents dans une vie, alors je suis content que l'occasion m'en soit donnée ici. Eux, cela fait 27 ans qu'ils me supporte alors je leur dois bien ça. Merci Maman, pour ta gentillesse et pour m'avoir forcé à apprendre mes tables de multiplications quand j'étais petit. Merci Papa pour ton soutien lors de mes études, pour ta confiance en moi et pour m'avoir prodigué tes conseils sans jamais m'imposer tes choix.

Merci à tous, soyez assurés que votre amitié m'est plus précieuse que mes travaux.

Alexandre
17 juillet 2006

Une fois qu'on a passé les bornes, il n'y a plus de limite.
Alphonse Allais

Résumé

Cette thèse porte sur l'étude des performances des estimateurs dans le cadre du traitement du signal et s'attache plus particulièrement à étudier et unifier les bornes minimales d'estimation. Nous caractérisons ainsi les limites des méthodes du maximum de vraisemblance en terme d'erreur quadratique moyenne (EQM).

La difficulté majeure provient du fait que l'EQM de l'estimateur d'un paramètre à support borné se divise en trois régions : la *plage asymptotique*, souvent caractérisée par un grand nombre d'observations ou un faible niveau de bruit, où l'erreur d'estimation est faible, la *plage de décrochement* où l'EQM se dégrade rapidement et la *zone a priori* où les observations se réduisent principalement à la seule contribution du bruit et donc, n'apportent pratiquement plus d'informations sur les paramètres à estimer. Beaucoup de résultats sont disponibles pour la zone asymptotique : distribution des estimées, biais, variance. En revanche, le comportement des estimateur dans les zones de décrochement et *a priori* a été beaucoup moins étudié. Pourtant ces zones non-asymptotiques constituent au même titre que le biais ou la variance une caractéristique fondamentale d'un estimateur puisque qu'elle délimite la plage acceptable de fonctionnement optimal.

Le but de cette thèse est, dans un premier temps, de compléter la caractérisation de la zone asymptotique (en particulier lorsque le rapport signal sur bruit est élevé et pour un nombre d'observations fini) pour les estimateurs au sens du maximum de vraisemblance dans un contexte traitement d'antenne. Dans un second temps, le but est de donner les limites fondamentales de l'EQM d'un estimateur sur ses trois plages de fonctionnement. Les outils utilisés ici sont les bornes minimales de l'EQM autres que les bornes de Cramér-Rao dont la validité n'est qu'asymptotique.

Les résultats obtenus sont appliqués à l'analyse spectrale et à l'estimation de porteuse dans le contexte des communications numériques et fournissent de surcroît des outils intéressants pour prédire la zone de décrochement d'un récepteur.

Mots clés : Traitement d'antenne, analyse spectrale, estimation de porteuse, théorie de l'estimation, estimateurs au sens du maximum de vraisemblance, étude des performances asymptotique et non-asymptotique, bornes minimales d'estimation.

Abstract

This thesis deals with the study of estimators performance in the statistical signal processing framework and, more particularly, with the minimal bounds on the mean square error (MSE). The main difficulty comes from the fact that the MSE of an estimator of a parameter with a finite support exhibits three areas : the *asymptotic area*, characterized by a large number of observations or a weak noise level, where the estimation error is weak, the *threshold area* where the MSE increases dramatically and the *a priori area* where the noise is high and where the observations do not bring informations. Many results concerning the asymptotic area are already available : distribution of the estimates, bias and variance. On the other hand, the non-asymptotic areas have been less studied.

The goal of this thesis is twofold. First, to pursue the investigation in the asymptotic area (particularly in terms of high signal to noise ratio with a finite number of observations) of direction of arrivals maximum likelihood estimators. Secondly, to give methods in order to predict the fundamental limits of an estimator on the three areas. The tool used here will be the minimal bounds on the MSE other than the Cramér-Rao bound for which the validity only holds in the asymptotic area. A main contribution is to propose a unified framework for the Bayesian bounds. The obtained results have been applied in the spectral analysis context and in the carrier estimation context. They bring an interesting tool in order to predict the threshold phenomena.

Keywords : Array processing, spectral analysis, estimation theory, maximum likelihood estimators, asymptotic and non-asymptotic performance analysis, minimal bounds on the mean square error.

Table des matières

Résumé	ix
1 Introduction	1
1.1 Généralités	1
1.2 But de la thèse et résultats	3
2 Performance asymptotique du MV en traitement d'antenne	9
2.1 Introduction	9
2.2 Méthodes du maximum de vraisemblance	11
2.2.1 Maximum de vraisemblance stochastique	12
2.2.2 Maximum de vraisemblance déterministe	13
2.3 Généralités sur les performances asymptotiques	15
2.3.1 Bornes de Cramér-Rao stochastique et déterministe	15
2.3.2 État de l'art	16
2.4 Performance asymptotique en terme de rapport signal sur bruit	17
2.4.1 Maximum de vraisemblance déterministe	17
2.4.1.1 Estimateur du maximum de vraisemblance	18
2.4.1.2 Rappel : le théorème des fonctions implicites	18
2.4.1.3 Application au gradient du critère du maximum de vraisemblance	19
2.4.1.4 Gaussianité asymptotique en RSB de l'estimateur du maximum	
de vraisemblance	20
2.4.1.5 Efficacité asymptotique en RSB de l'estimateur du maximum de	
vraisemblance	21
2.4.1.6 Simulations	21
2.4.2 Maximum de vraisemblance stochastique	23
2.4.2.1 Cas multi-sources	23
2.4.2.2 Cas particuliers	28
2.4.2.3 Simulations	32
3 Bornes minimales d'estimation	37
3.1 Introduction	37
3.2 Un outil pour l'unification des bornes minimales	38
3.3 Unification des bornes déterministes	40
3.3.1 Relations concernant le biais d'un estimateur	40
3.3.2 Borne triviale	41
3.3.3 Bornes de Cramér-Rao et de Bhattacharyya	41
3.3.4 Bornes de Chapman-Robbins et de Barankin	42
3.3.5 Borne d'Abel	44

3.3.6	Remarque	45
3.4	Unification des bornes bayésiennes	45
3.4.1	La famille Ziv-Zakai	46
3.4.1.1	Borne de Bell	46
3.4.1.2	Borne de Bell-Steinberg-Ephraim-VanTrees	49
3.4.1.3	Borne de Bellini-Tartara	49
3.4.1.4	Borne de Chazan-Zakai-Ziv	49
3.4.1.5	Borne de Weinstein	49
3.4.1.6	Remarques concernant les bornes de la famille Ziv-Zakai	50
3.4.2	La famille Weiss-Weinstein	50
3.4.2.1	Reformulations de l'EQM minimale	51
3.4.2.2	Application aux bornes bayésiennes	54
3.4.2.2.1	Borne de Cramér-Rao bayésienne et borne de Bhattacharyya bayésienne	55
3.4.2.2.2	Borne de Bobrovsky-MayerWolf-Zakai	56
3.4.2.2.3	Borne de Bobrovsky-Zakai et borne de Reuven-Messer	56
3.4.2.2.4	Borne de Weiss-Weinstein	58
3.4.2.3	Relations d'ordre	58
3.5	Vers de nouvelles bornes minimales	59
3.5.1	Borne d'Abel bayésienne	59
3.5.2	Borne de Weiss-Weinstein déterministe	61
4	Bornes minimales en traitement statistique du signal	63
4.1	Introduction	63
4.2	Une forme utile de la borne d'Abel	64
4.3	Modèle d'observations gaussiennes à moyenne paramétrée	66
4.3.1	Bornes de Cramér-Rao	67
4.3.2	Bornes de Bhattacharyya	68
4.3.3	Bornes de Chapman-Robbins et de Bobrovsky-Zakai	70
4.3.4	Bornes d'Abel	71
4.3.5	Bornes de Weiss-Weinstein	73
4.4	Analyse spectrale et synchronisation de fréquence porteuse	74
4.4.1	Modèles des observations	74
4.4.2	Formes analytiques des bornes déterministes et bayésiennes	75
4.4.2.1	Bornes de Cramér-Rao	75
4.4.2.2	Bornes de Bhattacharyya	76
4.4.2.3	Bornes de Chapman-Robbins et de Bobrovsky-Zakai	77
4.4.2.4	Bornes d'Abel	78
4.4.2.5	Bornes de Weiss-Weinstein	79
4.4.2.6	Simplification des bornes de Weiss-Weinstein	81
4.4.2.7	Simulations	81
4.4.3	Sur le point test optimal	84
4.4.3.1	Prédiction du décrochement	84
4.4.3.2	Réduction de la charge de calcul et approximations des bornes minimales	85
5	Conclusion	91

A Bornes minimales pour un modèle d'observations gaussiennes à moyenne paramétrée	93
B Bornes minimales en analyse spectrale et synchronisation de fréquence porteuse	95
Liste des figures	101
Liste des tableaux	103
Bibliographie	105

Chapitre 1

Introduction

1.1 Généralités

Un des objets du traitement du signal est l'extraction d'informations pertinentes "cachées" dans un jeu d'observations bruitées obtenu à partir de capteurs. Ces informations ou paramètres d'intérêts peuvent être : une fréquence, une phase, une direction d'arrivée, la puissance d'un bruit, un retard, etc... La théorie de l'estimation offre un cadre formel pour résoudre ce type de problèmes. Plus particulièrement, en estimation dite paramétrique, on se fixe un modèle d'observation dépendant des paramètres d'intérêts et, à partir d'une règle pré-établie basée sur ce modèle, on estime les paramètres d'intérêt (figure 1.1).

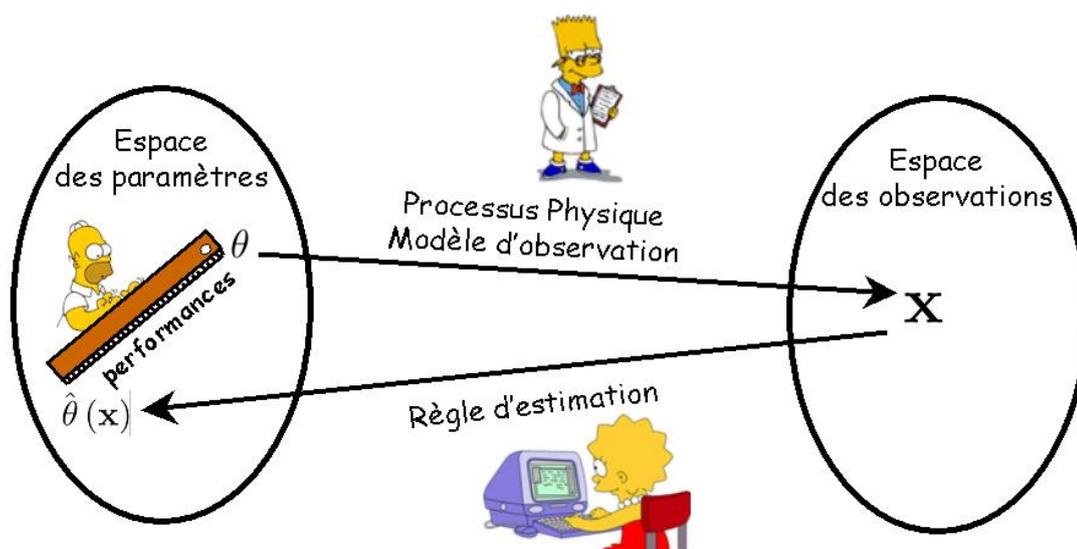


FIGURE 1.1 – Principe de l'estimation paramétrique.

Le modèle d'observation est basé sur les connaissances de l'utilisateur concernant le processus physique considéré. Il dépend généralement d'une fonction non-linéaire des paramètres et incorpore un modèle statistique. La règle d'estimation constitue la pierre angulaire du traitement du signal moderne. Elle conditionne les performances d'estimation, c'est-à-dire le fait que l'estimée soit "suffisamment proche" de la vraie valeur du paramètre et la charge de calcul à mettre en œuvre lors de l'estimation. L'expérience montre qu'un estimateur présentant des

performances proches de l’optimalité devient gourmand en charge de calcul alors qu’un estimateur sous-optimal demande une charge de calcul moindre ce qui présente un intérêt en terme d’implémentation pratique.

L’estimation d’un paramètre s’effectue généralement par la recherche d’un maximum global d’un critère dépendant de la règle d’estimation. Lorsque la recherche du paramètre s’effectue sur un support fini, comme par exemple lors de l’estimation d’un angle ou d’une fréquence normalisée, l’Erreur Quadratique Moyenne (EQM) de l’estimateur qui est l’indicateur de performance le plus répandu s’articule autour de trois régions [Van68] [RB74]. Lorsque le Rapport Signal sur Bruit (RSB) ou le nombre d’observations est élevé, la région est dite asymptotique et l’erreur d’estimation est généralement faible. Lorsque le RSB ou le nombre d’observations décroît, il apparaît un accroissement rapide de l’EQM, donc de l’erreur, dû à l’apparition d’outliers¹ dans le critère. On appelle cette région la zone de décrochement. Enfin, lorsque le RSB ou le nombre d’observations est très faible, le signal observé se réduit principalement à la composante de bruit, ce qui conduit à une distribution de l’estimée quasi-uniforme sur le support du paramètre. L’EQM exhibe alors un comportement plat : c’est la zone de non-information.

Dans le contexte de l’estimation au sens du maximum de vraisemblance de la fréquence d’une cisoïde bruitée, la figure 1.2 illustre deux allures de critère pour un RSB de 30dB et un RSB de -30dB. La vraie valeur de la fréquence normalisée est $\frac{1}{4}$. On remarque un critère “idéal” (non-bruité) avec un maximum global proche de la vraie valeur à fort RSB alors que, pour un faible RSB, le critère est composé de beaucoup de maxima locaux et d’un maximum global éloigné de la vraie valeur. La figure 1.3 donne l’évolution de l’EQM de l’estimateur pour les trois régions susmentionnées (10 observations et 10000 réalisations).

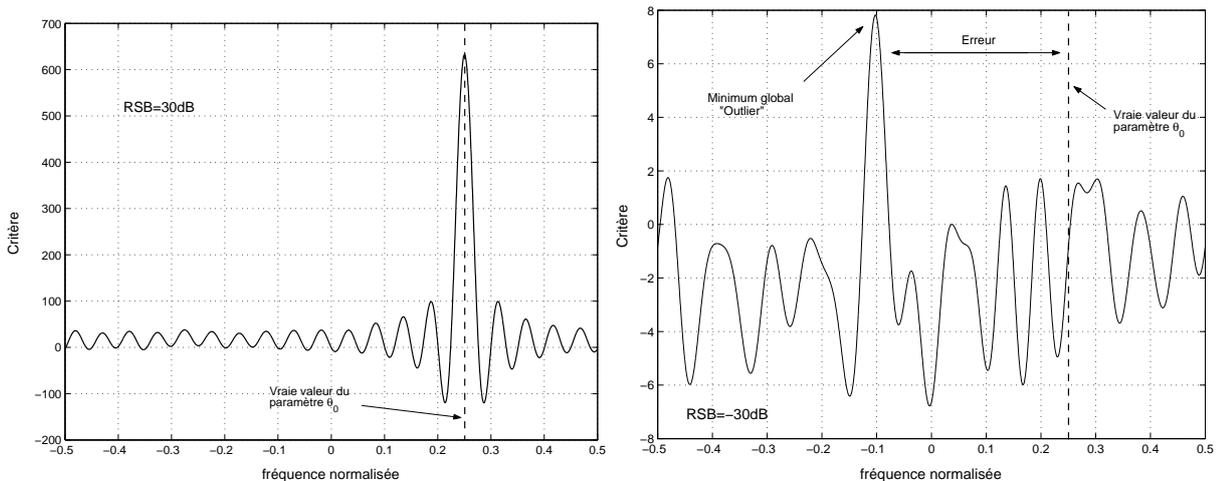


FIGURE 1.2 – Réalisations du critère du maximum de vraisemblance.

La zone de décrochement d’un estimateur délimite sa zone de fonctionnement optimale, c’est-à-dire la zone où l’on peut faire abstraction des outliers dans l’étude des performances. Avec le durcissement des cahiers des charges (temps d’acquisitions réduits, environnements électromagnétiques perturbés, etc...), les estimateurs sont désormais amenés à travailler dans une région de plus en plus proche de cette limite, voire en dessous. En conséquence, la connaissance de la valeur du RSB ou du nombre d’observations, pour un scénario donné, pour laquelle cette rupture brutale des performances apparaît est fondamentale dans l’étude des performances d’estimation et doit être considérée au même titre que l’étude du biais ou de la variance.

1. Terme anglo-saxon désignant une observation aberrante.

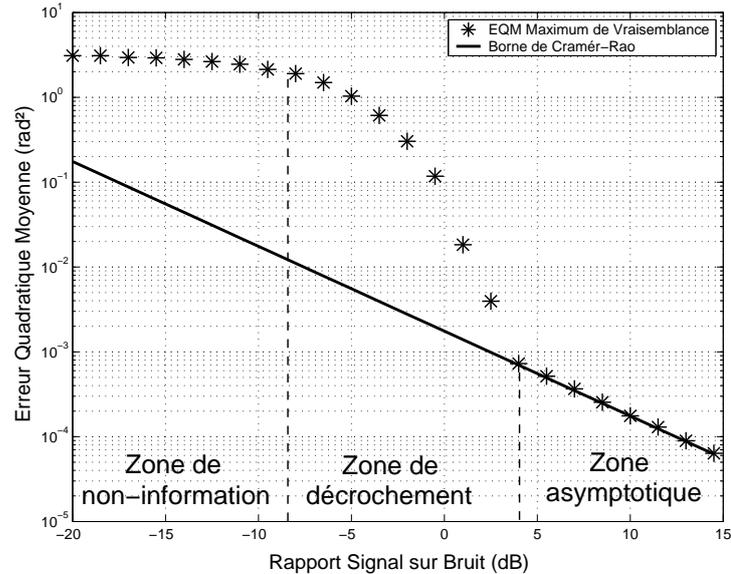


FIGURE 1.3 – Comportement de l'EQM d'un estimateur du maximum de vraisemblance.

La borne de Cramér-Rao (BCR) [Fis22] [Dug37] [Fre43] [Dar45] [Rao45] [Cra46] est également reportée sur la figure 1.3. Elle donne la variance minimale qu'un estimateur non biaisé peut espérer atteindre et est l'outil le plus utilisé par la communauté du traitement du signal pour porter un jugement sur les performances d'un estimateur. Un estimateur qui atteint la BCR est dit efficace. On pourra noter sur cet exemple l'efficacité asymptotique de l'estimateur du maximum de vraisemblance. Néanmoins, dans les zones non-asymptotiques, c'est-à-dire la zone de décrochement et la zone de non-information, la BCR n'est plus pertinente puisqu'elle repose sur un développement limité à l'ordre 1 de l'erreur alors que la présence d'outliers est un phénomène global.

De ces quelques remarques, émergent plusieurs questions! Comment calculer le biais et la variance d'un estimateur dans les zones asymptotique et non-asymptotique? A quoi est due l'apparition des outliers dans le critère et existe-t-il un cadre théorique pour analyser leur influence? Comment prédire le décrochement d'un estimateur pour un scénario donné en évitant une simulation de type Monte-Carlo souvent lourde en charge de calcul? Existe-t-il d'autres bornes pertinentes dans les trois zones de l'EQM?

Ces sujets ont intéressé la communauté du traitement du signal et plusieurs réponses, parfois partielles, à ces questions ont été apportées depuis les premiers travaux de Cramér [Cra46] et leurs extensions par Van Trees [Van68]. Il subsiste néanmoins des zones d'ombre qui résistent encore à l'heure actuelle. Le travaux présentés dans ce document s'inscrivent dans cette problématique.

1.2 But de la thèse et résultats

Le travail proposé dans ce document a été initialement motivé par les résultats obtenus dans la thèse d'Eric Boyer soutenue le 17 décembre 2002 à l'École Normale Supérieure de Cachan et intitulée : *Estimation paramétrique des moments spectraux d'échos Doppler : application aux radars strato-troposphériques* [Boy02]. L'un des résultats d'Eric concerne les performances asymptotiques et non-asymptotiques des estimateurs du maximum de vraisemblance stochastique (MVS) et déterministe (MVD) en terme d'EQM dans le contexte de l'estimation d'une direction d'arrivée.

L'approche était doublement originale :

- Premièrement et contrairement au cas standard, le terme asymptotique n'était pas pris au sens d'un grand nombre d'observations mais en terme de RSB élevé pour lequel peu de résultats étaient disponibles. Il a ainsi été démontré que, à fort RSB, dans le cas d'une seule source et d'une antenne linéaire uniforme (c'est-à-dire un seul angle d'arrivée à estimer) et pour un nombre fixe d'observations T , l'estimateur du MVS était distribué selon une loi de Student à $2T$ degrés de libertés avec une variance asymptotique égale à $T/(T - 1)$ fois la BCR. De manière surprenante, l'estimateur du MVS était donc non-gaussien et non-efficace (il n'atteignait pas la BCR) à fort RSB pour un nombre fini d'observations contrairement au sens classique du terme asymptotique, c'est-à-dire à nombre d'échantillons infini.
- Deuxièmement, une étude des performances non-asymptotique (faible nombre d'observation et/ou faible RSB) était conduite en suivant les travaux de Rife et Boorstyn [RB74] et un estimateur du décrochement de l'EQM de l'estimateur du MVS et du MVD était fourni toujours dans le cadre de l'estimation d'une direction d'arrivée.

Basées sur ces résultats, les motivations de cette thèse étaient donc doubles :

- Généraliser les résultats obtenus à un nombre quelconque de sources et pour une géométrie d'antenne arbitraire pour l'estimation de directions d'arrivées au sens du maximum de vraisemblance à fort RSB. En effet, les algorithmes du maximum de vraisemblance sont caractérisés par un pouvoir de résolution important, c'est-à-dire la capacité à séparer deux sources proches. Il semblait donc intéressant de connaître les performances statistiques de ces estimateurs dans le cas de sources multiples.
- Mettre en place des outils théoriques généraux pour l'étude des performances ultimes des estimateurs. En effet, la méthode de Rife et Boorstyn approxime l'EQM d'un estimateur $\hat{\theta}$ de la manière suivante :

$$EQM(\hat{\theta}) \simeq (1 - p)\sigma_{asymptotique}^2 + p\sigma_{a priori}^2, \quad (1.1)$$

où p représente la probabilité d'apparition d'un outlier, $\sigma_{asymptotique}^2$ est la variance asymptotique de l'estimateur (parfois approximée par la BCR) et $\sigma_{a priori}^2$ la variance dans la zone de non-information (généralement approximée par la variance d'une variable aléatoire uniforme sur le support du paramètre). Selon les hypothèses mises en jeu dans le calcul de la probabilité p , l'équation (1.1) fournit une approximation plus ou moins grossière de l'EQM d'un estimateur sur les trois zones susmentionnées et permet donc une caractérisation de la zone de décrochement. Cette méthode et ses variantes telles que la Method of Interval Errors (MIE) [Van68] ont été appliquées avec succès à divers problèmes d'estimation [RB74] [SB85] [TST94] [Xu01] [BFL04b] [BFL04a] [Ath05] [Ric05]. Il faut remarquer que ces méthodes fournissent une EQM approchée pour un estimateur particulier et, en conséquence, ne rendent pas compte des performances ultimes. De plus, cette méthode souffre d'un défaut majeur : le calcul de la probabilité d'apparition d'un outlier p est un problème non-trivial notamment dans un contexte multi-paramètres. Pour rendre compte des performances ultimes, il faut utiliser un autre outil : les bornes minimales de l'EQM ou bornes minimales d'estimation autres que la BCR. En effet, comme vu précédemment la BCR devient trop optimiste dans la zone non-asymptotique et ne rend pas compte du phénomène de décrochement. Il existe dans la littérature d'autres bornes minimales plus pertinentes que la BCR et présentant l'avantage de rendre compte

du phénomène de décrochement. On peut citer par exemple les bornes dites déterministes de Bhattacharyya [Bha46], Barankin [Bar49], et Abel [Abe93] lorsque les paramètres inconnus sont supposés déterministes et les bornes dites bayésiennes de la famille Ziv-Zakai [ZZ69] [BT74] [CZZ75] [BSET97] et de la famille Weiss-Weinstein [WW88] lorsque les paramètres inconnus sont supposés aléatoires. Si ces bornes dépendent du modèle d'observation considéré, elles présentent l'avantage d'être indépendantes de la règle d'estimation. De ce fait, elles fournissent les performances ultimes qu'un estimateur peut espérer atteindre en termes d'EQM. En outre, elles pourront être utilisées comme un outil pour prédire le décrochement des estimateurs ou optimiser les paramètres du modèle afin d'améliorer les performances d'estimation (par exemple optimiser la géométrie d'un réseau d'antenne).

Trois ans plus tard..., les résultats présentés dans ce document s'articulent autour des trois thèmes suivants :

- Premièrement, nous avons caractérisé le comportement statistique des estimateurs du maximum de vraisemblance à fort RSB pour un nombre fixe d'échantillons et comparé ces résultats aux BCR associées. Dans le contexte d'estimation de directions d'arrivées, lorsque les signaux sources sont modélisés par un processus aléatoire gaussien, nous avons démontré que l'estimateur du maximum de vraisemblance stochastique (MVS) est asymptotiquement (en RSB) non-gaussien et non-efficace quel que soit le nombre de sources et quelle que soit la géométrie de l'antenne [RFBL04] [FBLR03] [RFBL06]. Dans le cas d'un scénario à deux sources, nous avons dérivé la forme analytique pour la matrice de covariance de l'estimateur [RFB04] [RFBL06]. Lorsque les signaux sources sont supposés déterministes mais inconnus, nous avons établi la gaussianité et l'efficacité asymptotique en RSB pour un nombre fixe d'échantillons de l'estimateur du maximum de vraisemblance déterministe (MVD) quelque soit le nombre de sources et quelque soit la géométrie de l'antenne [RFCL06]. Ces résultats sont décrits dans le chapitre 2.
- La seconde partie de ce travail est dédiée à l'étude de bornes minimales d'estimation plus pertinentes que la BCR en vue de la prédiction de la zone de décrochement et de l'approximation des zones asymptotique et non-asymptotique. Ces travaux théoriques sur les bornes minimales sont basés sur les résultats préliminaires concernant l'unification des bornes déterministes [Gla72] [FL01] [FL02] [Cha04]. En effet, les bornes déterministes peuvent être exprimées comme la solution d'un problème de minimisation sous contraintes dont l'objet est de rendre compte le plus finement possible d'un idéal d'absence de biais. Cette formulation permet d'établir des familles de contraintes plus ou moins pertinentes, et par conséquent des classes d'estimateurs dont on cherche l'EQM minimale. On obtient ainsi la meilleure des bornes pour un jeu de contraintes données. Ceci nous permet d'établir des relations d'ordre entre les différentes bornes souvent inaccessibles dans les développements traditionnels exploitant une inégalité de covariance. En adoptant la même démarche, nous avons unifié les bornes Bayésiennes de la famille Weiss-Weinstein [RFL05a] [RFL05b], c'est-à-dire la borne de Cramér-Rao bayésienne [Van68], la borne de Bobrovsky-MayerWolf-Zakai [BMWZ87], la borne de Bhattacharyya bayésienne [Van68], la borne de Bobrovsky-Zakai [BZ76], la borne de Reuven-Messer [RM97] et la borne de Weiss-Weinstein [WW85]. Nous avons démontré que cette famille est également la solution d'un problème d'optimisation sous contraintes. De plus, il est apparu que ces contraintes sont les duales des contraintes intervenant dans l'unification des bornes déterministes. A la lumière d'une telle analogie, nous avons montré que toute les bornes déterministes ont une version Bayésienne dans la famille Weiss-Weinstein et vice versa. Ceci nous a conduit à proposer deux nouvelles bornes minimales : une version bayésienne de la borne d'Abel [RFLR06a] et une

version déterministe de la borne de Weiss-Weinstein. Cette dernière apparaît comme un outil performant dans l'étude de la prédiction des valeurs de RSB ou du nombre d'observations pour lesquelles le décrochement apparaît. Ces résultats sont décrits au chapitre 3.

- La troisième et dernière partie concerne l'application de ces bornes à différents problèmes d'estimation et leur utilisation dans l'étude de la prédiction du décrochement. Nous avons tout d'abord proposé une version simplifiée des bornes d'Abel (déterministe et bayésienne) [RAFL06]. A partir d'un modèle général d'observations gaussiennes à moyenne paramétrée, nous proposons différentes formes analytiques pour la borne d'Abel [RAFL06], la borne d'Abel bayésienne, la borne de Bobrovsky-Zakaï, la borne de Weiss-Weinstein [Ren06] et la borne de Weiss-Weinstein déterministe. Ces bornes sont alors appliquées dans un contexte de l'estimation d'une fréquence porteuse en communications numériques avec signaux pilotes. Toute ces bornes requièrent une maximisation par rapport à des paramètres appelés points tests. Nous avons étudié le comportement de ces points tests et nous en avons déduit des bornes moins gourmandes en charge de calcul [RAFL06] [Ren06]. Nous avons également montré comment l'évolution de la valeur optimale des points tests permet de prédire de manière automatique les valeurs de RSB ou du nombre d'observations pour lesquelles le décrochement des estimateurs apparaît [RAFL06] [Ren06]. Ces résultats sont décrits dans le chapitre 4.

La plupart des travaux présentés dans ce document a été effectuée dans le cadre du réseau d'excellence Européen NEWCOM (Network of Excellence in Wireless COMMunications) dont le but est de renforcer les liens entre les différents groupes de recherche dans le domaine des "systèmes de communications numériques sans fils après la 3G". Notre travail a permis de traiter l'un des problèmes majeurs de ce réseau d'excellence, à savoir la caractérisation des performances d'estimations dans le contexte du Département 2 (MIMO Radio Channel Modelling for Design Optimisation and Performance Assessment of Next Generation Communication Systems).

Les travaux reportés dans ce document ont donné lieu aux articles et communications suivants :

Revue internationale

- [RFCL06] A. Renaux, Ph. Forster, E. Chaumette and P. Larzabal, "On the High SNR Conditional Maximum Likelihood Estimator Full Statistical Characterization", accepté pour publication dans *IEEE Transactions on Signal Processing* Février 2006.
- [RFBL06] A. Renaux, Ph. Forster, E. Boyer and P. Larzabal, "Unconditional Maximum Likelihood Performance At Finite Number Of Samples And High Signal To Noise Ratio", accepté pour publication dans *IEEE Transactions on Signal Processing* Juillet 2006.
- [RAFL06] A. Renaux, L. Najjar-Atallah, P. Larzabal and Ph. Forster, "A Useful Form of the Abel Bound and its Application to Estimator Threshold Prediction", accepté pour publication dans *IEEE Transactions on Signal Processing* Juillet 2006.
- [Ren06] A. Renaux, "Weiss-Weinstein bound for Data Aided carrier estimation", accepté (sous réserve de modifications mineurs) pour publication dans *IEEE Signal Processing Letters* Juillet 2006.

- [RFLR06b] A. Renaux, Ph. Forster, P. Larzabal and C. Richmond, “A New Look at the Bayesian Bounds”, en préparation pour *IEEE Transactions on Information Theory*.

Congrès avec comité de lecture et actes

- [FBLR03] Ph. Forster, E. Boyer, P. Larzabal and A. Renaux, “Non-Efficacité Et Non-Gaussianité Asymptotiques D’un Estimateur Du Maximum de Vraisemblance A Fort Rapport Signal Sur Bruit”, Actes du Dix-Neuvième Colloque GRETSI 2003, Paris, France.
- [RFBL04] A. Renaux, E. Boyer, Ph. Forster and P. Larzabal, “Non-Efficiency And Non-Gaussianity Of A Maximum Likelihood Estimator At High Signal-To-Noise Ratio And Finite Number Of Samples”, in Proc. of *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-04*, Montreal, Canada.
- [RFB04] A. Renaux, Ph. Forster and E. Boyer, “Non Asymptotic Efficiency Of A Maximum Likelihood At Finite Number Of Samples”, in Proc. of European Signal Processing Conference, EUSIPCO-04, Vienna, Austria.
- [RFL05a] A. Renaux, Ph. Forster, and P. Larzabal, “A new derivation of the Bayesian bounds for parameter estimation”, in Proc. of *IEEE Workshop on Statistical Signal Processing 2005, SSP-05*, Bordeaux, France.
- [RFL05b] A. Renaux, Ph. Forster, and P. Larzabal, “Une Nouvelle Approche Des Bornes Bayésiennes”, Actes du Vingtième Colloque GRETSI 2005, Louvain-la-Neuve, Belgium.
- [RFLR06a] A. Renaux, Ph. Forster, P. Larzabal, and C. Richmond “The Bayesian Abel bound on the Mean Square Error”, in Proc. of *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-06*, Toulouse, France.

Séminaires et Workshop

- A. Renaux, “Bornes minimales d’estimation”, Séminaire École Normale Supérieure de Cachan, Septembre 2004, Cachan, France.
- A. Renaux, “Théorie de l’estimation et bornes minimales d’estimation : application à l’analyse spectrale et au Radar passif”, Séminaire ONERA, Mars 2005, Palaiseau, France.
- A. Renaux, “Estimation Lower Bounds and Synchronization Issue in Single Carrier System”, Newcom Autumn School on “Estimation Theory for wireless communications” 24-28 Octobre 2005, ENST Paris, France.
- A. Renaux, “Estimation Lower Bounds : links between deterministic and Bayesian bounds”, Newcom Workshop at Aalborg University, 18-19 Novembre 2005, Aalborg, Denmark.

Chapitre 2

Performance asymptotique du MV en traitement d'antenne

2.1 Introduction

L'estimation de directions d'arrivée à l'aide d'un réseau de capteurs est un problème qui a reçu une attention particulière depuis plusieurs années. Bien que la motivation initiale ait été d'ordre militaire avec le développement du radar pendant la seconde guerre mondiale, les applications civiles sont aujourd'hui nombreuses : positionnement par satellite, téléphonie mobile, etc...

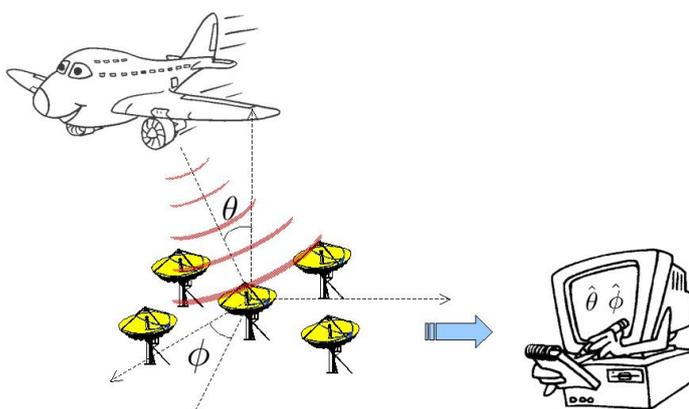


FIGURE 2.1 – Principe de la localisation.

Le but est de localiser angulairement une ou plusieurs sources rayonnantes à partir d'observations issues d'un réseau de capteurs appelé antenne (figure 2.1). Cette localisation peut se faire de manière active, c'est-à-dire qu'un signal connu est émis et différentes versions de ce signal réfléchies par l'environnement sont captées par l'antenne puis traitées. Les applications sont le radar de surveillance ou de poursuite, le GSM, ou encore la prospection sismique. Ce type de localisation est à opposer à la localisation passive où l'antenne écoute simplement les signaux émis spontanément par l'environnement. Les applications sont le sonar ou encore la radio-astronomie.

Le modèle paramétrique des observations reçues par une antenne composée de N capteurs est le suivant :

$$\mathbf{y}(t_i) = \mathbf{A}(\boldsymbol{\theta}) \mathbf{s}(t_i) + \mathbf{b}(t_i) \quad i = 1, \dots, T, \quad (2.1)$$

où :

- T est le nombre d'observations considérées ;
- $\mathbf{y}(t_i) \in \mathcal{M}_{N \times 1}(\mathbb{C})$ est un vecteur d'observations complexes ;
- $\mathbf{s}(t_i) \in \mathcal{M}_{M \times 1}(\mathbb{C})$ est le vecteur complexe des signaux rayonnés par les M sources ;
- $\mathbf{b}(t_i) \in \mathcal{M}_{N \times 1}(\mathbb{C})$ est un vecteur de bruit complexe additif ;
- $\boldsymbol{\theta} = [\boldsymbol{\theta}_1 \quad \boldsymbol{\theta}_2 \cdots \boldsymbol{\theta}_M] \in \mathcal{M}_{P \times M}(\mathbb{R})$ est le vecteur réel des paramètres angulaires des sources ; dans le cas général d'une localisation en 3 dimensions, $\boldsymbol{\theta}_m$ est issu de la concaténation de l'azimut et de l'élévation de chaque source et $P = 2$;
- $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\boldsymbol{\theta}_1) \quad \mathbf{a}(\boldsymbol{\theta}_2) \cdots \mathbf{a}(\boldsymbol{\theta}_M)] \in \mathcal{M}_{N \times M}(\mathbb{C})$ est la matrice déterministe des vecteurs directeurs $\mathbf{a}(\boldsymbol{\theta}_m) \in \mathcal{M}_{N \times 1}(\mathbb{C})$ dépendant de la géométrie de l'antenne, de la structure des capteurs et des directions d'arrivée. Le vecteur directeur $\mathbf{a}(\boldsymbol{\theta}_m)$ mesure la réponse du réseau (fonction de transfert). A titre d'exemple, pour une antenne linéaire uniforme (ALU) et une antenne circulaire uniforme (ACU), le $k^{\text{ième}}$ élément de $\mathbf{a}(\boldsymbol{\theta}_m)$ s'écrit

$$\begin{aligned} a_k^{ALU}(\theta_m) &= e^{j(k-1)\frac{2\pi d}{\lambda} \sin(\theta_m)}, \\ a_k^{ACU}(\theta_m, \phi_m) &= e^{j\frac{2\pi r}{\lambda} \cos(\theta_m) \cos(\phi_m - k\gamma)}, \end{aligned} \quad (2.2)$$

où λ est la longueur d'onde, d la distance inter-capteur pour l'ALU, γ et r respectivement l'angle inter-capteur et le rayon de l'ACU. ϕ_m est l'azimut et θ_m est l'élévation de la source m (figure 2.1).

Le processus $\mathbf{y}(t_i)$ est observé à T instants $\{t_1, t_2, \dots, t_T\}$. De la concaténation des vecteurs d'observations $\mathbf{y}(t_i)$, nous obtenons la matrice des données \mathbf{Y}_T

$$\boxed{\mathbf{Y}_T = \mathbf{A}(\boldsymbol{\theta}) \mathbf{S}_T + \mathbf{B}_T}, \quad (2.3)$$

avec

$$\mathbf{Y}_T = [\mathbf{y}(t_1) \quad \mathbf{y}(t_2) \quad \cdots \quad \mathbf{y}(t_T)], \quad (2.4)$$

$$\mathbf{S}_T = [\mathbf{s}(t_1) \quad \mathbf{s}(t_2) \quad \cdots \quad \mathbf{s}(t_T)], \quad (2.5)$$

$$\mathbf{B}_T = [\mathbf{b}(t_1) \quad \mathbf{b}(t_2) \quad \cdots \quad \mathbf{b}(t_T)]. \quad (2.6)$$

La construction du modèle (2.3) est détaillée dans [OVS93] en particulier concernant l'identifiabilité. Il faut cependant noter que ce modèle est établi selon l'hypothèse de signaux $\mathbf{s}(t_i)$ à bande étroite, c'est-à-dire que la fonction de transfert est invariante dans la bande instantanée des signaux reçus, et aussi selon l'hypothèse de signaux en champ lointain.

Les hypothèses effectuées dans ce chapitre sont les suivantes :

H1 Les signaux $\mathbf{s}(t_i)$ sont supposés circulaires ($\mathbb{E}[\mathbf{s}(t_i) \mathbf{s}^T(t_i)] = \mathbf{0}$), temporellement blancs, issus de processus aléatoires gaussiens de moyennes nulles et de matrice de covariance inconnue $\boldsymbol{\Sigma}_s = \mathbb{E}[\mathbf{s}(t_i) \mathbf{s}^H(t_i)]$. Notons qu'ils peuvent être partiellement corrélés ($\boldsymbol{\Sigma}_s$ n'est pas forcément diagonale mais de rang plein),

H'1 Les signaux $\mathbf{s}(t_i)$ sont supposés déterministes, mais inconnus.

H2 Le bruit $\mathbf{b}(t_i)$ est supposé indépendant des signaux $\mathbf{s}(t_i)$, circulaire ($\mathbb{E}[\mathbf{b}(t_i) \mathbf{b}^T(t_i)] = \mathbf{0}$), issu d'un processus aléatoire gaussien, temporellement blanc, à moyenne nulle et de matrice de covariance $\boldsymbol{\Sigma}_b = \mathbb{E}[\mathbf{b}(t_i) \mathbf{b}^H(t_i)] = \sigma^2 \mathbf{I}_N$ (hypothèse de blancheur spatiale) où σ^2 est la puissance du bruit supposée inconnue.

- H3** Tous les paramètres inconnus sont supposés déterministes, c'est-à-dire invariant par l'expérience.
- H4** $\mathbf{A}(\boldsymbol{\theta})$ est supposée de rang plein ($M < N$) et $\forall \theta_m, \|\mathbf{a}(\theta_m)\| = \sqrt{N}$. La géométrie de l'antenne est non-ambiguë.
- H5** On suppose une phase de détection préliminaire, en conséquence, le nombre de sources est connu.

Dans la suite selon la méthode du maximum de vraisemblance considérée, nous nous placerons dans le cadre de l'hypothèse $H1$ (MVS) ou dans le cadre de l'hypothèse $H'1$ (MVD).

Notons que d'après $H1$, $H'1$ et $H2$, les observations $\mathbf{y}(t_i)$ sont indépendantes. Il faut également noter que le modèle (2.3), classique en traitement d'antenne, nécessite une calibration en condition opérationnelle. En effet, le vecteur directionnel, $\mathbf{a}(\boldsymbol{\theta}_i)$, est supposé connu et déterministe. Cette hypothèse est souvent erronée dans le sens où, par exemple, la géométrie de l'antenne n'est généralement pas parfaitement connue. On pourra se reporter aux références suivantes pour de plus amples détails concernant la calibration : [Van02] [HLS93].

2.2 Méthodes du maximum de vraisemblance

La localisation angulaire de sources a donné lieu à une pléthore d'algorithmes d'estimation. Ces algorithmes peuvent se diviser en deux parties :

- **Les méthodes non-paramétriques.** Elles n'utilisent pas de modèle de signaux reçus. La plus célèbre est sans nul doute la formation de voies qui est une extension spatio-temporelle de la transformée de Fourier. L'extension la plus connue de la formation de voies est la méthode de Capon [Cap69]. Malheureusement, ces méthodes souffrent d'un désavantage majeur : elles sont limitées en résolution¹ par le diagramme d'ambiguïté de l'antenne. En effet, deux sources placées dans le lobe principale du diagramme d'ambiguïté ne seront pas séparées. On peut donner l'exemple de la formation de voies pour laquelle la résolution limite est définie comme la largeur, noté $2\theta_{3dB}$, du lobe principal à 3 dB en dessous du maximum. On obtient approximativement dans le cas du modèle ALU avec des capteurs distants d'une demi longueur d'onde :

$$2\theta_{3dB} = \frac{100}{N}, \quad (\text{en degrés}) \quad (2.7)$$

ce qui signifie qu'une antenne composée de deux capteurs ne pourra séparer deux sources dont l'écart angulaire est inférieur à 50 degrés ou qu'il faudra une centaine de capteurs pour séparer deux sources dont l'écart angulaire est égal à 1 degré, conduisant à un coût de fabrication prohibitif.

- **Les méthodes paramétriques.** Elles exploitent un modèle de signaux reçus, par exemple le modèle (2.3). Ces méthodes sont souvent dites à haute résolution puisqu'elles permettent, sous certaines conditions, d'avoir un pouvoir de résolution infini. Néanmoins, ces méthodes sont beaucoup moins robustes aux erreurs de modèle que les méthodes non-paramétriques. On peut classer ces techniques en deux catégories : les méthodes de sous-espaces et la méthode du maximum de vraisemblance. Les méthodes de sous-espaces s'appuient sur une décomposition de l'espace des observations, grâce à la matrice de covariance des observations, en deux sous-espaces : le sous-espace signal et le sous-espace bruit, et exploitent la propriété d'orthogonalité entre ceux-ci. On citera par exemple la

1. Le concept de résolution est l'aptitude à séparer deux sources proches et de puissance égale.

méthode MUSIC (Multiple Signal Classification) [Sch81] et ses nombreuses variantes : l'algorithme Min-Norm [KT83] [Red79] [LVT89], ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) [PRK86] [RPK86], WSF (Weighted Subspace Fitting) [VOK91], etc.... Les méthodes de sous-espaces requièrent généralement une simple optimisation mono-dimensionnelle. La méthode du maximum de vraisemblance s'appuie quant à elle sur une exploitation des connaissances statistiques des observations par l'intermédiaire de la fonction de vraisemblance. La méthode du maximum de vraisemblance conduit à l'optimisation d'un critère multi-dimensionnel. Elle est souvent qualifiée "d'optimale" pour ses bonnes propriétés statistiques en termes de consistance et d'efficacité. Le revers de la médaille est bien sûr une charge de calcul élevée pour la méthode du maximum de vraisemblance, mais l'augmentation de la puissance des calculateurs a conduit la communauté du traitement du signal à s'y intéresser.

Le principe de l'estimation au sens du maximum de vraisemblance est illustré figure 2.2.

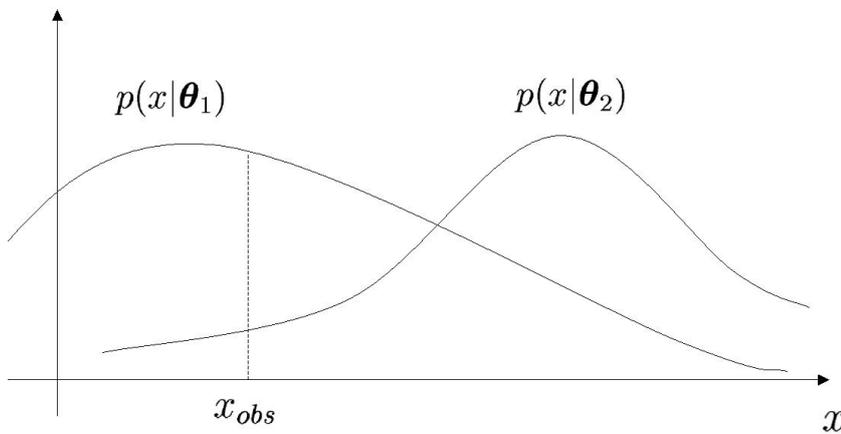


FIGURE 2.2 – Principe de l'estimation au sens du maximum de vraisemblance.

La densité de probabilité $p(x|\theta)$ d'une variable aléatoire x dépend d'un vecteur de paramètres inconnu θ . Sur l'exemple de la figure 2.2, ce vecteur de paramètres peut prendre deux valeurs θ_1 et θ_2 conduisant à deux densités de probabilité $p(x|\theta_1)$ et $p(x|\theta_2)$. Soit une réalisation de la variable aléatoire x appelée observation x_{obs} , alors le paramètre θ_1 est plus vraisemblable que θ_2 car il maximise la probabilité d'apparition de x_{obs} . En généralisant ceci à un vecteur aléatoire \mathbf{x} dont une réalisation sera le vecteur d'observation \mathbf{x}_{obs} et à un continuum de jeux de paramètres, on obtient l'estimateur au sens du maximum de vraisemblance

$$\hat{\theta}_{ML} = \arg \max_{\theta} p(\mathbf{x}_{obs}|\theta). \quad (2.8)$$

Dans le cadre du modèle paramétrique (2.3), deux méthodes du maximum de vraisemblance ont été développées. La méthode du maximum de vraisemblance déterministe (MVD) et la méthode du maximum de vraisemblance stochastique (MVS).

2.2.1 Maximum de vraisemblance stochastique

Le MVS s'appuie sur l'hypothèse $H1$, à savoir que les signaux sources sont issus d'un processus aléatoire gaussien stationnaire. Cette hypothèse trouve sa légitimité dans la multitude de

processus naturels régis par cette loi ainsi que dans le théorème de la limite centrale. Sous les hypothèses $H1$ et $H2$, les observations suivent une loi gaussienne complexe de moyenne nulle et, les matrices de covariance et de pseudo-covariance des observations s'expriment sous la forme :

$$\mathbb{E} [\mathbf{y}(t_i) \mathbf{y}^H(t_j)] = \boldsymbol{\Sigma}_{\mathbf{y}} \delta_{ij} = (\mathbf{A}(\boldsymbol{\theta}) \boldsymbol{\Sigma}_{\mathbf{s}} \mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2 \mathbf{I}_N) \delta_{ij}, \quad (2.9)$$

$$\mathbb{E} [\mathbf{y}(t_i) \mathbf{y}^T(t_j)] = \mathbf{0}. \quad (2.10)$$

Dans ce contexte, les $M^2 + PM + 1$ paramètres inconnus sont

$$\boldsymbol{\eta}_{MVS} = [\text{vec}^T(\boldsymbol{\Sigma}_{\mathbf{s}}) \text{vec}^T(\boldsymbol{\theta}) \sigma^2]^T. \quad (2.11)$$

Sous $H3$, la vraisemblance des observations est donnée par

$$p(\mathbf{Y}_T | \boldsymbol{\eta}_{MVS}) = \frac{1}{(\pi^N |\boldsymbol{\Sigma}_{\mathbf{y}}|)^T} e^{-\sum_{i=1}^T \mathbf{y}^H(t_i) \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{y}(t_i)}. \quad (2.12)$$

L'estimateur du MVS de $\boldsymbol{\eta}_{MVS}$ est obtenu par maximisation de la vraisemblance ou, de manière équivalente, par minimisation de la log-vraisemblance négative normalisée par T où les termes indépendants des paramètres ont été ignorés, c'est-à-dire :

$$\hat{\boldsymbol{\eta}}_{MVS} = \arg \min_{\boldsymbol{\eta}_{MVS}} \ln |\boldsymbol{\Sigma}_{\mathbf{y}}| + \text{tr} \left(\boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \hat{\boldsymbol{\Sigma}}_{\mathbf{y}} \right), \quad (2.13)$$

où $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}} = \frac{1}{T} \sum_{i=1}^T \mathbf{y}(t_i) \mathbf{y}^H(t_i)$ est la matrice de covariance empirique des observations. On peut démontrer que le problème est séparable et qu'il existe une solution analytique $\hat{\boldsymbol{\Sigma}}_{\mathbf{s}}$ et $\hat{\sigma}^2$ à $\boldsymbol{\theta}$ fixé. Après concentration du critère, c'est-à-dire après report de $\hat{\boldsymbol{\Sigma}}_{\mathbf{s}}$ et $\hat{\sigma}^2$ dans celui-ci, on obtient l'estimateur du MVS pour les paramètres d'intérêts $\boldsymbol{\theta}$ [Böh86] [Bre88] [Jaf88] :

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{MVS} &= \arg \min_{\boldsymbol{\theta}} \ln \left| \mathbf{A}(\boldsymbol{\theta}) \hat{\boldsymbol{\Sigma}}_{\mathbf{s}}(\boldsymbol{\theta}) \mathbf{A}^H(\boldsymbol{\theta}) + \hat{\sigma}^2(\boldsymbol{\theta}) \mathbf{I}_N \right| \\ &= \boxed{\arg \min_{\boldsymbol{\theta}} \left| \mathbf{A}(\boldsymbol{\theta}) \hat{\boldsymbol{\Sigma}}_{\mathbf{s}}(\boldsymbol{\theta}) \mathbf{A}^H(\boldsymbol{\theta}) + \hat{\sigma}^2(\boldsymbol{\theta}) \mathbf{I}_N \right|}, \end{aligned} \quad (2.14)$$

avec

$$\hat{\sigma}^2(\boldsymbol{\theta}) = \frac{1}{N-M} \text{tr} \left(\boldsymbol{\Pi}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta}) \hat{\boldsymbol{\Sigma}}_{\mathbf{y}} \right), \quad (2.15)$$

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{s}}(\boldsymbol{\theta}) = \mathbf{A}^{\dagger}(\boldsymbol{\theta}) \left(\hat{\boldsymbol{\Sigma}}_{\mathbf{y}} - \hat{\sigma}^2(\boldsymbol{\theta}) \mathbf{I}_N \right) \mathbf{A}^{\dagger H}(\boldsymbol{\theta}), \quad (2.16)$$

où $\boldsymbol{\Pi}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta}) = \mathbf{I}_N - \mathbf{A}(\boldsymbol{\theta}) \mathbf{A}^{\dagger}(\boldsymbol{\theta})$ est le projecteur orthogonal sur le noyau de $\mathbf{A}^H(\boldsymbol{\theta})$.

