

**UNIVERSITE PARIS XI
UFR SCIENTIFIQUE D'ORSAY**

THESE

Présentée

Pour obtenir

**Le GRADE de DOCTEUR EN SCIENCES
DE L'UNIVERSITE PARIS XI ORSAY**

PAR

Mr. Yann MEURISSE

**Sujet : Contributions à l'étude de performances statistiques
de réseaux d'antennes**

Soutenue le 21 Novembre 2002 devant la Commission d'examen

Mr Pascal LARZABAL	Président
Mr Olivier BESSON	Rapporteur
Mr Philippe FORSTER	Rapporteur
Mr Jean Pierre DELMAS	Examineur
Mr Phillip REGALIA	Examineur

Remerciements

Je tiens tout d'abord à remercier Monsieur Pascal Larzabal d'avoir bien voulu me faire l'honneur d'accepter la présidence du jury.

Je remercie particulièrement Monsieur Olivier Besson et Monsieur Philippe Forster d'avoir bien voulu accepter d'être les rapporteurs de cette thèse.

Je remercie également Monsieur Phillip Regalia d'avoir bien voulu être membre du jury et de m'avoir encouragé, durant ces dernières années, dans mon travail.

Je remercie l'Institut National des Télécommunications de m'avoir donné les moyens de poursuivre cette tâche et tout spécialement Monsieur Claude Gimenes de m'avoir incité à poursuivre mes études et de m'avoir ensuite soutenu indéfectiblement dans cette voie.

Je tiens également à remercier tout particulièrement mon directeur de thèse, Monsieur Jean Pierre Delmas, d'avoir consacré une bonne partie de son temps à m'aider, me soutenir et me guider avec exigence tout au long de mes travaux de recherche et pour l'attention constante dont il a fait preuve à mon égard.

Je tiens enfin à remercier mes collègues et amis de l'INT qui m'ont prodigué, de près ou de loin, leur aide et leur soutien.

Table des matières

1	Présentation générale	9
2	Formulaire et notations	13
3	Réseaux réguliers non uniformes mono et multidimensionnels	17
3.1	Introduction	17
3.2	Panorama des réseaux d’antennes non uniformes	20
3.2.1	Réseaux linéaires non uniformes	20
3.2.2	Réseaux non uniformes 2D et 3D	22
3.3	L’augmentation des réseaux linéaires non-uniformes	23
3.3.1	Réseaux complètement augmentables	23
3.3.2	Réseaux partiellement augmentables	25
3.4	Bornes pour les réseaux lacunaires plans ou de volume	25
3.5	Conclusion	28
	Annexe 1 : Preuve de la relation (4.4)	31
4	Sensibilité des algorithmes d’estimation de DOA “bande étroite” à la distribution des signaux et à la corrélation temporelle des signaux	35
4.1	Introduction	35
4.2	Modèle et notations	36
4.3	Distribution asymptotique de la matrice de covariance empirique $\mathbf{R}_x(n)$	37
4.4	Distribution asymptotique des estimées de DOA	39
4.4.1	Approche fonctionnelle	39
4.4.2	Algorithmes standards	40
4.5	Bruit temporellement corrélé	41
4.5.1	Covariance asymptotique des estimées et fréquence d’échantillonnage	41
4.5.2	Covariance asymptotique des estimées et densité spectrale de puissance du bruit	43
4.6	Bruit temporellement non corrélé	45
4.6.1	Antenne de géométrie quelconque	45
4.6.2	Antenne centrosymétrique	46
4.7	Algorithmes non standards	47
4.7.1	Algorithmes de Toeplization	47
4.7.2	Algorithmes d’augmentation	51
4.7.3	L’algorithme COMET en DOA bande étroite	52
4.8	Simulations	54
4.9	Conclusion	59
	Annexe 1 : Preuve du résultat (4.1)	61
	Annexe 2 : Expression de $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{MUSIC-2D}}$	66
	Annexe 3 : Preuve de l’expression (5.1)	67
	Annexe 4 : Preuve de l’expression (5.3)	67