2.2.2 Maximum de vraisemblance déterministe

Le MVD s'appuie sur l'hypothèse $H'1$, à savoir que les signaux sources sont supposés déterministes et sont donc autant d'inconnues à estimer. L'intérêt de cette méthode réside dans le fait qu'il existe des applications où les signaux sources ne sont pas gaussiens. On peut citer par exemple le cas des communications numériques ou du radar actif. Sous les hypothèses $H'1$ et $H2$, les observations suivent une loi gaussienne complexe de moyenne

$$\mathbb{E} [\mathbf{y}(t_i)] = \mathbf{A}(\boldsymbol{\theta}) \mathbf{s}(t_i), \quad (2.17)$$

et de matrices de covariance et de pseudo-covariance

$$\mathbb{E} \left[(\mathbf{y}(t_i) - E[\mathbf{y}(t_i)]) (\mathbf{y}(t_j) - E[\mathbf{y}(t_j)])^H \right] = \sigma^2 \mathbf{I}_N \delta_{ij}, \quad (2.18)$$

$$\mathbb{E} \left[(\mathbf{y}(t_i) - E[\mathbf{y}(t_i)]) (\mathbf{y}(t_j) - E[\mathbf{y}(t_j)])^T \right] = \mathbf{0}. \quad (2.19)$$

Dans ce contexte, les $2TM + PM + 1$ paramètres inconnus sont

$$\boldsymbol{\eta}_{MVD} = [\text{vec}^T(\mathbf{S}_T) \text{vec}^T(\boldsymbol{\theta}) \sigma^2]^T. \quad (2.20)$$

Sous $H3$, la vraisemblance des observations est donnée par

$$p(\mathbf{Y}_T | \boldsymbol{\eta}_{MVD}) = \frac{1}{(\pi\sigma^2)^{TN}} e^{-\frac{1}{\sigma^2} \sum_{i=1}^T (\mathbf{y}(t_i) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t_i))^H (\mathbf{y}(t_i) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t_i))}. \quad (2.21)$$

Par passage à la log-vraisemblance négative où les termes indépendants des paramètres ont été ignorés, on obtient l'estimateur du MVD pour $\boldsymbol{\eta}_{MVD}$:

$$\hat{\boldsymbol{\eta}}_{MVD} = \arg \min_{\boldsymbol{\eta}_{MVD}} TN \ln(\sigma^2) + \frac{1}{\sigma^2} \sum_{i=1}^T \|\mathbf{y}(t_i) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t_i)\|^2. \quad (2.22)$$

Pour tout $\boldsymbol{\theta}$ et \mathbf{S}_T , la minimisation de (2.22) par rapport à σ^2 conduit de manière triviale à :

$$\hat{\sigma}^2 = \frac{1}{TN} \sum_{i=1}^T \|\mathbf{y}(t_i) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t_i)\|^2. \quad (2.23)$$

En posant $\boldsymbol{\eta}'_{MVD} = [\text{vec}^T(\mathbf{S}_T) \text{vec}^T(\boldsymbol{\theta})]^T$ et en insérant (2.23) dans (2.22), il vient :

$$\hat{\boldsymbol{\eta}}'_{MVD} = \arg \min_{\boldsymbol{\eta}'_{MVD}} \sum_{i=1}^T \|\mathbf{y}(t_i) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t_i)\|^2. \quad (2.24)$$

Ce critère est quadratique en les $\mathbf{s}(t_i)$, donc

$$\hat{\mathbf{s}}(t_i) = \mathbf{A}^\dagger(\boldsymbol{\theta}) \mathbf{y}(t_i) \iff \hat{\mathbf{S}}_T = \mathbf{A}^\dagger(\boldsymbol{\theta}) \mathbf{Y}_T, \quad (2.25)$$

et par report de (2.25) dans (2.24) on obtient l'estimateur du MVD pour les paramètres d'intérêts $\boldsymbol{\theta}$ [Böh84] :

$$\boxed{\hat{\boldsymbol{\theta}}_{MVD} = \arg \min_{\boldsymbol{\theta}} \text{tr} \left(\mathbf{\Pi}_{\mathbf{A}}^\perp(\boldsymbol{\theta}) \hat{\boldsymbol{\Sigma}}_{\mathbf{y}} \right)}, \quad (2.26)$$

avec $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}} = \frac{1}{T} \sum_{i=1}^T \mathbf{y}(t_i) \mathbf{y}^H(t_i)$ et $\mathbf{\Pi}_{\mathbf{A}}^\perp(\boldsymbol{\theta}) = \mathbf{I}_N - \mathbf{A}(\boldsymbol{\theta}) \mathbf{A}^\dagger(\boldsymbol{\theta})$ est le projecteur orthogonal sur le noyau de $\mathbf{A}^H(\boldsymbol{\theta})$.

Remarque : il existe une variante du MVD où les signaux sources sont également modélisés par des quantités déterministes inconnues mais paramétrées par un jeu fini de fonctions connues. Cette variante est appelée maximum de vraisemblance paramétré (MVP) et est généralement utilisée dans un contexte trajets multiples ou radar actif [LC93] [LHSV95] [CM97].

2.3 Généralités sur les performances asymptotiques

La plupart des études sur les performances asymptotiques des estimateurs sont conduites à fort nombre d'observations ($T \rightarrow \infty$). Ceci est vraisemblablement dû au fait que les méthodes utilisées pour ces analyses asymptotiques proviennent de la communauté mathématique. On pourra citer, par exemple, le théorème limite centrale. Or la notion de RSB ou de nombre de capteurs sont des éléments physiques importants pour la communauté du traitement du signal. Concernant le domaine du traitement d'antenne qui nous intéresse ici, on peut comprendre le terme asymptotique selon trois axes différents : un nombre infini d'observations et/ou un nombre infini de capteurs et/ou un RSB infini. Nous présentons ici un résumé des résultats connus pour les estimateurs au sens du maximum de vraisemblance présentés ci-avant.

2.3.1 Bornes de Cramér-Rao stochastique et déterministe

La borne de Cramér-Rao est la référence la plus généralement utilisée pour juger des performances d'un estimateur. Elle donne la variance minimale qu'un estimateur non-biaisé peut espérer atteindre et est définie comme suit : soit $\hat{\phi}$ un estimateur d'un paramètre déterministe candidat ϕ de vraie valeur ϕ_0 . $\hat{\phi}$ est supposé non-biaisé en ϕ_0 et la dérivée de son biais est supposée nulle en ϕ_0 . Soit $\mathbf{x} \in \Omega$ un vecteur d'observation dépendant de ϕ au travers d'une fonction de vraisemblance $p(\mathbf{x}|\phi)$. Alors on a l'inégalité suivante dite inégalité de Cramér-Rao²

$$\text{cov}(\hat{\phi}) \geq \left(\mathbb{E} \left[\frac{\partial \ln p(\mathbf{x}|\phi)}{\partial \phi} \bigg|_{\phi_0} \frac{\partial \ln p(\mathbf{x}|\phi)}{\partial \phi^T} \bigg|_{\phi_0} \right] \right)^{-1}, \quad (2.27)$$

où la matrice de covariance de $\hat{\phi}$ est donnée par $\text{cov}(\hat{\phi}) = \mathbb{E}[\hat{\phi}\hat{\phi}^T] = \int_{\Omega} \hat{\phi}\hat{\phi}^T p(\mathbf{x}|\hat{\phi}) d\mathbf{x}$.

La quantité $\mathbb{E} \left[\frac{\partial \ln p(\mathbf{x}|\phi)}{\partial \phi} \bigg|_{\phi_0} \frac{\partial \ln p(\mathbf{x}|\phi)}{\partial \phi^T} \bigg|_{\phi_0} \right]$ est appelée matrice d'information de Fisher (MIF). L'inverse de la MIF donne la BCR.

Dans le contexte du modèle (2.3) sous l'hypothèse $H1$ ou $H'1$, on obtient en utilisant (2.12) ou (2.21) et après quelques efforts algébriques, la BCR stochastique (BCRS) et la BCR déterministe (BCRD) pour les directions d'arrivée θ_0 et pour $P = 1$, c'est-à-dire pour un seul paramètre angulaire à estimer par source [SN90b]

$$\boxed{\mathbf{BCRS}(\theta_0) = \frac{\sigma^2}{2T} [\text{Re} \{ \mathbf{H}(\theta_0) \odot \Sigma_s \mathbf{A}^H(\theta_0) \Sigma_y^{-1} \mathbf{A}(\theta_0) \Sigma_s \}]^{-1}}, \quad (2.28)$$

et [SN89] [CTO89]

$$\boxed{\mathbf{BCRD}(\theta_0) = \frac{\sigma^2}{2T} [\text{Re} \{ \mathbf{H}(\theta_0) \odot \Sigma_s \}]^{-1}}, \quad (2.29)$$

avec

$$\begin{aligned} \Sigma_s &= \frac{1}{T} \sum_{i=1}^T \mathbf{s}(t_i) \mathbf{s}^H(t_i), \\ \mathbf{H}(\theta) &= \mathbf{D}^H(\theta) \Pi_{\mathbf{A}}^{\perp}(\theta) \mathbf{D}(\theta), \\ \mathbf{D}(\theta_0) &= \begin{bmatrix} \frac{\partial \mathbf{a}(\theta)}{\partial \theta} \big|_{\theta_1} & \frac{\partial \mathbf{a}(\theta)}{\partial \theta} \big|_{\theta_2} & \cdots & \frac{\partial \mathbf{a}(\theta)}{\partial \theta} \big|_{\theta_M} \end{bmatrix}. \end{aligned} \quad (2.30)$$

On présentera dans la suite certains résultats concernant la relation d'ordre entre $\mathbf{BCRS}(\theta_0)$ et $\mathbf{BCRD}(\theta_0)$.

2. La relation d'ordre $\mathbf{A} \geq \mathbf{B}$ avec \mathbf{A} et \mathbf{B} appartenant à $\mathcal{M}_{Q \times Q}(\mathbb{R})$ signifie que la matrice $\mathbf{A} - \mathbf{B}$ est définie non-négative.

2.3.2 État de l'art

Nous présentons ici un certain nombre de résultats concernant les performances asymptotiques des estimateurs du maximum de vraisemblance.

- Les résultats asymptotiques les plus nombreux s'intéressent à un nombre d'observations T tendant vers l'infini (à nombre de capteur fini et à RSB fini). Dans ce contexte, l'estimateur du MVS est asymptotiquement gaussien et efficace [Leh83] [SN90b]. En d'autres termes

$$\boxed{\sqrt{T} \left(\hat{\boldsymbol{\theta}}_{MVS} - \boldsymbol{\theta}_0 \right) \overset{T \rightarrow \infty}{\underset{\sim}{\mathcal{N}}}_{\mathbb{R}} \left(\mathbf{0}, \mathbf{C}_{MVS} \right)}, \quad (2.31)$$

avec

$$\boxed{\frac{1}{T} \mathbf{C}_{MVS} = \mathbf{BCRS} \left(\boldsymbol{\theta}_0 \right)}. \quad (2.32)$$

Concernant l'estimateur du MVD dans le contexte d'un fort nombre d'observations, il a été établi qu'il était asymptotiquement gaussien mais non-efficace [SN90a] [VO91]. Il paraît en effet logique que l'estimation de la matrice des vecteurs signaux \mathbf{S}_T n'est pas consistante³ puisque le nombre d'inconnues augmente avec T . En d'autres termes

$$\boxed{\sqrt{T} \left(\hat{\boldsymbol{\theta}}_{MVD} - \boldsymbol{\theta}_0 \right) \overset{T \rightarrow \infty}{\underset{\sim}{\mathcal{N}}}_{\mathbb{R}} \left(\mathbf{0}, \mathbf{C}_{MVD} \right)}, \quad (2.33)$$

avec

$$\frac{1}{T} \mathbf{C}_{MVD} = \mathbf{BCRD} \left(\boldsymbol{\theta}_0 \right) + 2T \mathbf{BCRD} \left(\boldsymbol{\theta}_0 \right) \operatorname{Re} \left\{ \mathbf{H} \left(\boldsymbol{\theta}_0 \right) \odot \left(\mathbf{A}^T \left(\boldsymbol{\theta}_0 \right) \mathbf{A}^* \left(\boldsymbol{\theta}_0 \right) \right)^{-1} \right\} \mathbf{BCRD} \left(\boldsymbol{\theta}_0 \right), \quad (2.34)$$

$$\boxed{\frac{1}{T} \mathbf{C}_{MVD} > \mathbf{BCRD} \left(\boldsymbol{\theta}_0 \right)}. \quad (2.35)$$

Il a également été démontré que l'estimateur du MVD devenait efficace lorsque le nombre d'observations tendait vers l'infini et le nombre de capteurs tendait lui aussi vers l'infini (à RSB fini) [SN90b].

Concernant l'estimateur du MVP, dans le cas de signaux sources décorrélés, il a été établi qu'il était asymptotiquement (en nombre d'observations) gaussien et efficace [LC93] [LHSV95] [CM97].

Enfin, concernant les BCR et les matrices de covariance asymptotique (en nombre d'observations), on peut démontrer que [OVSN93]

$$\boxed{\frac{1}{T} \mathbf{C}_{MVD} \geq \frac{1}{T} \mathbf{C}_{MVS} = \mathbf{BCRS} \left(\boldsymbol{\theta}_0 \right) \geq \mathbf{BCRD} \left(\boldsymbol{\theta}_0 \right)}. \quad (2.36)$$

Il faut noter que l'inégalité de gauche provient de l'application de la méthode du maximum de vraisemblance déterministe sous l'hypothèse de signaux gaussiens.

- Lorsque le terme asymptotique correspond à un nombre de capteurs tendant vers l'infini (à nombre d'observations fini et à RSB fini), Viberg [VON95] a démontré la gaussianité et l'efficacité de l'estimateur du MVD et l'efficacité de l'estimateur du MVS.
- Lorsque le terme asymptotique correspond à un RSB tendant vers l'infini (à nombre d'observations fini et à nombre de capteurs fini), Forster, Boyer et Larzabal [Boy02] [FBL04] ont démontré que l'estimateur du MVS est non-gaussien et non-efficace à fort RSB pour un scénario à une seule source. Toujours dans le cas d'une seule source, Kay [Kay93] page 297 a démontré que l'estimateur du MVD était gaussien.

3. On peut néanmoins démontrer que les estimées de \mathbf{S}_T sont non-biaisées lorsque le nombre d'observations tend vers l'infini (à nombre de capteur fini et à RSB fini).

2.4 Performance asymptotique en terme de rapport signal sur bruit

Dans cette section, l'objectif est de caractériser les performances de $\hat{\boldsymbol{\theta}}_{MVD}$ et $\hat{\boldsymbol{\theta}}_{MVS}$ pour un nombre fini d'observations et de capteurs et un RSB infini et ainsi d'étendre les caractérisations asymptotiques rappelées ci-dessus.

Remarque concernant les notations : dans cette section, la convention suivante est adoptée : le vecteur (ou la matrice) quelconque noté $\mathbf{x}(\omega)$ est une réalisation d'un processus aléatoire noté \mathbf{x} au travers de l'événement ω appartenant à l'espace des événements Ω .

2.4.1 Maximum de vraisemblance déterministe

Nous formalisons et étendons ici le résultat obtenu par Kay dans [Kay93] page 297 démontrant que l'estimateur du MVD pour une seule source fonction d'un seul paramètre ($P = M = 1$) est gaussien à fort RSB pour un nombre fini d'échantillons et un nombre fini de capteurs. Ainsi, nous démontrons la gaussianité et l'efficacité asymptotique en RSB du MVD et du MVP quel que soit le nombre de sources et quelle que soit la géométrie de l'antenne (néanmoins non-ambïgue). Pour ce faire, nous utilisons un modèle plus général que le modèle du MVD et du MVP. Puis nous utilisons le théorème des fonctions implicites (voir par exemple, [Rud76] théorème 9.28) pour terminer notre démonstration.

Nous utiliserons ici un modèle d'observations gaussiennes à moyenne paramétrée plus général que le modèle (2.1) sous $H'1$,

$$\mathbf{y}(\omega) = \mathbf{m}(\mathbf{p}_0) + \mathbf{b}(\omega), \quad (2.37)$$

où $\mathbf{y}(\omega) \in \mathcal{M}_{N \times 1}(\mathbb{R})$ est une réalisation des observations, $\mathbf{p} = [p_1, \dots, p_P]^T \in \mathcal{M}_{P \times 1}(\mathbb{R})$ est le vecteur réel des P ($P \leq N$) paramètres d'intérêts déterministes mais inconnus avec pour vraie valeur le vecteur $\mathbf{p}_0 \in \mathcal{M}_{P \times 1}(\mathbb{R})$, $\mathbf{m} \in \mathcal{M}_{N \times 1}(\mathbb{R})$ est un vecteur réel déterministe dépendant généralement de manière non-linéaire de \mathbf{p} , $\mathbf{b}(\omega) \in \mathcal{M}_{N \times 1}(\mathbb{R})$ est une réalisation d'un bruit additif réel issu d'un processus gaussien \mathbf{b} de moyenne nulle et de matrice de covariance $\sigma^2 \mathbf{C}$. La matrice \mathbf{C} est supposée connue et σ^2 est supposé inconnu. En outre, on supposera le modèle $\mathbf{m}(\mathbf{p})$ identifiable, c'est-à-dire que $\mathbf{m}(\mathbf{p}_1) = \mathbf{m}(\mathbf{p}_2) \implies \mathbf{p}_1 = \mathbf{p}_2$. L'analyse à fort RSB reviendra à analyser le comportement statistique de $\hat{\mathbf{p}}$ lorsque $\sigma^2 \rightarrow 0$.

La fonction de log-vraisemblance $\ln f(\mathbf{y}(\omega) | \mathbf{p}, \sigma^2)$ des observations s'écrit :

$$\ln f(\mathbf{y}(\omega) | \mathbf{p}, \sigma^2) = -\frac{1}{2} \ln \left((2\pi)^N |\sigma^2 \mathbf{C}| \right) - \frac{1}{2\sigma^2} (\mathbf{y}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{y}(\omega) - \mathbf{m}(\mathbf{p})). \quad (2.38)$$

On remarquera que l'étude du modèle $\check{\mathbf{y}}(\omega) = \check{\mathbf{m}}(\mathbf{p}_0) + \check{\mathbf{b}}(\omega)$ où $\check{\mathbf{y}}(\omega)$, $\check{\mathbf{m}}(\mathbf{p}_0)$ et $\check{\mathbf{b}}(\omega)$ sont des vecteurs complexes peut être conduite, sans perte de généralité, avec le modèle réel (2.37). En effet, en concaténant les parties réelles et imaginaires de $\check{\mathbf{y}}(\omega)$, $\check{\mathbf{m}}(\mathbf{p}_0)$ et $\check{\mathbf{b}}(\omega)$, on obtient le modèle :

$$\begin{bmatrix} \text{Re} \{ \check{\mathbf{y}}(\omega) \} \\ \text{Im} \{ \check{\mathbf{y}}(\omega) \} \end{bmatrix} = \begin{bmatrix} \text{Re} \{ \check{\mathbf{m}}(\mathbf{p}_0) \} \\ \text{Im} \{ \check{\mathbf{m}}(\mathbf{p}_0) \} \end{bmatrix} + \begin{bmatrix} \text{Re} \{ \check{\mathbf{b}}(\omega) \} \\ \text{Im} \{ \check{\mathbf{b}}(\omega) \} \end{bmatrix}, \quad (2.39)$$

qui est similaire à (2.37). Les modifications concernant la fonction de log-vraisemblance des observations (2.38) sont triviales et conduisent à une matrice de covariance augmentée de taille $2N \times 2N$ incluant la possibilité d'un bruit circulaire.

Dans le contexte du traitement d'antenne, le modèle (2.37) rend compte des deux problèmes suivants :

– La méthode du MVD présentée section 2.2.2, en posant

$$\mathbf{m}(\mathbf{p}) = \left[\operatorname{Re} \{ \operatorname{vec}(\mathbf{A}(\boldsymbol{\theta}_0) \mathbf{S}_T) \}^T \quad \operatorname{Im} \{ \operatorname{vec}(\mathbf{A}(\boldsymbol{\theta}_0) \mathbf{S}_T) \}^T \right]^T, \quad (2.40)$$

où $\mathbf{p} = [\operatorname{vec}^T(\boldsymbol{\theta}) \quad \operatorname{vec}(\mathbf{S}_T)^T]^T$.

– Le modèle avec forme d'onde connue à la base de la méthode du MVP (voir [LC93]) (avec les notations de l'équation (8) de [LC93])

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}_0) \mathbf{P}(t) \boldsymbol{\alpha} + \mathbf{n}(t), \quad (2.41)$$

en posant

$$\mathbf{m}(\mathbf{p}) = \left[\operatorname{Re} \{ \operatorname{vec}(\mathbf{A}(\boldsymbol{\theta}_0) \mathbf{P}(t) \boldsymbol{\alpha}) \}^T \quad \operatorname{Im} \{ \operatorname{vec}(\mathbf{A}(\boldsymbol{\theta}_0) \mathbf{P}(t) \boldsymbol{\alpha}) \}^T \right]^T, \quad (2.42)$$

où $\mathbf{p} = [\boldsymbol{\theta}^T \quad \boldsymbol{\alpha}^T]^T$. Ici, $\mathbf{P}(t)$ représente la matrice des signaux de forme d'onde connue et $\boldsymbol{\alpha}$ est le vecteur inconnu des gains.

En conséquence, l'analyse statistique de $\hat{\mathbf{p}}$ pour le modèle (2.37) nous permettra de conclure sur les performances asymptotiques à fort RSB des estimateurs du MVD et du MVP.

2.4.1.1 Estimateur du maximum de vraisemblance

L'estimée au sens du maximum de vraisemblance du vecteur \mathbf{p} est donnée par :

$$\begin{cases} \hat{\mathbf{p}}(\omega) = \arg \min_{\mathbf{p}} J(\mathbf{p}), \\ J(\mathbf{p}) = (\mathbf{y}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{y}(\omega) - \mathbf{m}(\mathbf{p})). \end{cases} \quad (2.43)$$

La solution de l'équation (2.43) est obtenue par annulation du gradient du critère $J(\mathbf{p})$. Ceci conduit au système à P équations suivant (la dépendance par rapport à $\mathbf{b}(\omega)$ est laissée explicite)

$$\mathbf{g}(\mathbf{p}, \mathbf{b}(\omega))|_{\mathbf{p}=\hat{\mathbf{p}}(\omega)} = \mathbf{0}, \quad (2.44)$$

où les éléments du vecteur $\mathbf{g} = [g_1(\mathbf{p}, \mathbf{b}(\omega)) \quad \dots \quad g_P(\mathbf{p}, \mathbf{b}(\omega))]^T$, pour $i = 1, \dots, P$, sont donnés par :

$$\begin{aligned} g_i(\mathbf{p}, \mathbf{b}(\omega)) &= (\mathbf{y}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} \frac{\partial \mathbf{m}(\mathbf{p})}{\partial p_i} \\ &= (\mathbf{m}(\mathbf{p}_0) - \mathbf{m}(\mathbf{p}) + \mathbf{b}(\omega))^T \mathbf{C}^{-1} \frac{\partial \mathbf{m}(\mathbf{p})}{\partial p_i}. \end{aligned} \quad (2.45)$$

2.4.1.2 Rappel : le théorème des fonctions implicites

Soit $\boldsymbol{\vartheta}(\mathbf{u}, \mathbf{v}) = [\vartheta_1(\mathbf{u}, \mathbf{v}) \quad \dots \quad \vartheta_P(\mathbf{u}, \mathbf{v})]^T$ une fonction de $\mathbb{R}^P \times \mathbb{R}^N \rightarrow \mathbb{R}^P$. Nous supposons les hypothèses suivantes vérifiées

H1_{TFI} $\vartheta_i(\mathbf{u}, \mathbf{v})$ pour $i = 1, \dots, P$ est une fonction différentiable sur un voisinage de $(\mathbf{u}_0, \mathbf{v}_0)$ dans $\mathbb{R}^P \times \mathbb{R}^N$.

H2_{TFI} $\vartheta_i(\mathbf{u}_0, \mathbf{v}_0) = 0$ pour $i = 1, \dots, P$.

H3_{TFI} la matrice Jacobienne $\boldsymbol{\Phi} \in \mathcal{M}_{P \times P}(\mathbb{R})$ de $\vartheta_i(\mathbf{u}, \mathbf{v})$ par rapport à \mathbf{u} est supposée non-singulière en $(\mathbf{u}_0, \mathbf{v}_0)$.

Alors, il existe un voisinage V du point \mathbf{v}_0 dans \mathbb{R}^N , il existe un voisinage U du point \mathbf{u}_0 dans \mathbb{R}^P , et il existe une unique fonction $\varphi : V \rightarrow U$ telle que $\varphi(\mathbf{v}_0) = \mathbf{u}_0$ et $\vartheta_i(\varphi(\mathbf{v}), \mathbf{v}) = 0$ pour $i = 1, \dots, P$ et pour tout \mathbf{v} dans V . De plus, φ est différentiable en \mathbf{v}_0 et on a

$$\varphi(\mathbf{v}) - \mathbf{u}_0 = -\mathbf{\Phi}^{-1}\mathbf{\Psi}(\mathbf{v} - \mathbf{v}_0) + \mathbf{r}(\mathbf{v} - \mathbf{v}_0), \quad (2.46)$$

où le reste $\mathbf{r}(\mathbf{v} - \mathbf{v}_0) = o(\|\mathbf{v} - \mathbf{v}_0\|)$ et, où

$$\mathbf{\Phi} = \left[\frac{\partial \vartheta}{\partial u_1} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)}, \dots, \frac{\partial \vartheta}{\partial u_P} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} \right], \quad \mathbf{\Phi} \in \mathcal{M}_{P \times P}(\mathbb{R}), \quad (2.47)$$

$$\mathbf{\Psi} = \left[\frac{\partial \vartheta}{\partial v_1} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)}, \dots, \frac{\partial \vartheta}{\partial v_N} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} \right], \quad \mathbf{\Psi} \in \mathcal{M}_{P \times N}(\mathbb{R}), \quad (2.48)$$

en posant $\mathbf{u} = [u_1 \ \dots \ u_P]^T$ et $\mathbf{v} = [v_1 \ \dots \ v_N]^T$.

2.4.1.3 Application au gradient du critère du maximum de vraisemblance

Nous allons maintenant détailler la structure des matrices $\mathbf{\Phi}$ et $\mathbf{\Psi}$ dans le cadre de la fonction \mathbf{g} (2.45). Dans la suite, nous verrons que le vecteur \mathbf{u} est connecté aux paramètres \mathbf{p} et que le vecteur \mathbf{v} est connecté au vecteur du bruit \mathbf{b} . Concernant les éléments de $\mathbf{\Psi}$, on a

$$\Psi_{i,j} = \frac{\partial g_i(\mathbf{u}, \mathbf{v})}{\partial v_j} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} = \frac{\partial \mathbf{v}^T}{\partial v_j} \mathbf{C}^{-1} \frac{\partial \mathbf{m}}{\partial u_i} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)}, \quad 1 \leq i \leq P \text{ et } 1 \leq j \leq N. \quad (2.49)$$

Concernant les éléments de $\mathbf{\Phi}$, on a

$$\Phi_{i,j} = \frac{\partial g_i(\mathbf{u}, \mathbf{v})}{\partial u_j} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} = - \frac{\partial \mathbf{m}^T}{\partial u_j} \mathbf{C}^{-1} \frac{\partial \mathbf{m}}{\partial u_i} \Big|_{\mathbf{u}_0}, \quad 1 \leq i, j \leq P. \quad (2.50)$$

En d'autres mots,

$$\mathbf{\Psi} = \frac{\partial \mathbf{m}^T}{\partial \mathbf{u}} \mathbf{C}^{-1} \frac{\partial \mathbf{v}}{\partial \mathbf{v}^T} \text{ et } \mathbf{\Phi} = - \frac{\partial \mathbf{m}^T}{\partial \mathbf{u}} \mathbf{C}^{-1} \frac{\partial \mathbf{m}}{\partial \mathbf{u}^T}. \quad (2.51)$$

On remarquera que les matrices $\mathbf{\Phi}$ et $\mathbf{\Psi}$ sont reliées par la relation suivante

$$\mathbf{\Phi} = -\mathbf{\Psi} \mathbf{C} \mathbf{\Psi}^T, \quad (2.52)$$

puisque $\frac{\partial \mathbf{v}}{\partial \mathbf{v}^T} = \mathbf{I}_N$.

Appliquons maintenant le théorème des fonctions implicites à \mathbf{g} en posant $\mathbf{v} = \mathbf{b}(\omega)$, $\varphi(\mathbf{v}) = \hat{\mathbf{p}}(\omega)$, $\mathbf{u}_0 = \mathbf{p}_0$ et $\mathbf{v}_0 = \mathbf{0}$. On notera que l'hypothèse $H2_{TFI}$ est satisfaite puisque, dans le cas où le bruit est nul, le critère $(\mathbf{y}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{y}(\omega) - \mathbf{m}(\mathbf{p}))$ est minimum et égal à $\mathbf{0}$ pour $\mathbf{p} = \mathbf{p}_0$. Les hypothèses $H1_{TFI}$ et $H3_{TFI}$ seront supposées satisfaites dans la suite.

Le théorème des fonctions implicites conduit à

$$\hat{\mathbf{p}}(\omega) - \mathbf{p}_0 = -\mathbf{\Phi}^{-1}\mathbf{\Psi}\mathbf{b}(\omega) + \mathbf{r}(\mathbf{b}(\omega)), \quad \forall \omega \in \Omega. \quad (2.53)$$

2.4.1.4 Gaussianité asymptotique en RSB de l'estimateur du maximum de vraisemblance

Soit \mathbf{b}_k une suite de vecteurs de bruit aléatoires gaussiens de moyennes nulles et de matrices de covariance $\sigma_k^2 \mathbf{C}$. σ_k est une suite telle que $\sigma_k \rightarrow 0$ quand $k \rightarrow \infty$. Soient $\hat{\mathbf{p}}_k(\omega)$ et $\mathbf{b}_k(\omega)$ les réalisations de deux suites de vecteurs aléatoires $\hat{\mathbf{p}}_k$ et \mathbf{b}_k . $\hat{\mathbf{p}}_k(\omega)$ est la solution de (2.44). Soit $\Delta \mathbf{p}_k(\omega) = \frac{1}{\sigma_k} (\hat{\mathbf{p}}_k(\omega) - \mathbf{p}_0)$ la réalisation d'une suite de vecteurs aléatoires $\Delta \mathbf{p}_k$. Le comportement de l'estimateur du maximum de vraisemblance à fort RSB est alors donné par le comportement du vecteur aléatoire $\Delta \mathbf{p}_k$ lorsque $k \rightarrow \infty$. De (2.53) on obtient

$$\Delta \mathbf{p}_k(\omega) = -\Phi^{-1} \Psi \frac{\mathbf{b}_k(\omega)}{\sigma_k} + \frac{\mathbf{r}(\mathbf{b}_k(\omega))}{\sigma_k}, \quad \forall \omega \in \Omega, \quad (2.54)$$

où de manière équivalente en termes de vecteurs aléatoires \mathbf{b}_k and $\Delta \mathbf{p}_k$

$$\Delta \mathbf{p}_k = -\Phi^{-1} \Psi \frac{\mathbf{b}_k}{\sigma_k} + \frac{\mathbf{r}(\mathbf{b}_k)}{\sigma_k}. \quad (2.55)$$

Le reste $\mathbf{r}(\mathbf{b}_k)$ est donné par le lemme 2.12 (i) de [vdV98]

$$\mathbf{r}(\mathbf{b}_k) = o_p(\|\mathbf{b}_k\|). \quad (2.56)$$

Il s'agit maintenant d'étudier les deux termes du membre de droite de (2.55). Pour cela, on notera que le terme $\frac{\mathbf{b}_k}{\sigma_k}$ est distribué selon une loi gaussienne à moyenne nulle et de matrice de covariance \mathbf{C} . Le terme $\frac{\mathbf{r}(\mathbf{b}_k)}{\sigma_k} = \frac{o_p(\|\mathbf{b}_k\|)}{\sigma_k}$ peut être réécrit comme

$$\frac{o_p(\|\mathbf{b}_k\|)}{\sigma_k} = \frac{\|\mathbf{b}_k\|}{\sigma_k} \frac{o_p(\|\mathbf{b}_k\|)}{\|\mathbf{b}_k\|}, \quad (2.57)$$

où, par définition, $\frac{o_p(\|\mathbf{b}_k\|)}{\|\mathbf{b}_k\|}$ converge en probabilité vers zéro et où $\frac{\|\mathbf{b}_k\|}{\sigma_k}$ suit une loi du chi qui ne dépend pas de k . En conséquence, $\frac{o_p(\|\mathbf{b}_k\|)}{\sigma_k}$ converge en probabilité vers zéro ([Rao02] page 122).

Concernant le terme $-\Phi^{-1} \Psi \frac{\mathbf{b}_k}{\sigma_k}$, il suit une distribution gaussienne de moyenne nulle et de matrice de covariance $\mathbf{\Gamma}$ égale à

$$\begin{aligned} \mathbf{\Gamma} &= \Phi^{-1} \Psi \mathbf{C} (\Phi^{-1} \Psi)^T \\ &= \Phi^{-1} \Psi \mathbf{C} \Psi^T (\Phi^T)^{-1} \\ &= -\Phi^{-1} = (\Psi \mathbf{C} \Psi^T)^{-1}, \end{aligned} \quad (2.58)$$

grâce à (2.52). Sachant que la somme d'une variable aléatoire gaussienne et d'une variable aléatoire convergeant en probabilité vers 0 est une variable aléatoire gaussienne [Rao02], $\Delta \mathbf{p}_k$ converge en distribution vers un vecteur aléatoire gaussien de moyenne nulle et de matrice de covariance $\mathbf{\Gamma} = (\Psi \mathbf{C} \Psi^T)^{-1}$ lorsque $k \rightarrow \infty$ c'est-à-dire lorsque $\sigma_k \rightarrow 0$. Ceci démontre la gaussianité asymptotique en RSB de l'estimateur du maximum de vraisemblance pour un modèle d'observations à moyenne paramétrée, et, de ce fait la gaussianité asymptotique en RSB du MVD et du MVP quel que soit le nombre de sources. Ces conclusions sont résumées dans le théorème suivant.

Théorème 2.1 :

$$\Delta \boldsymbol{\theta}_{MVD} \stackrel{RSB \rightarrow \infty}{\underset{\sim}{\mathcal{N}}}_{\mathbb{R}}(\mathbf{0}, \mathbf{\Gamma}). \quad (2.59)$$

2.4.1.5 Efficacité asymptotique en RSB de l'estimateur du maximum de vraisemblance

La forme analytique de la matrice de covariance $\mathbf{\Gamma}$ obtenue à l'aide du théorème des fonctions implicites va nous permettre d'établir l'efficacité asymptotique en RSB de l'estimateur du maximum de vraisemblance par une comparaison directe avec la BCR. Regardons de plus près la structure de la MIF qui se réduit dans ce cas à une matrice bloc diagonale grâce au découplage entre les paramètres \mathbf{p} et σ^2 [ZN90]

$$\mathbf{MIF}(\mathbf{p}_0, \sigma^2) = \begin{pmatrix} \mathbf{MIF}(\mathbf{p}_0) & \mathbf{0} \\ \mathbf{0} & \mathbf{MIF}(\sigma^2) \end{pmatrix}. \quad (2.60)$$

Concernant \mathbf{p}_0 la BCR est donnée par

$$\begin{aligned} \mathbf{BCR}(\mathbf{p}_0) &= \mathbf{MIF}^{-1}(\mathbf{p}_0) \\ &= - \left(\mathbb{E} \left[\frac{\partial^2 \ln f(\mathbf{y} | \mathbf{p}, \sigma^2)}{\partial \mathbf{p} \partial \mathbf{p}^T} \right] \right)^{-1} \\ &= 2\sigma^2 \left(\mathbb{E} \left[\frac{\partial^2 (\mathbf{y} - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{y} - \mathbf{m}(\mathbf{p}))}{\partial \mathbf{p} \partial \mathbf{p}^T} \right] \right)^{-1} \\ &= -\sigma^2 \left(\mathbb{E} \left[\frac{\partial \mathbf{g}(\mathbf{p}, \mathbf{n})}{\partial \mathbf{p}^T} \right] \right)^{-1} \\ &= -\sigma^2 \mathbf{\Phi}^{-1} = \sigma^2 (\mathbf{\Psi} \mathbf{C} \mathbf{\Psi}^T)^{-1}. \end{aligned} \quad (2.61)$$

L'égalité des expressions (2.58) et (2.61) démontre le théorème suivant.

Théorème 2.2 : Les estimateurs du MVD et du MVP sont asymptotiquement efficaces en RSB quel que soit le nombre de sources

$$\boxed{\frac{1}{\sigma^2} \mathbf{BCR}(\mathbf{p}_0) = \mathbf{\Gamma}(\hat{\mathbf{p}})}. \quad (2.62)$$

2.4.1.6 Simulations

Afin d'illustrer cette analyse, nous présentons maintenant quelques résultats de simulations dans le contexte de l'estimation de directions d'arrivée par la méthode du MVD décrit section (2.2.2). La BCR est donnée par (2.29). Dans toute cette section, nous considérons une antenne linéaire uniforme (ALU) composée de $N = 10$ capteurs espacés d'une demi longueur d'onde. Les directions d'arrivées sont données par rapport au travers de l'antenne. $T = 10$ observations. L'algorithme du MVD est mis en oeuvre à l'aide d'un algorithme de Gauss-Newton précédé d'une recherche grossière sur une grille de valeurs.

Efficacité

On considère le cas de deux sources de même puissance situées en $\boldsymbol{\theta}_0 = [0^\circ \quad 3^\circ]$, c'est-à-dire dans le lobe principal de l'antenne de largeur 10 degrés. Les simulations de Monte-Carlo sont conduites à l'aide de $r = 1000$ réalisations indépendantes. La figure 2.3 donne le comportement de la variance empirique de l'estimateur du MVD et de la BCR associée pour les deux sources en fonction du RSB. On observe l'efficacité asymptotique en RSB.

On observe également le phénomène de décrochement évoqué au chapitre 1, lorsque le RSB devient faible (environ 5dB dans ce cas). Ce phénomène donne la plage en dehors de laquelle le terme "asymptotique en RSB" n'est plus valable (on pourra se reporter à [Ath05] où une analyse de la prédiction du décrochement est donnée pour ce type de scénario).

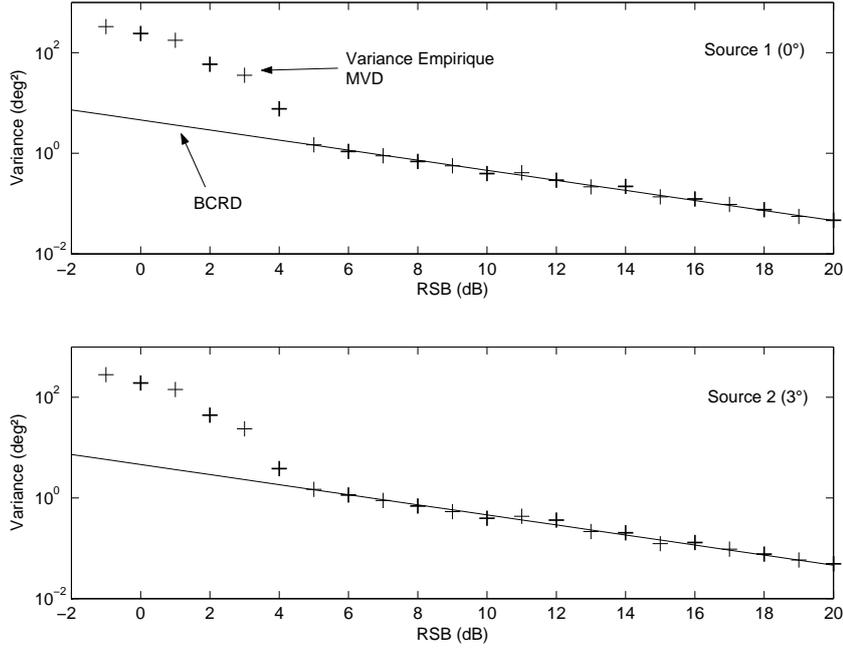


FIGURE 2.3 – MVD. Variance empirique de l'estimateur, BCR déterministe en fonction du RSB. $\theta_0 = [0^\circ \ 3^\circ]$. $T = 10$ observations. $N = 10$ capteurs. $r = 1000$ réalisations.

Gaussianité

La figure 2.4 représente l'histogramme des estimées pour un RSB égal à 30dB et une comparaison avec une distribution gaussienne de moyenne θ_0 et de matrice de covariance $\mathbf{BCR}(\mathbf{p}_0)$. Les simulations de Monte-Carlo sont conduites à l'aide de $r = 10000$ réalisations indépendantes. On constate une bonne adéquation entre les estimées et la distribution théorique.

Pour confirmer ce résultat "visuel", nous avons réalisé un test d'hypothèses de Lilliefors [Lil67]. Ce test évalue l'hypothèse de gaussianité d'un échantillon $[x_1 \cdots x_L]$. Ce test est proche du test de Kolmogorov-Smirnov, mais la moyenne et la variance *a priori* n'ont pas besoin d'être spécifiées au préalable.

La quantité T_{stat} issue du test de Lilliefors est la distance verticale maximale entre la fonction de répartition empirique $F(s)$ de la série

$$z_k = \frac{x_k - \frac{1}{L} \sum_{i=1}^L x_i}{\sqrt{\frac{1}{L} \sum_{j=1}^L \left(x_j - \frac{1}{L} \sum_{i=1}^L x_i \right)^2}}, \quad (2.63)$$

et la fonction de répartition $Q(s)$ d'une distribution gaussienne à moyenne nulle et de variance 1. En d'autres mots

$$T_{stat} = \sup_{s \in \mathbb{R}} |F(s) - Q(s)|. \quad (2.64)$$

L'hypothèse de gaussianité est rejetée pour un indice de confiance α (égale à 5% dans la suite) si T dépasse le quantile $1 - \alpha$ disponible dans les tables de quantiles de Lilliefors [Con80] égal à $\frac{0.886}{\sqrt{r}}$ pour $r > 30$. Ici, cette valeur est de $\frac{0.886}{\sqrt{r}} = 0.028$ puisque les simulations de Monte-Carlo sont conduites à l'aide de $r = 1000$ réalisations indépendantes.

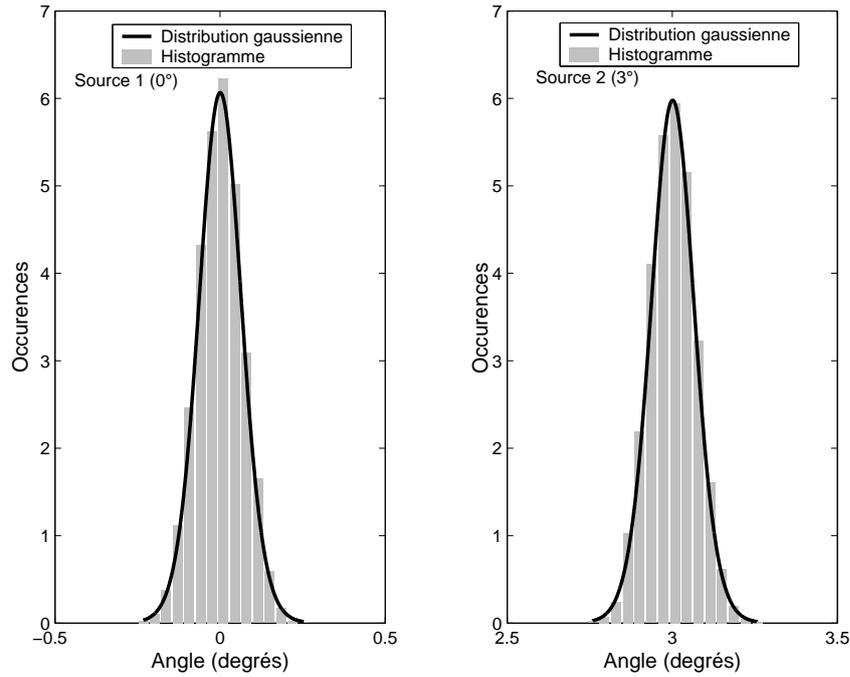


FIGURE 2.4 – MVD. Histogramme des estimées. $\theta_0 = [0^\circ \ 3^\circ]$. $RSB = 30dB$. $T = 10$ observations. $N = 10$ capteurs. $r = 10000$ réalisations.

Le test est réalisé dans le cadre du scénario susmentionné (deux sources de même puissance situées en $\theta_0 = [0^\circ \ 3^\circ]$, 10 observations, 10 capteurs, 1000 réalisations de Monte-Carlo). La figure 2.5 montre le comportement de T_{stat} en fonction du RSB.

La gaussianité de l'estimateur du MVD est confirmée pour les RSB supérieurs à approximativement 9 dB.

2.4.2 Maximum de vraisemblance stochastique

Dans cette section nous étendons les résultats obtenus par Forster, Boyer et Larzabal dans [Boy02] [FBL04] démontrant que l'estimateur du MVS est non-gaussien et non-efficace à fort RSB pour une seule source. Nous démontrons sa non-gaussianité et sa non-efficacité asymptotique en RSB quel que soit le nombre de sources et quelle que soit la géométrie de l'antenne (néanmoins non-ambigue). Nous établissons tout d'abord une forme d'équivalence entre les critères du MVS et du MVD pour des signaux stochastiques. Puis à partir de cette équivalence nous en déduisons la non-gaussianité du MVS. Finalement, nous établissons une extension matricielle du théorème de Jensen pour démontrer la non-efficacité.

On supposera dans toute cette section que $P = 1$, c'est-à-dire qu'il n'y a qu'un seul paramètre angulaire à estimer par source.

2.4.2.1 Cas multi-sources

Equivalence des critères du MVS et du MVD à fort RSB

Rappelons les critères $C_{MVS}(\theta)$ et $C_{MVD}(\theta)$ à minimiser pour obtenir les estimateurs du

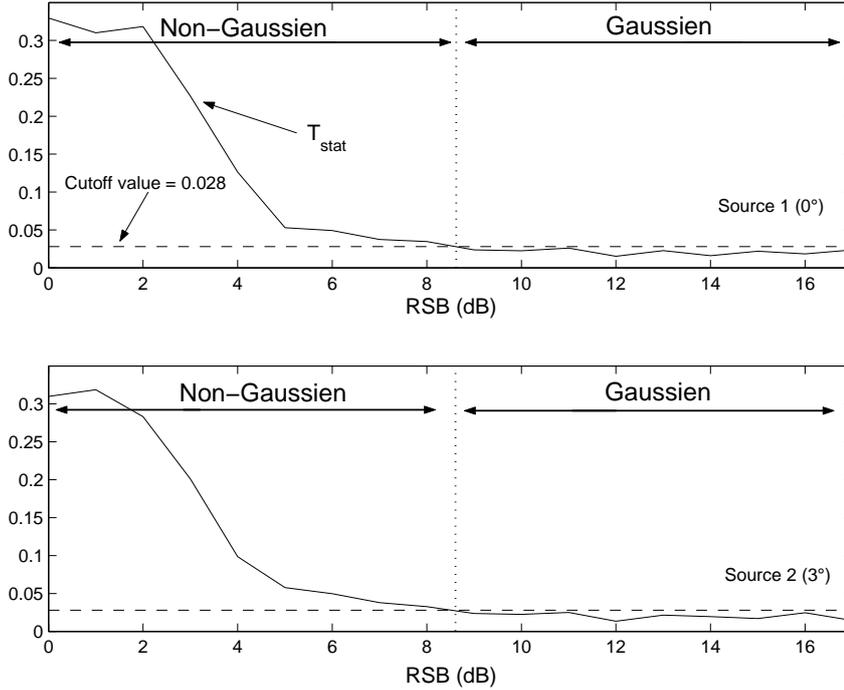


FIGURE 2.5 – MVD. Test statistique de Lilliefors : T_{stat} et quantile $1 - \alpha$ en fonction du RSB. $\theta_0 = [0^\circ \ 3^\circ]$, $T = 10$ observations, $N = 10$ capteurs, $r = 1000$ réalisations.

MVS et du MVD (équations (2.14), (2.26), (2.15) et (2.16))

$$C_{MVS}(\theta) = \left| \mathbf{A}(\theta) \hat{\Sigma}_s(\theta) \mathbf{A}^H(\theta) + \hat{\sigma}^2(\theta) \mathbf{I}_N \right|, \quad (2.65)$$

$$C_{MVD}(\theta) = \frac{1}{N-M} \text{tr} \left(\mathbf{\Pi}_A^\perp(\theta) \hat{\Sigma}_y \right). \quad (2.66)$$

En introduisant (2.66) dans (2.65) on obtient

$$C_{MVS}(\theta) = \left| \mathbf{\Pi}_A(\theta) \hat{\Sigma}_y \mathbf{\Pi}_A(\theta) + C_{MVD}(\theta) \mathbf{\Pi}_A^\perp(\theta) \right|, \quad (2.67)$$

où $\mathbf{\Pi}_A(\theta) = \mathbf{A}(\theta) \mathbf{A}^\dagger(\theta)$ est le projecteur sur l'espace engendré par $\mathbf{A}(\theta)$.