5	Robustesse des algorithmes du second ordre d'estimation de DOA face à la largeur de bande	71
5.1	Introduction	71
5.2	Modélisation du problème	72
5.3	Performance des méthodes de détection de l'ordre	74
5.4	Biais et covariance asymptotiques	77
5.4.1	Cas d'une seule source	79
5.4.2	Cas de deux sources	81
5.5	Borne de Cramer-Rao	83
5.5.1	Cas d'une seule source	83
5.5.2	Cas de plusieurs sources	88
5.6	Performances comparées d'algorithmes bande étroite et large bande	89
5.7	Conclusion	91
	Annexe 1 : Preuve des relations (2.7), (2.8) et (2.9)	93
	Annexe 2 : Preuve des relations (4.12), (4.13), (4.14) et (4.15)	93
	Annexe 3 : Preuve de la relation (4.16)	94
	Annexe 4 : Preuve de la relation (4.19)	94
	Annexe 5 : Preuve des relations (5.4), (5.5) et (5.6)	95
	Annexe 6 : Preuve de la relation (5.8)	96
	Annexe 7 : Autre justification de (5.7)	98
6	Application des techniques de "Covariance Matching Estimation" (COMET) à l'estimation de direction d'arrivée ou de fréquence de sinusoides	103
6.1	Introduction	103
6.2	Principe de l'algorithme COMET	104
6.2.1	Principe d'Invariance Etendue	104
6.2.2	Principe de l'algorithme COMET original	106
6.2.3	Formulation de l'algorithme COMET dans le cas réel	109
6.2.4	Performances asymptotiques de COMET	110
6.3	Reformulation de l'algorithme COMET original	111
6.3.1	Cas d'observations gaussiennes réelles indépendantes	114
6.3.2	Cas d'observations complexes gaussiennes circulaires indépendantes	116
6.4	Algorithme AMV dans le cas d'observations non indépendantes	118
6.4.1	Cas d'observations réelles	119
6.4.2	Cas d'observation complexes	120
6.5	Performances des algorithmes AMV	122
6.5.1	Cas optimum	123
6.5.2	Cas sous optimum	124
6.6	Application de COMET à l'estimation de DOA	126
6.6.1	Modèle et notations	126
6.6.2	Performances asymptotiques de COMET en DOA	128
6.6.3	Sensibilité de COMET à la distribution et à la corrélation temporelle des signaux sources	134
6.6.4	Robustesse par rapport au S/B de certains capteurs	136
6.7	Application de COMET à l'estimation de fréquence de sinusoides	139
6.7.1	Modèle et notations	139
6.7.2	L'algorithme AMV	140
6.7.3	Analyse des performances et de la robustesse	141
6.8	Conclusion	151
	Annexe 1 : Preuve de (2.9)	153

Annexe 2 : Preuve des résultats 6.2 et 6.3	153
Annexe 3 : Preuve des résultats 6.4 et 6.5	154
Annexe 4 : Preuve de (4.8)	156
Annexe 5 : Calcul des différentielles des estimateurs (5.3) et (5.7)	156
Annexe 6 : L'algorithme de Pisarenko modifié (MPD)	157
Annexe 7 :	161

Table des figures

3.1	Exemples de réseaux linéaires lacunaires	18
3.2	Diagrammes de rayonnement	18
3.3	Fonctions co-réseau	19
3.4	Exemples de réseaux carrés MR	22
3.5	Exemples de réseaux carrés MH	23
3.6	Performance des réseaux complètement augmentables	24
3.7	Nombre de capteurs et bornes pour les réseaux lacunaires carrés	29
4.1	EQM en fonction du coef. AR du signal et du bruit	44
4.2	EQM en fonction du coef. AR du signal et du bruit	45
4.3	Probabilité de résolution de deux sources	50
4.4	Coefficients de corrélation des estimées de deux directions	51
4.5	EQM en fonction du nb. d'observations	54
4.6	EQM en fonction du rapport S/B	55
4.7	EQM en fonction de la largeur de bande des signaux ARMA	56
4.8	EQM en fonction du nb. de capteurs	57
4.9	EQM en azimuth et élévation dans le cas d'un réseau lacunaire	58
5.1	Valeurs propres de la matrice de covariance spatiale	75
5.2	Nombre moyen de sources détectées	75
5.3	Nombre de sources détectées en fonction de la largeur de bande réduite (1 source)	76
5.4	Nombre de sources détectées en fonction de la largeur de bande réduite (2 sources)	76
5.5	Courbes d'équi-EQM en fonction de $\frac{b}{f_0}$ et de $\frac{f_m}{f_0}$	80
5.6	EQM en fonction de $\frac{b}{f_0}$ pour différentes valeurs de $\frac{f_m}{f_0}$	80
5.7	Biais asymptotique en fonction de $\frac{Mb}{f_0}$ (5 capteurs)	82
5.8	Biais asymptotique en fonction de $\frac{Mb}{f_0}$ (10 capteurs)	82
5.9	Evolution des coefficients c et c' en fonction du rapport S/B	85
5.10	$\text{Var}(\theta_1^b(T))$ et $\text{Var}(\theta_1^0(T))$ en fonction du rapport S/B	85
5.11	$\text{Var}(\theta_1^b(T))$ et $\text{Var}(\theta_1^0(T))$ en fonction du rapport S/B	86
5.12	Comparaison de l'EQM MUSIC et de la BCR dans le cas d'1 source large bande	87
5.13	Comparaison de l'EQM MUSIC et de la BCR dans le cas de 2 sources large bande	89
5.14	Performances comparées d'algo. de DOA dans le cas d'une source large bande	90
5.15	Performances comparées d'algo. de DOA dans le cas de 2 sources large bande	90
6.1	Perf. comparées de COMET large bande et bande étroite et de l'algo. de Friedlander	133
6.2	Robustesse de COMET large bande vis à vis de la distribution de la source	134
6.3	EQM en fonction de la largeur de la bande de la source	135
6.4	EQM en fonction de l'écartement spectral de 2 sources	135
6.5	Performance de COMET large bande en fonction du rapport S/B sur un capteur	137
6.6	Borne de Cramer-Rao en fonction du rapport S/B sur un capteur	138