Proposition 2.1 : à fort RSB, les directions d'arrivées $\hat{\theta}_{MVS}$ et $\hat{\theta}_{MVD}$ obtenues par minimisation de $C_{MVS}(\theta)$ et $C_{MVD}(\theta)$ sous $H1$ sont les mêmes.

Preuve : soit $\mathbf{E}_s(\theta) \in \mathcal{M}_{N \times M}(\mathbb{C})$ et $\mathbf{E}_n(\theta) \in \mathcal{M}_{N \times (N-M)}(\mathbb{C})$ les matrices construites avec les bases orthonormées des sous-espaces signal et bruit et soit $\mathbf{E}(\theta) = [\mathbf{E}_s(\theta) \ \mathbf{E}_n(\theta)] \in \mathcal{M}_{N \times N}(\mathbb{C})$. L'équation (2.67) devient :

$$\begin{aligned} C_{MVS}(\theta) &= \left| \mathbf{E}^H(\theta) \left(\mathbf{\Pi}_A(\theta) \hat{\Sigma}_y \mathbf{\Pi}_A(\theta) + C_{MVD}(\theta) \mathbf{\Pi}_A^\perp(\theta) \right) \mathbf{E}(\theta) \right| \\ &= \left| \begin{array}{cc} \mathbf{E}_s^H(\theta) \hat{\Sigma}_y \mathbf{E}_s(\theta) & \mathbf{0} \\ \mathbf{0} & C_{MVD}(\theta) \mathbf{I}_{N-M} \end{array} \right| \\ &= \left| \mathbf{E}_s^H(\theta) \hat{\Sigma}_y \mathbf{E}_s(\theta) \right| (C_{MVD}(\theta))^{N-M}, \end{aligned} \quad (2.68)$$

puisque $\mathbf{E}_s^H(\boldsymbol{\theta}) \boldsymbol{\Pi}_A(\boldsymbol{\theta}) = \mathbf{I}_N$, $\mathbf{E}_n^H(\boldsymbol{\theta}) \boldsymbol{\Pi}_A^\perp(\boldsymbol{\theta}) = \mathbf{I}_{N-M}$, $\mathbf{E}_s^H(\boldsymbol{\theta}) \boldsymbol{\Pi}_A^\perp(\boldsymbol{\theta}) = \mathbf{0}$ et $\mathbf{E}_n^H(\boldsymbol{\theta}) \boldsymbol{\Pi}_A(\boldsymbol{\theta}) = \mathbf{0}$.

On peut remarquer que la minimisation du critère $C_{MVS}(\boldsymbol{\theta})$ est équivalente à la minimisation de $\tilde{C}_{MVS}(\boldsymbol{\theta}) = (C_{MVS}(\boldsymbol{\theta}))^{\frac{1}{N-M}}$. En conséquence, nous allons étudier

$$\tilde{C}_{MVS}(\boldsymbol{\theta}) = \left| \mathbf{E}_s^H(\boldsymbol{\theta}) \widehat{\boldsymbol{\Sigma}}_y \mathbf{E}_s(\boldsymbol{\theta}) \right|^{\frac{1}{N-M}} C_{MVD}(\boldsymbol{\theta}) \quad (2.69)$$

Soit $\alpha(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) = \left| \mathbf{E}_s^H(\boldsymbol{\theta}) \widehat{\boldsymbol{\Sigma}}_y(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) \mathbf{E}_s(\boldsymbol{\theta}) \right|^{\frac{1}{N-M}}$ où nous écrirons explicitement la dépendance de chaque terme en la réalisation du bruit $\mathbf{B}_T(\omega)$. L'étude du MVS à fort RSB repose sur le comportement de $\tilde{C}_{MVS}(\boldsymbol{\theta})$ en $(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) = (\boldsymbol{\theta}_0, \mathbf{0})$. Un développement limité à l'ordre zéro de $\alpha(\boldsymbol{\theta}, \mathbf{B}_T(\omega))$ en $(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) = (\boldsymbol{\theta}_0, \mathbf{0})$ conduit à

$$\alpha(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) = \alpha(\boldsymbol{\theta}_0, \mathbf{0}) + o(1), \quad (2.70)$$

où

$$\alpha(\boldsymbol{\theta}_0, \mathbf{0}) = \left| \mathbf{E}_s^H(\boldsymbol{\theta}_0) \mathbf{A}(\boldsymbol{\theta}_0) \widehat{\boldsymbol{\Sigma}}_s \mathbf{A}^H(\boldsymbol{\theta}_0) \mathbf{E}_s(\boldsymbol{\theta}_0) \right|^{\frac{1}{N-M}} \neq 0. \quad (2.71)$$

En conséquence, le premier terme non nul du développement limité de $\alpha(\boldsymbol{\theta}, \mathbf{b}(\omega))$ en $(\boldsymbol{\theta}_0, \mathbf{0})$ est $\alpha(\boldsymbol{\theta}_0, \mathbf{0})$. Un développement limité à l'ordre deux de $C_{MVD}(\boldsymbol{\theta}, \mathbf{B}_T(\omega))$ en $(\boldsymbol{\theta}_0, \mathbf{0})$ conduit à

$$C_{MVD}(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) = C_{MVD}(\boldsymbol{\theta}_0, \mathbf{0}) + \boldsymbol{\Delta}^T \mathbf{G} + \frac{1}{2} \boldsymbol{\Delta}^T \ddot{\mathbf{H}} \boldsymbol{\Delta} + o(\|\boldsymbol{\Delta}\|^2), \quad (2.72)$$

où

$$\boldsymbol{\Delta} = \left[(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T, \text{vec}(\text{Re}\{\mathbf{B}_T(\omega)\})^T, \text{vec}(\text{Im}\{\mathbf{B}_T(\omega)\})^T \right]^T, \quad (2.73)$$

avec \mathbf{G} le gradient de $C_{MVD}(\boldsymbol{\theta}, \mathbf{B}_T(\omega))$ en $(\boldsymbol{\theta}_0, \mathbf{0})$ et où $\ddot{\mathbf{H}}$ est le Hessien de $C_{MVD}(\boldsymbol{\theta}, \mathbf{B}_T(\omega))$ en $(\boldsymbol{\theta}_0, \mathbf{0})$.

Le critère $C_{MVD}(\boldsymbol{\theta}, \mathbf{B}_T(\omega))$ est minimal et nul en $(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) = (\boldsymbol{\theta}_0, \mathbf{0})$. En conséquence,

$$C_{MVD}(\boldsymbol{\theta}_0, \mathbf{0}) = 0 \text{ et } \mathbf{G} = \mathbf{0}, \quad (2.74)$$

et

$$C_{MVD}(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) = \frac{1}{2} \boldsymbol{\Delta}^T \ddot{\mathbf{H}} \boldsymbol{\Delta} + o(\|\boldsymbol{\Delta}\|^2). \quad (2.75)$$

D'où

$$\begin{aligned} \tilde{C}_{MVS}(\boldsymbol{\theta}) &= \frac{1}{2} \alpha(\boldsymbol{\theta}_0, \mathbf{0}) \boldsymbol{\Delta}^T \ddot{\mathbf{H}} \boldsymbol{\Delta} + o(\|\boldsymbol{\Delta}\|^2) \\ &= \alpha(\boldsymbol{\theta}_0, \mathbf{0}) C_{MVD}(\boldsymbol{\theta}, \mathbf{B}_T(\omega)) + o(\|\boldsymbol{\Delta}\|^2), \end{aligned} \quad (2.76)$$

démontrant ainsi l'équivalence à fort RSB de $C_{MVS}(\boldsymbol{\theta})$ et $C_{MVD}(\boldsymbol{\theta})$. ■

Cette proposition nous servira à démontrer la non-gaussianité asymptotique en RSB du MVS. Néanmoins, on peut noter que ce résultat est d'ores et déjà intéressant puisque la charge de calcul du MVD est plus faible que celle du MVS.

Non-gaussianité asymptotique en RSB du MVS

Soit $\Delta\boldsymbol{\theta}_{MVS} = \frac{1}{\sigma} \left(\hat{\boldsymbol{\theta}}_{MVS} - \boldsymbol{\theta}_0 \right)$.

Théorème 2.3 : à fort RSB ($\sigma \rightarrow 0$), $\Delta\boldsymbol{\theta}_{MVS}$ est distribué comme le vecteur $\mathbf{C}(\boldsymbol{\theta}_0) \mathbf{q}$, où $\mathbf{q} \in \mathcal{M}_{M \times 1}(\mathbb{R})$ est un vecteur aléatoire gaussien réel de moyenne nulle et de matrice de covariance \mathbf{I}_M et $\mathbf{C} \in \mathcal{M}_{M \times M}(\mathbb{R})$ est une matrice aléatoire quelconque indépendante de \mathbf{q} satisfaisant

$$\mathbf{C}(\boldsymbol{\theta}_0) \mathbf{C}^T(\boldsymbol{\theta}_0) = \frac{1}{2T} \left(\text{Re} \left\{ \mathbf{H}(\boldsymbol{\theta}_0) \odot \hat{\boldsymbol{\Sigma}}_s^T \right\} \right)^{-1}, \quad (2.77)$$

où $\mathbf{H} \in \mathcal{M}_{M \times M}(\mathbb{C})$ est une matrice déterministe qui dépend des directions d'arrivées $\boldsymbol{\theta}_0$ et de la géométrie de l'antenne définie par (2.30) :

$$\mathbf{H}(\boldsymbol{\theta}_0) = \mathbf{D}^H(\boldsymbol{\theta}_0) \boldsymbol{\Pi}_A^\perp(\boldsymbol{\theta}_0) \mathbf{D}(\boldsymbol{\theta}_0), \quad (2.78)$$

$$\text{avec } \mathbf{D}(\boldsymbol{\theta}_0) = \begin{bmatrix} \left. \frac{\partial \mathbf{a}(\theta)}{\partial \theta} \right|_{\theta=\theta_1} & \left. \frac{\partial \mathbf{a}(\theta)}{\partial \theta} \right|_{\theta=\theta_2} & \dots & \left. \frac{\partial \mathbf{a}(\theta)}{\partial \theta} \right|_{\theta=\theta_M} \end{bmatrix}.$$

Preuve : Grâce à la proposition 2.1, on peut considérer que $\Delta\boldsymbol{\theta}_{MVS}$ est obtenu à partir de la minimisation de $C_{MVD}(\boldsymbol{\theta})$ bien que les signaux sources soient supposés gaussiens. Si l'on se focalise sur la densité de probabilité de $\Delta\boldsymbol{\theta}_{MVS}$ conditionnée aux signaux sources $f(\Delta\boldsymbol{\theta}_{MVS} | \mathbf{S}_T)$, d'après les théorèmes 2.1 et 2.2, cette densité de probabilité est gaussienne de covariance asymptotique en RSB donnée par la BCR déterministe

$$\mathbf{BCRD}(\boldsymbol{\theta}_0) = \frac{1}{2T} \left(\text{Re} \left\{ \mathbf{H}(\boldsymbol{\theta}_0) \odot \hat{\boldsymbol{\Sigma}}_s^T \right\} \right)^{-1}. \quad (2.79)$$

Soit la décomposition de Cholesky de cette covariance asymptotique en RSB : $\mathbf{BCRD}(\boldsymbol{\theta}_0) = \mathbf{C}(\boldsymbol{\theta}_0) \mathbf{C}^T(\boldsymbol{\theta}_0)$. Donc, $f(\Delta\boldsymbol{\theta}_{MVS} | \mathbf{S}_T)$ est la même distribution que celle du vecteur $\mathbf{C}(\boldsymbol{\theta}_0) \mathbf{q}$, où $\mathbf{q} \in \mathcal{M}_{M \times 1}(\mathbb{R})$ est un vecteur aléatoire gaussien de moyenne nulle et de matrice de covariance \mathbf{I}_M . En conséquence, la distribution marginale asymptotique en RSB $f(\Delta\boldsymbol{\theta}_{MVS})$ est la même que celle du vecteur $\mathbf{C}(\boldsymbol{\theta}_0) \mathbf{q}$, où $\mathbf{q} \in \mathcal{M}_{M \times 1}(\mathbb{R})$ est un vecteur aléatoire gaussien de moyenne nulle et de matrice de covariance \mathbf{I}_M et où $\mathbf{C}(\boldsymbol{\theta}_0)$ devient une matrice aléatoire puisque $T\hat{\boldsymbol{\Sigma}}_s \in \mathcal{M}_{M \times M}(\mathbb{C})$ est une matrice aléatoire suivant une distribution de Wishart complexe à T degrés de libertés avec comme matrice paramètre la matrice de covariance des signaux sources $\boldsymbol{\Sigma}_s$. ■

Corollaire 2.1 : à fort RSB l'estimateur du MVS est non-gaussien quel que soit le nombre de sources.

Preuve : d'après le théorème 2.3, puisque $\mathbf{C}(\boldsymbol{\theta}_0)$ est une matrice aléatoire (au travers de $\hat{\boldsymbol{\Sigma}}_s$) le produit $\mathbf{C}(\boldsymbol{\theta}_0) \mathbf{q}$ ne peut avoir une distribution gaussienne. ■

Non-efficacité asymptotique en RSB du MVS

Afin de démontrer la non-efficacité asymptotique en RSB du MVS, nous allons comparer la covariance asymptotique en RSB de $\Delta\boldsymbol{\theta}_{MVS}$ à la BCR stochastique asymptotique en RSB.

D'après [SN90b], la BCR stochastique s'écrit comme :

$$\mathbf{BCRS}(\boldsymbol{\theta}_0) = \frac{\sigma^2}{2T} \left(\text{Re} \left\{ \mathbf{H}(\boldsymbol{\theta}_0) \odot \left(\boldsymbol{\Sigma}_s \mathbf{A}^H(\boldsymbol{\theta}_0) \boldsymbol{\Sigma}_y^{-1} \mathbf{A}(\boldsymbol{\theta}_0) \boldsymbol{\Sigma}_s \right)^T \right\} \right)^{-1}. \quad (2.80)$$

On peut remarquer que $\mathbf{A}^H(\boldsymbol{\theta}_0) \boldsymbol{\Sigma}_y^{-1} \mathbf{A}(\boldsymbol{\theta}_0) = \left(\boldsymbol{\Sigma}_s + \sigma^2 (\mathbf{A}^H(\boldsymbol{\theta}_0) \mathbf{A}(\boldsymbol{\theta}_0))^{-1} \right)^{-1}$ qui tend vers $\boldsymbol{\Sigma}_s^{-1}$ lorsque $\sigma \rightarrow 0$. Donc, la BCR stochastique asymptotique en RSB s'écrit

$$\lim_{\sigma \rightarrow 0} \frac{\mathbf{BCRS}(\boldsymbol{\theta}_0)}{\sigma^2} = \frac{1}{2T} \left(\text{Re} \left\{ \mathbf{H}(\boldsymbol{\theta}_0) \odot \boldsymbol{\Sigma}_s^T \right\} \right)^{-1}. \quad (2.81)$$

Le corollaire suivant se déduit trivialement du théorème 2.3.

Corollaire 2.2 : la covariance asymptotique en RSB de $\Delta\boldsymbol{\theta}_{MVS}$, $cov_{as}(\Delta\boldsymbol{\theta}_{MVS})$ est donnée par

$$cov_{as}(\Delta\boldsymbol{\theta}_{MVS}) = \frac{1}{2T} \mathbb{E} \left[\left(\text{Re} \left\{ \mathbf{H}(\boldsymbol{\theta}_0) \odot \hat{\boldsymbol{\Sigma}}_s^T \right\} \right)^{-1} \right]. \quad (2.82)$$

Théorème 2.4 : Extension matricielle du théorème de Jensen.

Soit $\boldsymbol{\Theta} \in \mathcal{M}_{M \times M}(\mathbb{R})$ une matrice aléatoire réelle définie positive. Alors,

$$\boxed{\mathbb{E}[\boldsymbol{\Theta}^{-1}] \geq (\mathbb{E}[\boldsymbol{\Theta}])^{-1}}, \quad (2.83)$$

où l'égalité est atteinte si et seulement si $\boldsymbol{\Theta}$ est une matrice constante avec une probabilité de un.

Ce théorème constitue une extension de l'inégalité de Jensen [Leh83]. Il a été partiellement démontré dans [BMWZ87] mais sans donner la condition d'égalité.

Preuve du théorème 2.4 :

Lemme 2.1 : soit $\boldsymbol{\Omega} \in \mathcal{M}_{M \times M}(\mathbb{R})$ une matrice réelle définie positive. Alors, pour tout vecteur non nul \mathbf{q} appartenant à $\mathcal{M}_{M \times 1}(\mathbb{R})$:

$$\mathbf{q}^T \boldsymbol{\Omega} \mathbf{q} + \mathbf{q}^T \boldsymbol{\Omega}^{-1} \mathbf{q} - 2\mathbf{q}^T \mathbf{q} \geq 0, \quad (2.84)$$

où l'égalité est atteinte si et seulement si \mathbf{q} est un vecteur propre de $\boldsymbol{\Omega}$ avec comme valeur propre associée 1.

Preuve du lemme 2.1 : soit $\boldsymbol{\Omega} = \sum_{i=1}^M \lambda_i \mathbf{r}_i \mathbf{r}_i^T$ la décomposition en valeur propre de $\boldsymbol{\Omega}$ sur une base orthonormée $\{\mathbf{r}_i\}_{i=1..M}$ avec les valeurs propres associées λ_i . L'équation (2.84) se réécrit :

$$\sum_{i=1}^M \left(\lambda_i - 2 + \frac{1}{\lambda_i} \right) (\mathbf{q}^T \mathbf{r}_i)^2 \geq 0. \quad (2.85)$$

La preuve est triviale en remarquant que $\lambda_i - 2 + \frac{1}{\lambda_i} \geq 0$ pour $\lambda_i > 0$, et que $\lambda_i - 2 + \frac{1}{\lambda_i} = 0$ si et seulement si $\lambda_i = 1$. ■

Lemme 2.2 : soit $\boldsymbol{\Omega} \in \mathcal{M}_{M \times M}(\mathbb{R})$ une matrice aléatoire réelle définie positive avec une probabilité un telle que $\mathbb{E}[\boldsymbol{\Omega}] = \mathbf{I}_M$. Alors, il existe un vecteur $\mathbf{q} \in \mathcal{M}_{M \times 1}(\mathbb{R})$ tel que

$$\mathbf{q}^T \mathbb{E}[\boldsymbol{\Omega}^{-1}] \mathbf{q} - \mathbf{q}^T \mathbf{q} > 0, \quad (2.86)$$

si et seulement si $\Pr(\boldsymbol{\Omega} = \mathbf{I}_M) \neq 1$.

Preuve du lemme 2.2 : soit $\zeta_{\mathbf{q}} = \mathbf{q}^T \boldsymbol{\Omega} \mathbf{q} + \mathbf{q}^T \boldsymbol{\Omega}^{-1} \mathbf{q} - 2\mathbf{q}^T \mathbf{q}$. Puisque $\mathbb{E}[\boldsymbol{\Omega}] = \mathbf{I}_M$, nous avons $\mathbb{E}[\zeta_{\mathbf{q}}] = \mathbf{q}^T \mathbb{E}[\boldsymbol{\Omega}^{-1}] \mathbf{q} - \mathbf{q}^T \mathbf{q}$. En conséquence, la démonstration du lemme 2.2 revient à démontrer que $\exists \mathbf{q}$ tel que $\mathbb{E}[\zeta_{\mathbf{q}}] > 0$ si et seulement si $\Pr[\boldsymbol{\Omega} = \mathbf{I}_M] \neq 1$. Or, d'après (2.84) $\zeta_{\mathbf{q}}$ est une variable aléatoire non négative. Donc, $\mathbb{E}[\zeta_{\mathbf{q}}] > 0$ si et seulement si $\Pr[\zeta_{\mathbf{q}} = 0] \neq 1$. D'après le lemme 2.1, $\Pr[\zeta_{\mathbf{q}} = 0] = \Pr[\boldsymbol{\Omega} \mathbf{q} = \mathbf{q}]$. En conséquence, $\forall \mathbf{q} \Pr[\zeta_{\mathbf{q}} = 0] = 1 \iff \forall \mathbf{q} \Pr[\boldsymbol{\Omega} \mathbf{q} = \mathbf{q}] = 1 \iff \Pr[\boldsymbol{\Omega} = \mathbf{I}_M] = 1$. ■

Finalement, avec les notations du théorème 2.4, on pose $\boldsymbol{\Omega} = \mathbb{E}[\boldsymbol{\Theta}]^{-1/2} \boldsymbol{\Theta} \mathbb{E}[\boldsymbol{\Theta}]^{-1/2}$. Le lemme 2.2 démontre le théorème 2.4. ■

Corollaire 2.3 : l'estimateur du MVS est non-efficace quel que soit le nombre de sources.

Preuve : on pose $\boldsymbol{\Theta} = 2T \text{Re} \left\{ \mathbf{H}(\boldsymbol{\theta}_0) \odot \hat{\boldsymbol{\Sigma}}_s^T \right\}$ dans les équations (2.81) et (2.82). L'équation (2.83) provient de l'inégalité de Cramér-Rao, c'est-à-dire $cov_{as}(\Delta\boldsymbol{\theta}_{MVS}) \geq \lim_{\sigma \rightarrow 0} \frac{\mathbf{BCRS}(\boldsymbol{\theta}_0)}{\sigma^2}$. Puisque $\boldsymbol{\Theta}$ n'est pas une matrice constante avec une probabilité un, il vient du théorème 2.4 que l'inégalité est stricte. ■

2.4.2.2 Cas particuliers

Nous avons démontré ci-dessus la non-gaussianité et la non-efficacité asymptotiques en RSB de l'estimateur du MVS quel que soit le nombre de sources et quelle que soit la géométrie de l'antenne. Même si ces résultats ont un intérêt théorique certain, en terme d'application il serait intéressant d'établir la loi asymptotique en RSB de l'estimateur du MVS ainsi qu'une expression analytique de la variance asymptotique en RSB. Dans cette section nous traitons deux cas particuliers, à savoir un scénario à une source et un scénario à deux sources proches, c'est-à-dire situées dans le lobe principal de l'antenne. Concernant le cas une source, nous retrouvons directement les résultats de [Boy02] et [FBL04] grâce aux expressions établies précédemment. Nous nous focalisons ensuite sur le cas de deux sources proches. Ce problème trouve sa légitimité dans le fait que les méthodes du maximum de vraisemblance ont été développées pour obtenir une résolution plus performante que la formation de voie classique dont on peut rappeler que, pour un scénario à une source, elle est équivalente à la méthode du maximum de vraisemblance. Dans ce contexte, nous obtenons une forme analytique de la covariance asymptotique en RSB de l'estimateur du MVS ainsi que sa distance à la BCR.

Scénario une source

Dans le cas d'un scénario une source, $\Sigma_s = \Sigma$ et $\mathbf{H}(\theta_0) = h(\theta_0)$.

Corollaire 2.4 : $\Delta\theta_{MVS}$ est asymptotiquement distribué à fort RSB comme $\sqrt{k}S_{2T}$ où S_{2T} est une variable aléatoire suivant une loi de Student à $2T$ degrés de liberté et où k est donné par :

$$k = \lim_{\sigma \rightarrow 0} \frac{BCRS(\theta_0)}{\sigma^2} = \frac{1}{2Th(\theta_0)\Sigma}. \quad (2.87)$$

Preuve : d'après le théorème 2.3, $\Delta\theta_{MVS}$ est distribué asymptotiquement en RSB comme cq où q est une variable aléatoire réelle issue d'un processus gaussien de moyenne nulle et de variance un et où c satisfait :

$$c^2 = \frac{1}{2Th(\theta_0)\hat{\Sigma}}, \quad (2.88)$$

où $T\hat{\Sigma}$ est un cas particulier (scalaire) de distribution de Wishart réelle à T degrés de liberté et de paramètre Σ . Donc, $\hat{\Sigma}$ est distribué comme $\frac{\Sigma}{2T}$ fois une variable aléatoire réelle suivant une loi du chi deux avec $2T$ degrés de liberté notée X . En d'autres termes :

$$c = \frac{1}{\sqrt{2Th(\theta_0)\frac{\Sigma}{2T}X}}, \quad (2.89)$$

où $X \sim \chi^2(2T)$. Si $Z \sim \mathcal{N}(0, 1)$ est une variable aléatoire réelle issue d'un processus gaussien de moyenne nulle et de variance un et si $Y \sim \chi^2(Q)$ est une variable aléatoire réelle distribuée selon une loi de chi deux à Q degrés de liberté indépendante de Z , alors le rapport $\frac{Z}{\sqrt{\frac{Y}{Q}}} = S_Q$ suit une loi de Student avec Q degrés de liberté. Ici, puisque q est gaussien et indépendant de X , on a

$$\begin{aligned} cq &= \frac{1}{\sqrt{2Th(\theta_0)\Sigma}} \frac{q}{\sqrt{\frac{X}{2T}}}, \\ &= \frac{1}{\sqrt{2Th(\theta_0)\Sigma}} S_{2T}. \end{aligned} \quad (2.90)$$

Donc, $\Delta\theta_{MVS}$ est asymptotiquement en RSB distribué comme $\sqrt{k}S_{2T}$ avec $k = \frac{1}{2Th(\theta_0)\Sigma}$. D'après (2.81), la BCR stochastique à fort RSB est donnée dans le cas d'une seule source par :

$$\lim_{\sigma \rightarrow 0} \frac{\mathbf{BCRS}(\theta_0)}{\sigma^2} = \frac{1}{2Th(\theta_0)\Sigma} = k. \quad (2.91)$$

■

Corollaire 2.5 : la variance asymptotique en RSB, $var(\Delta\theta_{MVS})$, de $\Delta\theta_{MVS}$ est donnée par :

$$\boxed{var(\Delta\theta_{MVS}) = \frac{T}{T-1}k.} \quad (2.92)$$

Preuve : la variance d'un processus aléatoire suivant une loi de Student à P degrés de liberté est donnée par $var(S_P) = \frac{P}{P-2}$ (voir [AS72]). ■

Puisque $\frac{T}{T-1} > 1$, ceci confirme la non-efficacité asymptotique en RSB démontrée précédemment pour MVS dans le cas d'un nombre fini d'observations.

Scénario deux sources

Une grande variété d'antennes rencontrée en pratique possèdent un centre de symétrie. On pourra citer par exemple l'antenne linéaire uniforme.

lemme 2.3 : pour une antenne centro-symétrique, la matrice $\mathbf{H}(\boldsymbol{\theta}_0)$ (2.30) est réelle et symétrique quel que soit le nombre de sources.

Preuve : d'après (2.30) il est clair que $\mathbf{H}(\boldsymbol{\theta}_0)$ est une matrice hermitienne, c'est-à-dire $\mathbf{H}(\boldsymbol{\theta}_0) = \mathbf{H}^H(\boldsymbol{\theta}_0)$. Il reste donc à démontrer que dans le cas d'une antenne centro-symétrique, $\mathbf{H}(\boldsymbol{\theta}_0)$ est réelle $\forall \boldsymbol{\theta}_0$. En partant de (2.30), on a :

$$\begin{aligned} \mathbf{H}(\boldsymbol{\theta}) &= \mathbf{D}^H(\boldsymbol{\theta}) \boldsymbol{\Pi}_{\mathbf{A}}^\perp(\boldsymbol{\theta}) \mathbf{D}(\boldsymbol{\theta}) \\ &= \mathbf{D}^H(\boldsymbol{\theta}) \mathbf{D}(\boldsymbol{\theta}) - (\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{D}(\boldsymbol{\theta}))^H (\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta}))^{-1} (\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{D}(\boldsymbol{\theta})). \end{aligned} \quad (2.93)$$

Le $i^{\text{ième}}$ élément du vecteur directionnel $\mathbf{a}(\theta_k)$ est donné par :

$$a_i(\theta_k) = e^{j\frac{2\pi}{\lambda} \mathbf{v}_i^T \mathbf{u}(\theta_k)}, \quad (2.94)$$

où λ est la longueur d'onde des signaux émis, \mathbf{v}_i est le vecteur contenant les coordonnées géométriques du $i^{\text{ième}}$ capteur et $\mathbf{u}(\theta_k)$ est le vecteur unité pointant vers la $k^{\text{ième}}$ source. Donc, le $i^{\text{ième}}$ élément du vecteur $\mathbf{d}(\theta_k) = \left. \frac{d\mathbf{a}(\theta)}{d\theta} \right|_{\theta=\theta_k}$ s'écrit :

$$d_i(\theta_k) = j \frac{2\pi}{\lambda} \mathbf{v}_i^T \left. \frac{d\mathbf{u}(\theta)}{d\theta} \right|_{\theta=\theta_k} e^{j\frac{2\pi}{\lambda} \mathbf{v}_i^T \mathbf{u}(\theta_k)}. \quad (2.95)$$

L'élément correspondant à la $n^{\text{ième}}$ ligne et $m^{\text{ième}}$ colonne de chaque terme intervenant dans (2.93) est donné par

$$\left\{ \begin{array}{l} \mathbf{A}^H \mathbf{A} \Big|_{n,m} = \sum_{i=1}^N e^{j\frac{2\pi}{\lambda} \mathbf{v}_i^T (\mathbf{u}(\theta_m) - \mathbf{u}(\theta_n))}, \\ \mathbf{D}^H \mathbf{D} \Big|_{n,m} = \sum_{i=1}^N \left(\frac{2\pi}{\lambda} \right)^2 \left(\mathbf{v}_i^T \left. \frac{d\mathbf{u}(\theta)}{d\theta} \right|_{\theta=\theta_n} \right) \left(\mathbf{v}_i^T \left. \frac{d\mathbf{u}(\theta)}{d\theta} \right|_{\theta=\theta_m} \right) e^{j\frac{2\pi}{\lambda} \mathbf{v}_i^T (\mathbf{u}(\theta_m) - \mathbf{u}(\theta_n))}, \\ \mathbf{A}^H \mathbf{D} \Big|_{n,m} = \sum_{i=1}^N j \frac{2\pi}{\lambda} \mathbf{v}_i^T \left. \frac{d\mathbf{u}(\theta)}{d\theta} \right|_{\theta=\theta_m} e^{j\frac{2\pi}{\lambda} \mathbf{v}_i^T (\mathbf{u}(\theta_m) - \mathbf{u}(\theta_n))}. \end{array} \right. \quad (2.96)$$

Selon l'hypothèse de centro-symétrie concernant l'antenne, on peut indexer les capteurs de la façon suivante : $\mathbf{v}_i = -\mathbf{v}_{M-i+1}$. $\mathbf{A}^H \mathbf{A} \Big|_{n,m}$ est une somme d'éléments deux à deux complexes

conjugués avec le même module⁴. Donc, $\mathbf{A}^H \mathbf{A}|_{n,m} \in \mathbb{R}$. Dans le même esprit, on conclut que $\mathbf{D}^H \mathbf{D}|_{n,m} \in \mathbb{R}$ puisque

$$\left(\mathbf{v}_i^T \frac{d\mathbf{u}(\theta)}{d\theta} \Big|_{\theta=\theta_n} \right) \left(\mathbf{v}_i^T \frac{d\mathbf{u}(\theta)}{d\theta} \Big|_{\theta=\theta_m} \right) = \left(\mathbf{v}_{M-i+1}^T \frac{d\mathbf{u}(\theta)}{d\theta} \Big|_{\theta=\theta_n} \right) \left(\mathbf{v}_{M-i+1}^T \frac{d\mathbf{u}(\theta)}{d\theta} \Big|_{\theta=\theta_m} \right). \quad (2.97)$$

Finalement $\mathbf{A}^H \mathbf{D}|_{n,m} \in \mathbb{R}$ puisque :

$$\mathbf{A}^H \mathbf{D}|_{n,m} = j \frac{2\pi}{\lambda} \sum_{i=1}^{\frac{N}{2}} \left(\mathbf{v}_i^T \frac{d\mathbf{u}(\theta)}{d\theta} \Big|_{\theta=\theta_m} \right) \left(e^{j \frac{2\pi}{\lambda} \mathbf{v}_i^T (\mathbf{u}(\theta_m) - \mathbf{u}(\theta_n))} - e^{-j \frac{2\pi}{\lambda} \mathbf{v}_i^T (\mathbf{u}(\theta_m) - \mathbf{u}(\theta_n))} \right), \quad (2.98)$$

où $e^{j \frac{2\pi}{\lambda} \mathbf{v}_i^T (\mathbf{u}(\theta_m) - \mathbf{u}(\theta_n))} - e^{-j \frac{2\pi}{\lambda} \mathbf{v}_i^T (\mathbf{u}(\theta_m) - \mathbf{u}(\theta_n))}$ est un nombre imaginaire pur. ■

On se place maintenant dans le cas de deux sources décorréées.

Corollaire 2.6 : pour un scénario à deux sources, on obtient directement du lemme 2.3 la structure de $\mathbf{H}(\boldsymbol{\theta}_0)$ suivante :

$$\mathbf{H}(\boldsymbol{\theta}_0) = \begin{pmatrix} h_1 & h_3 \\ h_3 & h_2 \end{pmatrix}, \quad (2.99)$$

où $(h_1, h_2, h_3) \in \mathbb{R}^3$.

Corollaire 2.7 : dans le cas d'un scénario à deux sources non-corrélées, $\boldsymbol{\Sigma}_s = \text{Diag}\{\Sigma_1, \Sigma_2\}$ et la covariance asymptotique en RSB du MVS est donnée par :

$$\boxed{\text{cov}(\Delta\boldsymbol{\theta}_{MVS}) = \frac{1}{2(T-1)} {}_2F_1\left(1, 1; 2T; \frac{h_3^2}{h_1 h_2}\right) \text{Diag}\left\{\frac{1}{h_1 \Sigma_1}, \frac{1}{h_2 \Sigma_2}\right\}} \quad (2.100)$$

$$= \frac{T}{T-1} {}_2F_1\left(1, 1; 2T; \frac{h_3^2}{h_1 h_2}\right) \mathbf{K}, \quad (2.101)$$

où $\mathbf{K} = \lim_{\sigma \rightarrow 0} \frac{\text{BCRS}(\boldsymbol{\theta}_0)}{\sigma^2}$ et où ${}_2F_1(a, b; c; \omega)$ est la fonction hypergéométrique de Gauss définie par sa représentation intégrale (voir [AS72] page 558)

$${}_2F_1(a, b; c; \omega) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 z^{b-1} (1-z)^{c-b-1} (1-z\omega)^{-a} dz, \quad (2.102)$$

où $\Gamma(z)$ est la fonction Gamma : $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$.

Preuve : d'après (2.82) et sous l'hypothèse d'une antenne centro-symétrique :

$$\begin{aligned} \text{cov}(\Delta\boldsymbol{\theta}_{MVS}) &= \frac{1}{2T} \mathbb{E} \left[\left(\mathbf{H}(\boldsymbol{\theta}_0) \odot \text{Re} \left\{ \hat{\boldsymbol{\Sigma}}_s^T \right\} \right)^{-1} \right] \\ &= \frac{1}{2} \mathbb{E} \left[\left(\mathbf{H}(\boldsymbol{\theta}_0) \odot \mathbf{W}_R \right)^{-1} \right], \end{aligned} \quad (2.103)$$

où $\mathbf{W}_R = T \text{Re} \left\{ \hat{\boldsymbol{\Sigma}}_s^T \right\} \in \mathcal{M}_{M \times M}(\mathbb{R})$ est une matrice aléatoire distribué selon une loi de Wishart réelle avec $2T$ degrés de liberté et comme matrice paramètre $\frac{1}{2} \boldsymbol{\Sigma}_s = \frac{1}{2} \text{Diag}\{\Sigma_1, \Sigma_2\}$. Une factorisation de Cholesky donne $\mathbf{W}_R = \mathbf{Q}\mathbf{Q}^T$, avec :

$$\mathbf{Q} = \begin{pmatrix} \rho_1 & 0 \\ \alpha & \rho_2 \end{pmatrix}. \quad (2.104)$$

4. Si le nombre de capteur est impair, $\mathbf{v}_{\frac{N}{2}+1} = 0$.

Les éléments de la matrice triangulaire inférieure \mathbf{Q} sont indépendants et satisfont [And84] :

$$\begin{cases} \rho_1 \sim \sqrt{\frac{\Sigma_1}{2} \chi^2(2T)}, \\ \rho_2 \sim \sqrt{\frac{\Sigma_2}{2} \chi^2(2T-1)}, \\ \alpha \sim \mathcal{N}_{\mathbb{R}}\left(0, \frac{\Sigma_2}{2}\right). \end{cases} \quad (2.105)$$

La matrice de covariance asymptotique en RSB de $\Delta\theta_{MVS}$ est donnée par :

$$\text{cov}(\Delta\theta_{MVS}) = \begin{pmatrix} \text{var}(\Delta\theta_{MVS1}) & \Psi \\ \Psi & \text{var}(\Delta\theta_{MVS2}) \end{pmatrix} \quad (2.106)$$

$$\begin{aligned} &= \frac{1}{2} \mathbb{E} \left[(\mathbf{H}(\theta_0) \odot \mathbf{W}_R)^{-1} \right] \\ &= \frac{1}{2} \mathbb{E} \left[\frac{1}{\Phi} \begin{pmatrix} h_2(\rho_2^2 + \alpha^2) & -h_3\rho_1\alpha \\ -h_3\rho_1\alpha & h_1\rho_1^2 \end{pmatrix} \right], \end{aligned} \quad (2.107)$$

où $\text{var}(\Delta\theta_{MVS1})$ (respectivement $\text{var}(\Delta\theta_{MVS2})$) est la variance asymptotique en RSB de la première source (respectivement la variance asymptotique en RSB de la seconde source), Ψ est l'inter-corrélation et $\Phi = h_1 h_2 \rho_1^2 (\rho_2^2 + \alpha^2) - (h_3 \rho_1 \alpha)^2$.

Nous allons maintenant détailler $\text{var}(\Delta\theta_{MVS1})$,

$$\begin{aligned} \text{var}(\Delta\theta_{MVS1}) &= \frac{1}{2} \mathbb{E} \left[\frac{h_2(\rho_2^2 + \alpha^2)}{h_1 h_2 \rho_1^2 (\rho_2^2 + \alpha^2) - (h_3 \rho_1 \alpha)^2} \right] \\ &= \frac{1}{2h_1} \mathbb{E} \left[\frac{1}{\rho_1^2 \left(1 - \frac{h_3^2}{h_1 h_2} \frac{\alpha^2}{\alpha^2 + \rho_2^2} \right)} \right], \end{aligned}$$

où $\alpha^2 \sim \frac{\Sigma_2}{2} \chi^2(1)$. Le rapport $\frac{\alpha^2}{\alpha^2 + \rho_2^2} = Z$ est donc distribué selon une loi beta avec 1 et $2T - 1$ degrés de liberté indépendante de $Y = \rho_1^2$. Cette indépendance nous permet de réécrire $\text{var}(\Delta\theta_{MVS1})$:

$$\text{var}(\Delta\theta_{MVS1}) = \frac{1}{2h_1} \mathbb{E} \left[\frac{1}{Y} \right] \mathbb{E} \left[\frac{1}{1 - \frac{h_3^2}{h_1 h_2} Z} \right] = \frac{I_1 I_2}{2h_1}, \quad (2.108)$$

avec

$$\begin{cases} I_1 = \mathbb{E} \left[\frac{1}{Y} \right] = \int_0^\infty \frac{1}{y} \Pi_{\chi^2}(y) dy, \\ I_2 = \mathbb{E} \left[\frac{1}{1 - \frac{h_3^2}{h_1 h_2} Z} \right] = \int_0^1 \frac{1}{1 - \frac{h_3^2}{h_1 h_2} z} \Pi_\beta(z) dz, \end{cases} \quad (2.109)$$

où $\Pi_{\chi^2}(y)$ et $\Pi_\beta(z)$ sont les densités de probabilité d'une loi du chi deux $\frac{\Sigma_1}{2} \chi^2(2T)$ et d'une loi beta $\beta(1, 2T - 1)$, c'est-à-dire

$$\begin{cases} \Pi_{\chi^2}(y) = \frac{1}{2^{T-1} \Gamma(T) \Sigma_1^T} y^{T-1} e^{-\frac{y}{\Sigma_1}}, \\ \Pi_\beta(z) = (2T - 1) (1 - z)^{2(T-1)}. \end{cases} \quad (2.110)$$

Lorsque $T \geq 2$, I_1 converge : c'est une fonction Gamma. I_2 est la représentation intégrale d'une fonction hypergéométrique de Gauss ([AS72] page 558)

$${}_2F_1(a, b; c; \omega) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 z^{b-1} (1-z)^{c-b-1} (1-z\omega)^{-a} dz, \quad (2.111)$$

avec $a = 1$, $b = 1$, $c = 2T$ et $\omega = \frac{h_3^2}{h_1 h_2}$. Si $-1 < \omega < 1$, l'intégrale (2.111) est finie $\forall (a, b, c)$. D'un autre côté, si $\omega = \pm 1$, l'intégrale (2.111) est finie pour $c > a + b$. Dans notre cas, $\mathbf{H}(\boldsymbol{\theta}_0)$ est définie positive, donc $|\mathbf{H}(\boldsymbol{\theta}_0)| \geq 0 \Leftrightarrow \omega = \frac{h_3^2}{h_1 h_2} \leq 1$. Cela signifie que I_2 est finie pour $T \geq 2$.

Finalement :

$$\begin{cases} I_1 = \frac{1}{(T-1)\Sigma_1}, \\ I_2 = {}_2F_1\left(1, 1; 2T; \frac{h_3^2}{h_1 h_2}\right). \end{cases} \quad (2.112)$$

et

$$\text{var}(\Delta\theta_{MVS1}) = \frac{1}{2(T-1)h_1\Sigma_1} {}_2F_1\left(1, 1; 2T; \frac{h_3^2}{h_1 h_2}\right). \quad (2.113)$$

Par un raisonnement similaire on obtient :

$$\text{var}(\Delta\theta_{MVS2}) = \frac{1}{2(T-1)h_2\Sigma_2} {}_2F_1\left(1, 1; 2T; \frac{h_3^2}{h_1 h_2}\right). \quad (2.114)$$

Le coefficient d'inter-corrélation Ψ est nul, il s'agit d'une intégrale sur \mathbb{R} d'une fonction impaire de la variable α .

D'après (2.81), la BCR stochastique dans le cas d'un scénario deux sources non-corrélées s'écrit :

$$\lim_{\sigma \rightarrow 0} \frac{1}{\sigma^2} \mathbf{BCRS}(\boldsymbol{\theta}_0) = \frac{1}{2T} \text{Diag} \left\{ \frac{1}{h_1 \Sigma_1}, \frac{1}{h_2 \Sigma_2} \right\}. \quad (2.115)$$

D'où, en utilisant (2.106), (2.113), (2.114) et (2.115) on obtient (2.100). ■

Ceci confirme la non-efficacité asymptotique en RSB démontré précédemment pour le MVS dans le cas d'un nombre fini d'observations puisque $\frac{T}{T-1} > 1$ et ${}_2F_1\left(1, 1; 2T; \frac{h_3^2}{h_1 h_2}\right) \geq 1$.

2.4.2.3 Simulations

Afin d'illustrer l'analyse des performances asymptotiques en RSB de la méthode du MVS, nous présentons maintenant quelques résultats de simulations. Nous utilisons l'algorithme présenté section 2.2.1 dans lequel la minimisation du critère est conduite à l'aide d'un algorithme de Gauss-Newton précédé d'une recherche grossière au travers d'une grille de valeurs. Nous considérons une ALU dans laquelle les capteurs sont espacés d'une demi longueur d'onde. Les directions d'arrivées sont données par rapport au travers de l'antenne. La BCR stochastique est donnée par (2.81).

Scénario une source

On considère l'estimation de l'angle d'arrivée d'une source située en $\theta_0 = 0^\circ$. L'algorithme du MVS opère avec $T = 2$ observations et l'antenne est composée de $N = 4$ capteurs. La figure 2.6 superpose la variance empirique de l'estimateur du MVS, la BCR stochastique, et la variance asymptotique théorique donnée par (2.92) en fonction du RSB. La simulation de Monte-Carlo est conduite à l'aide de $r = 1000$ réalisations indépendantes. Dans la zone asymptotique ($RSB \geq 15\text{dB}$), on observe une variance empirique double par rapport à la BCR stochastique ($\frac{T}{T-1} = 2$). Pour les RSB inférieurs à 15dB, on observe le phénomène de décrochement introduit au chapitre 1. Ce phénomène donne la plage à partir de laquelle le terme "asymptotique en RSB" n'est plus valable (on pourra se reporter à [Ath05] pour une analyse de la prédiction du décrochement pour ce type de scénario).

La figure 2.7 représente l'histogramme des estimées du MVS pour un RSB égal à 30 dB. La simulation de Monte-Carlo est conduite à l'aide de $r = 10000$ réalisations indépendantes. Cet histogramme est comparé à une distribution de Student dérivée section 2.4.2.2 et à une

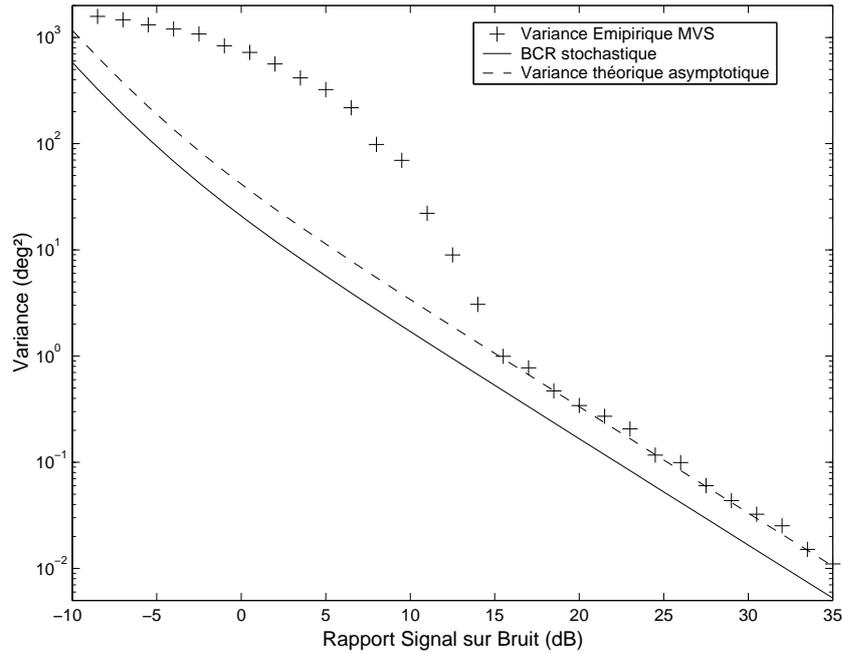


FIGURE 2.6 – MVS. Scénario une source. $\theta_0 = 0^\circ$. Variance empirique de l’estimateur, BCR stochastique, et variance asymptotique théorique en fonction du RSB. $T = 2$ observations $N = 4$ capteurs. $r = 1000$ réalisations.

distribution gaussienne de même moyenne et variance que les estimées. On observe la bonne adéquation de la distribution des estimées avec la loi de Student ce qui confirme la non-gaussianité du MVS pour un fort RSB et un nombre fini d’observations.

Nous avons également réalisé un test statistique du chi deux pour confirmer ce résultat “visuel”. Les données sont divisées en $k = 15$ classes et le test statistique requière le calcul de

$$\Delta = \sum_{i=1}^k \frac{(O_i - rp_i)^2}{rp_i}, \quad (2.116)$$

où O_i est le nombre d’occurrences observées pour la classe i et p_i est la probabilité candidate pour la classe i . L’hypothèse que les données soient issues d’une population suivant la distribution candidate, c’est-à-dire une distribution de Student ou gaussienne, est rejetée si

$$\Pr(X \geq \Delta) = \frac{\Gamma\left(\frac{k-1}{2}, \frac{\Delta}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)} \leq 5\%, \quad (2.117)$$

où $X \sim \chi^2(k-1)$. Les résultats donnés dans le tableau 2.1 confirment la non-gaussianité du MVS pour un fort RSB et un nombre fini d’observations.

Scénario deux sources

L’algorithme du MVS est conduit avec $T = 2$ observations et l’antenne est composée de $N = 10$ capteurs. On considère maintenant un scénario avec deux sources de même puissance situées en $\theta_0 = [0^\circ \ 5^\circ]$, c’est-à-dire dans le lobe principal de l’antenne de largeur 10 degrés. Nous avons reporté sur la figure 2.8 l’évolution de la variance empirique du MVS, la variance théorique donnée par (2.100) ainsi que la BCR stochastique en fonction du RSB. La simulation de Monte-Carlo

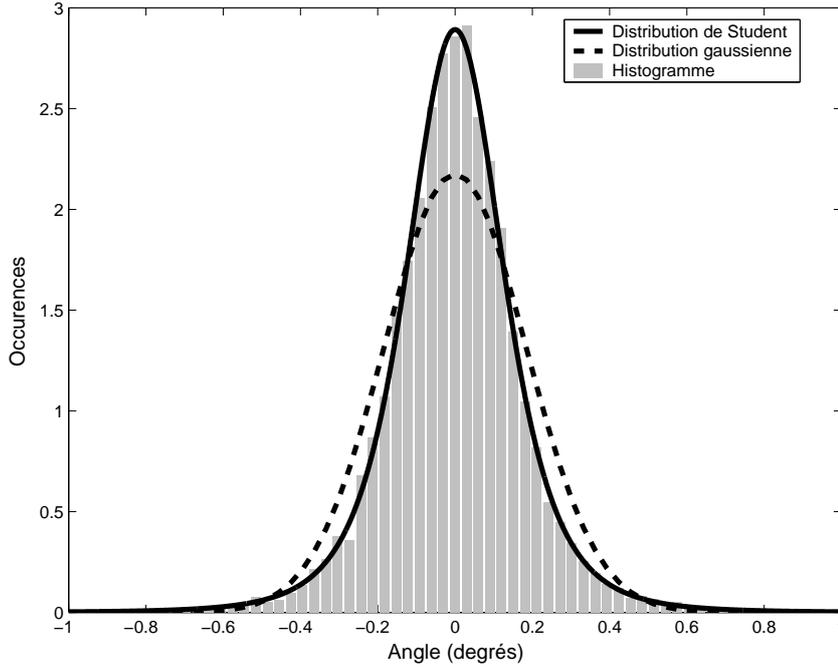


FIGURE 2.7 – MVS. Scénario une source. $\theta_0 = 0^\circ$. Histogramme des estimées, loi de Student théorique et loi gaussienne. $RSB = 30\text{dB}$. $T = 2$ observations $N = 4$ capteurs. $r = 10000$ réalisations.

Scénario une source	Distribution de Student	Distribution Gaussienne
Δ	10.349	$2.4 \cdot 10^{10}$
$\Pr(X \geq \Delta)$	80%	0%
Hypothèses	acceptée	rejetée

TABLE 2.1 – MVS. Scénario une source. $\theta_0 = 0^\circ$. Test statistique du chi deux. $RSB = 30\text{dB}$. $T = 2$ observations $N = 4$ capteurs. $r = 10000$ réalisations. $k = 15$ classes.

est conduite à l'aide de $r = 1000$ réalisations indépendantes. Ici, $\frac{T}{T-1} {}_2F_1\left(1, 1; 2T; \frac{h_3^2}{h_1 h_2}\right) = 2.9$. Dans la zone asymptotique ($RSB \geq 10\text{dB}$), on observe la bonne adéquation entre la variance du MVS et la variance théorique asymptotique en RSB ainsi que la non-efficacité asymptotique du MVS.

Pour le même scénario, la figure 2.9 donne les histogrammes des estimées ainsi qu'une distribution gaussienne de même moyenne et variance que les estimées. La simulation de Monte-Carlo est conduite à l'aide de $r = 10000$ réalisations indépendantes et le RSB est égal à 30dB. On observe encore une fois la non-gaussiannité du MVS.

Le tableau 2.2 donne les résultat du test du chi deux ($k = 15$ classes). La simulation de Monte-Carlo est conduite à l'aide de $r = 10000$ réalisations indépendantes et le RSB est égal à 30dB.

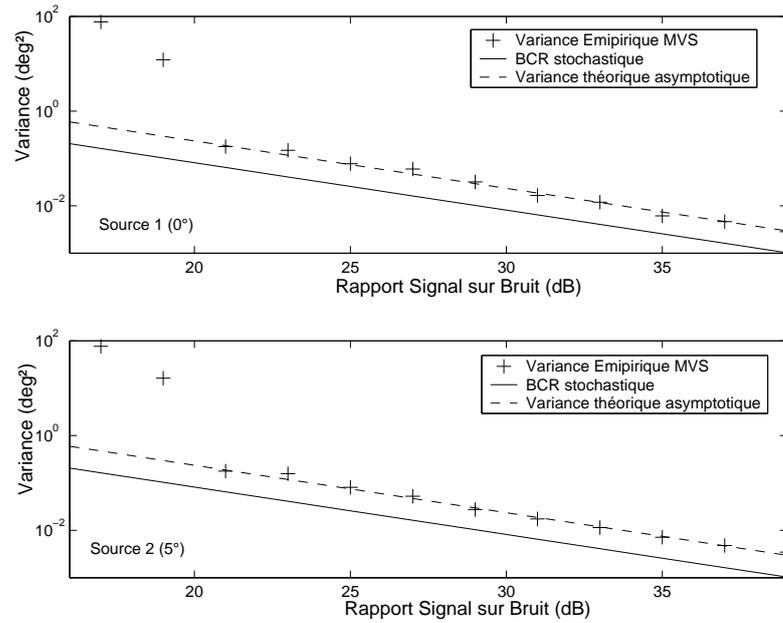


FIGURE 2.8 – MVS. Scénario deux sources. $\theta_0 = [0^\circ \ 5^\circ]$. Variance empirique de l'estimateur, BCR stochastique, et variance asymptotique théorique en fonction du RSB. $T = 2$ observations $N = 10$ capteurs. $r = 1000$ réalisations.

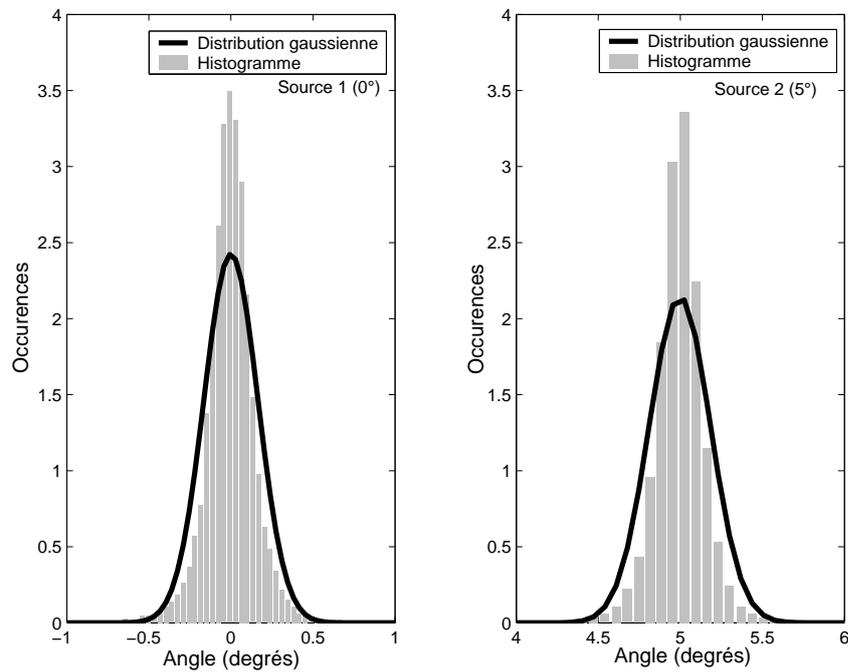


FIGURE 2.9 – MVS. Scénario deux sources. $\theta_0 = [0^\circ \ 5^\circ]$. Histogrammes des estimées et loi gaussienne. $RSB = 30\text{dB}$. $T = 2$ observations $N = 10$ capteurs. $r = 10000$ réalisations.

Scénario 2 sources	Distribution gaussienne : source 1	Distribution gaussienne : source 2
Δ	$1.32 \cdot 10^{22}$	$4 \cdot 10^{24}$
$\Pr(X \geq \Delta)$	0%	0%
Hypothèse	rejetée	rejetée

TABLE 2.2 – MVS. Scénario deux sources. $\theta_0 = [0^\circ \ 5^\circ]$. Test statistique du chi deux. $RSB = 30\text{dB}$. $T = 2$ observations $N = 10$ capteurs. $r = 10000$ réalisations. $k = 15$ classes.

Chapitre 3

Bornes minimales d'estimation

3.1 Introduction

Les bornes minimales d'estimation sont des minorants de l'EQM. Elles fournissent des indications sur les performances ultimes qu'un estimateur peut espérer atteindre pour un modèle d'observation donné. En conséquence, elles sont utilisées comme "benchmarks" afin de jauger les performances d'un estimateur et de savoir si, à partir d'un cahier des charges donné, une amélioration est possible. La borne de Cramér-Rao (BCR) est la borne la plus utilisée par la communauté du traitement du signal et fait encore, à l'heure actuelle, l'objet de nombreux travaux tant théoriques, en particulier au travers des variétés différentielles en géométrie Riemannienne (voir, par exemple, [SM93] [XB04b] [XB04a] [Man04] [Smi05b] [Smi05a]) que pratiques (voir, par exemple, [DA04] [DA05] [CLF05] [ZLO06] [YN06]). Le principal argument généralement avancé en faveur de la BCR est sa simplicité de calcul. Nous pensons que la BCR jouit surtout de son antériorité historique, c'est-à-dire qu'en tant que première borne minimale elle a souvent été préférée aux autres bornes minimales. Il existe en effet une pléthore d'autres bornes minimales d'estimation qui ont été dérivées depuis plus de soixante ans à l'aide de différentes inégalités mathématiques telles que : l'inégalité de Cauchy-Schwartz, l'inégalité de Kotelnikov, l'inégalité de Hölder, l'inégalité de Ibragimov-Hasminskii, l'inégalité de Bhattacharyya, l'inégalité de Kiefer, etc.... Cette profusion désordonnée a nuit à l'utilisation et à la dissémination de ces bornes puisqu'il devenait difficile pour le chercheur non spécialiste de discerner les différents concepts théoriques sous-jacents à leur l'établissement permettant de les retranscrire dans le contexte d'un problème d'estimation spécifique. Cependant, nous montrerons dans ce chapitre que toute ces bornes sont plus pertinentes que la BCR, justifiant largement le surcroît d'efforts induits.

Les bornes minimales d'estimation se classent en deux catégories distinctes : les bornes déterministes lorsque les paramètres θ_0 sont supposés déterministes et les bornes bayésiennes lorsque les paramètres θ sont supposés aléatoires avec une densité de probabilité *a priori* $p(\theta)$. Les bornes déterministes sont : la borne de Cramér-Rao [Fis22] [Dug37] [Fre43] [Dar45] [Rao45] [Cra46], la borne de Bhattacharyya [Bha46] [FG52], la borne de Chapman-Robbins [Ham50] [CR51] [Kie52], la borne de Barankin [Bar49] [MS69], et la borne d'Abel [Abe90] [Abe93]. Les bornes bayésiennes se sub-divisent en deux catégories : les bornes de la famille Ziv-Zakaï qui dérivent d'un problème de test d'hypothèse binaire (et plus généralement d'un test d'hypothèses M -aires) et les bornes de la famille Weiss-Weinstein qui dérivent, comme les bornes déterministes, d'un principe d'inégalité de covariance. La famille Ziv-Zakaï est composée de la borne de Ziv-Zakaï [ZZ69], de la borne de Bellini-Tartara [BT74], de la borne de Chazan-Zakaï-Ziv [CZZ75], de la borne de Weinstein [Wei88], de la borne de Bell-Steinberg-Ephraïm-VanTrees [BSET97] et de la borne de Bell [Bel95]. La famille Weiss-Weinstein est composée de la borne de Cramér-Rao bayé-

sienne [Van68], de la borne de Bobrovsky-MayerWolf-Zakaï [BMWZ87], de la borne de Bhattacharyya bayésienne [Van68], de la borne de Bobrovsky-Zakaï [BZ76], de la borne de Reuven-Messer [RM97] et de la borne de Weiss-Weinstein [WW85].

Il faut garder présent à l'esprit que les bornes déterministes sont des minorants de l'EQM locale en θ_0 c'est-à-dire de

$$\mathbf{EQM}_{locale}(\theta_0) = \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0) (\hat{\theta}(\mathbf{y}) - \theta_0)^H p(\mathbf{y}|\theta_0) d\mathbf{y}, \quad (3.1)$$

où $\mathbf{y} \in \Omega$ est un vecteur d'observations, $p(\mathbf{y}|\theta_0)$ représente la vraisemblance des observations paramétrée par la vraie valeur du paramètre θ_0 dont $\hat{\theta}(\mathbf{y})$ est un estimateur.

D'un autre côté, les bornes bayésiennes sont des minorants de l'EQM globale, c'est-à-dire de

$$\mathbf{EQM}_{globale} = \int_{\Theta} \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta) (\hat{\theta}(\mathbf{y}) - \theta)^H p(\mathbf{y}, \theta) d\mathbf{y} d\theta, \quad (3.2)$$

où $\theta \in \Theta$ est le vecteur de paramètres aléatoires muni d'une densité de probabilité *a priori* $p(\theta) = \frac{p(\mathbf{y}, \theta)}{p(\mathbf{y}|\theta)}$ où $p(\mathbf{y}, \theta)$ est la densité de probabilité jointe des observations et des paramètres. Nous verrons dans le chapitre suivant comment les bornes bayésiennes peuvent être utilisées afin de décrire les performances d'un estimateur dans un contexte déterministe.

L'objectif de ce chapitre est de présenter une unification des bornes déterministes et des bornes bayésiennes de la famille Weiss-Weinstein à l'aide d'un même formalisme s'appuyant sur un problème d'optimisation sous contraintes. Cette unification permettra non seulement de mieux comprendre les hypothèses sous-jacentes à l'établissement de chaque borne mais nous servira également d'outil pour présenter deux nouvelles bornes minimales d'estimation : la borne d'Abel bayésienne et la borne de Weiss-Weinstein déterministe. Nous présenterons également une unification des bornes de la famille Ziv-Zakaï issues des travaux de Kristine Bell [Bel95]. Les résultats proposés dans ce chapitre sont essentiellement théoriques, c'est-à-dire qu'ils exposent les fondements sur lesquels reposent les bornes minimales d'estimation. Nous verrons dans le chapitre suivant comment ces bornes peuvent être utilisées d'un point de vue opérationnel.

Nous nous focaliserons sur l'étude d'un paramètre réel scalaire θ , l'extension au cas d'un vecteur de paramètres étant généralement triviale.

3.2 Un outil pour l'unification des bornes minimales

Théorème 3.1 : Soit $\mathbf{x} \in \mathcal{M}_{N \times 1}(\mathbb{R})$ un vecteur réel. Soient $f(\mathbf{x})$ et $g(\mathbf{x})$ deux fonctions de $\mathbb{R}^N \rightarrow \mathbb{R}$. Soit

$$\langle f(\mathbf{x}), g(\mathbf{x}) \rangle = \int_{\mathbb{R}^N} f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x}, \quad (3.3)$$

un produit scalaire de ces deux fonctions et sa norme associée $\|f(\mathbf{x})\|^2 = \langle f(\mathbf{x}), f(\mathbf{x}) \rangle$. Soit $u(\mathbf{x})$, $g_0(\mathbf{x}), \dots$, et $g_K(\mathbf{x})$, un jeu de fonctions de $\mathbb{R}^N \rightarrow \mathbb{R}$ et c_0, c_1, \dots , et c_K , $K + 1$ nombres réels. Alors, le minimum de $\|u(\mathbf{x})\|^2$ avec les $K + 1$ contraintes

$$\langle u(\mathbf{x}), g_k(\mathbf{x}) \rangle = c_k \quad \text{pour } k = 0, \dots, K, \quad (3.4)$$

est donné par

$$\begin{cases} \min_u \|u(\mathbf{x})\|^2 = \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c}, \\ \text{sujet à (3.4),} \end{cases} \quad (3.5)$$

avec

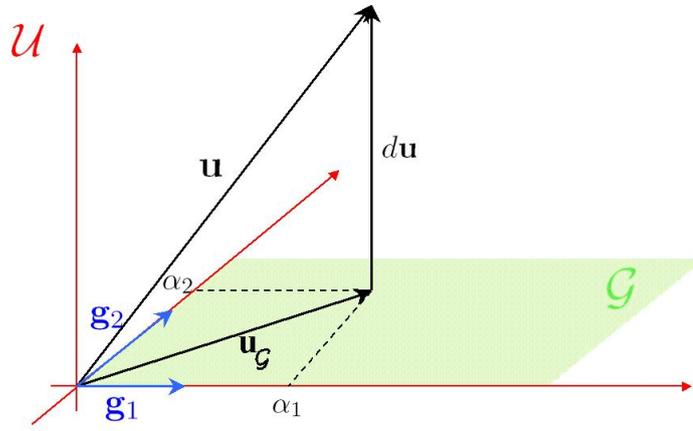
$$\mathbf{c} = [c_0 \ c_1 \ \cdots \ c_K]^T, \quad (3.6)$$

et

$$G_{m,n} = \langle g_m(\mathbf{x}), g_n(\mathbf{x}) \rangle. \quad (3.7)$$

Preuve : Cette démonstration est issue de [Cha04] page 83. Une autre démonstration utilisant les multiplicateurs de Lagrange est disponible dans [FL01] et [FL02].

Soit \mathcal{U} un espace vectoriel de dimension quelconque sur le corps des nombres réels \mathbb{R} muni d'un produit scalaire noté $\langle \mathbf{u}, \mathbf{w} \rangle$, où \mathbf{u} et \mathbf{w} sont deux vecteurs de \mathcal{U} . Soient $\{\mathbf{g}_1, \dots, \mathbf{g}_K\}$ une famille libre de K vecteurs de \mathcal{U} et $\mathbf{c} = [c_1 \ \cdots \ c_K]^T$ un vecteur de \mathbb{R}^K . Nous cherchons la solution du problème de minimisation de $\langle \mathbf{u}, \mathbf{u} \rangle$ sous les K contraintes linéaires $\langle \mathbf{u}, \mathbf{g}_k \rangle = c_k$, $k \in [1, K]$.



Appelons \mathcal{G} le sous-espace vectoriel de dimension K généré par les éléments $\{\mathbf{g}_1, \dots, \mathbf{g}_K\}$. Alors $\forall \mathbf{u} \in \mathcal{U}$, $\mathbf{u} = \mathbf{u}_G + d\mathbf{u}$ où \mathbf{u}_G est la projection orthogonale de \mathbf{u} sur \mathcal{G} , c'est à dire le vecteur $\mathbf{u}_G \in \mathcal{G}$ tel que $\langle \mathbf{u} - \mathbf{u}_G, \mathbf{g}_k \rangle = 0$, $k \in [1, K]$.

Soit $\alpha = [\alpha_1 \ \cdots \ \alpha_K]^T$ les coordonnées de \mathbf{u}_G dans la base $\{\mathbf{g}_1, \dots, \mathbf{g}_K\}$ de \mathcal{G} ($\mathbf{u}_G = \sum_{k=1}^K \alpha_k \mathbf{g}_k$), elles vérifient : $\langle \mathbf{u}, \mathbf{g}_k \rangle = \langle \mathbf{u}_G, \mathbf{g}_k \rangle$, $k \in [1, K]$. Si de plus \mathbf{u} satisfait les K contraintes $\langle \mathbf{u}, \mathbf{g}_k \rangle = c_k$, $k \in [1, K]$ alors

$$\langle \mathbf{u}, \mathbf{g}_k \rangle = c_k \Rightarrow \langle \mathbf{u}_G, \mathbf{g}_k \rangle = c_k \Rightarrow \left\langle \sum_{l=1}^K \alpha_l \mathbf{g}_l, \mathbf{g}_k \right\rangle = c_k \Rightarrow \sum_{l=1}^K \alpha_l \langle \mathbf{g}_l, \mathbf{g}_k \rangle = c_k, \quad (3.8)$$

c'est-à-dire sous forme matricielle $\mathbf{G}\alpha = \mathbf{c}$, où \mathbf{G} est la matrice de Gram associée à la famille $\{\mathbf{g}_1, \dots, \mathbf{g}_K\}$: $G_{k,l} = \langle \mathbf{g}_l, \mathbf{g}_k \rangle$. L'équation $\mathbf{G}\alpha = \mathbf{c}$ admet pour solution unique $\alpha = \mathbf{G}^{-1}\mathbf{c}$; notons alors $\mathbf{u}_{G,c}$ le vecteur de \mathcal{G} correspondant à cette solution. Par conséquent, $\forall \mathbf{u} \in \mathcal{U}$ et vérifiant les K contraintes susmentionnées alors $\langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{u}_{G,c}, \mathbf{u}_{G,c} \rangle + \langle d\mathbf{u}, d\mathbf{u} \rangle \geq \langle \mathbf{u}_{G,c}, \mathbf{u}_{G,c} \rangle$, et le minimum est atteint pour $d\mathbf{u} = 0$, ce qui signifie que $\mathbf{u}_{G,c}$ est la solution du problème, la norme minimale valant

$$\langle \mathbf{u}_{G,c}, \mathbf{u}_{G,c} \rangle = \left\langle \sum_{k=1}^K \alpha_k \mathbf{g}_k, \sum_{l=1}^K \alpha_l \mathbf{g}_l \right\rangle$$

$$\begin{aligned}
&= \sum_{k=1}^K \sum_{l=1}^K \alpha_k \alpha_l \langle \mathbf{g}_k, \mathbf{g}_l \rangle \\
&= \boldsymbol{\alpha}^T \mathbf{G} \boldsymbol{\alpha} \\
&= (\mathbf{G}^{-1} \mathbf{c})^T \mathbf{G} \mathbf{G}^{-1} \mathbf{c} \\
&= \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c}.
\end{aligned} \tag{3.9}$$

■

3.3 Unification des bornes déterministes

Dans toute cette section, $\mathbf{y} \in \Omega$ est le vecteur réel¹ des observations issues d'un processus aléatoire. $\Omega = \mathbb{R}^N$ représente l'espace des observations. $\theta \in \Theta$ est un paramètre candidat réel, déterministe et scalaire. La vraisemblance des observations est notée $p(\mathbf{y}|\theta_0)$. $\hat{\theta}(\mathbf{y})$ est un estimateur de la vraie valeur du paramètre θ_0 . Avec les notations du théorème 3.1, on pose $u(\mathbf{y}) = (\hat{\theta}(\mathbf{y}) - \theta_0) \sqrt{p(\mathbf{y}|\theta_0)}$ et $\mathbf{x} = \mathbf{y}$. En conséquence,

$$\|u(\mathbf{y})\|^2 = \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0)^2 p(\mathbf{y}|\theta_0) d\mathbf{y}, \tag{3.10}$$

qui est la définition de l'EQM locale en θ_0 . La solution (3.5) du théorème 3.1 donne l'EQM minimum en θ_0 qu'un estimateur vérifiant les contraintes de la forme (3.4) peut espérer atteindre. Nous allons voir dans la suite que les bornes déterministes sont obtenues comme solution de (3.5) lorsque les contraintes sont relatives au biais de l'estimateur.

3.3.1 Relations concernant le biais d'un estimateur

Dans le contexte déterministe, l'EQM locale se décompose en deux parties

$$EQM_{locale}(\theta_0) = \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0)^2 p(\mathbf{y}|\theta_0) d\mathbf{y} \tag{3.11}$$

$$= \mathbb{E}_{\mathbf{y}|\theta_0}^2 [\hat{\theta}(\mathbf{y}) - \theta_0] + \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\hat{\theta}(\mathbf{y}) - \mathbb{E}_{\mathbf{y}|\theta_0} [\hat{\theta}(\mathbf{y})] \right)^2 \right] \tag{3.12}$$

où le terme

$$\mathbb{E}_{\mathbf{y}|\theta_0}^2 [\hat{\theta}(\mathbf{y}) - \theta_0], \tag{3.13}$$

est le biais au carré de l'estimateur et où le terme

$$\mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\hat{\theta}(\mathbf{y}) - \mathbb{E}_{\mathbf{y}|\theta_0} [\hat{\theta}(\mathbf{y})] \right)^2 \right], \tag{3.14}$$

représente la variance de l'estimateur. L'une des qualités désirées pour un estimateur est un biais nul sur tout le support Θ . Pour se rapprocher de ceci, nous utiliserons les trois types de contraintes sur le biais de l'estimateur.

Un biais nul en la vraie valeur du paramètre θ_0 -ponctuellement sans biais en θ_0 -, c'est-à-dire

$$\mathbb{E}_{\mathbf{y}|\theta_0} [\hat{\theta}(\mathbf{y}) - \theta_0] = 0 \iff \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0) p(\mathbf{y}|\theta_0) d\mathbf{y} = 0, \tag{3.15}$$

1. L'extension a un vecteur d'observation complexe $\tilde{\mathbf{y}}$ est triviale en posant $\mathbf{y} = [\text{Re}^T \{\tilde{\mathbf{y}}\} \quad \text{Im}^T \{\tilde{\mathbf{y}}\}]^T$.

et/ou un biais nul en un point arbitraire $\theta_1 = \theta_0 + h$, c'est-à-dire

$$\begin{aligned} \mathbb{E}_{\mathbf{y}|\theta_1} [\hat{\theta}(\mathbf{y}) - \theta_1] = 0 &\iff \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0 - h) p(\mathbf{y}|\theta_0 + h) d\mathbf{y} = 0 \\ &\iff \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0) p(\mathbf{y}|\theta_0 + h) d\mathbf{y} = h, \end{aligned} \quad (3.16)$$

et/ou la dérivée $P^{\text{ième}}$ du biais nulle en la vraie valeur du paramètre, c'est-à-dire

$$\begin{aligned} \left. \frac{\partial^P \mathbb{E}_{\mathbf{y}|\theta} [\hat{\theta}(\mathbf{y}) - \theta]}{\partial \theta^P} \right|_{\theta_0} = 0 &\iff \left. \frac{\partial^P}{\partial \theta^P} \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta) p(\mathbf{y}|\theta) d\mathbf{y} \right|_{\theta_0} = 0 \\ &\iff \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0) \left. \frac{\partial^P p(\mathbf{y}|\theta)}{\partial \theta^P} \right|_{\theta_0} d\mathbf{y} = \delta_{1P}. \end{aligned} \quad (3.17)$$

3.3.2 Borne triviale

Imposons tout d'abord à $\hat{\theta}(\mathbf{y})$ d'avoir un biais nul en θ_0 . L'EQM minimale en θ_0 de $\hat{\theta}(\mathbf{y})$ notée $EQM_{triviale}(\theta_0)$ est donc donnée en utilisant (3.15) dans (3.4) et en appliquant le théorème 3.1 avec

$$\begin{cases} \mathbf{c} = 0, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)}, \end{cases} \quad (3.18)$$

conduisant à

$$\boxed{EQM_{triviale}(\theta_0) = \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c} = 0.} \quad (3.19)$$

On en déduit que l'EQM de tout estimateur sans biais est toujours positive, on le savait déjà...

3.3.3 Bornes de Cramér-Rao et de Bhattacharyya

Intéressons nous maintenant à un estimateur localement sans biais dans un voisinage de θ_0 . Ceci conduit dans un premier temps à imposer à $\hat{\theta}(\mathbf{y})$ un biais nul en θ_0 et une dérivée du biais nulle en θ_0 . Nous cherchons donc l'EQM minimale en θ_0 , notée $EQM_{CR}(\theta_0)$, d'un estimateur localement sans biais. Dans ce cas, l'utilisation de (3.15) et (3.17) à l'ordre un dans (3.4), c'est-à-dire

$$\begin{cases} \mathbf{c} = [0 \quad 1]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)}, \\ g_1(\mathbf{y}) = \frac{1}{\sqrt{p(\mathbf{y}|\theta_0)}} \left. \frac{\partial p(\mathbf{y}|\theta)}{\partial \theta} \right|_{\theta_0}, \end{cases} \quad (3.20)$$

et l'application du théorème 3.1 conduit à

$$EQM_{CR}(\theta_0) = [0 \quad 1] \left(\begin{array}{cc} 1 & 0 \\ 0 & \int_{\Omega} \left(\left. \frac{\partial \ln p(\mathbf{y}|\theta)}{\partial \theta} \right|_{\theta_0} \right)^2 p(\mathbf{y}|\theta_0) d\mathbf{y} \end{array} \right)^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (3.21)$$

$$\boxed{EQM_{CR}(\theta_0) = \left(\int_{\Omega} \left(\left. \frac{\partial \ln p(\mathbf{y}|\theta)}{\partial \theta} \right|_{\theta_0} \right)^2 p(\mathbf{y}|\theta_0) d\mathbf{y} \right)^{-1}}, \quad (3.22)$$

où l'on reconnaît la BCR [Fis22] [Dug37] [Fre43] [Dar45] [Rao45] [Cra46].

On peut maintenant étendre la notion d'estimateur localement sans biais. Imposons à cet effet à $\hat{\theta}(\mathbf{y})$ un biais nul en θ_0 et les K dérivées successives du biais nulle en θ_0 . Nous cherchons donc l'EQM minimale, notée $EQM_{Bhatt(K)}(\theta_0)$, d'un estimateur localement sans biais dans un voisinage de θ_0 croissant avec K . Dans ce cas, en utilisant (3.15) et (3.17) dans (3.4), c'est-à-dire

$$\begin{cases} \mathbf{c} = [0 & 1 & 0 & \cdots & 0]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)}, \\ g_k(\mathbf{y}) = \frac{1}{\sqrt{p(\mathbf{y}|\theta_0)}} \left. \frac{\partial^k p(\mathbf{y}|\theta)}{\partial \theta^k} \right|_{\theta_0} \quad k = 1, \dots, K, \end{cases} \quad (3.23)$$

et l'application du théorème 3.1 conduit à

$$EQM_{Bhatt(K)}(\theta_0) = [0 \quad 1 \quad 0 \quad \cdots \quad 0] \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & \mathbf{B} & & \\ 0 & & & \end{pmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (3.24)$$

$$EQM_{Bhatt(K)}(\theta_0) = [1 \quad 0 \quad \cdots \quad 0] \mathbf{B}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (3.25)$$

où $B_{i,j} = \int_{\Omega} \frac{1}{p(\mathbf{y}|\theta_0)} \left. \frac{\partial^i p(\mathbf{y}|\theta)}{\partial \theta^i} \right|_{\theta_0} \left. \frac{\partial^j p(\mathbf{y}|\theta)}{\partial \theta^j} \right|_{\theta_0} d\mathbf{y}$ avec $1 \leq i, j \leq K$. \mathbf{B} est généralement appelée

matrice de Bhattacharyya d'ordre K . En notant $\{\mathbf{B}^{-1}\}_{1,1}$ l'élément correspondant à la première ligne et à la première colonne de l'inverse de la matrice de Bhattacharyya d'ordre K on obtient la borne de Bhattacharyya [Bha46] [FG52] d'ordre K

$$\boxed{EQM_{Bhatt(K)}(\theta_0) = \{\mathbf{B}^{-1}\}_{1,1}} \quad (3.26)$$

De manière triviale

$$EQM_{Bhatt(K>1)}(\theta_0) \geq EQM_{Bhatt(1)}(\theta_0) = EQM_{CR}(\theta_0) \quad (3.27)$$

La borne de Cramér-Rao et la borne de Bhattacharyya sont également appelées bornes locales puisqu'elles n'autorisent une analyse de l'erreur que dans un voisinage de la vraie valeur du paramètre.

3.3.4 Bornes de Chapman-Robbins et de Barankin

Toujours dans le but d'obtenir l'EQM minimale d'un estimateur non-biaisé sur tout Θ , imposons maintenant à $\hat{\theta}(\mathbf{y})$ un biais nul en θ_0 et un biais nul en un autre point arbitraire $\theta_1 = \theta_0 + h$. Nous cherchons donc l'EQM minimale en θ_0 , notée $EQM_{ChR}(\theta_0, h)$, d'un estimateur ponctuellement sans biais en deux points. Dans ce cas, l'utilisation de (3.15) et (3.16) en un point dans (3.4), c'est-à-dire

$$\begin{cases} \mathbf{c} = [0 & h]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)}, \\ g_1(\mathbf{y}) = \frac{p(\mathbf{y}|\theta_0+h)}{\sqrt{p(\mathbf{y}|\theta_0)}}, \end{cases} \quad (3.28)$$

et l'application du théorème 3.1 conduit à

$$EQM_{ChR}(\theta_0, h) = \begin{bmatrix} 0 & h \end{bmatrix} \left(\begin{array}{c} 1 & 1 \\ 1 & \int_{\Omega} \frac{p^2(\mathbf{y}|\theta_0+h)}{p(\mathbf{y}|\theta_0)} d\mathbf{y} \end{array} \right)^{-1} \begin{bmatrix} 0 \\ h \end{bmatrix}, \quad (3.29)$$

$$\boxed{EQM_{ChR}(\theta_0, h) = \frac{h^2}{\int_{\Omega} \frac{p^2(\mathbf{y}|\theta_0+h)}{p(\mathbf{y}|\theta_0)} d\mathbf{y} - 1}}. \quad (3.30)$$

On retrouve la borne de Chapman-Robbins [Ham50] [CR51]. On peut remarquer que

$$EQM_{ChR}(\theta_0, h \neq 0) \geq \lim_{h \rightarrow 0} EQM_{ChR}(\theta_0, h) = EQM_{CR}(\theta_0). \quad (3.31)$$

L'intérêt de cette borne provient du fait que la valeur du paramètre h n'est pas imposée, c'est-à-dire que, contrairement à la borne de Cramér-Rao ou à la borne de Bhattacharyya, on ne reste pas dans un voisinage de θ_0 mais la contrainte peut se situer sur tout le support du paramètre.

Une extension naturelle, dans le but d'obtenir l'EQM minimale d'un estimateur non-biaisé sur tout Θ , revient à augmenter le nombre de points tests. Imposons maintenant à $\hat{\theta}(\mathbf{y})$ un biais nul en θ_0 et un biais nul en K autres points arbitraires $\theta_k = \theta_0 + h_k$ avec $k = 1, \dots, K$. Nous cherchons donc l'EQM minimale en θ_0 , notée $EQM_{Bar(K)}(\theta_0, \mathbf{h})$, d'un estimateur ponctuellement sans biais en $K + 1$ points. Dans ce cas, en posant $\mathbf{h} = [h_1 \ \dots \ h_K]^T$, l'utilisation de (3.15) et (3.17) dans (3.4), c'est-à-dire

$$\begin{cases} \mathbf{c} = [0 \ \mathbf{h}^T]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)}, \\ g_k(\mathbf{y}) = \frac{p(\mathbf{y}|\theta_0+h_k)}{\sqrt{p(\mathbf{y}|\theta_0)}}, \quad k = 1, \dots, K. \end{cases} \quad (3.32)$$

et l'application du théorème 3.1 conduit à

$$EQM_{Bar(K)}(\theta_0, \mathbf{h}) = \begin{bmatrix} 0 & \mathbf{h}^T \end{bmatrix} \left(\begin{array}{c} 1 & 1 & \dots & 1 \\ 1 & & & \\ \vdots & & \mathbf{D} & \\ 1 & & & \end{array} \right)^{-1} \begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix}, \quad (3.33)$$

$$\boxed{EQM_{Bar(K)}(\theta_0, \mathbf{h}) = \mathbf{h}^T (\mathbf{D} - \mathbf{1}\mathbf{1}^T)^{-1} \mathbf{h}}, \quad (3.34)$$

avec $D_{i,j} = \int_{\Omega} \frac{p(\mathbf{y}|\theta_0+h_i)p(\mathbf{y}|\theta_0+h_j)}{p(\mathbf{y}|\theta_0)} d\mathbf{y}$. \mathbf{D} est généralement appelée matrice de Barankin d'ordre

K . Cette borne présentée par McAulay et Seidman [MS69] est une approximation directe de la vraie borne de Barankin dont nous dirons quelques mots plus bas. $EQM_{Bar(K)}(\theta_0, \mathbf{h})$ est généralement appelée la borne de Barankin d'ordre K . Du fait de l'augmentation du nombre des contraintes dans l'application du théorème 3.1, nous obtenons la relation d'ordre suivante

$$EQM_{Bar(K>1)}(\theta_0, \mathbf{h}) \geq EQM_{Bar(1)}(\theta_0, h) = EQM_{ChR}(\theta_0, h). \quad (3.35)$$

Les bornes de Chapman-Robbins et de Barankin font intervenir un jeu de points tests qui peuvent couvrir l'ensemble du support du paramètre. De ce fait, elles sont également appelées bornes globales (à ne pas confondre avec le terme borne globale relatives à l'EQM globale intervenant dans les bornes bayésiennes) par opposition aux bornes de Cramér-Rao et de Bhattacharyya.

3.3.5 Borne d'Abel

Afin de combiner les caractéristiques des bornes locales et globales, Abel a proposé une borne hybride en utilisant l'inégalité de Ibragimov-Hasminskii ([IH81] ou [Leh83]). Ici nous proposons une nouvelle démonstration de cette borne en concaténant les contraintes de la bornes de Bhattacharyya d'ordre M et les contraintes de la borne de Barankin d'ordre R . C'est-à-dire, en concaténant (3.23) et (3.32) on obtient le nouveau jeu des $1 + M + R$ contraintes suivantes

$$\begin{cases} \mathbf{c} = [0 & 1 & 0 & \cdots & 0 & \mathbf{h}^T]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)}, \\ g_k(\mathbf{y}) = \frac{1}{\sqrt{p(\mathbf{y}|\theta_0)}} \left. \frac{\partial^k p(\mathbf{y}|\theta)}{\partial \theta^k} \right|_{\theta_0} & k = 1, \dots, M \\ g_k(\mathbf{y}) = \frac{p(\mathbf{y}|\theta_0 + h_k)}{\sqrt{p(\mathbf{y}|\theta_0)}} & k = M + 1, \dots, R. \end{cases} \quad (3.36)$$

L'application du théorème 3.1 conduit à l'EQM minimale en θ_0 notée $EQM_{Abel(M,R)}(\theta_0, \mathbf{h})$

$$EQM_{Abel(M,R)}(\theta_0, \mathbf{h}) = [0 \quad 1 \quad 0 \quad \cdots \quad 0 \quad \mathbf{h}^T] \mathbf{G}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ \mathbf{h} \end{bmatrix}, \quad (3.37)$$

avec

$$\mathbf{G} = \begin{pmatrix} 1 & \mathbf{0}^T & \mathbf{1}^T \\ \mathbf{0} & \mathbf{B} & \mathbf{\Gamma}^T \\ \mathbf{1} & \mathbf{\Gamma} & \mathbf{D} \end{pmatrix}, \quad (3.38)$$

où $\mathbf{0}$ est un vecteur colonne composé de M zéros, $\mathbf{1}$ est un vecteur colonne composé de R uns. Les matrices $\mathbf{B} \in \mathcal{M}_{M \times M}(\mathbb{R})$, $\mathbf{D} \in \mathcal{M}_{R \times R}(\mathbb{R})$ et $\mathbf{\Gamma} \in \mathcal{M}_{R \times M}(\mathbb{R})$ ont pour éléments

$$B_{i,j} = \int_{\Omega} \frac{1}{p(\mathbf{y}|\theta_0)} \left. \frac{\partial^i p(\mathbf{y}|\theta)}{\partial \theta^i} \right|_{\theta_0} \left. \frac{\partial^j p(\mathbf{y}|\theta)}{\partial \theta^j} \right|_{\theta_0} d\mathbf{y}, \quad (3.39)$$

$$D_{i,j} = \int_{\Omega} \frac{p(\mathbf{y}|\theta_0 + h_i) p(\mathbf{y}|\theta_0 + h_j)}{p(\mathbf{y}|\theta_0)} d\mathbf{y}, \quad (3.40)$$

$$\Gamma_{i,j} = \int_{\Omega} \frac{p(\mathbf{y}|\theta_0 + h_i)}{p(\mathbf{y}|\theta_0)} \left. \frac{\partial^j p(\mathbf{y}|\theta)}{\partial \theta^j} \right|_{\theta_0} d\mathbf{y}. \quad (3.41)$$

Soit $\boldsymbol{\beta} = [\mathbf{0}^T \quad \mathbf{1}^T]^T$, $\tilde{\mathbf{G}} = \begin{pmatrix} \mathbf{B} & \mathbf{\Gamma}^T \\ \mathbf{\Gamma} & \mathbf{D} \end{pmatrix}$, $\boldsymbol{\alpha} = [1 \quad 0 \quad \cdots \quad 0]^T$ et $\mathbf{c} = [0 \quad \boldsymbol{\alpha}^T \quad \mathbf{h}^T]^T$. Le vecteur $\boldsymbol{\alpha}$ est de taille $(M \times 1)$. Puisque le premier élément du vecteur \mathbf{c} est nul, seule la matrice $\mathbf{R} \in \mathcal{M}_{M+R \times M+R}(\mathbb{R})$ représentant le coin inférieur droit de la matrice \mathbf{G}^{-1} doit être étudiée. \mathbf{R} est donnée par (voir par exemple [Van02] équation A.68)

$$\mathbf{R} = \left(\tilde{\mathbf{G}} - \boldsymbol{\beta} \boldsymbol{\beta}^T \right)^{-1} = \begin{pmatrix} \mathbf{B} & \mathbf{\Gamma}^T \\ \mathbf{\Gamma} & \mathbf{D} - \mathbf{1} \mathbf{1}^T \end{pmatrix}^{-1}. \quad (3.42)$$

On posera $\mathbf{J} = \mathbf{D} - \mathbf{1} \mathbf{1}^T$. L'EQM minimale en θ_0 se réduit à

$$EQM_{Abel(M,R)}(\theta_0, \mathbf{h}) = [\boldsymbol{\alpha}^T \quad \mathbf{h}^T] \begin{pmatrix} \mathbf{B} & \mathbf{\Gamma}^T \\ \mathbf{\Gamma} & \mathbf{J} \end{pmatrix}^{-1} \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{h} \end{bmatrix}. \quad (3.43)$$

Après quelques efforts algébriques on obtient l'expression finale

$$\boxed{EQM_{Abel(M,R)}(\theta_0, \mathbf{h}) = \boldsymbol{\alpha}^T \mathbf{B}^{-1} \boldsymbol{\alpha} + \mathbf{u}^T \tilde{\mathbf{J}}^{-1} \mathbf{u},} \quad (3.44)$$

avec

$$\begin{cases} \mathbf{u} = \boldsymbol{\Gamma} \mathbf{B}^{-1} \boldsymbol{\alpha} - \mathbf{h}, \\ \tilde{\mathbf{J}} = \mathbf{J} - \boldsymbol{\Gamma} \mathbf{B}^{-1} \boldsymbol{\Gamma}^T. \end{cases} \quad (3.45)$$

On reconnaît la borne d'Abel [Abe90] [Abe93]. Du fait de l'augmentation du nombre des contraintes dans l'application du théorème 3.1, nous obtenons la relation d'ordre suivante

$$EQM_{Abel(M,R)}(\theta_0, \mathbf{h}) \geq \begin{cases} EQM_{Bhatt(M)}(\theta_0), \\ EQM_{Bar(R)}(\theta_0, \mathbf{h}). \end{cases} \quad (3.46)$$

3.3.6 Remarque

Toutes ces bornes sont des approximations de la "vraie" borne de Barankin, que nous noterons $EQM_{Barankin}(\theta_0)$, qui est l'EQM minimale en θ_0 d'un estimateur sans biais sur tout le support du paramètre. On obtient donc les relations suivantes

$$EQM_{Barankin}(\theta_0) = \begin{cases} \lim_{K \rightarrow \infty} EQM_{Bhatt(K)}(\theta_0), \\ \lim_{K \rightarrow \infty} \max_{\mathbf{h}} EQM_{Bar(K)}(\theta_0, \mathbf{h}), \\ \lim_{M \rightarrow \infty} \lim_{R \rightarrow \infty} \max_{\mathbf{h}} EQM_{Abel(M,R)}(\theta_0, \mathbf{h}). \end{cases} \quad (3.47)$$

Il est important de noter que la borne d'Abel converge plus rapidement vers $EQM_{Barankin}(\theta_0)$ par rapport aux bornes de Bhattacharyya et Barankin pour un de ses paramètres fixé (M, R). La pertinence des travaux de Barankin provient du fait que, dans le contexte de paramètres déterministes, la propriété la plus généralement attendue pour un estimateur est d'être sans biais sur tout le support du paramètre. Malheureusement, on peut démontrer de manière triviale que lorsque ce support du paramètre (continu) est borné, il n'existe pas d'estimateur prenant ces valeurs sur le même support borné à biais identiquement nul sur tout ce support, car la variance de celui-ci devrait être nulle aux extrémités. De plus, $EQM_{Barankin}(\theta_0)$ est la solution d'une équation intégrale et une forme analytique de cette solution n'existe que très rarement (voir, par exemple, [Mar97]). L'absence d'une telle expression analytique explique en partie la difficulté d'utilisation de la borne de Barankin en traitement du signal.

D'un point de vue opérationnel, nous ferons d'autres remarques concernant les contraintes mises en jeu dans l'établissement des bornes déterministes au chapitre suivant.

3.4 Unification des bornes bayésiennes

Les bornes bayésiennes donnent les performances ultimes des estimateurs en terme d'EQM globale (3.2). Dans l'approche bayésienne, les paramètres sont supposés aléatoires au travers d'une densité de probabilité *a priori* $p(\theta)$. Cette densité incorpore le support du paramètre contrairement aux bornes déterministes. Comme nous l'avons évoqué précédemment, les bornes bayésiennes se sub-divisent en deux catégories : les bornes de la famille Ziv-Zakaï qui dérivent d'un problème de test d'hypothèses binaires (et plus généralement d'un test d'hypothèses M -aires) et les bornes de la famille Weiss-Weinstein qui dérivent, comme les bornes déterministes, d'un principe d'inégalité de covariance.

3.4.1 La famille Ziv-Zakaï

Dans cette section, nous rappelons les résultats obtenus par Kristine Bell [Bel95] sur l'unification des bornes de la famille Ziv-Zakaï.

3.4.1.1 Borne de Bell

Les bornes de la famille Ziv-Zakaï se basent sur une réécriture de l'EQM globale et sur l'inégalité de Kotelnikov [Kot59]

$$\Pr \left(\text{abs} \left(\hat{\theta}(\mathbf{y}) - \theta \right) \geq \frac{\Delta}{2} \right) \geq \frac{2}{T_1 - T_0} \mathcal{V} \left\{ \int_{T_0}^{T_1 - \Delta} P_{\min}(\theta, \theta + \Delta) d\theta \right\}, \quad (3.48)$$

où θ est un paramètre scalaire aléatoire avec une densité *a priori* uniforme sur $[T_0, T_1]$, où $\mathcal{V}\{\}$ est la fonction "valley-filling" que nous détaillerons plus bas, et où $P_{\min}(\theta, \theta + \Delta)$ est la probabilité d'erreur minimale du test d'hypothèses binaires suivant

$$\begin{aligned} H_0 &: \theta; \Pr(H_0) = \frac{1}{2}; \mathbf{y} \sim p(\mathbf{y}|\theta) \\ H_1 &: \theta + \Delta; \Pr(H_1) = \frac{1}{2}; \mathbf{y} \sim p(\mathbf{y}|\theta + \Delta). \end{aligned} \quad (3.49)$$

Les travaux de Bell s'appuient sur une généralisation de cette inégalité. En effet, d'après [Cin75] page 24,

$$\begin{aligned} EQM_{\text{globale}} &= \int_{\Theta} \int_{\Omega} \left(\hat{\theta}(\mathbf{y}) - \theta \right)^2 p(\mathbf{y}, \theta) d\mathbf{y} d\theta \\ &= \int_0^{\infty} \frac{\Delta}{2} \Pr \left(\text{abs} \left(\hat{\theta}(\mathbf{y}) - \theta \right) \geq \frac{\Delta}{2} \right) d\Delta. \end{aligned} \quad (3.50)$$

On peut réécrire $\Pr \left(\text{abs} \left(\hat{\theta}(\mathbf{y}) - \theta \right) \geq \frac{\Delta}{2} \right)$ comme une somme de $M - 1$ termes identiques

$$\begin{aligned} \Pr \left(\text{abs} \left(\hat{\theta}(\mathbf{y}) - \theta \right) \geq \frac{\Delta}{2} \right) &= \frac{1}{M-1} \sum_{i=1}^{M-1} \left\{ \Pr \left(\hat{\theta}(\mathbf{y}) - \theta > \frac{\Delta}{2} \right) + \Pr \left(\hat{\theta}(\mathbf{y}) - \theta \leq -\frac{\Delta}{2} \right) \right\} \\ &= \frac{1}{M-1} \sum_{i=1}^{M-1} \left\{ \int_{-\infty}^{\infty} p(\varphi_{i-1}) \Pr \left(\hat{\theta}(\mathbf{y}) - \theta > \frac{\Delta}{2} \mid \theta = \varphi_{i-1} \right) d\varphi_{i-1} \right. \\ &\quad \left. + \int_{-\infty}^{\infty} p(\varphi_i) \Pr \left(\hat{\theta}(\mathbf{y}) - \theta \leq -\frac{\Delta}{2} \mid \theta = \varphi_i \right) d\varphi_i \right\} \\ &= \frac{1}{M-1} \sum_{i=1}^{M-1} \left\{ \int_{-\infty}^{\infty} p(\varphi_{i-1}) \Pr \left(\hat{\theta}(\mathbf{y}) > \varphi_{i-1} + \frac{\Delta}{2} \mid \theta = \varphi_{i-1} \right) d\varphi_{i-1} \right. \\ &\quad \left. + \int_{-\infty}^{\infty} p(\varphi_i) \Pr \left(\hat{\theta}(\mathbf{y}) \leq \varphi_i - \frac{\Delta}{2} \mid \theta = \varphi_i \right) d\varphi_i \right\}, \end{aligned} \quad (3.51)$$

et en posant $\varphi_0 = \varphi$ et $\varphi_i = \varphi + i\Delta$ pour $i = 1 \dots M - 1$,

$$\begin{aligned}
\Pr \left(\text{abs} \left(\hat{\theta}(\mathbf{y}) - \theta \right) \geq \frac{\Delta}{2} \right) &= \\
\frac{1}{M-1} \int_{-\infty}^{\infty} \sum_{i=1}^{M-1} &\left\{ p(\varphi + (i-1)\Delta) \Pr \left(\hat{\theta}(\mathbf{y}) > \varphi + \left(i - \frac{1}{2}\right)\Delta \middle| \theta = \varphi + (i-1)\Delta \right) \right. \\
&+ p(\varphi + i\Delta) \Pr \left(\hat{\theta}(\mathbf{y}) \leq \varphi + \left(i - \frac{1}{2}\right)\Delta \middle| \theta = \varphi + i\Delta \right) \left. \right\} d\varphi \\
&= \frac{1}{M-1} \int_{-\infty}^{\infty} \left(\sum_{n=0}^{M-1} p(\varphi + n\Delta) \right) \left[\frac{p(\varphi)}{\sum_{n=0}^{M-1} p(\varphi + n\Delta)} \Pr \left(\hat{\theta}(\mathbf{y}) > \varphi + \frac{\Delta}{2} \middle| \theta = \varphi \right) \right. \\
&+ \sum_{i=1}^{M-2} \frac{p(\varphi + i\Delta)}{\sum_{n=0}^{M-1} p(\varphi + n\Delta)} \left\{ \Pr \left(\hat{\theta}(\mathbf{y}) \leq \varphi + \left(i - \frac{1}{2}\right)\Delta \middle| \theta = \varphi + i\Delta \right) \right. \\
&\left. \left. + \Pr \left(\hat{\theta}(\mathbf{y}) > \varphi + \left(i + \frac{1}{2}\right)\Delta \middle| \theta = \varphi + i\Delta \right) \right\} \right. \\
&\left. + \frac{p(\varphi + (M-1)\Delta)}{\sum_{n=0}^{M-1} p(\varphi + n\Delta)} \Pr \left(\hat{\theta}(\mathbf{y}) \leq \varphi + \left(M - \frac{3}{2}\right)\Delta \middle| \theta = \varphi + (M-1)\Delta \right) \right] d\varphi. \quad (3.52)
\end{aligned}$$

On considère le test d'hypothèses M -aires suivant :

$$H_i : \theta = \varphi + i\Delta; \Pr(H_i) = \frac{p(\varphi + i\Delta)}{\sum_{n=0}^{M-1} p(\varphi + n\Delta)}; \mathbf{y} \sim p(\mathbf{y} | \theta = \varphi + i\Delta), \quad i = 0 \dots M - 1. \quad (3.53)$$

Alors le terme entre crochet dans (3.52) peut s'interpréter comme la probabilité d'erreur de ce test d'hypothèses lorsque la règle de décision est celle du "plus proche voisin", c'est-à-dire

$$\begin{cases} \text{On décide } H_0 \text{ si } \hat{\theta}(\mathbf{y}) \leq \varphi + \frac{\Delta}{2}, \\ \text{On décide } H_{i,i=1,\dots,M-2} \text{ si } \varphi + \left(i - \frac{1}{2}\right)\Delta < \hat{\theta}(\mathbf{y}) \leq \varphi + \left(i + \frac{1}{2}\right)\Delta, \\ \text{On décide } H_{M-1} \text{ si } \hat{\theta}(\mathbf{y}) > \varphi + \left(M - \frac{3}{2}\right)\Delta. \end{cases} \quad (3.54)$$

Cette règle de décision est sous-optimale, ainsi la probabilité d'erreur sera toujours supérieure ou égale à la probabilité d'erreur minimale d'un test d'hypothèses M -aires donné par le test du rapport de vraisemblance et notée $P_{\min}^{(M)}(\varphi, \dots, \varphi + (M-1)\Delta)$. On obtient

$$\Pr \left(\text{abs} \left(\hat{\theta}(\mathbf{y}) - \theta \right) \geq \frac{\Delta}{2} \right) \geq \frac{1}{M-1} \int_{-\infty}^{\infty} \left(\sum_{n=0}^{M-1} p(\varphi + n\Delta) \right) P_{\min}^{(M)}(\varphi, \dots, \varphi + (M-1)\Delta) d\varphi, \quad (3.55)$$

avec

$$\begin{aligned}
P_{\min}^{(M)}(\varphi, \dots, \varphi + (M-1)\Delta) &= \frac{1}{\sum_{n=0}^{M-1} p(\varphi + n\Delta)} \int_{-\infty}^{\infty} \min \{ p(\varphi) p(\mathbf{y} | \theta = \varphi), \dots \\
&\dots, p(\varphi + (M-1)\Delta) p(\mathbf{y} | \theta = \varphi + (M-1)\Delta) \} d\mathbf{y} \quad (3.56)
\end{aligned}$$

On peut remarquer que $\Pr\left(\text{abs}\left(\hat{\theta}(\mathbf{y}) - \theta\right) \geq \frac{\Delta}{2}\right)$ est une fonction monotone décroissante de Δ tandis que le terme de droite de l'équation (3.55) ne l'est pas forcément. En conséquence une borne plus précise de $\Pr\left(\text{abs}\left(\hat{\theta}(\mathbf{y}) - \theta\right) \geq \frac{\Delta}{2}\right)$ peut être obtenue en appliquant la fonction “valley-filling” représentée figure 3.1.