6.7	Variance de l'estimateur AMV en fonction du nombre d'observations	142
6.8	Variance des estimateurs MDP et AMV en fonction du rapport S/B	143
6.9	Variance asymptotique des algorithmes AMV et MUSIC en fonction du rapport S/B . . .	143
6.10	DSP du bruit MA et du bruit blanc	144
6.11	Variance des estimateurs AMV et MUSIC en fonction de la fréquence de la sinusoïde . . .	144
6.12	Variance des estimateurs MDP et AMV en fonction de Δf	145
6.13	Biais des estimateurs AMV et MDP en fonction du coefficient AR du bruit	146
6.14	Biais des estimateurs AMV et MDP en fonction du coefficient AR du bruit	147
6.15	Erreur quadratique moyenne des estimateurs AMV et MDP en fonction du coefficient AR du bruit	147
6.16	Variance asymptotique des algorithmes AMV optimal et AMV sous-optimaux en fonction du rapport S/B	149
6.17	Variance asymptotique des algorithmes AMV optimal et AMV sous-optimaux en fonction de M	149
6.18	Variance asymptotique des algorithmes AMV optimal et AMV sous-optimaux en fonction du rapport S/B	150
6.19	Variance asymptotique des algorithmes AMV optimal et AMV sous-optimaux en fonction de M	150

Liste des tableaux

3.1	Réseaux LMR	21
3.2	Réseaux LMR au sens large	21
3.3	Réseaux LMT	22
3.4	Nombre de configurations carrées MR	28
3.5	Nombre de configurations carrées MH	28

Chapitre 1

Présentation générale

Le traitement d'antenne est un domaine du traitement du signal qui a suscité beaucoup d'intérêt dans la communauté scientifique au cours de ces trois dernières décennies. Le problème de l'estimation de direction d'arrivée (DOA) de sources en particulier a fait l'objet de nombreuses études qui ont permis notamment l'élaboration de méthodes du second ordre performantes telles que les méthodes dites "haute résolution" comme MUSIC, ESPRIT ou MIN-NORM pour ne citer qu'elles.

Les performances de ces méthodes sont en général évaluées en supposant d'une part que les réseaux d'antennes utilisés sont linéaires, à répartition régulière uniforme des capteurs et d'autre part que les signaux sources et le bruit, sont gaussiens, temporellement non corrélés et indépendants entre eux.

Les signaux observés sont d'autre part supposés, soit "bande étroite" soit "large bande" selon que leur enveloppe complexe reste constante ou non durant la traversée du réseau. Des méthodes spécifiques, adaptées à chacune de ces situations, ont été développées. Les méthodes dites "bande étroite" sont en général plus simples à mettre en œuvre que les méthodes dites "large bande" dont la complexité de calcul est beaucoup importante.

Dans ce cadre général notre étude a porté sur les points suivants :

- l'étude de géométries de réseaux d'antennes, régulières non uniformes, optimales selon certains critères, pouvant permettre l'amélioration des performances des méthodes conventionnelles en DOA,
- l'étude des performances des méthodes du second ordre d'estimation de DOA utilisées dans des situations non conventionnelles, en particulier si les signaux sources et/ou le bruit sont des processus aléatoires éventuellement corrélés temporellement et éventuellement non gaussiens,
- l'étude de la robustesse des algorithmes bande étroite d'estimation de DOA vis à vis de la largeur de bande des signaux et de la symétrie de leurs spectres,
- l'étude enfin d'une méthode dite de "covariance matching estimation" (COMET) appliquée à l'estimation de DOA ainsi qu'à l'estimation de fréquences de sinusoides qui fournit, sous certaines conditions, des estimations asymptotiquement de variance minimale.

Plus précisément, ce document est organisé de la façon suivante :

Le chapitre 2 présente les notations et rappelle un certain nombre de définitions et de propriétés d'opérateurs utilisés dans la suite.

Le chapitre 3 est consacré à l'étude des géométries de réseaux d'antennes, régulières non-uniformes, mono ou multidimensionnelles. Après avoir rappelé l'intérêt que peut présenter ce type de réseaux, un panorama en est dressé en s'appuyant sur certains critères d'optimalité géométrique que l'on peut leur associer. Nous rappelons ensuite le principe de l'*augmentation* des réseaux linéaires non-uniformes c'est à dire comment de tels réseaux peuvent permettre d'améliorer les performances des méthodes d'estimation de DOA. Enfin, dans le cas des réseaux linéaires non-uniformes, nous proposons des bornes plus fines

que celles précédemment données par d'autres auteurs et nous étendons ces bornes au cas des réseaux multidimensionnels.

Le chapitre 4 est consacré à l'étude des performances asymptotiques des algorithmes du second ordre d'estimation de DOA de sources bande étroite placés dans des situations non conventionnelles. Contrairement aux hypothèses habituelles nous supposons en effet que les signaux des sources et/ou le bruit sont des processus aléatoires éventuellement corrélés temporellement et éventuellement non gaussiens. Avec nos hypothèses et grâce à une approche fonctionnelle, nous déduisons les lois asymptotiques gaussiennes des diverses estimées de DOA de la distribution asymptotique gaussienne de la matrice de covariance spatiale empirique. Nous montrons dans quelles conditions certains algorithmes ne sont pas sensibles à la distribution de probabilité et à la corrélation temporelle des sources. Nous examinons également l'influence de la corrélation temporelle du bruit et nous montrons en particulier que le paramètre d'intérêt dans l'étude de performance de DOA en présence de bruit additif temporellement blanc à temps continu est le temps d'observation et non le nombre de snapshots. Enfin, nous considérons certains algorithmes particuliers tels que les méthodes de "Toeplitzation" et d'"augmentation" et l'algorithme de "covariance matching estimation" (COMET) et nous montrons que les performances de ces algorithmes sont très sensibles à la corrélation temporelle des signaux sources.

Le chapitre 5 est consacré à l'étude de la robustesse des algorithmes bande étroite du second ordre d'estimation de DOA vis à vis de la largeur de bande du signal. Comme nous l'avons dit précédemment, dans le domaine de l'estimation de DOA on distingue deux types d'algorithmes selon que l'enveloppe complexe des signaux est considérée ou non comme constante durant la traversée du réseau, en d'autres termes, selon que les signaux sont considérés comme bande étroite ou large bande. Les méthodes associées à chacune de ces hypothèses n'ont pas le même degré de complexité, les méthodes large bande étant en général plus lourdes en charge de calcul que les méthodes bande étroite. Il est alors intéressant d'examiner où se situe la frontière entre ces deux domaines et d'étudier dans quelle mesure, en terme de dégradation de performance, des méthodes bande-étroite peuvent être appliquées à des signaux à bande étendue. Dans ce chapitre nous généralisons une étude faite par Sorelius *et al.* dans un contexte restreint aux signaux à spectre symétrique par rapport à la fréquence de démodulation, au cas plus général des signaux à spectre non symétrique et/ou non centrés sur cette fréquence. D'autre part, en considérant la détection d'ordre, la borne de Cramer-Rao et la comparaison des performances asymptotiques entre des algorithmes bande étroite et large bande, nous montrons que la définition empirique de l'hypothèse bande étroite souvent utilisée dans la littérature selon laquelle l'ouverture du réseau est beaucoup plus petite que l'inverse de la bande relative du signal (i.e., $\frac{Mb}{f_0} \ll 1$ où M représente le nombre de capteurs) est bien trop sévère dans le cas de sources à spectres symétriques par rapport à la fréquence de démodulation. Par contre, dans le cas de spectres dissymétriques par rapport à la fréquence de démodulation (en particulier pour les spectres symétriques mais démodulés avec une fréquence de démodulation différente de la fréquence centrale de leur spectre), cette propriété de robustesse des algorithmes bande étroite est perdue.