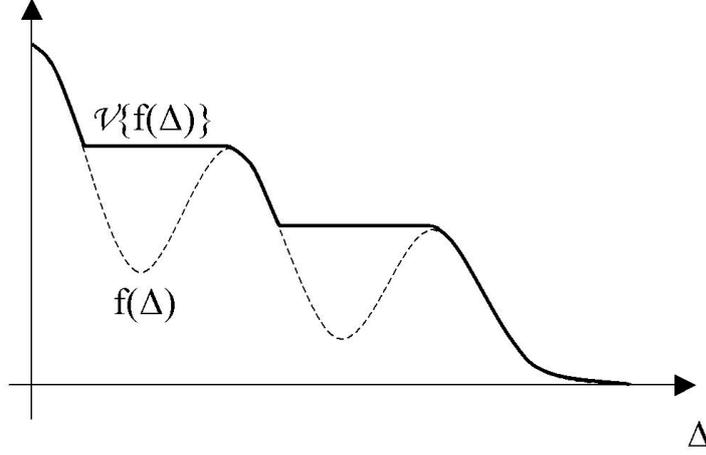


FIGURE 3.1 – Fonction Valley Filling.

On obtient l'inégalité de Kotelnikov généralisée

$$\Pr\left(\text{abs}\left(\hat{\theta}(\mathbf{y}) - \theta\right) \geq \frac{\Delta}{2}\right) \geq \mathcal{V}\left\{\frac{1}{M-1} \int_{-\infty}^{\infty} \left(\sum_{n=0}^{M-1} p(\varphi + n\Delta)\right) P_{\min}^{(M)}(\varphi, \dots, \varphi + (M-1)\Delta) d\varphi\right\}. \quad (3.57)$$

Finalement, en utilisant (3.57) dans (3.50), on obtient la borne de Bell [Bel95] à l'ordre M noté $EQM_{Bell}^{(M)}$

$$EQM_{Bell}^{(M)} = \int_0^{\infty} \frac{\Delta}{2} \mathcal{V}\left\{\frac{1}{M-1} \int_{-\infty}^{\infty} \left(\sum_{n=0}^{M-1} p(\varphi + n\Delta)\right) P_{\min}^{(M)}(\varphi, \dots, \varphi + (M-1)\Delta) d\varphi\right\} d\Delta. \quad (3.58)$$

Toutes les bornes de la famille Ziv-Zakai sont des cas particuliers de cette borne. D'un point de vue opérationnel, elle est difficilement applicable puisque le calcul analytique de $P_{\min}^{(M)}$ pour un modèle d'observation donné reste un problème non trivial. De plus, la fonction “valley-filling”, bien que mathématiquement séduisante, requiert une étude des variations d'une fonction complexe de Δ ce qui rend son utilisation difficile. Néanmoins, dans le cas de l'estimation d'un paramètre réel, scalaire, gaussien noyé additivement dans un bruit réel gaussien (indépendant du paramètre), il a été démontré [Bel95] que la borne de Bell est strictement égale à l'EQM de l'estimateur de la moyenne conditionnelle, qui, dans le contexte bayésien représente l'EQM minimale. Un problème théorique intéressant serait de démontrer que la borne de Bell tend vers l'EQM de l'estimateur de la moyenne conditionnelle lorsque $M \rightarrow \infty$.

3.4.1.2 Borne de Bell-Steinberg-Ephraim-VanTrees

La borne de Bell-Steinberg-Ephraim-VanTrees [BSET97], notée EQM_{BSET} , est obtenue en imposant $M = 2$, c'est-à-dire que l'on se réduit à un problème de test d'hypothèses binaires. Dans ce cas nous obtenons

$$EQM_{BSET} = \int_0^\infty \frac{\Delta}{2} \mathcal{V} \left\{ \int_{-\infty}^\infty (p(\varphi) + p(\varphi + \Delta)) P_{\min}^{(2)}(\varphi, \varphi + \Delta) d\varphi \right\} d\Delta, \quad (3.59)$$

avec

$$P_{\min}^{(2)}(\varphi, \varphi + \Delta) = \frac{1}{p(\varphi) + p(\varphi + \Delta)} \int_{-\infty}^\infty \min(p(\varphi) p(\mathbf{y}|\theta = \varphi), p(\varphi + \Delta) p(\mathbf{y}|\theta = \varphi + \Delta)) d\mathbf{y}, \quad (3.60)$$

la probabilité d'erreur minimale donnée par un test du rapport de vraisemblance pour le problème de détection suivant

$$\begin{cases} H_0 : \theta = \varphi; \Pr(H_0) = \frac{p(\varphi)}{p(\varphi) + p(\varphi + \Delta)}; \mathbf{y} \sim p(\mathbf{y}|\theta = \varphi), \\ H_1 : \theta = \varphi + \Delta; \Pr(H_1) = \frac{p(\varphi + \Delta)}{p(\varphi) + p(\varphi + \Delta)}; \mathbf{y} \sim p(\mathbf{y}|\theta = \varphi + \Delta). \end{cases} \quad (3.61)$$

3.4.1.3 Borne de Bellini-Tartara

A partir de la borne de Bell-Steinberg-Ephraim-VanTrees, si l'on suppose θ distribué uniformément sur $[0, T]$, on obtient la borne de Bellini-Tartara [BT74] notée EQM_{BT}

$$EQM_{BT} = \int_0^T \Delta \mathcal{V} \left\{ \int_0^{T-\Delta} \frac{1}{T} P_{\min}^{(2)}(\varphi, \varphi + \Delta) d\varphi \right\} d\Delta. \quad (3.62)$$

3.4.1.4 Borne de Chazan-Zakaï-Ziv

A partir de la borne de Bellini-Tartara, si l'on omet la fonction "valley-filling", on obtient la borne de Chazan-Zakaï-Ziv [CZZ75] notée EQM_{CZZ}

$$EQM_{CZZ} = \frac{1}{T} \int_0^T \Delta \int_0^{T-\Delta} P_{\min}^{(2)}(\varphi, \varphi + \Delta) d\varphi d\Delta. \quad (3.63)$$

3.4.1.5 Borne de Weinstein

La borne de Weinstein [Wei88], notée EQM_W , est une approximation de la borne de Bellini-Tartara où l'intégrale sur Δ est remplacée par une somme discrète

$$EQM_W = \max_{\{\Delta_i\}_{i=1 \dots K}} \sum_{i=1}^K \frac{\Delta_i^2 - \Delta_{i-1}^2}{2T} \int_0^{T-\Delta_i} P_{\min}^{(2)}(\varphi, \varphi + \Delta_i) d\varphi, \quad (3.64)$$

où $0 = \Delta_0 < \Delta_1 < \dots < \Delta_K \leq T$.

3.4.1.6 Remarques concernant les bornes de la famille Ziv-Zakai

On peut démontrer [Bel95] la relation d'ordre suivante

$$\boxed{EQM_{Bell}^{(M)} \geq EQM_{BSET} \geq EQM_{BT} \geq \begin{cases} EQM_{CZZ} \\ EQM_W \end{cases}} \quad (3.65)$$

Le point critique dans le calcul des bornes de la famille Ziv-Zakai est l'évaluation de P_{\min} pour lequel une expression analytique est rarement disponible même dans le cas général d'observations gaussiennes. L'approche généralement retenue consiste à approximer cette probabilité d'erreur ce qui conduit à des bornes moins pertinentes. L'approximation la plus connue consiste à utiliser l'inégalité de Pierce [Pie63] qui donne l'encadrement suivant :

$$\Phi \left(\frac{1}{2} \frac{\partial^2 \mu(s)}{\partial s^2} \Big|_{s=\frac{1}{2}} \right) e^{\mu(\frac{1}{2}) + \frac{1}{8} \frac{\partial^2 \mu(s)}{\partial s^2} \Big|_{s=\frac{1}{2}}} \leq P_{\min}^{(2)} \leq \frac{e^{\mu(\frac{1}{2})}}{2 \left(1 + \sqrt{\frac{1}{8} \frac{\partial^2 \mu(s)}{\partial s^2} \Big|_{s=\frac{1}{2}}} \right)}, \quad (3.66)$$

où $\mu(s)$ est la fonction génératrice des moments semi-invariants

$$\mu(s) = \ln \int_{-\infty}^{\infty} p^{1-s}(\mathbf{y} | H_0) p^s(\mathbf{y} | H_1) d\mathbf{y}, \quad (3.67)$$

et où

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-\frac{t^2}{2}} dt, \quad (3.68)$$

est la fonction d'erreur complémentaire. L'intérêt est ici qu'il existe plusieurs résultats concernant $\mu(s)$ dans la littérature (voir, en particulier, [Van01]). De plus, même s'il semble plus rigoureux d'utiliser la borne inférieure de P_{\min} , cet encadrement est suffisamment précis pour que l'utilisation de la borne supérieure ne donne pas une borne exagérément optimiste. On peut également utiliser l'inégalité de Bhattacharyya [Bha43]

$$P_{\min}^{(2)} \geq \Pr(H_0) \Pr(H_1) e^{2\mu(\frac{1}{2})}, \quad (3.69)$$

beaucoup plus simple que l'inégalité de Pierce mais également plus grossière. On pourra également se reporter à l'inégalité de Chernoff [Che52] ou de Shannon-Gallager-Berlekamp [SGB67].

Les bornes de la famille Ziv-Zakai ont été appliquées avec succès à divers problèmes d'estimation tels que l'estimation d'un retard [ZZ69] [WW83], l'estimation de directions d'arrivées [BET96], en acoustique sous-marine [XBR04], ainsi qu'en communications numériques [CG05]. Le texte de référence en la matière reste de loin la thèse de Kristine Bell [Bel95].

3.4.2 La famille Weiss-Weinstein

La famille Weiss-Weinstein repose sur un principe d'inégalité de covariance et de ce fait semble plus naturelle que les bornes de famille Ziv-Zakai pour lesquelles une réécriture astucieuse de l'EQM est nécessaire. Dans cette section, à l'aide du théorème 3.1, nous présentons une unification originale des bornes de la famille Weiss-Weinstein. Il est bon de noter qu'une première approche

pour l'unification des bornes de la famille Weiss-Weinstein à déjà été proposée par Weiss et Weinstein dans [WW88]. Cette approche se fonde sur l'inégalité suivante démontrée dans [WW88]

$$EQM_{globale} \geq \frac{\mathbb{E}_{\mathbf{y},\theta}^2 [\theta \psi(\mathbf{y}, \theta)]}{\mathbb{E}_{\mathbf{y},\theta} [\psi^2(\mathbf{y}, \theta)]}, \quad (3.70)$$

où l'espérance mathématique porte sur la loi jointe $p(\mathbf{y}, \theta)$ et où la fonction $\psi(\mathbf{y}, \theta)$ doit vérifier

$$\int_{\Theta} \psi(\mathbf{y}, \theta) p(\mathbf{y}, \theta) d\theta = 0. \quad (3.71)$$

Weiss et Weinstein donnent plusieurs fonctions $\psi(\mathbf{y}, \theta)$ vérifiant (3.71) pour lesquelles on retrouve la borne de Cramér-Rao bayésienne, la borne de Bhattacharyya bayésienne, la borne de Bobrovsky-Zakaï et la borne de Weiss-Weinstein. Malheureusement, il n'existe pas de règle générale pour trouver $\psi(\mathbf{y}, \theta)$ et cette étude ne permet pas de déduire systématiquement une relation d'ordre entre les différentes bornes. Ici, l'unification de la famille Weiss-Weinstein s'appuie sur la meilleure borne bayésienne, à savoir l'EQM de l'estimateur de la moyenne conditionnelle. En réécrivant cette EQM et en la comparant à l'EQM solution du théorème 3.1, nous unifions la borne de Cramér-Rao bayésienne, la borne de Bhattacharyya bayésienne, la borne de Bobrovsky-Zakaï, la borne de Reuven-Messer, la borne de Bobrovsky-MayerWolf-Zakaï et la borne de Weiss-Weinstein. L'intérêt de cette approche provient du fait qu'elle nous permet d'obtenir des relations d'ordre entre les différentes bornes proposées puisque nous les comparons directement à la borne ultime. De plus, cette approche offre un cadre théorique puissant pour établir de nouvelles bornes bayésiennes et déterministes. Ainsi, nous proposons deux nouvelles bornes qui seront appliquées à un problème d'estimation au chapitre suivant : la version bayésienne de la borne d'Abel et la version déterministe de la borne de Weiss-Weinstein.

3.4.2.1 Reformulations de l'EQM minimale

Dans le contexte bayésien, l'EQM globale minimale est obtenue par l'estimateur de la moyenne conditionnelle $\hat{\theta}(\mathbf{y}) = \int_{\Theta} \theta p(\theta|\mathbf{y}) d\theta$. L'obtention d'une forme analytique de cette EQM est un problème généralement non-trivial. L'estimateur de la moyenne conditionnelle est solution du problème de minimisation suivant

$$\min_{\hat{\theta}(\mathbf{y})} \int_{\Omega} \int_{\Theta} (\hat{\theta}(\mathbf{y}) - \theta)^2 p(\mathbf{y}, \theta) d\theta d\mathbf{y}. \quad (3.72)$$

On peut réécrire ce problème sous la forme d'un problème d'optimisation sous contraintes

$$\begin{cases} \min_v \int_{\Omega} \int_{\Theta} v^2(\mathbf{y}, \theta) p(\mathbf{y}, \theta) d\theta d\mathbf{y} \\ \text{sujet à } v(\mathbf{y}, \theta) = \hat{\theta}(\mathbf{y}) - \theta \end{cases} \quad (3.73)$$

D'un autre côté, en posant $\mathbf{x} = [\mathbf{y}^T \ \theta]^T$ et $v(\mathbf{y}, \theta) = \frac{u(\mathbf{y}, \theta)}{\sqrt{p(\mathbf{y}, \theta)}}$, le théorème 3.1 donne la solution du problème d'optimisation sous contraintes suivant

$$\begin{cases} \min_v \int_{\Omega} \int_{\Theta} v^2(\mathbf{y}, \theta) p(\mathbf{y}, \theta) d\theta d\mathbf{y} \\ \text{sujet à } \int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) g_k(\mathbf{y}, \theta) \sqrt{p(\mathbf{y}, \theta)} d\theta d\mathbf{y} = c_k \quad k = 0, \dots, K \end{cases} \quad (3.74)$$

Nous allons montrer que l'unique contrainte du problème d'optimisation (3.73) est équivalente à un continuum de contraintes particulières du problème d'optimisation (3.74).

Soit $f(\mathbf{y}, \theta)$ une fonction telle que

$$\lim_{\theta \rightarrow \pm\infty} v(\mathbf{y}, \theta) f(\mathbf{y}, \theta) = 0, \quad (3.75)$$

et soient les quatre propositions suivantes

$P1 :$

$$\forall f(\mathbf{y}, \theta), \quad \int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) \frac{\partial f(\mathbf{y}, \theta)}{\partial \theta} d\theta d\mathbf{y} = \int_{\Omega} \int_{\Theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y}. \quad (3.76)$$

Concernant $P2$ et $P3$, on supposera également $\int_{\Omega} \int_{\Theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y} = 1$

$P2 :$

$$\forall f(\mathbf{y}, \theta) \text{ et } \forall h, \quad \int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) (f(\mathbf{y}, \theta + h) - f(\mathbf{y}, \theta)) d\theta d\mathbf{y} = h. \quad (3.77)$$

$P3 :$

$\forall f(\mathbf{y}, \theta), \forall h$ et $\forall s,$

$$\begin{aligned} \int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) [L^s(\mathbf{y}, \theta + h) - L^{1-s}(\mathbf{y}, \theta - h)] f(\mathbf{y}, \theta) d\theta d\mathbf{y} = \\ h \int_{\Omega} \int_{\Theta} L^{1-s}(\mathbf{y}, \theta - h) f(\mathbf{y}, \theta) d\theta d\mathbf{y}, \end{aligned} \quad (3.78)$$

avec $L(\mathbf{y}, \eta) = \frac{f(\mathbf{y}, \eta)}{f(\mathbf{y}, \theta)}$ et $s \in [0, 1]$.

$P4 :$

$$v(\mathbf{y}, \theta) = q(\mathbf{y}) - \theta, \quad (3.79)$$

où $q(\mathbf{y})$ est une fonction de \mathbf{y} seulement.

Théorème 3.2 :

$$\boxed{P1 \Leftrightarrow P2 \Leftrightarrow P3 \Leftrightarrow P4.} \quad (3.80)$$

Preuve : la démonstration du théorème 3.2 se décompose en trois parties

Lemme 3.1 :

$$P1 \Leftrightarrow P4 \quad (3.81)$$

Preuve : supposons que $\forall f(\mathbf{y}, \theta), \int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) \frac{\partial f(\mathbf{y}, \theta)}{\partial \theta} d\theta d\mathbf{y} = \int_{\Omega} \int_{\Theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y}$, donc

$$\forall f(\mathbf{y}, \theta), \quad \int_{\Omega} \int_{\Theta} \frac{\partial v(\mathbf{y}, \theta)}{\partial \theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y} = - \int_{\Omega} \int_{\Theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y}, \quad (3.82)$$

puisque

$$\int_{\Theta} v(\mathbf{y}, \theta) \frac{\partial f(\mathbf{y}, \theta)}{\partial \theta} d\theta = \underbrace{\int_{\Theta} \frac{\partial v(\mathbf{y}, \theta)}{\partial \theta} f(\mathbf{y}, \theta) d\theta}_{=0} - \int_{\Theta} \frac{\partial v(\mathbf{y}, \theta)}{\partial \theta} f(\mathbf{y}, \theta) d\theta, \quad (3.83)$$

en utilisant (3.75).

L'expression (3.82) est valable pour toute fonction $f(\mathbf{y}, \theta)$, donc si on choisit $f(\mathbf{y}, \theta) = \delta(\mathbf{y} - \mathbf{y}_0, \theta - \theta_0)$, où $\delta(\cdot)$ est la distribution de Dirac, on obtient $\forall \mathbf{y}_0$ et $\forall \theta_0$

$$\left. \frac{\partial v(\mathbf{y}, \theta)}{\partial \theta} \right|_{\mathbf{y}_0, \theta_0} = -1 \implies v(\mathbf{y}, \theta) = q(\mathbf{y}) - \theta, \quad (3.84)$$

où $q(\mathbf{y})$ est une fonction de \mathbf{y} seulement.

D'un autre côté, si on suppose que $v(\mathbf{y}, \theta) = q(\mathbf{y}) - \theta$, alors $\forall f(\mathbf{y}, \theta)$

$$\begin{aligned} \int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) \frac{\partial f(\mathbf{y}, \theta)}{\partial \theta} d\theta d\mathbf{y} &= \int_{\Omega} \int_{\Theta} (q(\mathbf{y}) - \theta) \frac{\partial f(\mathbf{y}, \theta)}{\partial \theta} d\theta d\mathbf{y} \\ &= \int_{\Omega} \int_{\Theta} \frac{\partial (q(\mathbf{y}) - \theta) f(\mathbf{y}, \theta)}{\partial \theta} + f(\mathbf{y}, \theta) d\theta d\mathbf{y} \\ &= \int_{\Omega} \int_{\Theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y}, \end{aligned} \quad (3.85)$$

ce qui démontre le lemme 3.1. ■

Lemme 3.2 :

$$P2 \Leftrightarrow P4 \quad (3.86)$$

Preuve : supposons que $\forall f(\mathbf{y}, \theta)$ et $\forall h$,

$$\int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) (f(\mathbf{y}, \theta + h) - f(\mathbf{y}, \theta)) d\theta d\mathbf{y} = h. \quad (3.87)$$

Alors, lorsque $h \rightarrow 0$, (3.87) implique que

$$\int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) \frac{\partial f(\mathbf{y}, \theta)}{\partial \theta} d\theta d\mathbf{y} = 1 \implies v(\mathbf{y}, \theta) = q(\mathbf{y}) - \theta, \quad (3.88)$$

où la dernière implication est un résultat intermédiaire de la démonstration du lemme 3.1.

D'un autre côté, si on suppose $v(\mathbf{y}, \theta) = q(\mathbf{y}) - \theta$, alors en posant $\varphi = \theta + h$

$$\begin{aligned} \int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) f(\mathbf{y}, \theta + h) d\theta d\mathbf{y} &= \int_{\Omega} \int_{\Theta} (q(\mathbf{y}) - \theta) f(\mathbf{y}, \theta + h) d\theta d\mathbf{y} \\ &= \int_{\Omega} \int_{\Theta} (q(\mathbf{y}) - \varphi + h) f(\mathbf{y}, \varphi) d\varphi d\mathbf{y} \\ &= \int_{\Omega} \int_{\Theta} (q(\mathbf{y}) - \varphi) f(\mathbf{y}, \varphi) d\varphi d\mathbf{y} + h, \end{aligned} \quad (3.89)$$

conduisant à

$$\int_{\Omega} \int_{\Theta} (q(\mathbf{y}) - \theta) (f(\mathbf{y}, \theta + h) - f(\mathbf{y}, \theta)) d\theta d\mathbf{y} = h, \quad (3.90)$$

ce qui démontre le lemme 3.2. ■

Lemme 3.3 :

$$P3 \Leftrightarrow P4 \quad (3.91)$$

Preuve : en posant $L(\mathbf{y}, \eta) = \frac{f(\mathbf{y}, \eta)}{f(\mathbf{y}, \theta)}$ supposons que $\forall f(\mathbf{y}, \theta), \forall h$ et $\forall s$,

$$\int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) [L^s(\mathbf{y}, \theta + h) - L^{1-s}(\mathbf{y}, \theta - h)] f(\mathbf{y}, \theta) d\theta d\mathbf{y} = h \int_{\Omega} \int_{\Theta} L^{1-s}(\mathbf{y}, \theta - h) f(\mathbf{y}, \theta) d\theta d\mathbf{y}, \quad (3.92)$$

Alors, lorsque $s \rightarrow 1$, (3.92) implique

$$\int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) (f(\mathbf{y}, \theta + h) - f(\mathbf{y}, \theta)) d\theta d\mathbf{y} = h \implies v(\mathbf{y}, \theta) = q(\mathbf{y}) - \theta, \quad (3.93)$$

où la dernière implication est un résultat intermédiaire de la démonstration du lemme 3.2.

D'un autre côté, si on suppose $v(\mathbf{y}, \theta) = q(\mathbf{y}) - \theta$, alors en posant $\varphi = \theta + h$

$$\int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) L^s(\mathbf{y}, \theta + h) f(\mathbf{y}, \theta) d\theta d\mathbf{y} = \int_{\Omega} \int_{\Theta} (q(\mathbf{y}) - \varphi) L^{1-s}(\mathbf{y}, \varphi - h) f(\mathbf{y}, \varphi) d\varphi d\mathbf{y} + h \int_{\Omega} \int_{\Theta} L^{1-s}(\mathbf{y}, \varphi - h) f(\mathbf{y}, \varphi) d\varphi d\mathbf{y}, \quad (3.94)$$

conduisant à

$$\int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) [L^s(\mathbf{y}, \theta + h) - L^{1-s}(\mathbf{y}, \theta - h)] f(\mathbf{y}, \theta) d\theta d\mathbf{y} = h \int_{\Omega} \int_{\Theta} L^{1-s}(\mathbf{y}, \theta - h) f(\mathbf{y}, \theta) d\theta d\mathbf{y}, \quad (3.95)$$

ce qui démontre le lemme 3.3. ■

Les lemmes 3.1, 3.2, et 3.3 démontrent le théorème 3.2. ■

3.4.2.2 Application aux bornes bayésiennes

Le théorème 3.2 établit l'équivalence entre l'unique contrainte du problème d'optimisation (3.73) et un continuum de contraintes particulières. Ce continuum n'est pas utilisable en l'état dans le théorème 3.1, puisque par définition, il correspond à une infinité de contraintes qui conduirait à un vecteur \mathbf{c} et une matrice \mathbf{G} de dimensions infinies. Néanmoins, en réduisant ce continuum, c'est-à-dire en choisissant des valeurs particulières pour $f(\mathbf{y}, \theta)$, h , et s en nombre fini, on obtiendra, par le théorème 3.1, une EQM minimale globale inférieure à l'EQM de l'estimateur de la moyenne conditionnelle, donc, par définition, à des bornes minimales d'estimation. Le choix des paramètres $f(\mathbf{y}, \theta)$, h , et s donne une appréciation de la "distance" à l'EQM de l'estimateur de la moyenne conditionnelle, c'est-à-dire que plus on choisira de valeurs, plus on se rapprochera du continuum et donc de la meilleure borne minimale. De plus, dû fait de la restriction à des valeurs particulières de $f(\mathbf{y}, \theta)$, h , et s , il nous est permis de rajouter des contraintes sur les connaissances *a priori* concernant l'estimateur de la moyenne conditionnelle afin d'obtenir des bornes plus pertinentes. Ici, nous utiliserons la contrainte naturelle de biais nul au sens de la loi jointe.

3.4.2.2.1 Borne de Cramér-Rao bayésienne et borne de Bhattacharyya bayésienne

On suppose ici $f(\mathbf{x}, \theta) = p(\mathbf{x}, \theta)$, la densité jointe du problème. D'où $\int_{\Omega} \int_{\Theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y} = 1$.

En utilisant l'expression $P1$ on obtient le jeu des deux contraintes suivantes

$$\begin{cases} \mathbf{c} = [0 \ 1]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)}, \\ g_1(\mathbf{y}) = \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \frac{\partial}{\partial \theta} p(\mathbf{y}, \theta). \end{cases} \quad (3.96)$$

L'application du théorème 3.1 conduit à l'EQM globale minimale notée $EQM_{CR}^{globale}$

$$EQM_{CR}^{globale} = [0 \ 1] \begin{pmatrix} 1 & 0 \\ 0 & \int_{\Omega} \int_{\Theta} \left(\frac{\partial}{\partial \theta} \ln p(\mathbf{y}, \theta) \right)^2 p(\mathbf{y}, \theta) d\theta d\mathbf{y} \end{pmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (3.97)$$

$$\boxed{EQM_{CR}^{globale} = \left(\int_{\Omega} \int_{\Theta} \left(\frac{\partial \ln p(\mathbf{y}, \theta)}{\partial \theta} \right)^2 p(\mathbf{y}, \theta) d\theta d\mathbf{y} \right)^{-1}}. \quad (3.98)$$

Il s'agit de la borne de Cramér-Rao bayésienne [Van68].

On peut étendre la contrainte $P1$ par dérivation successive de l'expression

$$\int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) (f(\mathbf{y}, \theta + h) - f(\mathbf{y}, \theta)) d\theta d\mathbf{y} = h, \quad (3.99)$$

utilisée dans $P2$ par rapport à h en $h = 0$, on obtient

$$\int_{\Omega} \int_{\Theta} v(\mathbf{y}, \theta) \frac{\partial^k f(\mathbf{y}, \theta)}{\partial \theta^k} d\theta d\mathbf{x} = \delta_{1k}. \quad (3.100)$$

Nous allons donc utiliser le jeu des $K + 1$ contraintes suivantes

$$\begin{cases} \mathbf{c} = [0 \ 1 \ 0 \ \dots \ 0]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)}, \\ g_k(\mathbf{y}) = \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \frac{\partial^k}{\partial \theta^k} p(\mathbf{y}, \theta) \quad k = 1, \dots, K. \end{cases} \quad (3.101)$$

L'application du théorème 3.1 conduit à l'EQM globale minimale notée $EQM_{Bhatt(K)}^{globale}$

$$EQM_{Bhatt(K)}^{globale} = [0 \ 1 \ 0 \ \dots \ 0] \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & \mathbf{B} & & \\ 0 & & & \end{pmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (3.102)$$

$$EQM_{Bhatt(K)}^{globale} = [1 \ 0 \ \dots \ 0] \mathbf{B}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (3.103)$$

où $B_{i,j} = \int \int_{\Omega \Theta} \frac{1}{p(\mathbf{y}, \theta)} \frac{\partial^i}{\partial \theta^i} p(\mathbf{y}, \theta) \frac{\partial^j}{\partial \theta^j} p(\mathbf{y}, \theta) d\theta d\mathbf{y}$. En notant $\{\mathbf{B}^{-1}\}_{1,1}$ l'élément correspondant à la première ligne et à la première colonne de l'inverse de \mathbf{B} , on obtient la borne de Bhattacharyya bayésienne [Van68] d'ordre K

$$\boxed{EQM_{Bhatt(K)}^{globale} = \{\mathbf{B}^{-1}\}_{1,1}}. \quad (3.104)$$

3.4.2.2 Borne de Bobrovsky-MayerWolf-Zakaï On suppose ici $f(\mathbf{x}, \theta) = q(\mathbf{x}, \theta) p(\mathbf{x}, \theta)$, où $p(\mathbf{x}, \theta)$ est la densité jointe du problème et $q(\mathbf{x}, \theta)$ est une fonction vérifiant (3.75).

En utilisant directement l'expression $P1$, on obtient le jeu de deux contraintes suivantes

$$\begin{cases} \mathbf{c} = \left[0 \quad \int \int_{\Omega \Theta} q(\mathbf{x}, \theta) p(\mathbf{x}, \theta) d\theta d\mathbf{y} \right]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)}, \\ g_1(\mathbf{y}) = \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \frac{\partial p(\mathbf{y}, \theta) q(\mathbf{x}, \theta)}{\partial \theta}. \end{cases} \quad (3.105)$$

L'application du théorème 3.1 conduit à l'EQM globale minimale notée $EQM_{BMZ}^{globale}(q(\mathbf{x}, \theta))$

$$EQM_{BMZ}^{globale}(q(\mathbf{x}, \theta)) = \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c}, \quad (3.106)$$

avec

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ 0 & \int \int_{\Omega \Theta} \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \frac{\partial p(\mathbf{y}, \theta) q(\mathbf{x}, \theta)}{\partial \theta} d\theta d\mathbf{y} \end{pmatrix}. \quad (3.107)$$

Finalement,

$$\boxed{EQM_{BMZ}^{globale}(q(\mathbf{x}, \theta)) = \frac{\left(\int \int_{\Omega \Theta} q(\mathbf{x}, \theta) p(\mathbf{x}, \theta) d\theta d\mathbf{y} \right)^2}{\int \int_{\Omega \Theta} \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \left(\frac{\partial p(\mathbf{y}, \theta) q(\mathbf{x}, \theta)}{\partial \theta} \right)^2 d\theta d\mathbf{y}}}. \quad (3.108)$$

Il s'agit ici de la borne de Bobrovsky-MayerWolf-Zakaï [BMWZ87] qui constitue une extension généralisée de la borne de Cramér-Rao bayésienne puisque

$$EQM_{BMZ}^{global}(1) = EQM_{CR}^{globale}. \quad (3.109)$$

3.4.2.2.3 Borne de Bobrovsky-Zakaï et borne de Reuven-Messer On suppose ici $f(\mathbf{x}, \theta) = p(\mathbf{x}, \theta)$, la densité jointe du problème. Donc la contrainte supplémentaire sur $P2$, à savoir $\int \int_{\Omega \Theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y} = 1$ est satisfaite.

En utilisant directement l'expression $P2$, on obtient le jeu de deux contraintes suivantes

$$\begin{cases} \mathbf{c} = [0 \quad h]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)}, \\ g_1(\mathbf{y}) = \frac{p(\mathbf{y}, \theta + h)}{\sqrt{p(\mathbf{y}, \theta)}}. \end{cases} \quad (3.110)$$

L'application du théorème 3.1 conduit à l'EQM globale minimale notée $EQM_{BZ}^{globale}(h)$

$$\begin{aligned} EQM_{BZ}^{globale}(h) &= [0 \ h] \left(\begin{array}{c} 1 \\ 1 \int \int_{\Omega \Theta} \frac{1}{p(\mathbf{y}, \theta+h)} d\theta d\mathbf{y} \end{array} \right)^{-1} \begin{bmatrix} 0 \\ h \end{bmatrix} \\ &= \frac{h^2}{\int \int_{\Omega \Theta} \frac{p^2(\mathbf{y}, \theta+h)}{p(\mathbf{y}, \theta)} d\theta d\mathbf{y} - 1}. \end{aligned} \quad (3.111)$$

Puisque h est un paramètre libre, c'est-à-dire laissé au choix de l'utilisateur, l'EQM minimale la plus précise qui peut être obtenue à partir de (3.111) sera notée $EQM_{BZ}^{globale}$

$$\boxed{EQM_{BZ}^{globale} = \max_h EQM_{BZ}^{globale}(h) = \max_h \frac{h^2}{\int \int_{\Omega \Theta} \frac{p^2(\mathbf{y}, \theta+h)}{p(\mathbf{y}, \theta)} d\theta d\mathbf{y} - 1}.} \quad (3.112)$$

On reconnaît la borne de Bobrovsky-Zakaï [BZ76] qui n'est autre que la version bayésienne de la borne de Chapman-Robbins (3.30).

Pour obtenir la meilleure borne bayésienne, l'expression $P2$ doit être valable $\forall h$. Il paraît donc naturel de ne pas se restreindre à un seul point test. Si l'on choisit K points test $\{h_1, h_2, \dots, h_K\}$ on obtient le jeu de $K + 1$ contraintes suivant en posant $\mathbf{h} = [h_1 \ \dots \ h_K]^T$

$$\begin{cases} \mathbf{c} = [0 \ \mathbf{h}^T]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)}, \\ g_k(\mathbf{y}) = \frac{p(\mathbf{y}, \theta+h_k)}{\sqrt{p(\mathbf{y}, \theta)}} \quad k = 1, \dots, K. \end{cases} \quad (3.113)$$

L'application du théorème 3.1 conduit à l'EQM globale minimale notée $EQM_{RM(K)}^{globale}(\mathbf{h})$

$$EQM_{RM(K)}^{globale}(\mathbf{h}) = [0 \ \mathbf{h}^T] \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & & & \\ \vdots & & \mathbf{D} & \\ 1 & & & \end{pmatrix}^{-1} \begin{bmatrix} 0 \\ \mathbf{h} \end{bmatrix} \quad (3.114)$$

$$= \mathbf{h}^T (\mathbf{D} - \mathbf{1}\mathbf{1}^T)^{-1} \mathbf{h}, \quad (3.115)$$

avec $D_{i,j} = \int \int_{\Omega \Theta} \frac{p(\mathbf{y}, \theta+h_i)p(\mathbf{y}, \theta+h_j)}{p(\mathbf{y}, \theta)} d\theta d\mathbf{y}$. Le vecteur \mathbf{h} est un vecteur de paramètres libres, donc

la borne la plus précise notée $EQM_{RM}^{globale}$ est donnée par

$$\boxed{EQM_{RM}^{globale} = \lim_{K \rightarrow \infty} \max_{\mathbf{h}} EQM_{RM(K)}^{globale}(\mathbf{h}) = \lim_{K \rightarrow \infty} \max_{\mathbf{h}} \mathbf{h}^T (\mathbf{D} - \mathbf{1}\mathbf{1}^T)^{-1} \mathbf{h}.} \quad (3.116)$$

De part sa structure, cette borne est la version bayésienne de la borne de Barankin. En 1997, Reuven et Messer ont établi une borne de Barankin hybride [RM97], c'est-à-dire pour un vecteur de paramètres concaténant des paramètres déterministes et aléatoires. Il s'agit donc ici d'un cas particulier de la borne de Reuven-Messer pour laquelle tous les paramètres sont supposés aléatoires.

3.4.2.2.4 Borne de Weiss-Weinstein On suppose ici $f(\mathbf{x}, \theta) = p(\mathbf{x}, \theta)$, la densité jointe du problème. Donc la contrainte supplémentaire sur $P\mathfrak{B}$, à savoir $\int_{\Omega} \int_{\Theta} f(\mathbf{y}, \theta) d\theta d\mathbf{y} = 1$ est satisfaite.

Si l'on se place dans le cadre de $P\mathfrak{B}$ et d'après le théorème 3.2, la meilleure borne bayésienne est obtenue $\forall h$ et $\forall s$. Si l'on se restreint à un jeu de K paramètres h_k et à un jeu de K paramètres s_k , on obtient le jeux des $K + 1$ contraintes suivantes

$$\begin{cases} \mathbf{c} = [0 & h_1 \mathbb{E}_{\mathbf{y}, \theta} [L^{1-s_1}(\mathbf{y}, \theta - h_1)] & \cdots & h_K \mathbb{E}_{\mathbf{y}, \theta} [L^{1-s_K}(\mathbf{y}, \theta - h_K)]]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)}, \\ g_k(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)} (L^{s_k}(\mathbf{y}, \theta + h_k) - L^{1-s_k}(\mathbf{y}, \theta - h_k)) \quad k = 1, \dots, K. \end{cases} \quad (3.117)$$

On posera

$$\boldsymbol{\xi} = [h_1 \mathbb{E}_{\mathbf{y}, \theta} [L^{1-s_1}(\mathbf{y}, \theta - h_1)] \quad \cdots \quad h_K \mathbb{E}_{\mathbf{y}, \theta} [L^{1-s_K}(\mathbf{y}, \theta - h_K)]]^T, \quad (3.118)$$

$$\mathbf{h} = [h_1 \quad \cdots \quad h_K]^T, \quad (3.119)$$

$$\mathbf{s} = [s_1 \quad \cdots \quad s_K]^T. \quad (3.120)$$

L'application du théorème 3.1 conduit à l'EQM globale minimale notée $EQM_{WW(K)}^{globale}(\mathbf{h}, \mathbf{s})$

$$EQM_{WW(K)}^{globale}(\mathbf{h}, \mathbf{s}) = \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c} = \boldsymbol{\xi}^T \mathbf{W}^{-1} \boldsymbol{\xi}, \quad (3.121)$$

où

$$W_{i,j} = \mathbb{E}_{\mathbf{y}, \theta} [(L^{s_i}(\mathbf{y}, \theta + h_i) - L^{1-s_i}(\mathbf{y}, \theta - h_i)) (L^{s_j}(\mathbf{y}, \theta + h_j) - L^{1-s_j}(\mathbf{y}, \theta - h_j))]. \quad (3.122)$$

Les vecteurs \mathbf{h} et \mathbf{s} sont des vecteurs de paramètres libres, donc la borne la plus précise notée $EQM_{WW}^{globale}$ est donnée par

$$EQM_{WW}^{globale} = \lim_{K \rightarrow \infty} \max_{\mathbf{h}, \mathbf{s}} \boldsymbol{\xi}^T \mathbf{W}^{-1} \boldsymbol{\xi}. \quad (3.123)$$

On reconnaît la borne de Weiss-Weinstein [WW85].

3.4.2.3 Relations d'ordre

Dues à l'augmentation du nombre des contraintes dans le théorème 3.1, on peut en déduire les relations d'ordre suivantes

$$EQM_{CR}^{globale} = EQM_{BMZ}^{globale}(1) \leq \begin{cases} EQM_{BMZ}^{globale}(q(\mathbf{x}, \theta)) \\ EQM_{Bhatt(K)}^{globale} \end{cases} \quad (3.124)$$

et

$$EQM_{CR}^{globale} = \lim_{h \rightarrow 0} EQM_{BZ}^{globale}(h) \leq EQM_{BZ}^{globale} = EQM_{RM(1)}^{globale}, \quad (3.125)$$

$$EQM_{RM(1)}^{globale} \leq EQM_{RM(K)}^{globale} = \begin{cases} EQM_{WW(K)}^{globale}(\mathbf{h}, \mathbf{0}) \\ EQM_{WW(K)}^{globale}(\mathbf{h}, \mathbf{1}) \end{cases} \leq EQM_{WW}^{globale}. \quad (3.126)$$

3.5 Vers de nouvelles bornes minimales

Nous avons vu ci-avant que la majorité des bornes déterministes disposaient d'une version bayésienne dans la famille Weiss-Weinstein et vice versa. Cette dualité provient du fait que ces deux catégories de bornes minimales dérivent du principe d'inégalité de covariance. Ici, grâce au théorème général 3.1, on observe cette dualité au travers des contraintes mises en jeu lors de la minimisation de l'EQM locale ou globale. En effet, lorsque qu'une borne déterministe repose sur des contraintes de type

$$\int_{\Omega} \mathcal{L}_k(p(\mathbf{y}|\theta)) d\mathbf{y} = 0, \quad k = 1, \dots, K, \quad (3.127)$$

la borne bayésienne associée utilise une contrainte faisant apparaître le même opérateur \mathcal{L}_k

$$\int_{\Theta} \int_{\Omega} \mathcal{L}_k(p(\mathbf{y},\theta)) d\mathbf{y}d\theta = 0, \quad k = 1, \dots, K. \quad (3.128)$$

Nous utilisons dans cette section ce principe de mise en correspondance pour proposer deux nouvelles bornes minimales. En effet, à notre connaissance, la version bayésienne de la borne d'Abel et la version déterministe de la borne de Weiss-Weinstein que nous nous proposons d'établir n'existent pas.

3.5.1 Borne d'Abel bayésienne

Dans le contexte déterministe, la borne d'Abel est obtenue à partir du jeu de contraintes suivantes

$$\left\{ \begin{array}{l} \mathbf{c} = [0 \quad 1 \quad 0 \quad \dots \quad 0 \quad \mathbf{h}^T]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)}, \\ g_k(\mathbf{y}) = \frac{1}{\sqrt{p(\mathbf{y}|\theta_0)}} \left. \frac{\partial^k p(\mathbf{y}|\theta)}{\partial \theta^k} \right|_{\theta_0}, \quad k = 1, \dots, M \\ g_k(\mathbf{y}) = \frac{p(\mathbf{y}|\theta_0+h_k)}{\sqrt{p(\mathbf{y}|\theta_0)}}, \quad k = M+1, \dots, R. \end{array} \right. \quad (3.129)$$

Donc, dans un contexte bayésien, en appliquant le principe de mise en correspondance susmentionné, c'est-à-dire en concaténant les contraintes (3.101) et (3.113), on obtient le nouveau jeu des $M+R+1$ contraintes suivantes

$$\left\{ \begin{array}{l} \mathbf{c} = [0 \quad 1 \quad 0 \quad \dots \quad 0 \quad h_1 \quad \dots \quad h_R]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y},\theta)}, \\ g_k(\mathbf{y}) = \frac{1}{\sqrt{p(\mathbf{y},\theta)}} \frac{\partial^k p(\mathbf{y},\theta)}{\partial \theta^k}, \quad k = 1, \dots, M, \\ g_k(\mathbf{y}) = \frac{p(\mathbf{y},\theta+h_k)}{\sqrt{p(\mathbf{y},\theta)}}, \quad k = M+1, \dots, R. \end{array} \right. \quad (3.130)$$

L'application du théorème 3.1 conduit à l'EQM globale minimale notée $EQM_{Abel(M,R)}^{globale}(\mathbf{h})$

$$EQM_{Abel(M,R)}^{globale}(\mathbf{h}) = \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c}, \quad (3.131)$$

avec

$$\mathbf{G} = \begin{pmatrix} 1 & \mathbf{0}^T & \mathbf{1}^T \\ \mathbf{0} & \mathbf{B} & \mathbf{\Gamma}^T \\ \mathbf{1} & \mathbf{\Gamma} & \mathbf{D} \end{pmatrix}. \quad (3.132)$$

où $\mathbf{0}$ est un vecteur colonne composé de M zéros, $\mathbf{1}$ est un vecteur colonne composé de R uns, les éléments $B_{i,j}$ et $D_{i,j}$ des matrices $\mathbf{B} \in \mathcal{M}_{M \times M}(\mathbb{R})$ et $\mathbf{D} \in \mathcal{M}_{R \times R}(\mathbb{R})$ intervenant dans le

calcul de la borne de Bhattacharyya bayésienne et dans le calcul de la borne de Reuven-Messer sont donnés par

$$B_{i,j} = \int_{\Omega} \int_{\Theta} \frac{1}{p(\mathbf{y}, \theta)} \frac{\partial^i}{\partial \theta^i} p(\mathbf{y}, \theta) \frac{\partial^j}{\partial \theta^j} p(\mathbf{y}, \theta) d\theta d\mathbf{y}, \quad (3.133)$$

$$D_{i,j} = \int_{\Omega} \int_{\Theta} \frac{p(\mathbf{y}, \theta + h_i) p(\mathbf{y}, \theta + h_j)}{p(\mathbf{y}, \theta)} d\theta d\mathbf{y}, \quad (3.134)$$

et où l'élément $\Gamma_{i,j}$ de la matrice $\mathbf{\Gamma} \in \mathcal{M}_{R \times M}(\mathbb{R})$ est donné par

$$\Gamma_{i,j} = \int_{\Omega} \int_{\Theta} \frac{p(\mathbf{y}, \theta + h_i)}{p(\mathbf{y}, \theta)} \frac{\partial^j p(\mathbf{y}, \theta)}{\partial \theta^j} d\theta d\mathbf{y}. \quad (3.135)$$

Soit $\boldsymbol{\beta} = [\mathbf{0}^T \mathbf{1}^T]^T$, $\tilde{\mathbf{G}} = \begin{pmatrix} \mathbf{B} & \mathbf{\Gamma}^T \\ \mathbf{\Gamma} & \mathbf{D} \end{pmatrix}$ et $\mathbf{c} = [0 \ \boldsymbol{\alpha}^T \ \mathbf{h}^T]^T$ où le vecteur de taille $(M \times 1)$ $\boldsymbol{\alpha} = [1 \ 0 \ \dots \ 0]^T$. Puisque le premier élément du vecteur \mathbf{c} est nul, seule la matrice $\mathbf{R} \in \mathcal{M}_{M+R \times M+R}(\mathbb{R})$ représentant le coin inférieur droit de la matrice \mathbf{G}^{-1} a besoin d'être étudiée. \mathbf{R} est donnée par (voir par exemple [Van02] equation A.68)

$$\mathbf{R} = \left(\tilde{\mathbf{G}} - \boldsymbol{\beta} \boldsymbol{\beta}^T \right)^{-1} = \begin{pmatrix} \mathbf{B} & \mathbf{\Gamma}^T \\ \mathbf{\Gamma} & \mathbf{D} - \mathbf{1} \mathbf{1}^T \end{pmatrix}^{-1}. \quad (3.136)$$

On posera $\mathbf{J} = \mathbf{D} - \mathbf{1} \mathbf{1}^T$. L'EQM globale minimale se réduit à

$$EQM_{Abel(M,R)}^{globale}(\mathbf{h}) = [\boldsymbol{\alpha}^T \ \mathbf{h}^T] \begin{pmatrix} \mathbf{B} & \mathbf{\Gamma}^T \\ \mathbf{\Gamma} & \mathbf{J} \end{pmatrix}^{-1} \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{h} \end{bmatrix}. \quad (3.137)$$

Après quelques efforts algébriques on obtient l'expression finale

$$EQM_{Abel(M,R)}^{globale}(\mathbf{h}) = \boldsymbol{\alpha}^T \mathbf{B}^{-1} \boldsymbol{\alpha} + \mathbf{u}^T \tilde{\mathbf{J}}^{-1} \mathbf{u}, \quad (3.138)$$

avec

$$\begin{cases} \mathbf{u} = \mathbf{\Gamma} \mathbf{B}^{-1} \boldsymbol{\alpha} - \mathbf{h}, \\ \tilde{\mathbf{J}} = \mathbf{J} - \mathbf{\Gamma} \mathbf{B}^{-1} \mathbf{\Gamma}^T. \end{cases} \quad (3.139)$$

Puisque $EQM_{Abel(M,R)}^{globale}(\mathbf{h})$ est une fonction du vecteur \mathbf{h} laissé à la discrétion de l'utilisateur, la borne maximale notée $EQM_{Abel(M,R)}^{globale}$ sera donnée par

$$\boxed{EQM_{Abel(M,R)}^{globale} = \max_{\mathbf{h}} \boldsymbol{\alpha}^T \mathbf{B}^{-1} \boldsymbol{\alpha} + \mathbf{u}^T \tilde{\mathbf{J}}^{-1} \mathbf{u}.} \quad (3.140)$$

Dû fait de l'augmentation du nombre des contraintes dans l'application du théorème 3.1, à M et R fixés, nous obtenons la relation d'ordre suivante

$$EQM_{Abel(M,R)}^{globale} \geq \begin{cases} EQM_{Bhatt(M)}^{globale} \\ EQM_{RM(R)}^{globale} \end{cases} \quad (3.141)$$

Par contre, on ne peut pas conclure sur une éventuelle relation d'ordre avec la borne de Weiss-Weinstein.