Le chapitre 6 enfin est consacré à l'étude des techniques dites de "Covariance Matching Estimation" (COMET) qui sont des méthodes d'estimation de type "moindres carrés pondérés non linéaires du deuxième ordre" pour lesquelles l'existence, parmi l'ensemble des paramètres inconnus, de paramètres intervenant linéairement permet une simplification du problème d'optimisation à résoudre. Ces techniques ont été présentées dans le cadre d'observations complexes circulaires gaussiennes et indépendantes où elles s'avèrent asymptotiquement efficaces. Dans ce chapitre, nous étudions les performances de l'algorithme COMET, appliqué à l'estimation de paramètres réels, lorsque *les observations ne sont plus nécessairement gaussiennes et indépendantes* et nous les relierons à la notion d'estimateur asymptotiquement de variance minimale (AMV). Les résultats généraux de cette étude sont illustrés en considérant l'application de la méthode COMET à deux problèmes classiques en traitement du signal : l'estimation de direction d'arrivée de sources et l'estimation de fréquences de sinusoides en présence de bruit additif MA.

Publications de l'auteur

Le travail effectué dans le cadre de cette thèse a été l'occasion de proposer les contributions suivantes :

Articles de revue

J.P.Delmas, Y.Meurisse, "Asymptotic performance analysis of DOA finding algorithms with temporally correlated narrowband signals," *IEEE Transactions on Signal Processing*, vol. 48, no. 9, pp. 2669-2674, September 2000.

Y.Meurisse, J.P.Delmas, "Bounds for sparse planar and volume arrays," *IEEE Transactions on Information theory*, vol. 47, no. 1, pp. 464-468, January 2001.

J.P.Delmas, Y.Meurisse, "Robustness of narrowband DOA algorithms with respect to signal bandwidth," soumis à *Signal Processing (Elsevier)*, February 2002 et en cours de review.

J.P.Delmas, Y.Meurisse, "Asymptotically minimum variance second-order estimation of frequencies for mixed spectra time series," soumis à *IEEE Transactions on Signal Processing*, June 2002.

Communications dans un congrès

J.P.Delmas, Y.Meurisse, "Performances statistiques d'algorithmes de DOA en présence de sources temporellement corrélées," *Proc. GRETSI, Vannes*, pp. 753-756, Septembre 1999.

Y.Meurisse, J.P.Delmas, "Bounds for sparse planar and volume arrays with redundancies and holes," *ESA SP-444 Proceedings of millennium conference on Antennas and Propagation (AP2000)*, Davos, April 2000.

Y.Meurisse, J.P.Delmas, "Robustness of narrowband DOA algorithms with respect to signal bandwidth," accepted at *Conference on Mathematics in Signal Processing, Warwick*, December 2000.

Y.Meurisse, J.P.Delmas, "Robustesse de la localisation de sources bande étroite et large bande par covariance matching estimation," *Proc. GRETSI Toulouse*, Septembre 2001.

J.P.Delmas, Y.Meurisse, "Asymptotically minimum variance second-order estimation of frequencies for mixed spectra time series," *EUSIPCO 2002*, September 2002.

Communications dans journées d'Etude SEE

Y.Meurisse, J.P.Delmas, "Réseaux lacunaires mono et multidimensionnels," *Conférence SEE : Antennes non standard, techniques et traitements*, Paris, pp.19-24, Mars 2001.

Chapitre 2

Formulaire et notations

D'une façon générale, dans ce document, les matrices sont représentées par des majuscules grasses, les vecteurs par des minuscules grasses et les scalaires par des minuscules non grasses. Les vecteurs, par défaut, sont orientés "colonne". Nous utiliserons les notations suivantes :

$\mathbf{A}^T$	: la matrice transposée de $\mathbf{A}$
$\mathbf{A}^H$	: la matrice transposée-conjuguée de $\mathbf{A}$
a^*	: complexe conjugué de a
$\mathbf{A}^\#$	: la matrice pseudo-inverse de $\mathbf{A}$ au sens de Moore-Penrose
$\text{Tr}(\mathbf{A})$	: trace de la matrice $\mathbf{A}$
$\text{Det}(\mathbf{A})$	: déterminant de la matrice $\mathbf{A}$
$\text{Vec}(\mathbf{A})$	: vecteur obtenu en "empilant" les colonnes de la matrice $\mathbf{A}$ les unes sous les autres
$\text{Vet}(\mathbf{A})$	: vecteur obtenu en "empilant" les colonnes du bloc triangulaire inférieur, diagonale principale comprise, de la matrice $\mathbf{A}$ les unes sous les autres
$\mathbf{A} \otimes_r \mathbf{B}$	: produit de Kronecker "réel", matrice dont le bloc (i, j) est $a_{i,j} \mathbf{B}$
$\mathbf{A} \otimes_c \mathbf{B}$	: produit de Kronecker "complexe", matrice dont le bloc (i, j) est $b_{i,j}^* \mathbf{A}$

Le produit de Kronecker complexe sera souvent utilisé dans la suite, aussi nous en rappelons les principales propriétés.

Propriétés du produit de Kronecker complexe et de l'opérateur de vectorisation

En adaptant les propriétés du produit de Kronecker réel et de l'opérateur de vectorisation, démontrées notamment dans [1], au produit de Kronecker complexe, nous avons les propriétés suivantes (dans ces différentes expressions, les dimensions des matrices sont supposées telles que ces expressions aient un sens) :

$$(\mathbf{A} \otimes_c \mathbf{B}) (\mathbf{C} \otimes_c \mathbf{D}) = \mathbf{AC} \otimes_c \mathbf{BD} \quad (0.1)$$

$$\text{Vec}(\mathbf{ABC}) = (\mathbf{A} \otimes_c \mathbf{C}^H) \text{Vec}(\mathbf{B}) \quad (0.2)$$

$$(\mathbf{A} \otimes_c \mathbf{B})^H = \mathbf{A}^H \otimes_c \mathbf{B}^H \quad (0.3)$$

$$\text{Vec}(\mathbf{a}\mathbf{b}^H) = \mathbf{a} \otimes_c \mathbf{b} \quad (0.4)$$

$$\text{Vec}(\mathbf{A} \mathbf{b} \mathbf{c}^H) = (\mathbf{A} \mathbf{b}) \otimes_c \mathbf{c} = (\mathbf{A} \otimes_c \mathbf{c}) \mathbf{b} \quad (0.5)$$

$$\text{Vec}^H(\mathbf{c} \mathbf{b}^H \mathbf{A}^H) = \mathbf{c}^H \otimes_c (\mathbf{b}^H \mathbf{A}^H) = \mathbf{b}^H (\mathbf{c}^H \otimes_c \mathbf{A}^H) \quad (0.6)$$

$$\text{Tr}(\mathbf{A}\mathbf{B}) = \text{Vec}^T(\mathbf{A}^T) \text{Vec}(\mathbf{B}). \quad (0.7)$$

Matrice de commutation $\mathbf{K}_n^c$

Soit $\mathbf{A}$ une matrice $n \times n$, $\text{Vec}(\mathbf{A})$ et $\text{Vec}(\mathbf{A}^T)$ contiennent les mêmes termes mais dans un ordre différent. Il existe donc une matrice de permutation $n^2 \times n^2$ unique $\mathbf{K}_n^c$, appelée *matrice de commutation*, telle que

$$\text{Vec}(\mathbf{A}^T) = \mathbf{K}_n^c \text{Vec}(\mathbf{A}). \quad (0.8)$$

Propriétés de $\mathbf{K}_n^c$

Puisque $\mathbf{K}_n^c$ est une matrice de permutation, elle est orthogonale et $(\mathbf{K}_n^c)^T = (\mathbf{K}_n^c)^{-1}$. D'autre part, en multipliant à gauche (0.8) par $\mathbf{K}_n^c$ on obtient $\text{Vec}(\mathbf{A}) = \mathbf{K}_n^c \mathbf{K}_n^c \text{Vec}(\mathbf{A})$ donc

$$(\mathbf{K}_n^c)^T = (\mathbf{K}_n^c)^{-1} = \mathbf{K}_n^c. \quad (0.9)$$

Matrice de duplication $\mathbf{K}_n^d$

Soit $\mathbf{A}$ une matrice $(n \times n)$. Nous noterons $\text{Vet}(\mathbf{A})$ le $\frac{n(n+1)}{2} \times 1$ vecteur obtenu en “vectorisant” le bloc triangulaire inférieur de la matrice $\mathbf{A}$, diagonale principale comprise.