3.5.2 Borne de Weiss-Weinstein déterministe

Dans le contexte bayésien, la borne de Weiss-Weinstein est obtenue à partir du jeu de contraintes suivantes

$$\begin{cases} \mathbf{c} = \left[0 \quad h_1 \mathbb{E}_{\mathbf{y}, \theta} \left[\left(\frac{p(\mathbf{y}, \theta - h_1)}{p(\mathbf{y}, \theta)} \right)^{1-s_1} \right] \quad \cdots \quad h_K \mathbb{E}_{\mathbf{y}, \theta} \left[\left(\frac{p(\mathbf{y}, \theta - h_K)}{p(\mathbf{y}, \theta)} \right)^{1-s_K} \right] \right]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)}, \\ g_k(\mathbf{y}) = \sqrt{p(\mathbf{y}, \theta)} \left(\left(\frac{p(\mathbf{y}, \theta + h_k)}{p(\mathbf{y}, \theta)} \right)^{s_k} - \left(\frac{p(\mathbf{y}, \theta - h_k)}{p(\mathbf{y}, \theta)} \right)^{1-s_k} \right), \quad k = 1, \dots, K. \end{cases} \quad (3.142)$$

Donc, dans le contexte déterministe, en appliquant le principe de mise en correspondance susmentionné, on obtient le jeux des $K + 1$ contraintes suivantes

$$\begin{cases} \mathbf{c} = \left[0 \quad h_1 \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{p(\mathbf{y}|\theta_0 - h_1)}{p(\mathbf{y}|\theta_0)} \right)^{1-s_1} \right] \quad \cdots \quad h_K \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{p(\mathbf{y}|\theta_0 - h_K)}{p(\mathbf{y}|\theta_0)} \right)^{1-s_K} \right] \right]^T, \\ g_0(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)}, \\ g_k(\mathbf{y}) = \sqrt{p(\mathbf{y}|\theta_0)} \left(\left(\frac{p(\mathbf{y}|\theta_0 + h_k)}{p(\mathbf{y}|\theta_0)} \right)^{s_k} - \left(\frac{p(\mathbf{y}|\theta_0 - h_k)}{p(\mathbf{y}|\theta_0)} \right)^{1-s_k} \right), \quad k = 1, \dots, K. \end{cases} \quad (3.143)$$

On posera

$$\boldsymbol{\xi} = \left[h_1 \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{p(\mathbf{y}|\theta_0 - h_1)}{p(\mathbf{y}|\theta_0)} \right)^{1-s_1} \right] \quad \cdots \quad h_K \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{p(\mathbf{y}|\theta_0 - h_K)}{p(\mathbf{y}|\theta_0)} \right)^{1-s_K} \right] \right]^T, \quad (3.144)$$

$$\mathbf{h} = [h_1 \quad \cdots \quad h_K]^T, \quad (3.145)$$

$$\mathbf{s} = [s_1 \quad \cdots \quad s_K]^T. \quad (3.146)$$

L'application du théorème 3.1 conduit à l'EQM minimale en θ_0 notée $EQM_{WWD(K)}(\mathbf{h}, \mathbf{s}, \theta_0)$

$$EQM_{WWD(K)}(\mathbf{h}, \mathbf{s}, \theta_0) = \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c} = \boldsymbol{\xi}^T \mathbf{W}^{-1} \boldsymbol{\xi}, \quad (3.147)$$

où

$$W_{i,j} = \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\left(\frac{p(\mathbf{y}|\theta_0 + h_i)}{p(\mathbf{y}|\theta_0)} \right)^{s_i} - \left(\frac{p(\mathbf{y}|\theta_0 - h_i)}{p(\mathbf{y}|\theta_0)} \right)^{1-s_i} \right) \times \left(\left(\frac{p(\mathbf{y}|\theta_0 + h_j)}{p(\mathbf{y}|\theta_0)} \right)^{s_j} - \left(\frac{p(\mathbf{y}|\theta_0 - h_j)}{p(\mathbf{y}|\theta_0)} \right)^{1-s_j} \right) \right]. \quad (3.148)$$

Les vecteurs \mathbf{h} et \mathbf{s} sont des vecteurs de paramètres libres, donc la borne la plus précise notée EQM_{WWD} est donnée par

$$EQM_{WWD}(\theta_0) = \lim_{K \rightarrow \infty} \max_{\mathbf{h}, \mathbf{s}} \boldsymbol{\xi}^T \mathbf{W}^{-1} \boldsymbol{\xi}, \quad (3.149)$$

que nous appellerons borne de Weiss-Weinstein déterministe du fait de la correspondance avec la borne de Weiss-Weinstein. Puisqu'il s'agit d'une borne déterministe, il s'agit de l'EQM minimale en θ_0 d'un estimateur sujet à la contrainte de biais nul en θ_0 (d'après le premier élément de \mathbf{c} et $g_0(\mathbf{y})$) ainsi que des contraintes du type

$$\int_{\Omega} \left(\hat{\theta}(\mathbf{y}) - \theta_0 \right) p^{s_k}(\mathbf{y}|\theta_0 + h_k) p^{1-s_k}(\mathbf{y}|\theta_0) d\mathbf{y} = \int_{\Omega} \left(\hat{\theta}(\mathbf{y}) - \theta_0 + h_k \right) p^{s_k}(\mathbf{y}|\theta_0) p^{1-s_k}(\mathbf{y}|\theta_0 - h_k) d\mathbf{y}, \quad (3.150)$$

pour $k = 1, \dots, K$.

Malheureusement, nous ne comprenons pas encore les véritables implications de (3.150). Néanmoins, sachant que $s_k \in [0, 1]$, on peut étudier les deux cas particuliers suivants :

$$\lim_{s_k \rightarrow 0} (3.150) \Rightarrow \underbrace{\int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0) p(\mathbf{y} | \theta_0) d\mathbf{y}}_{=0} = \quad (3.151)$$

$$\int_{\Omega} (\hat{\theta}(\mathbf{y}) - (\theta_0 - h_k)) p(\mathbf{y} | \theta_0 - h_k) d\mathbf{y}, \quad (3.152)$$

$$\lim_{s_k \rightarrow 1} (3.150) \Rightarrow \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta_0) p(\mathbf{y} | \theta_0 + h_k) d\mathbf{y} =$$

$$\underbrace{\int_{\Omega} (\hat{\theta}(\mathbf{y}) - (\theta_0 - h_k)) p(\mathbf{y} | \theta_0) d\mathbf{y}}_{=h}$$

$$\Rightarrow \int_{\Omega} (\hat{\theta}(\mathbf{y}) - (\theta_0 + h_k)) p(\mathbf{y} | \theta_0 + h_k) d\mathbf{y} = 0. \quad (3.153)$$

Les deux jeux de contraintes (3.151) et (3.153) sont similaires aux contraintes intervenant dans la borne de Barankin (3.32), à savoir un estimateur dont le biais est nul en plusieurs points tests. Il paraît donc légitime de penser que cette borne est une extension de la borne de Barankin.

Chapitre 4

Bornes minimales en traitement statistique du signal

4.1 Introduction

Le but de ce chapitre est d'illustrer l'utilité des bornes minimales d'estimation et de les exploiter dans le cadre du traitement du signal. Dans le contexte déterministe, les bornes minimales présentées au chapitre précédent sont utilisées pour prédire la zone de décrochement des estimateurs, en particulier la borne de Chapman-Robbins et la borne de Barankin. En effet, ces bornes explorent le support du paramètre dans sa globalité et sont donc à même de représenter le phénomène de décrochement dû aux outliers. La borne de Cramér-Rao restant dans un voisinage très localisé du paramètre ne peut pas prendre en compte ce phénomène car les outliers de part leur nature peuvent apparaître n'importe où. La borne de Bhattacharyya peut également, en théorie, prendre en compte ce phénomène. Néanmoins ceci nécessiterait de la calculer à un ordre très élevé. La borne de Chapman-Robbins et la borne de Barankin ont été appliquées avec succès à plusieurs problèmes d'estimation [MS69] [Bag69] [MH71] [RM95] [Kno97] [TK97] [Mar97] [RM99] [TK99] [FT03] [CGFL05] [CFL04] [ABL04b] [ABL04a] [QCL06]. L'utilisation de la borne d'Abel en traitement du signal est restée très marginale.

Les bornes bayésiennes, du fait de la prise en compte du support du paramètre au travers de sa densité *a priori* $p(\boldsymbol{\theta})$, donnent les performances ultimes d'un estimateur sur les trois zones de l'EQM globale. Ces bornes permettent non seulement de connaître les performances des estimateurs bayésiens tels que l'estimateur au sens du Maximum *A Posteriori* ou l'estimateur de la moyenne conditionnelle, mais elles sont également utilisées pour connaître les performances globales d'estimateurs déterministes puisque

$$\begin{aligned} \mathbf{EQM}_{globale} &= \int_{\Theta} \int_{\Omega} (\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}) (\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta})^H p(\mathbf{y}, \boldsymbol{\theta}) d\mathbf{y} d\boldsymbol{\theta} \\ &= \int_{\Theta} \mathbf{EQM}_{locale}(\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}. \end{aligned} \quad (4.1)$$

On peut citer l'exemple des travaux de Wen Xu [Xu01] [XBB04] [XBR04], où les performances de l'estimateur du maximum de vraisemblance sont analysées dans le contexte de l'acoustique sous-marine à l'aide des bornes de Ziv-Zakai et Weiss-Weinstein. Dans ce chapitre, nous suivrons cette philosophie dans le contexte de l'analyse spectrale et de l'estimation d'une porteuse en communication numérique avec symboles pilotes. Les bornes de la famille Ziv-Zakai ont été

appliquées dans plusieurs autres domaines du traitement du signal : principalement pour l'estimation d'un retard [WW83] et pour l'estimation de directions d'arrivées [BET96] [BET95]. A notre connaissance, les bornes de la famille Weiss-Weinstein ont en revanche été très peu utilisées mis à part les travaux de Wen Xu susmentionnés.

De plus, les bornes les plus pertinentes (Barankin, Abel, Ziv-Zakaï et Weiss-Weinstein) s'appuient sur un jeu de points tests pour lequel elles sont optimisées. Plus le nombre de points tests est grand, plus les bornes déterministes se rapprochent de la vraie borne de Barankin et plus les bornes bayésiennes se rapprochent de l'EQM de l'estimateur de la moyenne conditionnelle. Ceci conduit généralement à une lourde charge de calcul. C'est pour cette raison que la communauté utilise généralement ces bornes avec un seul point test (par exemple la borne de Chapman-Robbins) afin de limiter cette charge de calcul. Nous pensons que c'est également pour cette raison que la borne d'Abel n'a pas été utilisée en traitement du signal alors qu'elle est plus pertinente que la borne de Bhattacharyya ou la borne de Barankin. En effet, la forme générale de la borne d'Abel est donnée dans le papier original pour un nombre quelconque de points test et un nombre quelconque de dérivées successives de la fonction score ce qui a pu faire peur aux chercheurs désirant l'utiliser.

Dans ce chapitre, nous montrons que pour un seul point test et une seule dérivée de la fonction score, la borne d'Abel (déterministe et bayésienne) se met sous une forme particulièrement intéressante. Puis, nous proposons différentes formes analytiques de bornes minimales pour un modèle d'observations gaussiennes à moyenne paramétrée. Ensuite, dans le contexte de l'estimation d'une fréquence porteuse en communications numériques utilisant des symboles pilotes (une extension de l'analyse spectrale) nous appliquons les différentes bornes pour les comparer aux performances de l'estimateur du maximum de vraisemblance. Enfin, nous étudions le comportement du point test optimal, c'est-à-dire celui qui maximise les bornes, pour en déduire un prédicteur de la zone de décrochement des estimateurs et des bornes dont la charge de calcul reste proche de celle de Cramér-Rao.

4.2 Une forme utile de la borne d'Abel

Dans cette section, nous étudions la borne d'Abel (4.6) dans le cas où $M = R = 1$, c'est-à-dire pour un point test et une seule dérivée de la fonction score. La borne ainsi obtenue dépend directement de la borne de Cramér-Rao et de la borne de Chapman-Robbins. La version bayésienne de cette borne est également proposée. Ce sont ces bornes qui seront utilisées dans la suite.

Rappelons l'expression de la borne d'Abel (déterministe) :

$$EQM_{Abel(M,R)}(\theta_0, \mathbf{h}) = \boldsymbol{\alpha}^T \mathbf{B}^{-1} \boldsymbol{\alpha} + \mathbf{u}^T \tilde{\mathbf{J}}^{-1} \mathbf{u}, \quad (4.2)$$

avec

$$\left\{ \begin{array}{l} \mathbf{u} = \mathbf{\Gamma} \mathbf{B}^{-1} \boldsymbol{\alpha} - \mathbf{h}, \\ \tilde{\mathbf{J}} = \mathbf{J} - \mathbf{\Gamma} \mathbf{B}^{-1} \mathbf{\Gamma}^T, \\ \boldsymbol{\alpha} = [1 \ 0 \ \dots \ 0]^T \quad (M \times 1), \\ \mathbf{h} = [h_1 \ \dots \ h_K]^T, \\ \mathbf{J} = \mathbf{D} - \mathbf{1} \mathbf{1}^T \\ B_{i,j} = \int_{\Omega} \frac{1}{p(\mathbf{y}|\theta_0)} \frac{\partial^i p(\mathbf{y}|\theta)}{\partial \theta^i} \Big|_{\theta_0} \frac{\partial^j p(\mathbf{y}|\theta)}{\partial \theta^j} \Big|_{\theta_0} d\mathbf{y} \quad \mathbf{B} \in \mathcal{M}_{M \times M}(\mathbb{R}), \\ D_{i,j} = \int_{\Omega} \frac{p(\mathbf{y}|\theta_0+h_i)p(\mathbf{y}|\theta_0+h_j)}{p(\mathbf{y}|\theta_0)} d\mathbf{y} \quad \mathbf{D} \in \mathcal{M}_{R \times R}(\mathbb{R}), \\ \Gamma_{i,j} = \int_{\Omega} \frac{p(\mathbf{y}|\theta_0+h_i)}{p(\mathbf{y}|\theta_0)} \frac{\partial^j p(\mathbf{y}|\theta)}{\partial \theta^j} \Big|_{\theta_0} d\mathbf{y} \quad \mathbf{\Gamma} \in \mathcal{M}_{R \times M}(\mathbb{R}). \end{array} \right. \quad (4.3)$$

Si l'on choisit $M = R = 1$, on obtient pour tout ces éléments

$$\left\{ \begin{array}{l} u = \alpha \frac{\Gamma}{B} - h, \\ \tilde{J} = J - \frac{\Gamma^2}{B}, \\ \alpha = 1, \\ \mathbf{h} = h, \\ J = D - 1, \\ B = \int_{\Omega} \frac{1}{p(\mathbf{y}|\theta_0)} \left(\frac{\partial p(\mathbf{y}|\theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 d\mathbf{y} = \frac{1}{EQM_{CR}(\theta_0)}, \\ D = \int_{\Omega} \frac{p^2(\mathbf{y}|\theta_0+h)}{p(\mathbf{y}|\theta_0)} d\mathbf{y} = \frac{1}{EQM_{ChR}(\theta_0, h)} + 1, \\ \Gamma = \int_{\Omega} p(\mathbf{y}|\theta_0+h) \frac{\partial \ln p(\mathbf{y}|\theta)}{\partial \theta} \Big|_{\theta_0} d\mathbf{y}, \end{array} \right. \quad (4.4)$$

ce qui donne pour la borne d'Abel $EQM_{Abel(1,1)}(\theta_0, h)$

$$\begin{aligned} EQM_{Abel(1,1)}(\theta_0, h) &= B^{-1} + \frac{u^2}{J - \frac{\Gamma^2}{B}} \\ &= EQM_{CR}(\theta_0) + \frac{(\Gamma EQM_{CR}(\theta_0) - h)^2}{\frac{1}{EQM_{ChR}(\theta_0, h)} - \Gamma^2 EQM_{CR}(\theta_0)} \\ &= \frac{EQM_{CR}^{-1}(\theta_0) + EQM_{ChR}^{-1}(\theta_0, h) - 2\psi(\theta_0, h)}{EQM_{CR}^{-1}(\theta_0) EQM_{ChR}^{-1}(\theta_0, h) - \psi^2(\theta_0, h)}, \end{aligned} \quad (4.5)$$

où l'on a posé $\psi(\theta_0, h) = \frac{\Gamma}{h}$. La borne la plus pertinente notée $EQM_{Abel(1,1)}(\theta_0)$ est obtenue en maximisant $EQM_{Abel(1,1)}(\theta_0, h)$ par rapport à h

$$\boxed{EQM_{Abel(1,1)}(\theta_0) = \max_h \frac{EQM_{CR}(\theta_0) + EQM_{ChR}(\theta_0, h) - 2\psi(\theta_0, h) EQM_{CR}(\theta_0) EQM_{ChR}(\theta_0, h)}{1 - \psi^2(\theta_0, h) EQM_{CR}(\theta_0) EQM_{ChR}(\theta_0, h)}}. \quad (4.6)$$

L'intérêt opérationnel de cette borne est triple.

– Premièrement, grâce à la relation d'ordre (3.46), on obtient

$$EQM_{Abel(1,1)}(\theta_0) \geq \begin{cases} EQM_{CR}(\theta_0) \\ EQM_{ChR}(\theta_0) \end{cases} \quad (4.7)$$

où

$$EQM_{ChR}(\theta_0) = \max_h EQM_{ChR}(\theta_0, h). \quad (4.8)$$

Il apparaît donc que cette borne est plus pertinente que les bornes de Cramér-Rao et de Chapman-Robbins.

- Deuxièmement, lorsque la borne de Cramér-Rao et la borne de Chapman-Robbins sont disponibles pour un problème d'estimation donné, obtenir la borne d'Abel requiert simplement le calcul de $\psi(\theta_0, h)$.
- Enfin, la charge de calcul de $EQM_{Abel(1,1)}(\theta_0)$ n'excède pas celle de la borne de Chapman-Robbins. Nous verrons dans la suite comment obtenir une charge de calcul équivalente à la borne de Cramér-Rao.

Dans le contexte bayésien, en se basant sur le même principe, on obtient trivialement la borne d'Abel bayésienne notée $EQM_{Abel(1,1)}^{globale}$

$$EQM_{Abel(1,1)}^{globale} = \max_h \frac{EQM_{CR}^{globale} + EQM_{BZ}^{globale}(h) - 2\psi(h)EQM_{CR}^{globale}EQM_{BZ}^{globale}(h)}{1 - \psi^2(h)EQM_{CR}^{globale}EQM_{BZ}^{globale}(h)}, \quad (4.9)$$

où

$$\psi(h) = \frac{1}{h} \int_{\Theta} \int_{\Omega} p(\mathbf{y}, \theta + h) \frac{\partial \ln p(\mathbf{y}, \theta)}{\partial \theta} d\mathbf{y} d\theta. \quad (4.10)$$

Les mêmes remarques que précédemment restent valides.

4.3 Modèle d'observations gaussiennes à moyenne paramétrée

Nous nous proposons dans cette section d'établir certaines formes analytiques des bornes minimales d'estimation pour un modèle général d'observations gaussiennes à moyenne paramétrée. Le résumé des résultats est reporté en annexe A.

Considérons le modèle d'observations gaussiennes à moyenne paramétrée suivant :

$$\mathbf{y} = \mathbf{m}(\theta) + \mathbf{b}, \quad (4.11)$$

où $\mathbf{y} \in \mathbf{M}_{N \times 1}(\mathbb{C})$ est le vecteur complexe des observations supposées indépendantes, θ est le paramètre scalaire d'intérêt ayant pour vraie valeur θ_0 , $\mathbf{m} \in \mathbf{M}_{N \times 1}(\mathbb{C})$ est un vecteur complexe déterministe dépendant (généralement non-linéairement) de θ , et $\mathbf{b} \in \mathbf{M}_{N \times 1}(\mathbb{C})$ est le vecteur complexe du bruit. Le bruit est supposé circulaire, gaussien de moyenne nulle et de matrice de covariance $\sigma^2 \mathbf{I}_M$. Dans le contexte bayésien, nous supposons θ gaussien de moyenne μ et de variance σ_θ^2

$$p(\theta) = \frac{1}{\sqrt{2\pi\sigma_\theta}} e^{-\frac{1}{2\sigma_\theta^2}(\theta - \mu)^2}. \quad (4.12)$$

Pour ce modèle général, la fonction de vraisemblance est donnée par

$$p(\mathbf{y}|\theta) = \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(\mathbf{y} - \mathbf{m}(\theta))^H (\mathbf{y} - \mathbf{m}(\theta))}. \quad (4.13)$$

A notre connaissance, seule la forme analytique de la borne de Cramér-Rao est connue dans ce cas (voir, par exemple, [Kay93]).

4.3.1 Bornes de Cramér-Rao

Plaçons-nous tout d'abord dans le contexte déterministe. La BCR est donnée par (3.22)

$$EQM_{CR}(\theta_0) = \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{\partial \ln p(\mathbf{y}|\theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 \right]^{-1}, \quad (4.14)$$

En posant $\beta(\mathbf{y}, \theta) = -\frac{1}{\sigma^2} (\mathbf{y} - \mathbf{m}(\theta))^H (\mathbf{y} - \mathbf{m}(\theta))$ on obtient trivialement les relations suivantes

$$\begin{cases} \frac{\partial p(\mathbf{y}|\theta)}{\partial \theta} \Big|_{\theta_0} = \frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} p(\mathbf{y}|\theta_0), \\ \frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} = \frac{2}{\sigma^2} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} (\mathbf{y} - \mathbf{m}(\theta_0)) \right\}, \\ \frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \sim \mathcal{N}_{\mathbb{R}} \left(0, \frac{2}{\sigma^2} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2 \right). \end{cases} \quad (4.15)$$

D'où

$$EQM_{CR}(\theta_0) = \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 \right]^{-1},$$

$$\boxed{EQM_{CR}(\theta_0) = \frac{\sigma^2}{2 \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2}.} \quad (4.16)$$

Dans le contexte bayésien, on peut démontrer que [Van68]

$$EQM_{CR}^{globale} = \left(\int_{\Theta} EQM_{CR}^{-1}(\theta) p(\theta) d\theta - \int_{\Theta} \frac{\partial^2 \ln p(\theta)}{\partial \theta^2} p(\theta) d\theta \right)^{-1}. \quad (4.17)$$

La borne de Cramér-Rao bayésienne se divise donc en deux parties : une partie issue des observations au travers de $EQM_{CR}^{-1}(\theta)$ et une partie issue des informations *a priori*. Le second terme de (4.17) s'écrit ici

$$\begin{aligned} \int_{\Theta} \frac{\partial^2 \ln p(\theta)}{\partial \theta^2} p(\theta) d\theta &= \int_{\Theta} \frac{\partial^2 \left(-\ln(\sqrt{2\pi}\sigma\theta) - \frac{1}{2\sigma_\theta^2} (\theta - \mu)^2 \right)}{\partial \theta^2} p(\theta) d\theta \\ &= -\frac{1}{2\sigma_\theta^2} \int_{\Theta} \frac{\partial^2 (\theta - \mu)^2}{\partial \theta^2} p(\theta) d\theta \\ &= -\frac{1}{\sigma_\theta^2} \int_{\Theta} p(\theta) d\theta \\ &= -\frac{1}{\sigma_\theta^2}. \end{aligned} \quad (4.18)$$

Finalement, la borne de Cramér-Rao bayésienne est donnée par

$$\boxed{EQM_{CR}^{globale} = \left(\frac{2}{\sigma^2} \int_{\Theta} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right\|^2 p(\theta) d\theta + \frac{1}{\sigma_\theta^2} \right)^{-1}.} \quad (4.19)$$

4.3.2 Bornes de Bhattacharyya

Nous détaillons ici les bornes de Bhattacharyya d'ordre 2.

Dans le contexte d'un paramètre supposé déterministe et à l'ordre 2, la borne de Bhattacharyya est donnée par (3.26)

$$EQM_{Bhatt(2)}(\theta_0) = \{\mathbf{B}^{-1}\}_{1,1}, \quad (4.20)$$

où les éléments de la matrice de Bhattacharyya d'ordre 2 sont donnés par

$$\begin{cases} B_{1,1} = \mathbb{E}_{\mathbf{y}|\theta_0} \left[\frac{1}{p^2(\mathbf{y}|\theta_0)} \left(\frac{\partial p(\mathbf{y}|\theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 \right], \\ B_{2,2} = \mathbb{E}_{\mathbf{y}|\theta_0} \left[\frac{1}{p^2(\mathbf{y}|\theta_0)} \left(\frac{\partial^2 p(\mathbf{y}|\theta)}{\partial \theta^2} \Big|_{\theta_0} \right)^2 \right], \\ B_{1,2} = B_{2,1} = \mathbb{E}_{\mathbf{y}|\theta_0} \left[\frac{1}{p^2(\mathbf{y}|\theta_0)} \frac{\partial^2 p(\mathbf{y}|\theta)}{\partial \theta^2} \Big|_{\theta_0} \frac{\partial p(\mathbf{y}|\theta)}{\partial \theta} \Big|_{\theta_0} \right]. \end{cases} \quad (4.21)$$

En posant $\beta(\mathbf{y}, \theta) = -\frac{1}{\sigma^2} (\mathbf{y} - \mathbf{m}(\theta))^H (\mathbf{y} - \mathbf{m}(\theta))$ on obtient les relations suivantes

$$\begin{cases} \frac{\partial p(\mathbf{y}|\theta)}{\partial \theta} \Big|_{\theta_0} = \frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} p(\mathbf{y}|\theta_0), \\ \frac{\partial^2 p(\mathbf{y}|\theta)}{\partial \theta^2} \Big|_{\theta_0} = \frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} p(\mathbf{y}|\theta_0) + \left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 p(\mathbf{y}|\theta_0), \\ \frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} = \frac{2}{\sigma^2} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} (\mathbf{y} - \mathbf{m}(\theta_0)) \right\}, \\ \frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} = \frac{2}{\sigma^2} \left(\operatorname{Re} \left\{ \frac{\partial^2 \mathbf{m}^H(\theta)}{\partial \theta^2} \Big|_{\theta_0} (\mathbf{y} - \mathbf{m}(\theta_0)) \right\} - \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2 \right), \\ \frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \sim \mathcal{N}_{\mathbb{R}} \left(0, \frac{2}{\sigma^2} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2 \right), \\ \frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} \sim \mathcal{N}_{\mathbb{R}} \left(-\frac{2}{\sigma^2} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2, \frac{2}{\sigma^2} \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \Big|_{\theta_0} \right\|^2 \right). \end{cases} \quad (4.22)$$

On posera $x_n = 2 \operatorname{Re} \left\{ \frac{\partial^n \mathbf{m}^H(\theta)}{\partial \theta^n} \Big|_{\theta_0} (\mathbf{y} - \mathbf{m}(\theta_0)) \right\}$. La variable aléatoire réelle x_n est distribuée comme

$$x_n \sim \mathcal{N}_{\mathbb{R}} \left(0, 2\sigma^2 \left\| \frac{\partial^n \mathbf{m}(\theta)}{\partial \theta^n} \Big|_{\theta_0} \right\|^2 \right). \quad (4.23)$$

De plus,

$$\begin{cases} \mathbb{E}_{\mathbf{y}|\theta_0} [x_n^2] = 2\sigma^2 \left\| \frac{\partial^n \mathbf{m}^H(\theta)}{\partial \theta^n} \Big|_{\theta_0} \right\|^2, \\ \mathbb{E}_{\mathbf{y}|\theta_0} [x_n x_m] = 2\sigma^2 \operatorname{Re} \left\{ \frac{\partial^n \mathbf{m}^H(\theta)}{\partial \theta^n} \Big|_{\theta_0} \frac{\partial^m \mathbf{m}(\theta)}{\partial \theta^m} \Big|_{\theta_0} \right\}, \\ \mathbb{E}_{\mathbf{y}|\theta_0} [x_n x_m^2] = 0. \end{cases} \quad (4.24)$$

Concernant le terme $B_{1,1}$, on reconnaît la matrice d'information de Fisher

$$B_{1,1} = \frac{2}{\sigma^2} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2. \quad (4.25)$$

Concernant le terme $B_{1,2}$

$$\begin{aligned}
B_{1,2} &= \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} + \left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 \right) \left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right) \right] \\
&= \mathbb{E}_{\mathbf{y}|\theta_0} \left[\frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} \frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right] + \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^3 \right] \\
&= \mathbb{E}_{\mathbf{y}|\theta_0} \left[\frac{1}{\sigma^4} \left(x_2 - 2 \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2 \right) x_1 \right] + \frac{1}{\sigma^2} \underbrace{\mathbb{E}_{\theta_0} [(x_1)^3]}_{=0} \\
&= \frac{1}{\sigma^4} \mathbb{E}_{\theta_0} [x_2 x_1] = \frac{2}{\sigma^2} \operatorname{Re} \left\{ \frac{\partial^2 \mathbf{m}^H(\theta)}{\partial \theta^2} \Big|_{\theta_0} \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\}. \tag{4.26}
\end{aligned}$$

Concernant le terme $B_{2,2}$

$$\begin{aligned}
B_{2,2} &= \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} + \left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 \right)^2 \right] \\
&= \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} \right)^2 + \left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^4 + 2 \frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} \left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 \right] \\
&= \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} \right)^2 \right] + \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^4 \right] \\
&\quad + 2 \mathbb{E}_{\mathbf{y}|\theta_0} \left[\frac{\partial^2 \beta(\mathbf{y}, \theta)}{\partial \theta^2} \Big|_{\theta_0} \left(\frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} \right)^2 \right] \\
&= \frac{2}{\sigma^2} \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \Big|_{\theta_0} \right\|^2 + \frac{12}{\sigma^4} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^4 \\
&\quad + 2 \mathbb{E}_{\mathbf{y}|\theta_0} \left[\frac{1}{\sigma^6} \left(x_2 - 2 \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2 \right) (x_1)^2 \right] \\
&= \frac{2}{\sigma^2} \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \Big|_{\theta_0} \right\|^2 + \frac{4}{\sigma^4} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^4. \tag{4.27}
\end{aligned}$$

En conséquence, la borne de Bhattacharyya à l'ordre 2 sera donnée par

$$EQM_{Bhatt(2)}(\theta_0) = \frac{B_{2,2}}{B_{1,1}B_{2,2} - B_{1,2}^2}$$

$$EQM_{Bhatt(2)}(\theta_0) = \frac{\frac{2}{\sigma^2} \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \Big|_{\theta_0} \right\|^2 + \frac{4}{\sigma^4} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^4}{\frac{4}{\sigma^4} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2 \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \Big|_{\theta_0} \right\|^2 + \frac{2}{\sigma^6} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^6 - \frac{4}{\sigma^4} \left(\operatorname{Re} \left\{ \frac{\partial^2 \mathbf{m}^H(\theta)}{\partial \theta^2} \Big|_{\theta_0} \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\} \right)^2}. \tag{4.28}$$

Dans le contexte bayésien, la borne de Bhattacharyya bayésienne s'écrit (3.104)

$$EQM_{Bhatt(K)}^{globale} = \{\mathbf{B}^{-1}\}_{1,1}, \quad (4.29)$$

où les éléments de \mathbf{B} sont donnés par

$$B_{i,j} = \int_{\Omega} \int_{\Theta} \frac{1}{p(\mathbf{y}, \theta)} \frac{\partial^i}{\partial \theta^i} p(\mathbf{y}, \theta) \frac{\partial^j}{\partial \theta^j} p(\mathbf{y}, \theta) d\theta d\mathbf{y}. \quad (4.30)$$

Du fait des dérivations successives, il n'existe pas de relation du type (4.17) pour écrire la borne de Bhattacharyya bayésienne en fonction des éléments de la borne de Bhattacharyya.

4.3.3 Bornes de Chapman-Robbins et de Bobrovsky-Zakaï

Dans le contexte déterministe, la borne de Chapman-Robbins est donnée par (3.30)

$$EQM_{ChR}(\theta_0) = \max_h \frac{h^2}{\int_{\Omega} \frac{p^2(\mathbf{y}|\theta_0+h)}{p(\mathbf{y}|\theta_0)} d\mathbf{y} - 1}. \quad (4.31)$$

On peut réécrire le terme intervenant dans l'intégrale ci-dessus comme suit :

$$\begin{aligned} \frac{p^2(\mathbf{y}|\theta_0+h)}{p(\mathbf{y}|\theta_0)} &= \frac{1}{(\pi\sigma^2)^N} \frac{e^{-\frac{2}{\sigma^2}(\mathbf{y}-\mathbf{m}(\theta_0+h))^H(\mathbf{y}-\mathbf{m}(\theta_0+h))}}{e^{-\frac{1}{\sigma^2}(\mathbf{y}-\mathbf{m}(\theta_0))^H(\mathbf{y}-\mathbf{m}(\theta_0))}} \\ &= \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(2(\mathbf{y}-\mathbf{m}(\theta_0+h))^H(\mathbf{y}-\mathbf{m}(\theta_0+h)) - (\mathbf{y}-\mathbf{m}(\theta_0))^H(\mathbf{y}-\mathbf{m}(\theta_0)))} \\ &= \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(\|\mathbf{y}\|^2 + 2\|\mathbf{m}(\theta_0+h)\|^2 - \|\mathbf{m}(\theta_0)\|^2 - 2\operatorname{Re}\{\mathbf{y}^H(2\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0))\})} \end{aligned} \quad (4.32)$$

On pose $\mathbf{x} = \mathbf{y} - (2\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0))$. Remarquons que

$$\|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 + \|2\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2 - 2\operatorname{Re}\{\mathbf{y}^H(2\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0))\}. \quad (4.33)$$

En reportant (4.33) dans (4.32), on obtient

$$\begin{aligned} \int_{\Omega} \frac{p^2(\mathbf{y}|\theta_0+h)}{p(\mathbf{y}|\theta_0)} d\mathbf{y} &= \frac{1}{(\pi\sigma^2)^N} \int_{\Omega} e^{-\frac{1}{\sigma^2}(\|\mathbf{x}\|^2 + 2\|\mathbf{m}(\theta_0+h)\|^2 - \|\mathbf{m}(\theta_0)\|^2 - \|2\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2)} d\mathbf{y} \\ &= \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(2\|\mathbf{m}(\theta_0+h)\|^2 + \|\mathbf{m}(\theta_0)\|^2 - \|2\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2)} \underbrace{\int_{\Omega} e^{-\frac{1}{\sigma^2}\|\mathbf{x}\|^2} d\mathbf{x}}_{=(\pi\sigma^2)^N} \\ &= e^{\frac{2}{\sigma^2}\|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2}. \end{aligned} \quad (4.34)$$

Finalement on obtient l'expression de la borne de Chapman-Robbins pour un modèle d'observation gaussienne à moyenne paramétrée

$$EQM_{ChR}(\theta_0) = \max_h \frac{h^2}{e^{\frac{2}{\sigma^2}\|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2} - 1}. \quad (4.35)$$

Dans le contexte bayésien, la borne de Bobrovsky-Zakaï s'écrit (3.112)

$$EQM_{BZ}^{globale} = \max_h \frac{h^2}{\int_{\Omega} \int_{\Theta} \frac{p^2(\mathbf{y}, \theta+h)}{p(\mathbf{y}, \theta)} d\theta d\mathbf{y}} - 1. \quad (4.36)$$

On peut réécrire le terme intégrale comme suit :

$$\begin{aligned} \int_{\Omega} \int_{\Theta} \frac{p^2(\mathbf{y}, \theta+h)}{p(\mathbf{y}, \theta)} d\theta d\mathbf{y} &= \int_{\Theta} \frac{p^2(\theta+h)}{p(\theta)} \int_{\Omega} \frac{p^2(\mathbf{y}|\theta+h)}{p(\mathbf{y}|\theta)} d\mathbf{y} d\theta \\ &= \int_{\Theta} \frac{p^2(\theta+h)}{p(\theta)} e^{\frac{2}{\sigma^2} \|\mathbf{m}(\theta+h) - \mathbf{m}(\theta)\|^2} d\theta \end{aligned} \quad (4.37)$$

D'où

$$EQM_{BZ}^{globale} = \max_h \frac{h^2}{\int_{\Theta} \frac{p^2(\theta+h)}{p(\theta)} e^{\frac{2}{\sigma^2} \|\mathbf{m}(\theta+h) - \mathbf{m}(\theta)\|^2} d\theta - 1}. \quad (4.38)$$

4.3.4 Bornes d'Abel

Dans le contexte de paramètres déterministes, la borne d'Abel à l'ordre un est donnée par (4.6)

$$EQM_{Abel(1,1)}(\theta_0) = \max_h \frac{EQM_{CR}(\theta_0) + EQM_{ChR}(\theta_0, h) - 2\psi(\theta_0, h) EQM_{CR}(\theta_0) EQM_{ChR}(\theta_0, h)}{1 - \psi^2(\theta_0, h) EQM_{CR}(\theta_0) EQM_{ChR}(\theta_0, h)}, \quad (4.39)$$

où $EQM_{CR}(\theta_0)$ et $EQM_{ChR}(\theta_0, h)$ sont données par

$$\begin{cases} EQM_{CR}(\theta_0) = \frac{\sigma^2}{2 \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2}, \\ EQM_{ChR}(\theta_0, h) = \frac{h^2}{e^{\frac{2}{\sigma^2} \|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2} - 1}. \end{cases} \quad (4.40)$$

Afin d'obtenir une forme analytique de la borne d'Abel, il suffit maintenant de calculer $\psi(\theta_0, h)$ donné par

$$\psi(\theta_0, h) = \frac{1}{h} \int_{\Omega} p(\mathbf{y}|\theta_0+h) \frac{\partial \ln p(\mathbf{y}|\theta)}{\partial \theta} \Big|_{\theta_0} d\mathbf{y}. \quad (4.41)$$

En posant, $\beta(\mathbf{y}, \theta) = -\frac{1}{\sigma^2} (\mathbf{y} - \mathbf{m}(\theta))^H (\mathbf{y} - \mathbf{m}(\theta))$ on obtient

$$\begin{aligned} \psi(\theta_0, h) &= \frac{1}{h} \int_{\Omega} p(\mathbf{y}|\theta_0+h) \frac{\partial \beta(\mathbf{y}, \theta)}{\partial \theta} \Big|_{\theta_0} d\mathbf{y} \\ &= \frac{2}{h\sigma^2} \int_{\Omega} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} (\mathbf{y} - \mathbf{m}(\theta_0)) \right\} p(\mathbf{y}|\theta_0+h) d\mathbf{y} \\ &= \frac{2}{h\sigma^2} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} \int_{\Omega} \mathbf{y} p(\mathbf{y}|\theta_0+h) d\mathbf{y} - \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} \mathbf{m}(\theta_0) \right\} \\ &= \frac{2}{h\sigma^2} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} \mathbf{m}(\theta_0+h) - \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} \mathbf{m}(\theta_0) \right\} \end{aligned}$$

$$\psi(\theta_0, h) = \frac{2}{h\sigma^2} \operatorname{Re} \left\{ \left. \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \right|_{\theta_0} (\mathbf{m}(\theta_0 + h) - \mathbf{m}(\theta_0)) \right\}. \quad (4.42)$$

Dans le contexte bayésien, la borne d'Abel s'écrit (4.9)

$$EQM_{Abel(1,1)}^{globale} = \max_h \frac{EQM_{CR}^{globale} + EQM_{BZ}^{globale}(h) - 2\psi(h) EQM_{CR}^{globale} EQM_{BZ}^{globale}(h)}{1 - \psi^2(h) EQM_{CR}^{globale} EQM_{BZ}^{globale}(h)}, \quad (4.43)$$

où $EQM_{CR}^{globale}$, $EQM_{BZ}^{globale}(h)$ et $\psi(h)$ sont données par

$$\begin{cases} EQM_{CR}^{globale} = \frac{1}{\frac{2}{\sqrt{2\pi}\sigma^2\sigma_\theta} \int_{\Theta} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right\|^2 e^{-\frac{1}{2\sigma_\theta^2}(\theta-\mu)^2} d\theta + \frac{1}{\sigma_\theta^2}}, \\ EQM_{BZ}^{globale}(h) = \frac{h^2}{\frac{1}{\sqrt{2\pi}\sigma_\theta} \int_{\Theta} e^{-\frac{2}{\sigma_\theta^2} \|\mathbf{m}(\theta+h) - \mathbf{m}(\theta)\|^2 - \frac{1}{2\sigma_\theta^2}(\theta-\mu+h(2-\sqrt{2}))(\theta-\mu+h(2+\sqrt{2}))} d\theta - 1}}, \end{cases} \quad (4.44)$$

et

$$\begin{aligned} \psi(h) &= \frac{1}{h} \int_{\Theta} \int_{\Omega} p(\mathbf{y}, \theta + h) \frac{\partial \ln p(\mathbf{y}, \theta)}{\partial \theta} d\mathbf{y} d\theta \\ &= \frac{1}{h} \int_{\Theta} p(\theta + h) \int_{\Omega} p(\mathbf{y} | \theta + h) \left(\frac{\partial \ln p(\mathbf{y} | \theta) + \ln p(\theta)}{\partial \theta} \right) d\mathbf{y} d\theta \\ &= \frac{1}{h} \int_{\Theta} p(\theta + h) \int_{\Omega} p(\mathbf{y} | \theta + h) \frac{\partial \ln p(\mathbf{y} | \theta)}{\partial \theta} d\mathbf{y} d\theta \\ &\quad + \frac{1}{h} \int_{\Theta} p(\theta + h) \frac{\partial \ln p(\theta)}{\partial \theta} d\theta \\ &= \int_{\Theta} \psi(\theta_0, h) p(\theta + h) d\theta + \frac{1}{h} \int_{\Theta} p(\theta + h) \frac{\partial \ln p(\theta)}{\partial \theta} d\theta. \end{aligned} \quad (4.45)$$

Le deuxième terme de $\psi(h)$ peut se réécrire

$$\begin{aligned} \frac{1}{h} \int_{\Theta} p(\theta + h) \frac{\partial \ln p(\theta)}{\partial \theta} d\theta &= \frac{1}{h} \int_{\Theta} \left(-\frac{\theta - \mu}{\sigma_\theta^2} \right) p(\theta + h) d\theta \\ &= -\frac{1}{h\sigma_\theta^2} \int_{\Theta} (\theta - \mu) p(\theta + h) d\theta \\ &= \frac{1}{\sigma_\theta^2}. \end{aligned} \quad (4.46)$$

Finalement

$$\psi(h) = \frac{1}{\sigma_\theta^2} + \frac{2}{h\sigma^2} \int_{\Theta} \operatorname{Re} \left\{ \left. \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \right|_{\theta_0} (\mathbf{m}(\theta_0 + h) - \mathbf{m}(\theta_0)) \right\} p(\theta + h) d\theta. \quad (4.47)$$

4.3.5 Bornes de Weiss-Weinstein

Nous utiliserons ici les bornes de Weiss-Weinstein pour un point test h . Dans le contexte déterministe, nous avons pour la borne de Weiss-Weinstein déterministe (3.149)

$$EQM_{WWD(1)}(h, s, \theta_0) = \max_{h, s} \frac{\left(h \mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\frac{p(\mathbf{y}|\theta_0-h)}{p(\mathbf{y}|\theta_0)} \right)^{1-s} \right] \right)^2}{\mathbb{E}_{\mathbf{y}|\theta_0} \left[\left(\left(\frac{p(\mathbf{y}|\theta_0+h)}{p(\mathbf{y}|\theta_0)} \right)^s - \left(\frac{p(\mathbf{y}|\theta_0-h)}{p(\mathbf{y}|\theta_0)} \right)^{1-s} \right)^2 \right]}, \quad (4.48)$$

En posant

$$\eta(\alpha|\beta) = \ln \int_{\Omega} \frac{p^\alpha(\mathbf{y}|\theta_0 + \beta)}{p^{\alpha-1}(\mathbf{y}|\theta_0)} d\mathbf{y}, \quad (4.49)$$

on peut réécrire $EQM_{WWD(1)}(h, s, \theta_0)$ comme

$$EQM_{WWD(1)}(h, s, \theta_0) = \max_{h, s} \frac{h^2 e^{2\eta(s|h)}}{e^{\eta(2s|h)} + e^{\eta(2-2s|h)} - 2e^{\eta(s|2h)}}. \quad (4.50)$$

Étudions le terme $\eta(\alpha|\beta)$ pour le modèle d'observations gaussiennes à moyenne paramétrée

$$\begin{aligned} \frac{p^\alpha(\mathbf{y}|\theta_0 + \beta)}{p^{\alpha-1}(\mathbf{y}|\theta_0)} &= \frac{1}{(\pi\sigma^2)^N} \frac{e^{-\frac{\alpha}{\sigma^2}(\mathbf{y}-\mathbf{m}(\theta_0+\beta))^H(\mathbf{y}-\mathbf{m}(\theta_0+\beta))}}{e^{-\frac{\alpha-1}{\sigma^2}(\mathbf{y}-\mathbf{m}(\theta))^H(\mathbf{y}-\mathbf{m}(\theta))}} \\ &= \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(\alpha(\mathbf{y}-\mathbf{m}(\theta_0+\beta))^H(\mathbf{y}-\mathbf{m}(\theta_0+\beta)) - (\alpha-1)(\mathbf{y}-\mathbf{m}(\theta))^H(\mathbf{y}-\mathbf{m}(\theta)))} \\ &= \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(\|\mathbf{y}\|^2 + \alpha\|\mathbf{m}(\theta_0+\beta)\|^2 - (\alpha-1)\|\mathbf{m}(\theta_0)\|^2 - 2\operatorname{Re}\{\mathbf{y}^H(\alpha\mathbf{m}(\theta_0+\beta) - (\alpha-1)\mathbf{m}(\theta_0))\})}. \end{aligned} \quad (4.51)$$

On pose $\mathbf{x} = \mathbf{y} - (\alpha\mathbf{m}(\theta_0 + \beta) - (\alpha - 1)\mathbf{m}(\theta_0))$. Remarquons que

$$\|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 + \|\alpha\mathbf{m}(\theta_0 + \beta) - (\alpha - 1)\mathbf{m}(\theta_0)\|^2 - 2\operatorname{Re}\{\mathbf{y}^H(\alpha\mathbf{m}(\theta_0 + \beta) - (\alpha - 1)\mathbf{m}(\theta_0))\}. \quad (4.52)$$

En reportant (4.52) dans (4.51), on obtient

$$\begin{aligned} \int_{\Omega} \frac{p^\alpha(\mathbf{y}|\theta_0 + \beta)}{p^{\alpha-1}(\mathbf{y}|\theta_0)} d\mathbf{y} &= \frac{1}{(\pi\sigma^2)^N} \int_{\Omega} e^{-\frac{1}{\sigma^2}(\|\mathbf{x}\|^2 - \|\alpha\mathbf{m}(\theta_0+\beta) - (\alpha-1)\mathbf{m}(\theta_0)\|^2 + \alpha\|\mathbf{m}(\theta_0+\beta)\|^2 - (\alpha-1)\|\mathbf{m}(\theta_0)\|^2)} d\mathbf{x} \\ &= \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(-\|\alpha\mathbf{m}(\theta_0+\beta) - (\alpha-1)\mathbf{m}(\theta_0)\|^2 + \alpha\|\mathbf{m}(\theta_0+\beta)\|^2 - (\alpha-1)\|\mathbf{m}(\theta_0)\|^2)} \underbrace{\int_{\Omega} e^{-\frac{\|\mathbf{x}\|^2}{\sigma^2}} d\mathbf{x}}_{=(\pi\sigma^2)^N} \\ &= e^{\frac{\alpha(\alpha-1)}{\sigma^2}\|\mathbf{m}(\theta_0+\beta) - \mathbf{m}(\theta_0)\|^2}. \end{aligned} \quad (4.53)$$

Donc

$$\eta(\alpha|\beta) = \ln \int_{\Omega} \frac{p^\alpha(\mathbf{y}|\theta_0 + \beta)}{p^{\alpha-1}(\mathbf{y}|\theta_0)} d\mathbf{y}$$

$$\eta(\alpha|\beta) = \frac{\alpha(\alpha-1)}{\sigma^2} \|\mathbf{m}(\theta_0 + \beta) - \mathbf{m}(\theta_0)\|^2, \quad (4.54)$$

et la borne de Weiss-Weinstein déterministe s'écrit

$$EQM_{WWD(1)}(h, s, \theta_0) = \max_{h, s} \frac{h^2 e^{\frac{2s(s-1)}{\sigma^2} \|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2}}{e^{\frac{2s(2s-1)}{\sigma^2} \|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2} + e^{\frac{2(s-1)(2s-1)}{\sigma^2} \|\mathbf{m}(\theta_0-h) - \mathbf{m}(\theta_0)\|^2} - 2e^{\frac{s(s-1)}{\sigma^2} \|\mathbf{m}(\theta_0+2h) - \mathbf{m}(\theta_0)\|^2}}. \quad (4.55)$$

Dans le contexte bayésien, la borne de Weiss-Weinstein est donnée par

$$EQM_{WW(1)}(h, s) = \max_{h, s} \frac{h^2 e^{2\eta(s, h)}}{e^{\eta(2s, h)} + e^{\eta(2-2s, -h)} - 2e^{\eta(s, 2h)}}, \quad (4.56)$$

où $\eta(\alpha, \beta)$ est donnée par

$$\eta(\alpha, \beta) = \ln \int_{\Theta} \int_{\Omega} \frac{p^\alpha(\mathbf{y}, \theta + \beta)}{p^{\alpha-1}(\mathbf{y}, \theta)} d\mathbf{y} d\theta. \quad (4.57)$$

Le terme intégrale peut se réécrire

$$\begin{aligned} \int_{\Theta} \int_{\Omega} \frac{p^\alpha(\mathbf{y}, \theta + \beta)}{p^{\alpha-1}(\mathbf{y}, \theta)} d\mathbf{y} d\theta &= \int_{\Theta} \frac{p^\alpha(\theta + \beta)}{p^{\alpha-1}(\theta)} \int_{\Omega} \frac{p^\alpha(\mathbf{y}|\theta + \beta)}{p^{\alpha-1}(\mathbf{y}|\theta)} d\mathbf{y} d\theta \\ &= \int_{\Theta} \frac{p^\alpha(\theta + \beta)}{p^{\alpha-1}(\theta)} e^{\frac{\alpha(\alpha-1)}{\sigma^2} \|\mathbf{m}(\theta+\beta) - \mathbf{m}(\theta)\|^2} d\theta \\ &= \frac{1}{\sqrt{2\pi\sigma_\theta}} \int_{\Theta} e^{\frac{\alpha(\alpha-1)}{\sigma^2} \|\mathbf{m}(\theta+\beta) - \mathbf{m}(\theta)\|^2 - \frac{1}{2\sigma_\theta^2} (\theta - (\sqrt{\alpha(\alpha-1)} - \alpha)h - \mu)(\theta + (\sqrt{\alpha(\alpha-1)} + \alpha)h - \mu)} d\theta. \end{aligned} \quad (4.58)$$

En conséquence, le terme $\eta(\alpha, \beta)$ est donné par

$$\eta(\alpha, \beta) = \ln \frac{1}{\sqrt{2\pi\sigma_\theta}} \int_{\Theta} e^{\frac{\alpha(\alpha-1)}{\sigma^2} \|\mathbf{m}(\theta+\beta) - \mathbf{m}(\theta)\|^2 - \frac{1}{2\sigma_\theta^2} (\theta - (\sqrt{\alpha(\alpha-1)} - \alpha)h - \mu)(\theta + (\sqrt{\alpha(\alpha-1)} + \alpha)h - \mu)} d\theta. \quad (4.59)$$

En reportant (4.59) dans (4.56) on obtient la borne de Weiss-Weinstein.

4.4 Analyse spectrale et synchronisation de fréquence porteuse

Dans cette section, les résultats obtenus précédemment pour le modèle d'observations gaussiennes à moyenne paramétrée sont appliqués à un problème d'estimation de porteuse dans un contexte de communications numériques avec symboles pilotes. Le résumé des résultats est reporté en annexe B.

4.4.1 Modèles des observations

Le modèle des observations est le suivant :

$$y_k = a_k e^{j2\pi k\theta} + b_k, \quad k = 0, \dots, N-1, \quad (4.60)$$

où la $k^{\text{ième}}$ observation complexe bruitée y_k , échantillonnée au rythme symbole, est un signal modulé linéairement, obtenu en appliquant une séquence de symboles pilotes connus $\mathbf{a} =$

$[a_0 \ a_1 \ \dots \ a_{N-1}]^T$ à un filtre de Nyquist. Les observations sont supposées indépendantes. $\{b_k\}$ est une suite de variables aléatoires complexes circulaires, i.i.d, gaussiennes de moyenne nulle et de variance σ^2 . Le paramètre réel inconnu $\theta \in \Theta =]-\frac{1}{2}, \frac{1}{2}]$ correspond à la fréquence porteuse. Dans le contexte bayésien, θ sera supposé aléatoire avec une densité de probabilité *a priori* $p(\theta)$ gaussienne de moyenne μ et de variance σ_θ^2 :

$$p(\theta) = \frac{1}{\sqrt{2\pi\sigma_\theta}} e^{-\frac{1}{2\sigma_\theta^2}(\theta-\mu)^2}. \quad (4.61)$$

Du fait de l'indépendance des observations, la fonction de vraisemblance s'écrit

$$p(\mathbf{y}|\theta) = \prod_{k=0}^{N-1} p(y_k|\theta) = \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2} \left(\|\mathbf{y}\|^2 + \|\mathbf{a}\|^2 - 2\operatorname{Re} \left\{ \sum_{k=0}^{N-1} y_k^* a_k e^{j2\pi k\theta} \right\} \right)}. \quad (4.62)$$

Ce modèle est un cas particulier du modèle d'observations gaussiennes à moyenne paramétrée (4.11) où

$$\mathbf{m}(\theta) = \mathbf{a} \odot \mathbf{s}(\theta), \quad (4.63)$$

avec

$$\mathbf{s}(\theta) = [1 \ e^{j(2\pi\theta)} \ \dots \ e^{j(2\pi(N-1)\theta)}]^T. \quad (4.64)$$

4.4.2 Formes analytiques des bornes déterministes et bayésiennes

À notre connaissance, seule la forme analytique de la borne de Cramér-Rao et de la borne de Chapman-Robbins sont connues dans ce cas (voir [ABL04b]). Nous donnons ici les résultats des différentes formes analytiques pour les bornes minimales considérées précédemment.

4.4.2.1 Bornes de Cramér-Rao

Dans le contexte d'un paramètre supposé déterministe, la borne de Cramér-Rao est donnée par (4.16)

$$EQM_{CR}(\theta_0) = \frac{\sigma^2}{2 \left\| \left. \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right|_{\theta_0} \right\|^2}. \quad (4.65)$$

Le terme $\left\| \left. \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right|_{\theta_0} \right\|^2$ s'écrit

$$\begin{aligned} \left\| \left. \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right|_{\theta_0} \right\|^2 &= \left\| \mathbf{a} \odot \left. \frac{\partial \mathbf{s}(\theta)}{\partial \theta} \right|_{\theta_0} \right\|^2 \\ &= \sum_{k=0}^{N-1} \|a_k\|^2 \left(j2\pi k e^{j2\pi k\theta} \right) \left(-j2\pi k e^{-j2\pi k\theta} \right) \\ &= 4\pi^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2. \end{aligned} \quad (4.66)$$

La borne de Cramér-Rao s'écrit par conséquent

$$EQM_{CR}(\theta_0) = \frac{\sigma^2}{8\pi^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2}. \quad (4.67)$$

Dans le contexte bayésien, en remarquant que l'expression ci-dessus ne dépend pas de θ_0 , nous obtenons d'après (4.19)

$$\begin{aligned}
 EQM_{CR}^{globale} &= \left(\frac{2}{\sigma^2} \int_{\Theta} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right\|^2 p(\theta) d\theta + \frac{1}{\sigma_\theta^2} \right)^{-1} \\
 &= \frac{1}{\frac{8\pi^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + \frac{1}{\sigma_\theta^2}}.
 \end{aligned}$$

$$EQM_{CR}^{globale} = \frac{\sigma_\theta^2}{8\pi^2 \frac{\sigma_\theta^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + 1}. \quad (4.68)$$

4.4.2.2 Bornes de Bhattacharyya

Dans le contexte déterministe, la borne de Bhattacharyya à l'ordre 2 s'écrit (4.28)

$$EQM_{Bhatt(2)}(\theta_0) = \frac{B_{2,2}}{B_{1,1}B_{2,2} - B_{1,2}^2}$$

$$EQM_{Bhatt(2)}(\theta_0) = \frac{\frac{2}{\sigma^2} \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \right\|_{\theta_0}^2 + \frac{4}{\sigma^4} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right\|_{\theta_0}^4}{\frac{4}{\sigma^4} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right\|_{\theta_0}^2 \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \right\|_{\theta_0}^2 + \frac{2}{\sigma^6} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right\|_{\theta_0}^6 - \frac{4}{\sigma^4} \left(\text{Re} \left\{ \frac{\partial^2 \mathbf{m}^H(\theta)}{\partial \theta^2} \right\|_{\theta_0} \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right\|_{\theta_0} \right)^2}.$$

(4.69)

Ici, le terme $B_{1,2}$ est nul puisque

$$\begin{aligned}
 B_{1,2} &= \frac{2}{\sigma^2} \text{Re} \left\{ \frac{\partial^2 \mathbf{m}^H(\theta)}{\partial \theta^2} \Big|_{\theta_0} \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\} \\
 &= \frac{2}{\sigma^2} \text{Re} \left\{ \mathbf{a}^H \odot \frac{\partial^2 \mathbf{s}^H(\theta)}{\partial \theta^2} \Big|_{\theta_0} \mathbf{a} \odot \frac{\partial \mathbf{s}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\} \\
 &= \frac{2}{\sigma^2} \text{Re} \left\{ -j8\pi^3 \sum_{k=0}^{N-1} k^3 \|a_k\|^2 \right\} = 0.
 \end{aligned} \quad (4.70)$$

D'où

$$EQM_{Bhatt(2)}(\theta_0) = \frac{\sigma^2}{8\pi^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2}. \quad (4.71)$$

Pour ce problème d'estimation, on retrouve la borne de Cramér-Rao. Dans le contexte bayésien, un calcul similaire conduit à la borne de Cramér-Rao bayésienne.

Il faudrait donc augmenter l'ordre de la borne de Bhattacharyya afin d'obtenir une borne plus pertinente. La borne à l'ordre 3 est donnée dans [FL01] [FL02] pour le cas particulier où $a_k = 1 \forall k$. Au vue de la faible différence entre cette borne et la borne de Cramér-Rao (voir figure 1 dans [FL01]), il nous a semblé judicieux de ne pas continuer dans cette voie.

4.4.2.3 Bornes de Chapman-Robbins et de Bobrovsky-Zakai

Dans le contexte d'un paramètre supposé déterministe, la borne de Chapman-Robbins est donnée par (4.35)

$$EQM_{ChR}(\theta_0) = \max_h \frac{h^2}{e^{\frac{2}{\sigma^2} \|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2} - 1}. \quad (4.72)$$

Le terme $\|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2$ s'écrit

$$\begin{aligned} \|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2 &= \sum_{k=0}^{N-1} \|a_k\|^2 \left(e^{j2\pi k(\theta_0+h)} - e^{j2\pi k\theta_0} \right) \left(e^{-j2\pi k(\theta_0+h)} - e^{-j2\pi k\theta_0} \right) \\ &= \sum_{k=0}^{N-1} \|a_k\|^2 \left(2 - 2 \operatorname{Re} \left\{ e^{j2\pi kh} \right\} \right) \\ &= 2 \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh)). \end{aligned} \quad (4.73)$$

La borne de Chapman-Robbins s'écrit par conséquent

$$EQM_{ChR}(\theta_0) = \max_h \frac{h^2}{e^{\frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1}. \quad (4.74)$$

Dans le contexte bayésien, du fait que la borne de Chapman-Robbins ne dépend pas de θ_0 , on obtient (4.38)

$$\begin{aligned} EQM_{BZ}^{globale} &= \max_h \frac{h^2}{\int_{\Omega} \int_{\Theta} \frac{p^2(\mathbf{y}, \theta+h)}{p(\mathbf{y}, \theta)} d\theta d\mathbf{y} - 1} \\ &= \max_h \frac{h^2}{\int_{\Theta} \frac{p^2(\theta+h)}{p(\theta)} \int_{\Omega} \frac{p^2(\mathbf{y}|\theta+h)}{p(\mathbf{y}|\theta)} d\mathbf{y} d\theta - 1} \\ &= \max_h \frac{h^2}{e^{\frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} \int_{\Theta} \frac{p^2(\theta+h)}{p(\theta)} d\theta - 1} \end{aligned} \quad (4.75)$$

Le terme $\int_{\Theta} \frac{p^2(\theta+h)}{p(\theta)} d\theta$ devient

$$\begin{aligned} \int_{\Theta} \frac{p^2(\theta+h)}{p(\theta)} d\theta &= \frac{1}{\sqrt{2\pi}\sigma_{\theta}} e^{-\frac{1}{2\sigma_{\theta}^2} [2h^2 - 4h\mu + \mu^2]} \int_{\Theta} e^{-\frac{1}{2\sigma_{\theta}^2} [\theta^2 + 2\theta(2h-\mu)]} d\theta \\ &= e^{-\frac{1}{2\sigma_{\theta}^2} [2h^2 - 4h\mu + \mu^2] + \frac{(2h-\mu)^2}{2\sigma_{\theta}^2}} = e^{\frac{h^2}{\sigma_{\theta}^2}}, \end{aligned} \quad (4.76)$$

où l'intégrale $\int_{\Theta} e^{-\frac{1}{2\sigma_{\theta}^2} [\theta^2 + 2\theta(2h-\mu)]} d\theta$ est obtenue grâce à [GR94] page 355, équation (BI((28))(1),

à savoir

$$\int_{-\infty}^{\infty} e^{-p^2 x^2 \pm qx} dx = \frac{\sqrt{\pi}}{\operatorname{abs}(p)} e^{\frac{q^2}{4p^2}}. \quad (4.77)$$

Finalement, la borne de Bobrovsky-Zakaï est donnée par

$$\boxed{EQM_{BZ}^{globale} = \max_h \frac{h^2}{e^{\frac{h^2}{\sigma^2} + \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1}}. \quad (4.78)$$

4.4.2.4 Bornes d'Abel

Dans le contexte d'un paramètre supposé déterministe, la borne d'Abel est donnée par (A.6)

$$EQM_{Abel(1,1)}(\theta_0) = \max_h \frac{EQM_{CR}(\theta_0) + EQM_{ChR}(\theta_0, h) - 2\psi(\theta_0, h) EQM_{CR}(\theta_0) EQM_{ChR}(\theta_0, h)}{1 - \psi^2(\theta_0, h) EQM_{CR}(\theta_0) EQM_{ChR}(\theta_0, h)}, \quad (4.79)$$

où $EQM_{CR}(\theta_0)$ et $EQM_{ChR}(\theta_0, h)$ sont données par (4.67) et (4.74)

$$\begin{cases} EQM_{CR}(\theta_0) = \frac{\sigma^2}{8\pi^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2}, \\ EQM_{ChR}(\theta_0, h) = \frac{h^2}{e^{\frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1}. \end{cases} \quad (4.80)$$

Le terme $\psi(\theta_0, h)$, pour le modèle d'observations gaussiennes à moyenne paramétrée est donnée par (4.42)

$$\begin{aligned} \psi(\theta_0, h) &= \frac{2}{h\sigma^2} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} (\mathbf{m}(\theta_0 + h) - \mathbf{m}(\theta_0)) \right\} \\ &= \frac{2}{h\sigma^2} \operatorname{Re} \left\{ \mathbf{a}^H \odot \frac{\partial \mathbf{s}^H(\theta)}{\partial \theta} \Big|_{\theta_0} \mathbf{a} \odot (\mathbf{s}(\theta_0 + h) - \mathbf{s}(\theta_0)) \right\} \\ &= \frac{2}{h\sigma^2} \operatorname{Re} \left\{ \sum_{k=0}^{N-1} \|a_k\|^2 \left(e^{j2\pi k(\theta_0 + h)} - e^{j2\pi k\theta_0} \right) \frac{\partial e^{-j2\pi k\theta}}{\partial \theta} \Big|_{\theta=\theta_0} \right\} \\ &= -\frac{4\pi}{h\sigma^2} \operatorname{Re} \left\{ \sum_{k=0}^{N-1} jk \|a_k\|^2 \left(e^{j2\pi kh} - 1 \right) \right\} \\ &= -\frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \operatorname{Re} \left\{ j \left(e^{j2\pi kh} - 1 \right) \right\} \end{aligned}$$

$$\boxed{\psi(\theta_0, h) = \frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh)}. \quad (4.81)$$

Finalement, la borne d'Abel est donnée par

$$\boxed{EQM_{Abel(1,1)}(\theta_0) = \max_h \frac{\frac{8\pi^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + \frac{1}{h^2} \left(e^{\frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1 \right) - \frac{8\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh)}{\frac{8\pi^2}{h^2\sigma^2} \left(e^{\frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1 \right) \sum_{k=0}^{N-1} \|a_k\|^2 k^2 - \frac{16\pi^2}{h^2\sigma^4} \left(\sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh) \right)^2}}. \quad (4.82)$$

Dans le contexte bayésien, la borne d'Abel bayésienne s'écrit (4.43)

$$EQM_{Abel(1,1)}^{globale} = \max_h \frac{EQM_{CR}^{globale} + EQM_{BZ}^{globale}(h) - 2\psi(h) EQM_{CR}^{globale} EQM_{BZ}^{globale}(h)}{1 - \psi^2(h) EQM_{CR}^{globale} EQM_{BZ}^{globale}(h)}, \quad (4.83)$$

où $EQM_{CR}^{globale}$ et $EQM_{BZ}^{globale}(h)$ sont données par (4.68) et (4.78)

$$\begin{cases} EQM_{CR}^{globale} = \frac{\sigma_\theta^2}{8\pi^2 \frac{\sigma_\theta^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + 1}, \\ EQM_{BZ}^{globale}(h) = \frac{h^2}{e^{\frac{h^2}{\sigma_\theta^2} + \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1}. \end{cases} \quad (4.84)$$

D'après (4.47), le terme $\psi(h)$ s'écrit pour le modèle d'observations gaussiennes à moyenne paramétrée

$$\begin{aligned} \psi(h) &= \frac{1}{\sigma_\theta^2} + \frac{2}{h\sigma^2} \int_{\Theta} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} (\mathbf{m}(\theta_0 + h) - \mathbf{m}(\theta_0)) \right\} p(\theta + h) d\theta \\ &= \frac{1}{\sigma_\theta^2} + \frac{4\pi}{h\sigma^2} \int_{\Theta} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh) p(\theta + h) d\theta \end{aligned}$$

$$\boxed{\psi(h) = \frac{1}{\sigma_\theta^2} + \frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh)}. \quad (4.85)$$

On en déduit la borne d'Abel bayésienne

$$EQM_{Abel(1,1)}^{globale} = \max_h \frac{\frac{8\pi^2 \frac{\sigma_\theta^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + 1}{\sigma_\theta^2} + \frac{\frac{h^2}{\sigma_\theta^2} + \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))}{h^2} - 1 - 2 \left(\frac{1}{\sigma_\theta^2} + \frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh) \right)}{\frac{1}{\sigma_\theta^2 h^2} \left(8\pi^2 \frac{\sigma_\theta^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + 1 \right) \left(e^{\frac{h^2}{\sigma_\theta^2} + \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1 \right) - \left(\frac{1}{\sigma_\theta^2} + \frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh) \right)^2}. \quad (4.86)$$

4.4.2.5 Bornes de Weiss-Weinstein

La borne de Weiss-Weinstein déterministe s'écrit pour le modèle d'observations gaussiennes à moyenne paramétrée

$$EQM_{WWD(1)}(\theta_0) = \max_{h,s} \frac{h^2 e^{2\eta(s|h)}}{e^{\eta(2s|h)} + e^{\eta(2-2s|h)} - 2e^{\eta(s|2h)}}, \quad (4.87)$$

avec

$$\eta(\alpha|\beta) = \frac{\alpha(\alpha-1)}{\sigma^2} \|\mathbf{m}(\theta_0 + \beta) - \mathbf{m}(\theta_0)\|^2$$

$$\boxed{\eta(\alpha|\beta) = \frac{2\alpha(\alpha-1)}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi k\beta))}, \quad (4.88)$$

d'après (4.73). On pose $\xi(h) = \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))$ et on en déduit la borne de Weiss-Weinstein déterministe

$$EQM_{WW(1)}(\theta_0) = \max_{h,s} \frac{h^2 e^{s(s-1)\xi(h)}}{e^{s(2s-1)\xi(h)} + e^{(1-s)(1-2s)\xi(h)} - 2e^{\frac{1}{2}s(s-1)\xi(2h)}}. \quad (4.89)$$

Dans le contexte bayésien, la borne de Weiss-Weinstein est donnée par

$$EQM_{WW(1)}^{globale} = \max_{h,s} \frac{h^2 e^{2\eta(s,h)}}{e^{\eta(2s,h)} + e^{\eta(2-2s,-h)} - 2e^{\eta(s,2h)}}, \quad (4.90)$$

où $\eta(\alpha, \beta)$ est donnée par

$$\eta(\alpha, \beta) = \ln \int_{\Theta} \int_{\Omega} \frac{p^\alpha(\mathbf{y}, \theta + \beta)}{p^{\alpha-1}(\mathbf{y}, \theta)} d\mathbf{y} d\theta. \quad (4.91)$$

Du fait que la borne de Weiss-Weinstein déterministe ne dépend pas de θ_0 , on obtient

$$\begin{aligned} \int_{\Theta} \int_{\Omega} \frac{p^\alpha(\mathbf{y}, \theta + \beta)}{p^{\alpha-1}(\mathbf{y}, \theta)} d\mathbf{y} d\theta &= \int_{\Theta} \frac{p^\alpha(\theta + \beta)}{p^{\alpha-1}(\theta)} \int_{\Omega} \frac{p^\alpha(\mathbf{y}|\theta + \beta)}{p^{\alpha-1}(\mathbf{y}|\theta)} d\mathbf{y} d\theta \\ &= e^{\frac{2\alpha(\alpha-1)}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi k\beta))} \int_{\Theta} \frac{p^\alpha(\theta + \beta)}{p^{\alpha-1}(\theta)} d\theta \end{aligned} \quad (4.92)$$

Détaillons le terme *a priori*

$$\begin{aligned} \int_{\Theta} \frac{p^\alpha(\theta + \beta)}{p^{\alpha-1}(\theta)} d\theta &= \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2\sigma_\theta^2}[\alpha\beta^2 - 2\alpha\beta\mu + \mu^2]} \int_{\Theta} e^{-\frac{1}{2\sigma_\theta^2}[\theta^2 + 2\theta(\alpha\beta - \mu)]} d\theta \\ &= e^{-\frac{1}{2\sigma_\theta^2}[\alpha\beta^2 - 2\alpha\beta\mu + \mu^2] + \frac{(\alpha\beta - \mu)^2}{2\sigma_\theta^2}} = e^{-\frac{\alpha\beta^2}{2\sigma_\theta^2}(1-\alpha)}, \end{aligned} \quad (4.93)$$

où l'intégrale $\int_{\Theta} e^{-\frac{1}{2\sigma_\theta^2}[\theta^2 + 2\theta(\alpha\beta - \mu)]} d\theta$ est obtenue grâce à [GR94] page 355, équation (BI((28))(1) (voir équation (4.77)). Nous obtenons donc pour $\eta(\alpha, \beta)$

$$\eta(\alpha, \beta) = \alpha(\alpha - 1) \left(\frac{2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi k\beta)) + \frac{\beta^2}{2\sigma_\theta^2} \right). \quad (4.94)$$

On pose $\xi(h) = \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))$. Et la borne de Weiss-Weinstein devient

$$EQM_{WW(1)}^{globale} = \max_{h,s} \frac{h^2 e^{s(s-1)\left(\xi(h) + \frac{h^2}{\sigma_\theta^2}\right)}}{e^{s(2s-1)\left(\xi(h) + \frac{h^2}{\sigma_\theta^2}\right)} + e^{(1-s)(1-2s)\left(\xi(h) + \frac{h^2}{\sigma_\theta^2}\right)} - 2e^{s(s-1)\left(\frac{1}{2}\xi(2h) + \frac{2h^2}{\sigma_\theta^2}\right)}}. \quad (4.95)$$

4.4.2.6 Simplification des bornes de Weiss-Weinstein

Les bornes de Weiss-Weinstein requièrent une optimisation selon deux paramètres : s et h . Si l'obtention de la valeur optimale de h semble difficile, Weiss et Weinstein ont remarqué, sans le démontrer, que la valeur s_{opt} qui maximise la borne quelque soit h était généralement

$$s_{opt} = \frac{1}{2}. \quad (4.96)$$

En posant $\xi(h) = \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))$, on peut réécrire la borne de Weiss-Weinstein déterministe comme

$$EQM_{WWD(1)}(\theta_0) = \max_{h,s} \frac{h^2 e^{s(s-1)\xi(h)}}{e^{s(2s-1)\xi(h)} + e^{(1-s)(1-2s)\xi(h)} - 2e^{\frac{1}{2}s(s-1)\xi(2h)}}. \quad (4.97)$$

Pour cette valeur particulière de s on obtient

$$EQM_{WWD(1)}(\theta_0) = \max_h EQM_{WWD(1)}\left(h, \frac{1}{2}, \theta_0\right) = \max_h \frac{h^2}{2} \frac{e^{-\frac{1}{4}\xi(h)}}{1 - e^{-\frac{1}{8}\xi(2h)}}. \quad (4.98)$$

Dans le contexte bayésien, la borne de Weiss-Weinstein obéit à la même règle, à savoir que le point s optimal correspond

$$s = \frac{1}{2}. \quad (4.99)$$

Pour cette valeur particulière de s on obtient

$$EQM_{WW(1)}^{globale} = \max_h EQM_{WW(1)}^{globale}\left(h, \frac{1}{2}\right) = \max_h \frac{h^2}{2} \frac{e^{-\frac{1}{4}\left(\xi(h) + \frac{h^2}{\sigma^2}\right)}}{1 - e^{-\left(\frac{1}{8}\xi(2h) + \frac{h^2}{2\sigma^2}\right)}}. \quad (4.100)$$

On obtient donc des bornes avec une charge de calcul approximativement identique à celle des bornes de Chapman-Robbins (Bobrovsky-Zakaï), et d'Abel.

4.4.2.7 Simulations

Afin d'illustrer les différents résultats concernant les bornes minimales établis ci-dessus ainsi qu'au chapitre précédent, nous présentons ici différents résultats de simulations.

Le scénario est le suivant : on considère une modulation QPSK et $N = 20$ observations. De plus, on peut remarquer que toutes les bornes dépendant du point test h sont paires en h . En conséquence, la recherche du maximum sur le point test h s'effectue sur $\left[0, \frac{1}{2}\right]$. Dans le contexte déterministe, la vraie valeur du paramètre est $\theta_0 = 0$.

On considère l'estimateur du maximum de vraisemblance donné par

$$\hat{\theta}_{MV} = \arg \min_{\theta} \left[\|\mathbf{y}\|^2 + \|\mathbf{a}\|^2 - 2 \operatorname{Re} \left\{ \sum_{k=0}^{N-1} y_k^* a_k e^{j2\pi k\theta} \right\} \right]. \quad (4.101)$$

Dans le contexte bayésien, l'EQM globale de cet estimateur est donnée par (4.1) et on choisit pour la densité *a priori* $\mu = 0$ et $\sigma_\theta^2 = \frac{1}{36}$.

Les figures (4.1) et (4.2) représentent, en fonction du RSB, dans les contextes déterministe et bayésien, les bornes de Weiss-Weinstein avec une optimisation sur h et s , les bornes de Weiss-Weinstein avec optimisation sur h et $s = \frac{1}{2}$, $s = \frac{1}{4}$, et $s = 0$ (identique aux bornes de Weiss-Weinstein avec optimisation sur h et $s = 1$), les bornes de Weiss-Weinstein avec $h = 0$ et $s = 0$

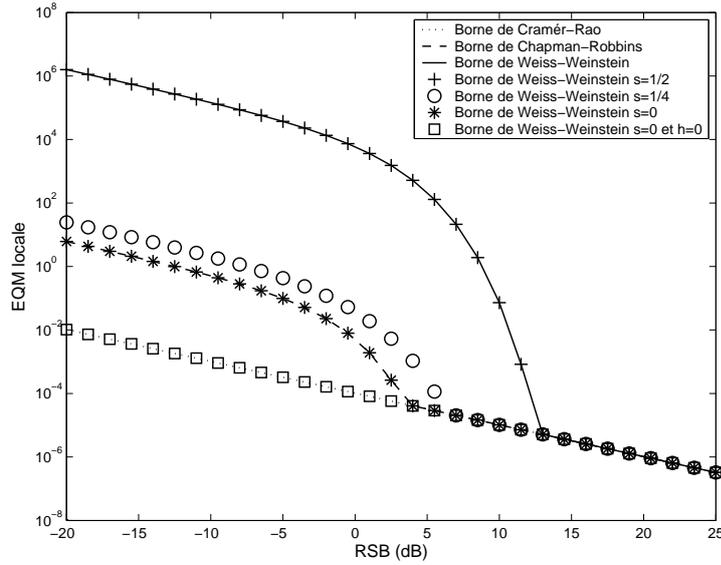


FIGURE 4.1 – Borne de Weiss-Weinstein déterministe avec une optimisation sur h et s , borne de Weiss-Weinstein avec optimisation sur h et $s = \frac{1}{2}$, $s = \frac{1}{4}$, et $s = 0$, borne de Weiss-Weinstein avec $h = 0$ et $s = 0$, la borne de Chapman-Robbins et borne de Cramér-Rao en fonction du RSB. Modulation QPSK et $N = 20$ observations.

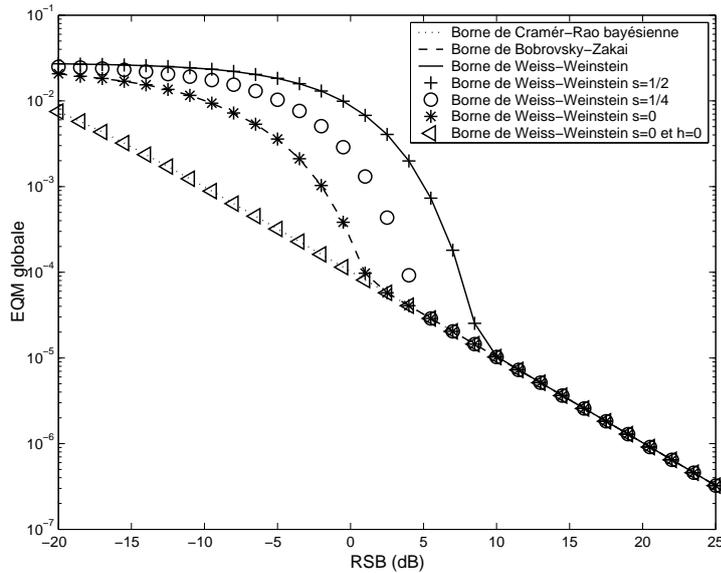


FIGURE 4.2 – Borne de Weiss-Weinstein avec une optimisation sur h et s , borne de Weiss-Weinstein avec optimisation sur h et $s = \frac{1}{2}$, $s = \frac{1}{4}$, et $s = 0$, borne de Weiss-Weinstein avec $h = 0$ et $s = 0$, la borne de Bobrovsky-Zakai et borne de Cramér-Rao bayésienne en fonction du RSB. Modulation QPSK et $N = 20$ observations.

(identique aux bornes de Weiss-Weinstein avec $h = 0$ et $s = 1$), la borne de Chapman-Robbins (Bobrovsky-Zakai dans le contexte bayésien) ainsi que la borne de Cramér-Rao (bayésienne).

On remarque que la valeur $s = \frac{1}{2}$ est bien la valeur optimale pour laquelle on obtient la borne la plus pertinente. Lorsque l'on effectue une optimisation sur h et que l'on impose $s = 0$ ou 1, les bornes de Weiss-Weinstein se ramènent à la borne de Chapman-Robbins (Bobrovsky-Zakaï dans le contexte bayésien). Lorsque l'on impose $h = 0$ et $s = 0$ (ou $h = 0$ et $s = 1$), les bornes de Weiss-Weinstein se ramènent à la borne de Cramér-Rao (bayésienne). Enfin, lorsque l'on optimise sur h et que l'on choisit pour s une valeur arbitraire entre 0 et $\frac{1}{2}$, on obtient une borne sous-optimale par rapport à la vraie borne de Weiss-Weinstein, mais néanmoins supérieure à la borne de Chapman-Robbins (Bobrovsky-Zakaï dans le contexte bayésien).

Les figures (4.3) et (4.4) représentent, en fonction du RSB, les performances locale et globale de l'estimateur du maximum de vraisemblance considéré ainsi que les bornes de Cramér-Rao déterministe et bayésienne, les bornes de Chapman-Robbins dans le contexte déterministe et de Bobrovsky-Zakaï dans le contexte bayésien, les bornes d'Abel déterministe et bayésienne et les bornes de Weiss-Weinstein déterministe et bayésienne. Concernant l'EQM (locale ou globale) de l'estimateur du maximum de vraisemblance, les simulations de Monte-Carlo sont obtenues à partir de 1000 réalisations indépendantes.

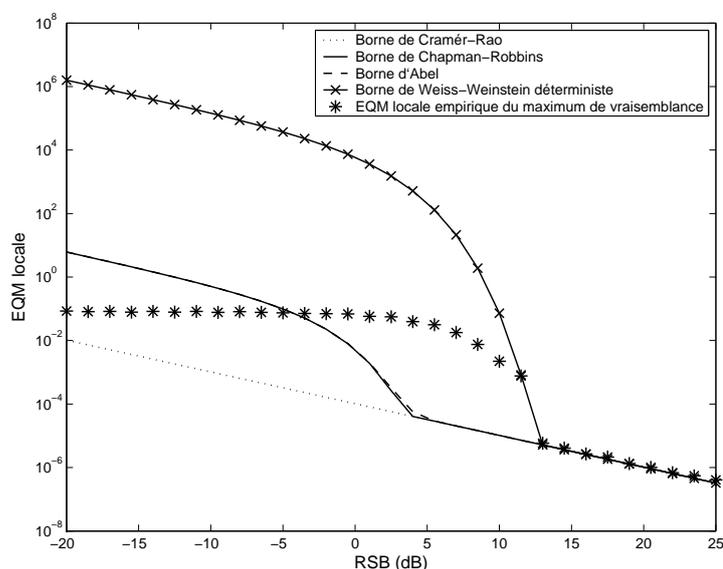


FIGURE 4.3 – EQM locale empirique de l'estimateur du maximum de vraisemblance, borne de Cramér-Rao, borne de Chapman-Robbins, borne d'Abel et borne de Weiss-Weinstein déterministe en fonction du RSB. Modulation QPSK et $N = 20$ observations.

Comme attendu, la borne de Cramér-Rao est la seule à ne pas exhiber un décrochement. On retrouve les relations d'ordre établies au chapitre précédent, à savoir que les bornes de Weiss-Weinstein sont plus pertinentes que les bornes d'Abel, elles mêmes légèrement plus pertinentes que la borne de Chapman-Robbins et la borne de Bobrovsky-Zakaï. En outre, dans le contexte déterministe la borne de Weiss-Weinstein semble être un outil particulièrement intéressant pour prédire précisément le décrochement de l'estimateur du maximum de vraisemblance. Dans le contexte bayésien, la borne de Weiss-Weinstein est la mieux à même de représenter l'EQM globale sur les trois zones de fonctionnement de l'estimateur.

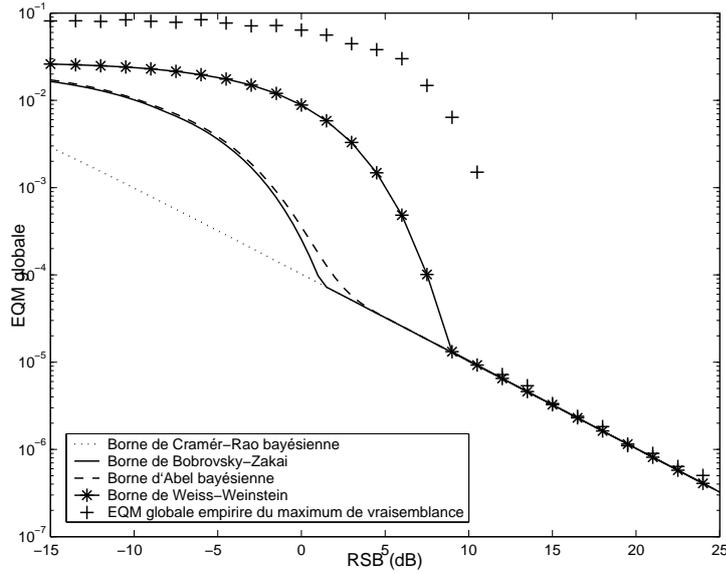


FIGURE 4.4 – EQM globale empirique de l’estimateur du maximum de vraisemblance, borne de Cramér-Rao bayésienne, borne de Bobrovsky-Zakai, borne d’Abel bayésienne et borne de Weiss-Weinstein en fonction du RSB. Modulation QPSK et $N = 20$ observations.

4.4.3 Sur le point test optimal

Outre le fait de fournir les performances ultimes en terme d’EQM, les bornes minimales d’estimations sont un outil attrayant afin de prédire la zone de décrochement des estimateurs.

Les simulations présentées dans cette section ont été effectuées pour le même scénario que ci-avant.

4.4.3.1 Prédiction du décrochement

Basée sur les bornes utilisant des points tests, nous présentons dans cette section un méthode pour prédire la zone de décrochement. Ce prédicteur est basé sur le comportement du point test optimal intervenant dans les bornes de Chapman-Robbins (Bobrovsky-Zakai), Abel, et Weiss-Weinstein. Pour une borne de la forme $f(\theta_0) = \max_h f(\theta_0, h)$, le point test optimal est défini comme la valeur h_{opt} telle que

$$f(\theta_0) = \max_h f(\theta_0, h) = f(\theta_0, h_{opt}). \quad (4.102)$$

Le comportement de h_{opt} en fonction du RSB pour les bornes de Chapman-Robbins (Bobrovsky-Zakai), Abel, et Weiss-Weinstein est reporté figure (4.5).

On observe un saut brusque de h_{opt} lorsque les bornes s’écartent de la borne de Cramér-Rao. Ce comportement est connu pour être relatif à la fonction d’ambiguïté du problème. En effet, les valeurs des points tests qui maximisent la borne sont celles où la fonction d’ambiguïté exhibe des maxima locaux [RM95] [XBR04] [RM97] [TK99] [Xu01] [RAFL06].

En conséquence, à partir du comportement du point test optimal, on peut en déduire le prédicteur suivant

$$\widehat{RSB} = \arg \max_{RSB} \left| \frac{\partial h_{opt}}{\partial RSB} \right|, \quad (4.103)$$

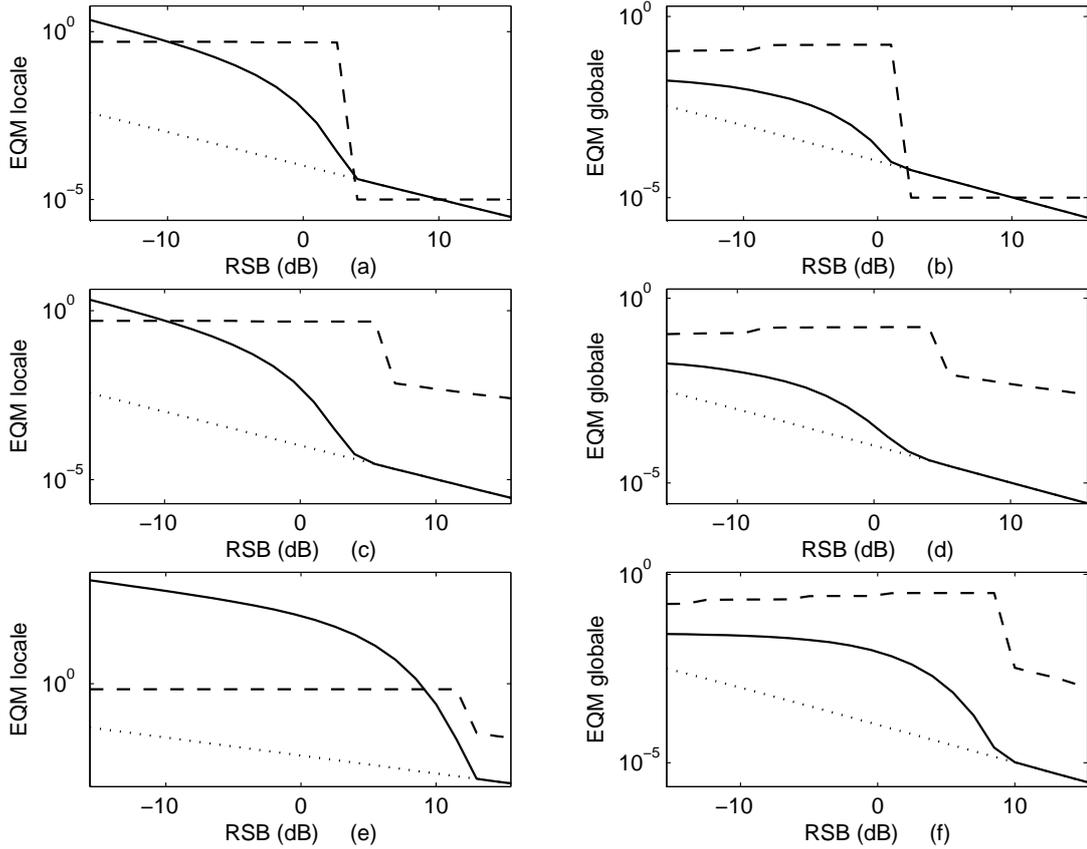


FIGURE 4.5 – Comportement de h_{opt} pour les bornes de Chapman-Robbins (a), Bobrovsky-Zakai (b), Abel déterministe (c) et bayésien (d), Weiss-Weinstein déterministe (e) et bayésien (f) en fonction du RSB. Modulation QPSK et $N = 20$ observations.

où la dérivée est calculée rapidement de manière numérique. Les résultats obtenus à l'aide de \widehat{RSB} sont reportés figure (4.6) et (4.7) et sont comparés à l'EQM locale et globale de l'estimateur du maximum de vraisemblance.

On observe encore une fois la bonne précision des bornes de Weiss-Weinstein concernant l'estimation du seuil de décrochement. On remarque également que les bornes d'Abel permettent de gagner en pertinence par rapport aux bornes de Chapman-Robbins et de Bobrovsky-Zakai.

4.4.3.2 Réduction de la charge de calcul et approximations des bornes minimales

Lorsque l'on utilise les bornes minimales dans le but de prédire la zone décrochement, l'un des objectifs est d'éviter une charge de calcul élevée due à des simulations de Monte-Carlo. Or les bornes les plus pertinentes sont aussi celles qui requièrent une charge de calcul plus importante. Nous présentons maintenant quelques résultats dans le but de réduire cette charge à celle de la BCR.

Cette section concerne uniquement les bornes déterministes et ces résultats sont également basés sur le comportement du point test optimal. Dans la zone asymptotique, ce point test est proche de 0, et, lorsque le phénomène de décrochement apparaît, il passe brusquement à une valeur proche de $\frac{1}{2}$. En conséquence, on peut obtenir une approximation des bornes utilisant un point test en optimisant celles-ci sur deux valeurs $\{0, \frac{1}{2}\}$. Le cas des bornes bayésiennes est plus

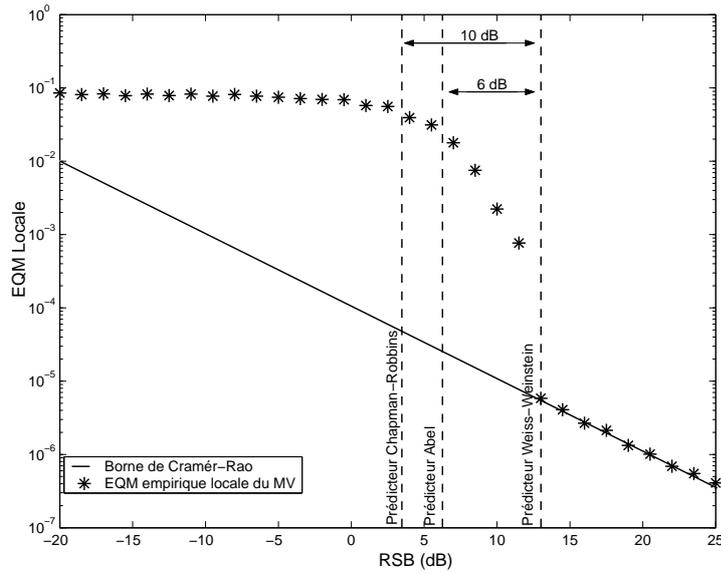


FIGURE 4.6 – EQM empirique locale de l'estimateur du maximum de vraisemblance en fonction du RSB et prédicteurs du décrochement obtenus à partir des bornes de Chapman-Robbins, Abel et Weiss-Weinstein déterministe. Modulation QPSK et $N = 20$ observations.

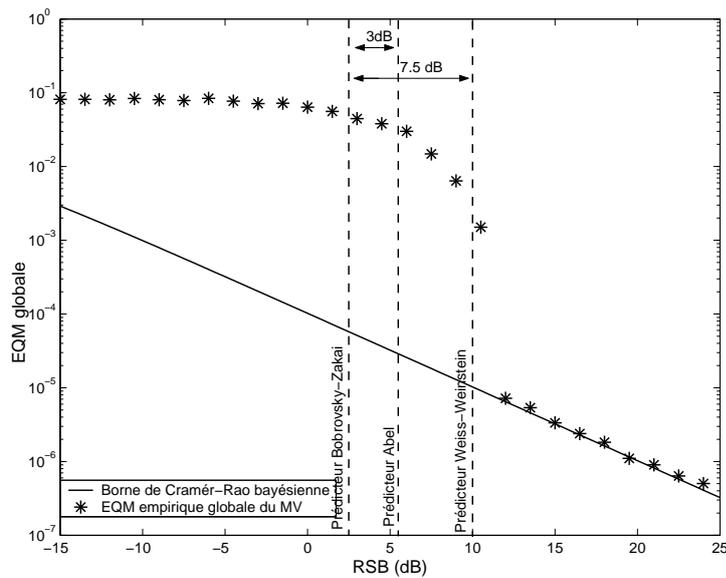


FIGURE 4.7 – EQM empirique globale de l'estimateur du maximum de vraisemblance en fonction du RSB et prédicteurs du décrochement obtenus à partir des bornes de Bobrovsky-Zakai, Abel bayésien et Weiss-Weinstein. Modulation QPSK et $N = 20$ observations.

délicat puisque le point test ne passe pas brusquement à la valeur $\frac{1}{2}$ mais reste plutôt autour de $\frac{1}{10}$ (voir figure (4.5)). Les figures (4.8), (4.9) et (4.10) comparent les bornes de Chapman-Robbins, Abel, et Weiss-Weinstein déterministe lorsque l'optimisation est faite sur le continuum $[0, \frac{1}{2}]$ ou sur les deux valeurs discrètes $\{0, \frac{1}{2}\}$.

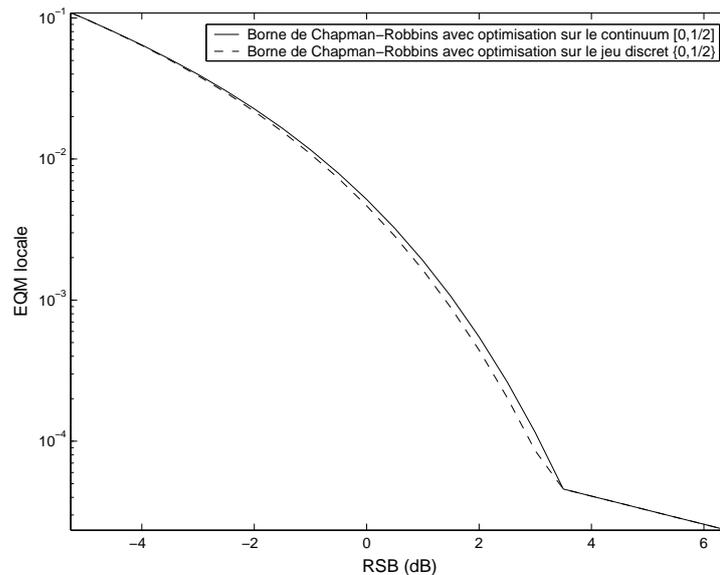


FIGURE 4.8 – Comparaison entre la borne de Chapman-Robbins obtenue à partir d’une optimisation sur le continuum $[0, \frac{1}{2}]$ et la borne de Chapman-Robbins obtenue à partir d’une optimisation sur 2 valeurs $\{0, \frac{1}{2}\}$ en fonction du RSB. Modulation QPSK et $N = 20$ observations.

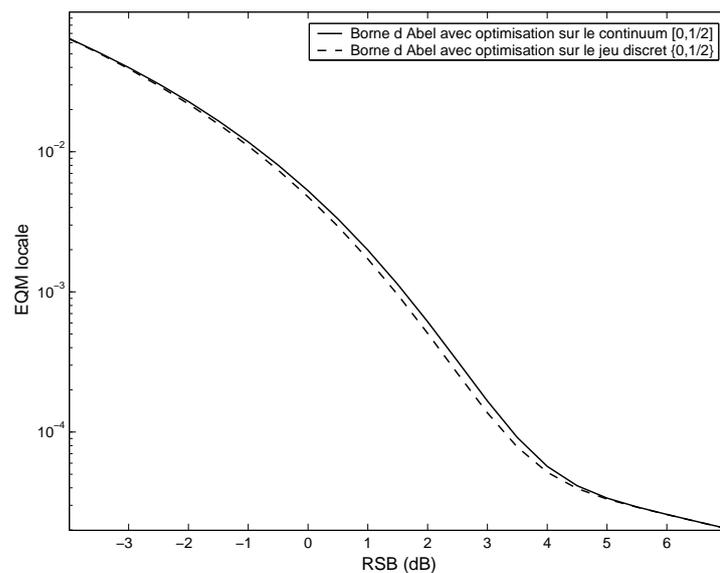


FIGURE 4.9 – Comparaison entre la borne d’Abel obtenue à partir d’une optimisation sur le continuum $[0, \frac{1}{2}]$ et la borne d’Abel obtenue à partir d’une optimisation sur 2 valeurs $\{0, \frac{1}{2}\}$ en fonction du RSB. Modulation QPSK et $N = 20$ observations.

Même si les bornes approximées restent en dessous des bornes optimales, la différence est suffisamment faible pour que leurs utilisations restent pertinente.

Concernant la borne d’Abel, on peut encore réduire la charge de calcul en remarquant que cette borne incorpore naturellement la borne de Cramér-Rao. En conséquence, la zone où $h \approx 0$

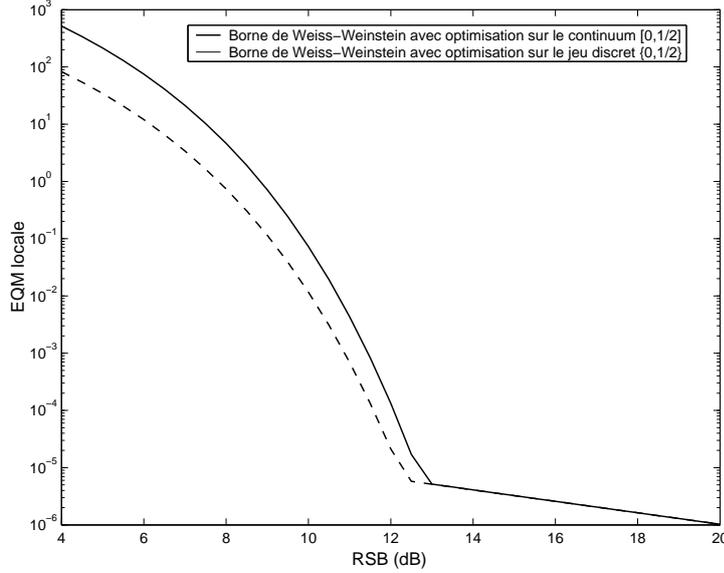


FIGURE 4.10 – Comparaison entre la borne de Weiss-Weinstein déterministe obtenue à partir d’une optimisation sur le continuum $[0, \frac{1}{2}]$ et la borne de Weiss-Weinstein déterministe obtenue à partir d’une optimisation sur 2 valeurs $\{0, \frac{1}{2}\}$ en fonction du RSB. Modulation QPSK et $N = 20$ observations.

est déjà prise en compte dans le calcul de la borne. On peut donc “forcer” le point test à être toujours égal à $\frac{1}{2}$. Dans ce cas, la fonction $\psi(\theta_0, h)$ intervenant dans la borne d’Abel devient

$$\psi\left(\theta_0, \frac{1}{2}\right) = \frac{8\pi}{\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(\pi k) = 0, \quad (4.104)$$

Ainsi, on peut réécrire la borne d’Abel comme

$$\begin{aligned} EQM_{Abel(1,1)}\left(\theta_0, \frac{1}{2}\right) &= EQM_{CR}(\theta_0) + EQM_{ChR}\left(\theta_0, \frac{1}{2}\right) \\ &= \frac{\sigma^2}{8\pi^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2} + \frac{1}{4 \left(e^{\frac{8}{\sigma^2} \sum_{k=1, k \text{ impaire}}^{N-1} \|a_k\|^2} - 1 \right)}, \end{aligned} \quad (4.105)$$

La figure (4.11) compare la borne d’Abel à la borne (4.105). On a également reporté la borne d’Abel avec optimisation sur les deux valeurs discrètes $\{0, \frac{1}{2}\}$. Elle coïncide parfaitement avec la borne obtenue avec (4.105).