Si la matrice $\mathbf{A}$ est symétrique, $\text{Vec}(\mathbf{A})$ et $\text{Vet}(\mathbf{A})$ contiennent les mêmes éléments. Il existe donc une matrice $n^2 \times \frac{n(n+1)}{2}$ unique $\mathbf{K}_n^d$ dite *matrice de duplication* telle que :

$$\text{Vec}(\mathbf{A}) = \mathbf{K}_n^d \text{Vet}(\mathbf{A}). \quad (0.10)$$

Propriétés de $\mathbf{K}_n^d$

Les colonnes de $\mathbf{K}_n^d$ sont linéairement indépendantes ce qui implique $\mathbf{K}_n^{d\#} = (\mathbf{K}_n^{dT} \mathbf{K}_n^d)^{-1} \mathbf{K}_n^{dT}$ et par suite, pour toute matrice $\mathbf{A}$ symétrique

$$\text{Vet}(\mathbf{A}) = \mathbf{K}_n^{d\#} \text{Vec}(\mathbf{A}). \quad (0.11)$$

D'autre part

$$\mathbf{K}_n^d \mathbf{K}_n^{d\#} = \frac{1}{2} (\mathbf{I}_{n^2} + \mathbf{K}_n^p). \quad (0.12)$$

Bibliographie

- [1] J.R. Magnus, H. Neudecker *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley and Sons, 1999.

Chapitre 3

Réseaux réguliers non uniformes mono et multidimensionnels

3.1 Introduction

Le domaine du traitement d'antenne a largement évolué au cours des trois dernières décennies, depuis l'apparition des toutes premières techniques utilisant principalement le filtrage spatial jusqu'à l'avènement des dernières méthodes dites "à haute résolution" fondées sur des approches statistiques optimales. Ces différentes techniques, bien que de principes différents, ont toutes à la base comme point commun d'utiliser des réseaux de capteurs, linéaires ou plans, à répartition uniforme. Les raisons de ce choix initial proviennent principalement de la facilité de conception et d'étude de tels réseaux et du fait qu'ils permettent l'implémentation efficaces de certains algorithmes. Dans un 2ème temps et pratiquement pour chacune de ces techniques s'est ensuite posée la question de savoir s'il existait d'autres types de réseaux, à répartition non-uniforme, qui soient plus intéressants car offrant des performances comparables voire supérieures à celles des réseaux uniformes pour un nombre inférieur de capteurs. Dans le domaine du filtrage spatial (formation de voie), les performances d'un réseau sont liées à son diagramme de rayonnement (largeur du lobe principal, hauteur des lobes secondaires) [1] alors que dans le domaine de la localisation de sources par des méthodes du second ordre, les performances (en nombre maximum de sources détectables et en variance des estimées) dépendent essentiellement, pour une méthode donnée, de la manière dont le réseau "échantillonne" la fonction de corrélation spatiale du champ d'onde.

Rappelons brièvement les relations existant entre la géométrie du réseau et les notions respectives de diagramme de rayonnement et de matrice de corrélation spatiale. Considérons pour cela un réseau quelconque de M capteurs dont les positions sont repérées, dans un référentiel donné, par leurs coordonnées cartésiennes $\mathbf{s}_m = [x_m, y_m, z_m]^T$ $m = 1, \dots, M$. Dans ce même référentiel la direction d'arrivée d'une onde plane de fréquence f_0 , de longueur d'onde λ_0 , sera repérée par son azimuth ϕ et son élévation θ ou encore par les fréquences spatiales associées : $f_x \stackrel{\text{def}}{=} \frac{\cos\phi \sin\theta}{\lambda_0}$, $f_y \stackrel{\text{def}}{=} \frac{\sin\phi \sin\theta}{\lambda_0}$ et $f_z \stackrel{\text{def}}{=} \frac{\cos\theta}{\lambda_0}$. Le gain complexe du réseau est alors

$$g(f_x, f_y, f_z) = \sum_{m=1}^M e^{i2\pi(f_x x_m + f_y y_m + f_z z_m)}$$

c'est à dire la transformée de Fourier de la distribution $i(x, y, z) = \sum_{m=1}^M \delta(x - x_m, y - y_m, z - z_m)$ qui peut être considérée comme la "distribution indicatrice" du réseau.