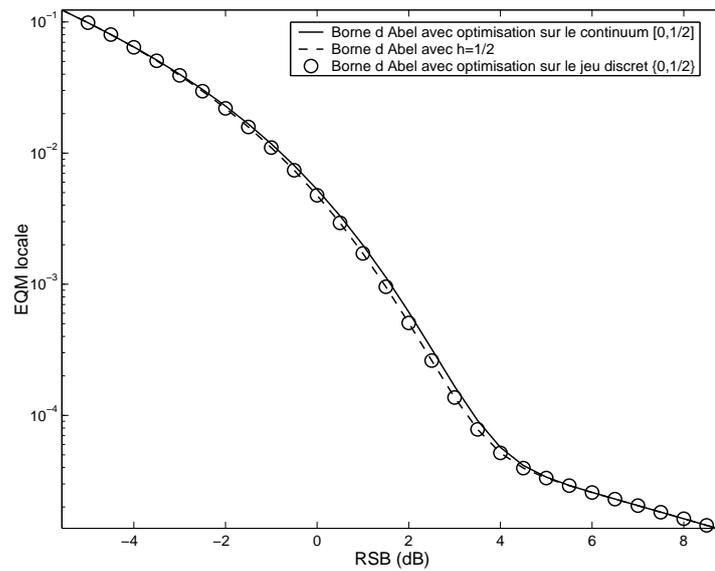


FIGURE 4.11 – Comparaison entre la borne d’Abel obtenue à partir d’une optimisation sur le continuum $[0, \frac{1}{2}]$, la borne d’Abel obtenue à partir d’une optimisation sur 2 valeurs $\{0, \frac{1}{2}\}$ et la borne d’Abel obtenue en forçant $h = \frac{1}{2}$ en fonction du RSB. Modulation QPSK et $N = 20$ observations.

Chapitre 5

Conclusion

Le travail rapporté dans ce document a concerné l'étude des performances d'estimation en terme d'Erreur Quadratique Moyenne (EQM) dans le cadre du traitement du signal.

- Nous nous sommes placés dans un premier temps en condition asymptotique. Alors que beaucoup de résultats sont disponibles lorsque le nombre d'observations est infini, nous nous sommes intéressés au cas d'un nombre d'observations fini et d'un Rapport Signal sur Bruit (RSB) infini. Dans ce contexte, nous avons caractérisé les performances des méthodes du maximum de vraisemblance pour l'estimation d'un nombre quelconque de directions d'arrivées. Dans le cas de sources déterministes, nous avons démontré la gaussianité et l'efficacité asymptotique (en RSB) de l'estimateur du maximum de vraisemblance déterministe. Dans le cas de sources gaussiennes, nous avons démontré la non-gaussianité et la non-efficacité asymptotiques (en RSB) de l'estimateur du maximum de vraisemblance stochastique. Pour ce dernier cas, nous avons, de plus, établi une forme analytique de la variance des estimées dans le contexte d'un scénario à deux sources décorréélées situées dans le lobe principal de l'antenne.
- Dans le but d'étudier les performances ultimes des estimateurs en condition asymptotique et non-asymptotique, nous avons étudié les bornes minimales d'estimation plus pertinentes que la borne de Cramér-Rao. Nous avons unifié les bornes bayésiennes de la famille Weiss-Weinstein, c'est-à-dire la borne de Cramér-Rao bayésienne, la borne de Bobrovsky-MayerWolf-Zakai, la borne de Bhattacharyya bayésienne, la borne de Bobrovsky-Zakai, la borne de Reuven-Messer, et la borne de Weiss-Weinstein. Cette unification a été réalisée à l'aide d'un formalisme général permettant de se ramener à un problème d'optimisation sous contraintes. L'intérêt de cette approche a été double. Premièrement, elle a établi un pont entre les bornes susmentionnées et la "meilleure" borne bayésienne : l'EQM de l'estimateur de la moyenne conditionnelle. Ceci nous a permis d'établir une relation d'ordre entre les différentes bornes souvent inaccessible dans les développements traditionnels exploitant une inégalité de covariance. Deuxièmement, cette approche se veut être duale de celle utilisée dans [Gla72] [FL01] [FL02] [Cha04] dans le cadre de l'unification des bornes déterministes. De ce fait, nous avons également établi des connexions entre les bornes déterministes et les bornes bayésiennes de la famille Weiss-Weinstein. Ce cadre offre une méthodologie propice au développement de nouvelles bornes minimales, nous permettant ainsi de proposer deux nouvelles bornes d'estimation : la borne d'Abel bayésienne et la borne de Weiss-Weinstein déterministe.
- Afin de tirer partie des résultats théoriques obtenus, nous avons proposé une forme simplifiée des bornes d'Abel (déterministe et bayésienne). Cette forme dépend explicitement de la borne de Cramér-Rao (bayésienne) et de la borne de Chapman-Robbins (Bobrovsky-Zakai).

Elle est plus pertinente que ces deux dernières. Pour un modèle d'observations gaussiennes à moyenne paramétrée et pour un problème d'estimation de fréquence porteuse avec symboles pilotes, nous avons proposé différentes formes analytiques pour l'ensemble des bornes déterministes et bayésiennes de la famille Weiss-Weinstein. Les bornes de Weiss-Weinstein (déterministe et bayésienne) se sont révélées être un outil intéressant pour prédire le décrochement des estimateurs. A partir de la recherche d'un point test optimal intervenant dans les différentes bornes minimales, nous avons proposé un prédicteur de la zone de décrochement des estimateurs. Nous avons également proposé différentes approximations des bornes minimales ayant une charge de calcul similaire à la borne de Cramér-Rao.

A l'issue de ce travail, les perspectives sont diverses.

- Concernant les performances asymptotiques en RSB de l'estimateur du maximum de vraisemblance stochastique, il convient de poursuivre l'étude afin d'obtenir une expression de la variance asymptotique et de la densité de probabilité des estimées dans le cas multi-sources.
- Concernant les bornes minimales d'estimation, les travaux que nous avons déjà effectués laissent entrevoir plusieurs développements théoriques possibles.
 - Il faut continuer l'étude de la borne de Weiss-Weinstein déterministe afin d'interpréter les contraintes mises en jeu dans cette borne prometteuse.
 - Nous avons relié les bornes de la famille Weiss-Weinstein à l'EQM de l'estimateur conditionnelle. Il s'agit maintenant d'interpréter les contraintes mises en jeu dans cette famille afin de définir les classes d'estimateurs pour lesquels les bornes proposées sont licites.
 - Enfin, un challenge intéressant serait de relier les bornes de la famille Ziv-Zakaï aux bornes de la famille Weiss-Weinstein ce qui permettrait de finaliser totalement l'unification des bornes minimales d'estimation.
- Concernant les applications des bornes minimales d'estimation en traitement du signal, nous pouvons envisager un certain nombre de compléments et d'améliorations.
 - Il convient d'appliquer désormais les bornes minimales à d'autres problèmes d'estimation, en particulier dans un contexte multi-paramètres. Par exemple, en traitement d'antenne, un tel travail permettrait de disposer d'une méthodologie pour optimiser une géométrie d'antenne.
 - Afin de diminuer la charge de calcul mise en jeu, il faudrait disposer de méthode systématique de positionnement de multiples points tests. Par exemple, utiliser le diagramme d'ambiguïté $\|\mathbf{A}^H(\boldsymbol{\theta}_0)\mathbf{A}(\boldsymbol{\theta})\|^2$ en traitement d'antenne.

Annexe A

Bornes minimales pour un modèle d'observations gaussiennes à moyenne paramétrée

Bornes de Cramér-Rao

$$EQM_{CR}(\theta_0) = \frac{\sigma^2}{2 \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2}. \quad (\text{A.1})$$

$$EQM_{CR}^{globale} = \left(\frac{2}{\sigma^2} \int_{\Theta} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \right\|^2 p(\theta) d\theta + \frac{1}{\sigma_{\theta}^2} \right)^{-1}. \quad (\text{A.2})$$

Borne de Bhattacharyya à l'ordre 2

$$EQM_{Bhatt(2)}(\theta_0) = \frac{\frac{2}{\sigma^2} \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \Big|_{\theta_0} \right\|^2 + \frac{4}{\sigma^4} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^4}{\frac{4}{\sigma^4} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^2 \left\| \frac{\partial^2 \mathbf{m}(\theta)}{\partial \theta^2} \Big|_{\theta_0} \right\|^2 + \frac{2}{\sigma^6} \left\| \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\|^6 - \frac{4}{\sigma^4} \left(\text{Re} \left\{ \frac{\partial^2 \mathbf{m}^H(\theta)}{\partial \theta^2} \Big|_{\theta_0} \frac{\partial \mathbf{m}(\theta)}{\partial \theta} \Big|_{\theta_0} \right\} \right)^2}. \quad (\text{A.3})$$

Bornes de Chapman-Robbins et de Bobrovsky-Zakaï

$$EQM_{ChR}(\theta_0) = \max_h \frac{h^2}{e^{\frac{2}{\sigma^2} \|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2} - 1}. \quad (\text{A.4})$$

$$EQM_{BZ}^{globale} = \max_h \frac{h^2}{\int_{\Theta} \frac{p^2(\theta+h)}{p(\theta)} e^{\frac{2}{\sigma^2} \|\mathbf{m}(\theta+h) - \mathbf{m}(\theta)\|^2} d\theta - 1}. \quad (\text{A.5})$$

Fonctions $\psi(\theta_0, h)$ et $\psi(h)$ pour les bornes d'Abel

$$EQM_{Abel(1,1)}(\theta_0) = \max_h \frac{EQM_{CR}(\theta_0) + EQM_{ChR}(\theta_0, h) - 2\psi(\theta_0, h) EQM_{CR}(\theta_0) EQM_{ChR}(\theta_0, h)}{1 - \psi^2(\theta_0, h) EQM_{CR}(\theta_0) EQM_{ChR}(\theta_0, h)}. \quad (A.6)$$

$$\psi(\theta_0, h) = \frac{2}{h\sigma^2} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} (\mathbf{m}(\theta_0 + h) - \mathbf{m}(\theta_0)) \right\}. \quad (A.7)$$

$$EQM_{Abel(1,1)}^{globale} = \max_h \frac{EQM_{CR}^{globale} + EQM_{BZ}^{globale}(h) - 2\psi(h) EQM_{CR}^{globale} EQM_{BZ}^{globale}(h)}{1 - \psi^2(h) EQM_{CR}^{globale} EQM_{BZ}^{globale}(h)}. \quad (A.8)$$

$$\psi(h) = \frac{1}{\sigma_\theta^2} + \frac{2}{h\sigma^2} \int_{\Theta} \operatorname{Re} \left\{ \frac{\partial \mathbf{m}^H(\theta)}{\partial \theta} \Big|_{\theta_0} (\mathbf{m}(\theta_0 + h) - \mathbf{m}(\theta_0)) \right\} p(\theta + h) d\theta. \quad (A.9)$$

Bornes de Weiss-Weinstein et fonctions $\eta(\alpha|\beta)$ et $\eta(\alpha, \beta)$

$$EQM_{WWD(1)}(h, s, \theta_0) = \max_{h,s} \frac{h^2 e^{2\frac{s(s-1)}{\sigma^2} \|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2}}{e^{\frac{2s(2s-1)}{\sigma^2} \|\mathbf{m}(\theta_0+h) - \mathbf{m}(\theta_0)\|^2} + e^{\frac{2(s-1)(2s-1)}{\sigma^2} \|\mathbf{m}(\theta_0-h) - \mathbf{m}(\theta_0)\|^2} - 2e^{\frac{s(s-1)}{\sigma^2} \|\mathbf{m}(\theta_0+2h) - \mathbf{m}(\theta_0)\|^2}}. \quad (A.10)$$

$$\eta(\alpha|\beta) = \frac{\alpha(\alpha-1)}{\sigma^2} \|\mathbf{m}(\theta_0 + \beta) - \mathbf{m}(\theta_0)\|^2. \quad (A.11)$$

$$EQM_{WW(1)}(h, s) = \max_{h,s} \frac{h^2 e^{2\eta(s,h)}}{e^{\eta(2s,h)} + e^{\eta(2-2s,-h)} - 2e^{\eta(s,2h)}} \quad (A.12)$$

$$\eta(\alpha, \beta) = \ln \frac{1}{\sqrt{2\pi}\sigma_\theta} \int_{\Theta} e^{\frac{\alpha(\alpha-1)}{\sigma^2} \|\mathbf{m}(\theta+\beta) - \mathbf{m}(\theta)\|^2 - \frac{1}{2\sigma_\theta^2} (\theta - (\sqrt{\alpha(\alpha-1)} - \alpha)h - \mu) (\theta + (\sqrt{\alpha(\alpha-1)} + \alpha)h - \mu)} d\theta. \quad (A.13)$$

Annexe B

Bornes minimales en analyse spectrale et synchronisation de fréquence porteuse

Bornes de Cramér-Rao

$$EQM_{CR}(\theta_0) = \frac{\sigma^2}{8\pi^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2}. \quad (\text{B.1})$$

$$EQM_{CR}^{globale} = \frac{\sigma_\theta^2}{8\pi^2 \frac{\sigma_\theta^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + 1}. \quad (\text{B.2})$$

Bornes de Bhattacharyya

$$EQM_{Bhatt(2)}(\theta_0) = EQM_{CR}(\theta_0) = \frac{\sigma^2}{8\pi^2 \sum_{k=0}^{N-1} \|a_k\|^2 k^2}. \quad (\text{B.3})$$

$$EQM_{Bhatt(2)}^{globale} = EQM_{CR}^{globale} = \frac{\sigma_\theta^2}{8\pi^2 \frac{\sigma_\theta^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + 1}. \quad (\text{B.4})$$

Bornes de Chapman-Robbins et de Bobrovsky-Zakai

$$EQM_{ChR}(\theta_0) = \max_h \frac{h^2}{e^{\frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1}. \quad (\text{B.5})$$

$$EQM_{BZ}^{globale} = \max_h \frac{h^2}{e^{\frac{h^2}{\sigma_\theta^2} + \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))} - 1}. \quad (\text{B.6})$$

Bornes d'Abel

$$\psi(\theta_0, h) = \frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh). \quad (\text{B.7})$$

$$EQM_{Abel(1,1)}(\theta_0) = \max_h \frac{\frac{8\pi^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + \frac{1}{h^2} \left(e^{\frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1-\cos(2\pi kh))} - 1 \right) - \frac{8\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh)}{\frac{8\pi^2}{h^2\sigma^2} \left(e^{\frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1-\cos(2\pi kh))} - 1 \right) \sum_{k=0}^{N-1} \|a_k\|^2 k^2 - \frac{16\pi^2}{h^2\sigma^4} \left(\sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh) \right)^2}. \quad (\text{B.8})$$

$$\psi(h) = \frac{1}{\sigma_\theta^2} + \frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh). \quad (\text{B.9})$$

$$EQM_{Abel(1,1)}^{globale} = \max_h \quad (\text{B.10})$$

$$\frac{\frac{8\pi^2 \frac{\sigma_\theta^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + 1}{\sigma_\theta^2} + \frac{e^{\frac{h^2}{\sigma_\theta^2} + \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1-\cos(2\pi kh))}}{h^2} - 1 - 2 \left(\frac{1}{\sigma_\theta^2} + \frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh) \right)}{\frac{1}{\sigma_\theta^2 h^2} \left(8\pi^2 \frac{\sigma_\theta^2}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 k^2 + 1 \right) \left(e^{\frac{h^2}{\sigma_\theta^2} + \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1-\cos(2\pi kh))} - 1 \right) - \left(\frac{1}{\sigma_\theta^2} + \frac{4\pi}{h\sigma^2} \sum_{k=0}^{N-1} k \|a_k\|^2 \sin(2\pi kh) \right)^2}.$$

Bornes de Weiss-Weinstein

On pose $\xi(h) = \frac{4}{\sigma^2} \sum_{k=0}^{N-1} \|a_k\|^2 (1 - \cos(2\pi kh))$.

$$EQM_{WWD(1)}(\theta_0) = \max_{h,s} \frac{h^2 e^{s(s-1)\xi(h)}}{e^{s(2s-1)\xi(h)} + e^{(1-s)(1-2s)\xi(h)} - 2e^{\frac{1}{2}s(s-1)\xi(2h)}}. \quad (\text{B.11})$$

$$EQM_{WW(1)}^{globale} = \max_{h,s} \frac{h^2 e^{s(s-1)\left(\xi(h) + \frac{h^2}{\sigma_\theta^2}\right)}}{e^{s(2s-1)\left(\xi(h) + \frac{h^2}{\sigma_\theta^2}\right)} + e^{(1-s)(1-2s)\left(\xi(h) + \frac{h^2}{\sigma_\theta^2}\right)} - 2e^{s(s-1)\left(\frac{1}{2}\xi(2h) + \frac{2h^2}{\sigma_\theta^2}\right)}}. \quad (\text{B.12})$$

Notations

Acronymes

- **ALU** : antenne linéaire uniforme
- **BCR** : borne de Cramér-Rao.
- **EQM** : erreur quadratique moyenne.
- **MIF** : matrice d'information de Fisher.
- **MVD** : maximum de vraisemblance déterministe.
- **MVP** : maximum de vraisemblance paramétré.
- **MVS** : maximum de vraisemblance stochastique.
- **QPSK** : quaternary phase shift keying
- **RSB** : rapport signal sur bruit.

Symboles mathématiques généraux

- \mathbb{C} indique le corps des complexes.
- \mathbb{R} indique le corps des réels.
- $\text{Re}\{z\}$ indique la partie réelle d'un nombre complexe z .
- $\text{Im}\{z\}$ indique la partie imaginaire d'un nombre complexe z .
- Le symbole $*$ indique l'opérateur de conjugaison.
- $\text{abs}(\cdot)$ indique la valeur absolue.
- $o(\cdot)$ est la notation de Landau signifiant que pour deux fonctions $f(x)$ et $g(x)$, $f = o(g)$ est équivalent à $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0$.
- $\delta(\cdot)$ indique la distribution de Dirac.
- δ_{ij} indique le symbole de Kronecker.
- $\min_{\boldsymbol{\eta}} f(\boldsymbol{\eta})$, $\max_{\boldsymbol{\eta}} f(\boldsymbol{\eta})$, indique la valeur minimale ou maximale de l'expression $f(\boldsymbol{\eta})$.
- $\arg \min_{\boldsymbol{\eta}} f(\boldsymbol{\eta})$, $\arg \max_{\boldsymbol{\eta}} f(\boldsymbol{\eta})$, indique la valeur du vecteur $\boldsymbol{\eta}$ minimisant ou maximisant l'expression $f(\boldsymbol{\eta})$.
- $\min\{a_1, \dots, a_N\}$ indique le plus petit élément du jeu $\{a_1, \dots, a_N\}$.
- $\Gamma(z)$ et $\Gamma(z, a)$ indiquent, respectivement, la fonction Gamma et la fonction Gamma incomplète telle que $\Gamma(z) = \Gamma(z, 0)$ avec

$$\Gamma(z, a) = \int_a^{\infty} t^{z-1} e^{-t} dt.$$

- ${}_2F_1(a, b; c; \omega)$ indique la fonction hypergéométrique de Gauss définie par :

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

Symboles et opérateurs matriciels

- a, A , les lettres en italiques représentent une quantité scalaire.
- \mathbf{a} , les lettres minuscules en gras représentent une quantité vectorielle (vecteur colonne).
- \mathbf{A} , les lettres majuscules en gras représentent une quantité matricielle.
- L'ensemble des matrices de dimension $M \times N$ à coefficients sur un corps \mathbb{K} est noté $\mathcal{M}_{M \times N}(\mathbb{K})$.
- \mathbf{A}^T , le symbole T indique l'opérateur transposé.
- \mathbf{A}^H , le symbole H indique l'opérateur Hermitien (transposé conjugué).
- \mathbf{A}^\dagger , le symbole \dagger indique l'opérateur pseudo-inverse tel que $\mathbf{A}^\dagger = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ où \mathbf{A} est de rang plein colonne.
- \mathbf{I}_N indique la matrice identité de taille $N \times N$.
- $|\mathbf{A}|$ indique le déterminant de la matrice \mathbf{A} .
- $\|\cdot\|$ indique la norme L^2 .
- $A_{i,j}$ est l'élément correspondant à la $i^{\text{ième}}$ ligne et $j^{\text{ième}}$ colonne de la matrice \mathbf{A} .
- $tr(\mathbf{A}) = \sum_{i=1}^N A_{i,i}$ est la trace de la matrice carrée \mathbf{A} de taille $N \times N$.
- \odot est l'opérateur de Hadamard.
- $vec(\cdot)$ est l'opérateur qui vectorise une matrice en concaténant ses colonnes.
- Soient deux matrices \mathbf{A} et \mathbf{B} appartenant à $\mathcal{M}_{N \times N}(\mathbb{R})$, alors la relation d'ordre

$$\mathbf{A} \geq \mathbf{B},$$

signifie que la matrice $\mathbf{A} - \mathbf{B}$ est définie non-négative.

- $Diag\{a_1, a_2, \dots, a_N\} = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_N \end{pmatrix}.$

Symboles relatifs aux probabilités

- $\mathcal{N}_{\mathbb{K}}(\mathbf{m}, \mathbf{C})$ indique une loi gaussienne multivariée sur le corps \mathbb{K} (circulaire si $\mathbb{K} = \mathbb{C}$) de moyenne \mathbf{m} et de matrice de covariance \mathbf{C} .
- $\chi^2(N)$ indique une loi du chi2 réelle à N degrés de libertés basée sur une somme au carré de N variables aléatoires indépendantes suivant une loi gaussienne de moyenne nulle et de variance 1.
- $\Pr(\cdot)$ indique une probabilité.

- $p(x)$ indique une densité de probabilité.
- $p(x, y)$ indique une densité de probabilité jointe.
- $p(x|y)$ indique une densité de probabilité conditionnelle ou une vraisemblance et sera précisée selon le contexte.
- $\mathbb{E}[\cdot]$ indique l'opérateur d'espérance mathématique.
- $\mathbb{E}_{\mathbf{y}, \theta}[\cdot]$ indique l'opérateur d'espérance mathématique par rapport à la loi jointe $p(\mathbf{y}, \theta)$.
- $\mathbb{E}_{\mathbf{y}|\theta}[\cdot]$ indique l'opérateur d'espérance mathématique par rapport à la loi conditionnelle $p(\mathbf{y}|\theta)$.
- $o_p(\cdot)$ est la notation statistique de Landau signifiant que pour deux suites de variables aléatoires p_i et q_i , $p_i = o_p(q_i)$ est équivalent à $\frac{p_i}{q_i}$ tend en probabilité vers zéro.
- \sim indique *suivre en loi*.
- $\overset{a \rightarrow \infty}{\sim}$ indique *suivre asymptotiquement en loi* (lorsque le paramètre a tend vers l'infini).

Table des figures

1.1	Principe de l'estimation paramétrique.	1
1.2	Réalisations du critère du maximum de vraisemblance.	2
1.3	Comportement de l'EQM d'un estimateur du maximum de vraisemblance.	3
2.1	Principe de la localisation.	9
2.2	Principe de l'estimation au sens du maximum de vraisemblance.	12
2.3	MVD. Variance empirique de l'estimateur, BCR déterministe en fonction du RSB. $\theta_0 = [0^\circ \ 3^\circ]$. $T = 10$ observations. $N = 10$ capteurs. $r = 1000$ réalisations.	22
2.4	MVD. Histogramme des estimées. $\theta_0 = [0^\circ \ 3^\circ]$. $RSB = 30dB$. $T = 10$ observations. $N = 10$ capteurs. $r = 10000$ réalisations.	23
2.5	MVD. Test statistique de Lilliefors : T_{stat} et quantile $1 - \alpha$ en fonction du RSB. $\theta_0 = [0^\circ \ 3^\circ]$, $T = 10$ observations, $N = 10$ capteurs, $r = 1000$ réalisations.	24
2.6	MVS. Scénario une source. $\theta_0 = 0^\circ$. Variance empirique de l'estimateur, BCR stochastique, et variance asymptotique théorique en fonction du RSB. $T = 2$ observations $N = 4$ capteurs. $r = 1000$ réalisations.	33
2.7	MVS. Scénario une source. $\theta_0 = 0^\circ$. Histogramme des estimées, loi de Student théorique et loi gaussienne. $RSB = 30dB$. $T = 2$ observations $N = 4$ capteurs. $r = 10000$ réalisations.	34
2.8	MVS. Scénario deux sources. $\theta_0 = [0^\circ \ 5^\circ]$. Variance empirique de l'estimateur, BCR stochastique, et variance asymptotique théorique en fonction du RSB. $T = 2$ observations $N = 10$ capteurs. $r = 1000$ réalisations.	35
2.9	MVS. Scénario deux sources. $\theta_0 = [0^\circ \ 5^\circ]$. Histogrammes des estimées et loi gaussienne. $RSB = 30dB$. $T = 2$ observations $N = 10$ capteurs. $r = 10000$ réalisations.	35
3.1	Fonction Valley Filling.	48
4.1	Borne de Weiss-Weinstein déterministe avec une optimisation sur h et s , borne de Weiss-Weinstein avec optimisation sur h et $s = \frac{1}{2}$, $s = \frac{1}{4}$, et $s = 0$, borne de Weiss-Weinstein avec $h = 0$ et $s = 0$, la borne de Chapman-Robbins et borne de Cramér-Rao en fonction du RSB. Modulation QPSK et $N = 20$ observations.	82
4.2	Borne de Weiss-Weinstein avec une optimisation sur h et s , borne de Weiss-Weinstein avec optimisation sur h et $s = \frac{1}{2}$, $s = \frac{1}{4}$, et $s = 0$, borne de Weiss-Weinstein avec $h = 0$ et $s = 0$, la borne de Bobrovsky-Zakai et borne de Cramér-Rao bayésienne en fonction du RSB. Modulation QPSK et $N = 20$ observations.	82

4.3	EQM locale empirique de l'estimateur du maximum de vraisemblance, borne de Cramér-Rao, borne de Chapman-Robbins, borne d'Abel et borne de Weiss-Weinstein déterministe en fonction du RSB. Modulation QPSK et $N = 20$ observations.	83
4.4	EQM globale empirique de l'estimateur du maximum de vraisemblance, borne de Cramér-Rao bayésienne, borne de Bobrovsky-Zakaï, borne d'Abel bayésienne et borne de Weiss-Weinstein en fonction du RSB. Modulation QPSK et $N = 20$ observations.	84
4.5	Comportement de h_{opt} pour les bornes de Chapman-Robbins (a), Bobrovsky-Zakaï (b), Abel déterministe (c) et bayésien (d), Weiss-Weinstein déterministe (e) et bayésien (f) en fonction du RSB. Modulation QPSK et $N = 20$ observations.	85
4.6	EQM empirique locale de l'estimateur du maximum de vraisemblance en fonction du RSB et prédicteurs du décrochement obtenus à partir des bornes de Chapman-Robbins, Abel et Weiss-Weinstein déterministe. Modulation QPSK et $N = 20$ observations.	86
4.7	EQM empirique globale de l'estimateur du maximum de vraisemblance en fonction du RSB et prédicteurs du décrochement obtenus à partir des bornes de Bobrovsky-Zakaï, Abel bayésien et Weiss-Weinstein. Modulation QPSK et $N = 20$ observations.	86
4.8	Comparaison entre la borne de Chapman-Robbins obtenue à partir d'une optimisation sur le continuum $[0, \frac{1}{2}]$ et la borne de Chapman-Robbins obtenue à partir d'une optimisation sur 2 valeurs $\{0, \frac{1}{2}\}$ en fonction du RSB. Modulation QPSK et $N = 20$ observations.	87
4.9	Comparaison entre la borne d'Abel obtenue à partir d'une optimisation sur le continuum $[0, \frac{1}{2}]$ et la borne d'Abel obtenue à partir d'une optimisation sur 2 valeurs $\{0, \frac{1}{2}\}$ en fonction du RSB. Modulation QPSK et $N = 20$ observations.	87
4.10	Comparaison entre la borne de Weiss-Weinstein déterministe obtenue à partir d'une optimisation sur le continuum $[0, \frac{1}{2}]$ et la borne de Weiss-Weinstein déterministe obtenue à partir d'une optimisation sur 2 valeurs $\{0, \frac{1}{2}\}$ en fonction du RSB. Modulation QPSK et $N = 20$ observations.	88
4.11	Comparaison entre la borne d'Abel obtenue à partir d'une optimisation sur le continuum $[0, \frac{1}{2}]$, la borne d'Abel obtenue à partir d'une optimisation sur 2 valeurs $\{0, \frac{1}{2}\}$ et la borne d'Abel obtenue en forçant $h = \frac{1}{2}$ en fonction du RSB. Modulation QPSK et $N = 20$ observations.	89

Liste des tableaux

- 2.1 MVS. Scénario une source. $\theta_0 = 0^\circ$. Test statistique du chi deux. $RSB = 30\text{dB}$.
 $T = 2$ observations $N = 4$ capteurs. $r = 10000$ réalisations. $k = 15$ classes. . . . 34
- 2.2 MVS. Scénario deux sources. $\theta_0 = [0^\circ \ 5^\circ]$. Test statistique du chi deux. $RSB = 30\text{dB}$. $T = 2$ observations $N = 10$ capteurs. $r = 10000$ réalisations. $k = 15$ classes. 36

Bibliographie

- [Abe90] J. S. Abel. A bound on mean square estimate error. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, volume 3, pages 1345–1348, Albuquerque, New Mexico, 1990.
- [Abe93] J. S. Abel. A bound on mean square estimate error. *IEEE Trans. Inform. Theory*, 39 :1675–1680, September 1993.
- [ABL04a] L. Atallah, J. P. Barbot, and P. Larzabal. From Chapman Robbins bound towards Barankin bound in threshold behaviour prediction. *Electronic Letters*, 40 :279–280, February 2004.
- [ABL04b] L. Atallah, J. P. Barbot, and P. Larzabal. SNR threshold indicator in data aided frequency synchronization. *IEEE Signal Processing Lett.*, 11 :652–654, August 2004.
- [And84] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley, New York, 1984.
- [AS72] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Dover Publications, inc., New York, 1972.
- [Ath05] F. Athley. Threshold region performance of Maximum Likelihood direction of arrival estimators. *IEEE Trans. Signal Processing*, 53 :1359–1373, April 2005.
- [Bag69] A. B. Baggeroer. Barankin bound on the variance of estimates of Gaussian random process. *Technical Report, MIT Lincoln Laboratory, Lexington, Massachusetts*, January 1969.
- [Bar49] E. W. Barankin. Locally best unbiased estimates. *Ann. Math. Stat.*, 20 :477–501, 1949.
- [Bel95] K. Bell. *Performance bounds in parameter estimation with application to bearing estimation*. PhD thesis, George Mason University, Fairfax, VA, 1995.
- [BET95] K. Bell, Y. Ephraim, and H. L. Van Trees. Ziv Zakai lower bounds in bearing estimation. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, volume 5, pages 2852–2855, Detroit, Michigan, 1995.
- [BET96] K. Bell, Y. Ephraim, and H. L. Van Trees. Explicit Ziv Zakai lower bound for bearing estimation. *IEEE Trans. Signal Processing*, 44 :2810–2824, November 1996.
- [BFL04a] E. Boyer, P. Forster, and P. Larzabal. Non asymptotic performances analysis of beamforming with stochastic signals. *IEEE Signal Processing Lett.*, 11(1) :23–25, January 2004.
- [BFL04b] E. Boyer, P. Forster, and P. Larzabal. Non asymptotic statistical performances of beamforming for deterministic signals. *IEEE Signal Processing Lett.*, 11(1) :20–22, January 2004.

- [Bha43] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.*, 35 :99–109, 1943.
- [Bha46] A. Bhattacharyya. On some analogues of the amount of information and their use in statistical estimation. *Sankhya Indian J. of Stat.*, 8 :1–14, 201–218, 315–328, 1946.
- [BMWZ87] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai. Some classes of global Cramer Rao bounds. *Ann. Statistics*, 15 :1421–1438, 1987.
- [Böh84] J. F. Böhme. Estimation of source parameters by maximum likelihood and non linear regression. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pages 731–734, 1984.
- [Böh86] J. F. Böhme. Estimation of spectral parameters of correlated signals in wavefields. *Signal Processing*, 10 :329–337, 1986.
- [Boy02] E. Boyer. *Estimation paramétrique des moments spectraux d'échos Doppler : application aux radars strato-troposphériques*. PhD thesis, Ecole Normale Supérieure de Cachan, Cachan, FR, December 2002.
- [Bre88] Y. Bresler. Maximum likelihood of linearly structured covariance with application to antenna array processing. In *Proc. ASSP Workshop on Spectrum Estimation and Modeling*, pages 172–175, Minneapolis, MN, 1988.
- [BSET97] K. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees. Extended Ziv Zakai lower bound for vector parameter estimation. *IEEE Trans. Signal Processing*, 43 :624–638, March 1997.
- [BT74] S. Bellini and G. Tartara. Bounds on error in signal parameter estimation. *IEEE Trans. Commun.*, 22 :340–342, March 1974.
- [BZ76] B. Z. Bobrovsky and M. Zakai. A lower bound on the estimation error for certain diffusion processes. *IEEE Trans. Inform. Theory*, 22 :45–52, January 1976.
- [Cap69] J. Capon. High resolution frequency wavenumber spectrum analysis. *Proc. IEEE*, 57 :1408–1418, August 1969.
- [CFL04] P. Ciblat, Ph. Forster, and P. Larzabal. Harmonic retrieval in noncircular complex-valued multiplicative noise : Barankin bound. In *EUSIPCO*, Vienne, AU, September 2004.
- [CG05] P. Ciblat and M. Ghogho. Ziv Zakai bound for harmonic retrieval in multiplicative and additive Gaussian noise. In *Proc. of IEEE Statistical Signal Processing Workshop - SSP05*, Bordeaux, FR, July 2005.
- [CGFL05] P. Ciblat, M. Ghogho, Ph. Forster, and P. Larzabal. Harmonic retrieval in the presence of non-circular gaussian multiplicative noise : Performance bounds. *Signal Processing*, 85 :737–749, April 2005.
- [Cha04] E. Chaumette. *Contribution à la caractérisation des performances des problèmes conjoints de détection et d'estimation*. PhD thesis, Ecole Normale Supérieure de Cachan, Cachan, FR, December 2004.
- [Che52] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations. *Ann. Math. Stat.*, 23 :493–507, 1952.
- [Cin75] E. Cinlar. *Introduction to Stochastic Process*. Englewood Cliffs, NJ : Prentice Hall, Inc., 1975.

- [CLF05] E. Chaumette, P. Larzabal, and Ph. Forster. On the influence of a detection step on lower bounds for deterministic parameter estimation. *IEEE Trans. Signal Processing*, 53 :4080–4090, November 2005.
- [CM97] M. Cedervall and R. L. Moses. Efficient maximum likelihood DOA estimation for signals with known waveforms in presence of multipath. *IEEE Trans. Signal Processing*, 45 :808–811, March 1997.
- [Con80] W. J. Conover. *Practical Nonparametric Statistics*. Wiley, New York, 1980.
- [CR51] D. G. Chapman and H. Robbins. Minimum variance estimation without regularity assumptions. *Ann. Math. Stat.*, 22 :581–586, 1951.
- [Cra46] H. Cramér. *Mathematical Methods of Statistics*. Princeton University, Press, New York, 1946.
- [CTO89] H. Clergeot, S. Tressens, and A. Ouamri. Performance of high resolution frequencies estimation methods compared to the Cramer Rao bounds. *IEEE Trans. Acoust., Speech, Signal Processing*, 37(11) :1703–1720, November 1989.
- [CZZ75] D. Chazan, M. Zakai, and J. Ziv. Improved lower bounds on signal parameter estimation. *IEEE Trans. Inform. Theory*, 21 :90–93, January 1975.
- [DA04] J.-P. Delmas and H. Abeida. Stochastic Cramér-Rao bound for noncircular signals with application to DOA estimation. *IEEE Trans. Signal Processing*, 52 :3192–3199, November 2004.
- [DA05] J.-P. Delmas and H. Abeida. Cramer Rao bounds of DOA estimates for BPSK and QPSK modulated signals. *IEEE Trans. Signal Processing*, 54 :117–126, January 2005.
- [Dar45] G. Darmois. Sur les lois limites de la dispersion de certaines estimations. *Rev. Inst. Int. Statist.*, 13 :9–15, 1945.
- [Dug37] D. Dugué. Application des propriétés de la limite au sens du calcul des probabilités à l'étude des diverses questions d'estimation. *Ecol. Poly.*, 3 :305–372, 1937.
- [FBL04] P. Forster, E. Boyer, and P. Larzabal. Nonefficiency of stochastic beamforming bearing estimates at high SNR and finite number of samples. *IEEE Signal Processing Lett.*, 11(5) :509–512, May 2004.
- [FBLR03] Ph. Forster, E. Boyer, P. Larzabal, and A. Renaux. Non-efficacité et non-Gaussianité asymptotiques d'un estimateur du maximum de vraisemblance à fort rapport signal sur bruit. In *Proceedings GRETSI*, volume 1, pages 125–128, Paris, FR, September 2003.
- [FG52] D. A. S. Fraser and I. Guttman. Bhattacharyya bounds without regularity assumptions. *Ann. Math. Stat.*, 23 :629–632, 1952.
- [Fis22] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Phil. Trans. Royal Soc.*, 222 :309, 1922.
- [FL01] Ph. Forster and P. Larzabal. Sur les bornes minimales pour l'estimation de paramètres déterministes. In *Proceedings GRETSI*, pages 707–710, Toulouse, France, 2001.
- [FL02] Ph. Forster and P. Larzabal. On lower bounds for deterministic parameter estimation. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Orlando, FL, 2002.
- [Fre43] M. Frechet. Sur l'extension de certaines évaluations statistiques au cas de petit échantillons. *Rev. Inst. Int. Statist.*, 11 :182–205, 1943.

- [FT03] A. Ferrari and J-Y Tournet. Barankin lower bound for change points in independent sequences. In *Proc. of IEEE Statistical Signal Processing Workshop - SSP03*, St. Louis, Missouri, September 2003.
- [Gla72] F. E. Glave. A new look at the Barankin lower bound. *IEEE Trans. Inform. Theory*, 18(3) :349–356, May 1972.
- [GR94] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, San Diego, 1994.
- [Ham50] J. M. Hammersley. On estimating restricted parameters. *J. R. Soc. Ser. B*, 12 :192–240, 1950.
- [HLS93] S. Haykin, J. Litva, and T. J. Shepherd. *Radar Array Processing*. Springer-Verlag, Berlin, 1993.
- [IH81] I. A. Ibragimov and R. Z. Has'minski. *Statistical estimation*. Springer-Verlag, New-York, 1981.
- [Jaf88] A. G. Jaffer. Maximum likelihood direction finding of stochastic sources : a separable solution. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, New York, 1988.
- [Kay93] S. M. Kay. *Fundamentals of Statistical Signal Processing*, volume 1. Prentice Hall, NJ, 1993.
- [Kie52] J. Kiefer. On minimum variance estimators. *Ann. Math. Stat.*, 23 :627–629, 1952.
- [Kno97] L. Knockaert. The Barankin bound and threshold behavior in frequency estimation. *IEEE Trans. Signal Processing*, 45 :2398–2401, September 1997.
- [Kot59] V. A. Kotel'nikov. *The Theory of Optimum Noise Immunity*. McGraw-Hill, New York, 1959.
- [KT83] R. Kumaresan and D. W. Tufts. Estimating the angles of arrival of multiple plane waves. *IEEE Trans. Aerosp. Electron. Syst.*, 19(1) :134–138, 1983.
- [LC93] J. Li and R. T. Compton. Maximum likelihood angle estimation for signals with known waveforms. *IEEE Trans. Signal Processing*, 41 :2850–2862, September 1993.
- [Leh83] E. L. Lehmann. *Theory of Point Estimation*. Wiley, New York, 1983.
- [LHSV95] J. Li, B. Halder, P. Stoica, and M. Viberg. Computationally efficient angle estimation for signals with known waveforms. *IEEE Trans. Signal Processing*, 43 :2154–2163, September 1995.
- [Lil67] H. Lilliefors. On the Kolmogorov Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62 :399–402, 1967.
- [LVT89] F. Li, J. Vaccaro, and D. W. Tuft. Min-Norm linear prediction for arbitrary sensor arrays. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pages 2613–2616, Glasgow, GB, May 1989.
- [Man04] A. Manikas. *Differential geometry in array processing*. Imperial College Press, London, UK, 2004.
- [Mar97] T. L. Marzetta. Computing the Barankin bound by solving an unconstrained quadratic optimization problem. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Munich, GR, April 1997.
- [MH71] R. J. McAulay and E. M. Hofstetter. Barankin bounds on parameter estimation. *IEEE Trans. Inform. Theory*, 17 :669–676, November 1971.

- [MS69] R. J. McAulay and L. P. Seidman. A useful form of the Barankin lower bound and its application to ppm threshold analysis. *IEEE Trans. Inform. Theory*, 15 :273–279, March 1969.
- [OVSN93] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai. Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing. In S. Haykin, J. Litva, and T. J. Shepherd, editors, *Radar Array Processing*, chapter 4, pages 99–151. Springer-Verlag, Berlin, 1993.
- [Pie63] J. N. Pierce. Approximate error probabilities for optimal diversity combining. *IEEE Trans. Comm. Syst.*, 1963.
- [PRK86] A. Paulraj, R. Roy, and T. Kailath. A subspace rotation approach to signal parameter estimation. *Proceeding of the IEEE*, pages 1044–1045, July 1986.
- [QCL06] A. Quinlan, E. Chaumette, and P. Larzabal. A direct method to generate approximations of the Barankin bound. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Toulouse, FR, May 2006.
- [RAFL06] A. Renaux, L. N. Atallah, Ph. Forster, and P. Larzabal. A useful form of the Abel bound and its application to estimator threshold prediction. *to appear in IEEE Trans. on Signal Processing*, April 2006.
- [Rao45] C. R. Rao. Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, 37 :81–91, 1945.
- [Rao02] C. R. Rao. *Linear Statistical Inference and Its Applications*. Wiley, New York, 2002.
- [RB74] D. C. Rife and R. R. Boorstyn. Single tone parameter estimation from discrete time observations. *IEEE Trans. Inform. Theory*, 20 :591–598, 1974.
- [Red79] S. Reddi. Multiple source location. A digital approach. *IEEE Trans. Aerosp. Electron. Syst.*, pages 95–105, 1979.
- [Ren06] A. Renaux. Weiss-Weinstein bound for data aided carrier estimation. *accepted (minor revision) for IEEE Signal Processing Letters*, June 2006.
- [RFB04] A. Renaux, Ph. Forster, and E. Boyer. Asymptotic non efficiency of a maximum likelihood estimator at finite number of samples. In *Proceedings EUSIPCO*, Vienna, AT, September 2004.
- [RFBL04] A. Renaux, Ph. Forster, E. Boyer, and P. Larzabal. Non efficiency and non Gaussianity of a maximum likelihood estimator at high signal to noise ratio and finite number of samples. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, volume 2, pages 121–124, Montreal, QC, May 2004.
- [RFBL06] A. Renaux, Ph. Forster, E. Boyer, and P. Larzabal. Unconditional maximum likelihood performance at finite number of samples and high signal to noise ratio. *to appear in IEEE Trans. on Signal Processing*, April 2006.
- [RFCL06] A. Renaux, Ph. Forster, E. Chaumette, and P. Larzabal. On the high SNR conditional maximum-likelihood estimator full statistical characterization. *to appear in IEEE Trans. on Signal Processing*, 2006.
- [RFL05a] A. Renaux, Ph. Forster, and P. Larzabal. A new derivation of the Bayesian bounds for parameter estimation. In *Proc. of IEEE Statistical Signal Processing Workshop - SSP05*, Bordeaux, FR, July 2005.
- [RFL05b] A. Renaux, Ph. Forster, and P. Larzabal. Une nouvelle approche des bornes bayésiennes. In *Proceedings GRETSI*, Louvain la Neuve, BE, September 2005.

- [RFLR06a] A. Renaux, Ph. Forster, P. Larzabal, and C.D. Richmond. The Bayesian Abel bound on the mean square error. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Toulouse, FR, May 2006.
- [RFLR06b] A. Renaux, Ph. Forster, P. Larzabal, and C.D. Richmond. A new derivation of the Bayesian bounds for parameter estimation. *in preparation for IEEE Trans. on Information Theory*, 2006.
- [Ric05] C. D. Richmond. Capon algorithm mean squared error threshold SNR prediction and probability of resolution. *IEEE Trans. Signal Processing*, 53(8) :2748–2764, August 2005.
- [RM95] I. Reuven and H. Messer. The use of the Barankin bound for determining the threshold snr in estimating the bearing of a source in the presence of another. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, volume 3, pages 1645–1648, Detroit, USA, May 1995.
- [RM97] I. Reuven and H. Messer. A Barankin-type lower bound on the estimation error of a hybrid parameter vector. *IEEE Trans. Inform. Theory*, 43(3) :1084–1093, May 1997.
- [RM99] I. Reuven and H. Messer. On the effect of nuisance parameters on the threshold snr value of the Barankin bound. *IEEE Trans. Signal Processing*, 47(2) :523–527, February 1999.
- [RPK86] R. Roy, A. Paulraj, and T. Kailath. ESPRIT a subspace rotation approach to estimation of parameters of cisoids in noise. *IEEE Trans. Acoust., Speech, Signal Processing*, 34 :1340–1342, 1986.
- [Rud76] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 1976.
- [SB85] A. O. Steinhardt and C. Bretherton. Threshold in frequency estimation. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pages 1273–1276, Tampa, FL, March 1985.
- [Sch81] R. O. Schmidt. *A signal subspace approach to multiple emitter location and spectral estimation*. PhD thesis, Stanford University, Stanford, CA, November 1981.
- [SGB67] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp. Lower bounds to error probability for coding on discrete memoryless channels i. *Information and Control*, 10 :65–103, February 1967.
- [SM93] L. L. Scharf and T. McWhorter. Geometry of the Cramer Rao bound. *Signal Processing*, 31 :301–311, 1993.
- [Smi05a] S. T. Smith. Covariance, subspace, and intrinsic Cramer Rao bounds. *IEEE Trans. Signal Processing*, 53 :1610–1630, May 2005.
- [Smi05b] S. T. Smith. Statistical resolution limits and the complexified Cramer Rao bound. *IEEE Trans. Signal Processing*, 53 :1597–1609, May 2005.
- [SN89] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and the Cramer Rao bound. *IEEE Trans. Acoust., Speech, Signal Processing*, 37 :720–741, May 1989.
- [SN90a] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and the Cramer Rao bound : further results and comparisons. *IEEE Trans. Acoust., Speech, Signal Processing*, 38 :2140–2150, 1990.
- [SN90b] P. Stoica and A. Nehorai. Performances study of conditional and unconditional direction of arrival estimation. *IEEE Trans. Acoust., Speech, Signal Processing*, 38 :1783–1795, October 1990.

- [TK97] J. Tabrikian and J. L. Krolik. Barankin bound for source localization in shallow water. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Munich, GR, April 1997.
- [TK99] J. Tabrikian and J.L. Krolik. Barankin bounds for source localization in an uncertain ocean environment. *IEEE Trans. Signal Processing*, 47 :2917–2927, November 1999.
- [TST94] J. K. Thomas, L. L. Scharf, and D. W. Tufts. The probability of a subspace swap in the SVD. *IEEE Trans. Signal Processing*, 43 :730–736, March 1994.
- [Van68] H. L. VanTrees. *Detection, Estimation and Modulation Theory*, volume 1. Wiley, New York, 1968.
- [Van01] H. L. VanTrees. *Detection, Estimation and Modulation Theory : Radar-Sonar Signal Processing and Gaussian Signals in Noise*, volume 3. Wiley, New York, 2001.
- [Van02] H. L. VanTrees. *Detection, Estimation and Modulation theory : Optimum Array Processing*, volume 4. Wiley, New York, 2002.
- [vdV98] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, Cambridge, UK, 1998.
- [VO91] M. Viberg and B. Ottersten. Sensor array processing based on subspace fitting. *IEEE Trans. Signal Processing*, 39 :1110–1121, May 1991.
- [VOK91] M. Viberg, B. Ottersten, and T. Kailath. Detection and estimation in sensor array processing using Weighted Subspace Fitting. *IEEE Trans. Signal Processing*, 39(11) :2436–2449, November 1991.
- [VON95] M. Viberg, B. Ottersten, and A. Nehorai. Performance analysis of direction finding with large arrays and finite data. *IEEE Trans. Signal Processing*, 43(2) :469–477, February 1995.
- [Wei88] E. Weinstein. Relations between Belini-Tartara, Chazan-Zakai-Ziv, and Wax-Ziv lower bounds. *IEEE Trans. Inform. Theory*, 34 :342–343, March 1988.
- [WW83] A. J. Weiss and E. Weinstein. Fundamental limitation in passive time delay estimation part i : Narrow-band systems. *IEEE Trans. Acoust., Speech, Signal Processing*, 31 :472–486, April 1983.
- [WW85] A. J. Weiss and E. Weinstein. A lower bound on the mean square error in random parameter estimation. *IEEE Trans. Inform. Theory*, 31 :680–682, September 1985.
- [WW88] E. Weinstein and A. J. Weiss. A general class of lower bounds in parameter estimation. *IEEE Trans. Inform. Theory*, 34 :338–342, March 1988.
- [XB04a] J. Xavier and V. Barroso. Intrinsic variance lower bound (IVLB) : an extension of the Cramer Rao bound to Riemannian manifolds. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, volume 5, pages 1033–1036, Hong Kong, CH, March 2004.
- [XB04b] J. Xavier and V. Barroso. The Riemannian geometry of certain parameter estimation problems with singular Fisher information matrices. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, volume 2, pages 1021–1024, Montreal, QC, May 2004.
- [XBB04] W. Xu, A. B. Baggeroer, and K. Bell. A bound on mean-square estimation with background parameter mismatch. *IEEE Trans. Inform. Theory*, 50 :621–632, April 2004.
- [XBR04] W. Xu, A. B. Baggeroer, and C. D. Richmond. Bayesian bounds for matched-field parameter estimation. *IEEE Trans. Signal Processing*, 52 :3293–3305, December 2004.

- [Xu01] W. Xu. *Performances bounds on matched-field methods for source localization and estimation of ocean environmental parameters*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, June 2001.
- [YN06] I.S. Yetik and A. Nehorai. Performance bounds on image registration. *IEEE Trans. Signal Processing*, 54 :1737–1736, May 2006.
- [ZLO06] Q. Zou, Z. Lin, and R. J. Ober. The Cramer Rao lower bound for bilinear systems. *IEEE Trans. Signal Processing*, 54 :1666–1680, May 2006.
- [ZN90] A. Zeira and A. Nehorai. Frequency domain Cramer Rao bound for Gaussian processes. *IEEE Trans. Signal Processing*, 38 :1063–1066, June 1990.
- [ZZ69] J. Ziv and M. Zakai. Some lower bounds on signal parameter estimation. *IEEE Trans. Inform. Theory*, 15 :386–391, May 1969.