Le gain du réseau, définissant son diagramme de rayonnement, est finalement donné par $G(f_x, f_y, f_z) = |g(f_x, f_y, f_z)|^2$, c'est à dire par la transformée de Fourier de $R_i(x, y, z)$: la fonction d'autocorrélation de $i(x, y, z)$. Ceci montre que la largeur du lobe principal du diagramme de rayonnement est inversement proportionnelle au support de la fonction $R_i(x, y, z)$ et par suite à l'ouverture du réseau. Pour illustrer ceci considérons 3 répartitions linéaires de 5 capteurs (FIG.3.1). Chacune de ces répartitions est "régulière"

en ce sens que les capteurs se situent sur des graduations entières de l'axe (l'unité de longueur étant égale à $\lambda_0/2$)¹. La répartition A_1 est uniforme, d'ouverture égale à 4 par opposition aux répartitiones A_2 et A_3 , dont les propriétés géométriques seront étudiées dans le paragraphe 3.2 et qui offrent des ouvertures respectivement égales à 9 et 11. On constate (FIG.3.2) que la largeur du lobe principal du diagramme de rayonnement correspondant est d'autant plus faible que l'ouverture du réseau est grande.

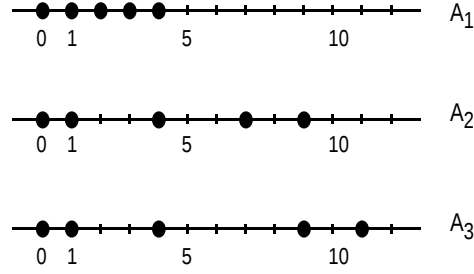


FIG. 3.1: Trois exemples de répartition linéaire, régulière de 5 capteurs, d'ouvertures respectives 4, 9 et 11.

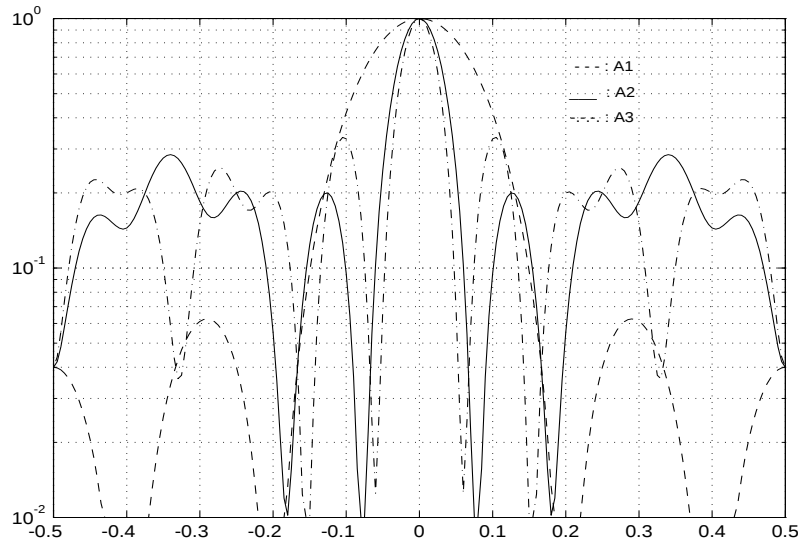


FIG. 3.2: Diagrammes de rayonnement associés aux 3 réseaux précédents.

Concernant maintenant la matrice de corrélation spatiale, si nous utilisons un réseau régulier pour échantillonner spatialement un champ d'onde aléatoire $e(\mathbf{s}, t)$ résultant de K sources spatialement et temporellement stationnaires, de fonction de corrélation $\mathbf{R}_e(\chi, \tau) \stackrel{\text{def}}{=} E[e(\mathbf{s}, t)e^*(\mathbf{s} - \chi, t - \tau)]$, le terme général de la matrice de covariance du signal de sortie $\mathbf{y}(t)$ du réseau sera donné par

$$[\mathbf{R}_y(\tau)]_{m,n} = \mathbf{R}_e(\mathbf{s}_m - \mathbf{s}_n, \tau)$$

qui ne dépend spatialement que de l'écart $\mathbf{d}_{m,n} = \mathbf{s}_m - \mathbf{s}_n$ entre les capteurs. L'ensemble des points $\mathbf{d}_{m,n}$ en lesquels le réseau fournit une estimée de $\mathbf{R}_e(\chi, \tau)$ constitue ce que l'on appelle le *co-réseau* [2] et, un même écart $\mathbf{d}$ pouvant apparaître plusieurs fois dans le réseau, la fonction $\Lambda(\mathbf{d})$, représentant le nombre de fois où un écart $\mathbf{d}$ apparaît dans le réseau, est appelée *fonction co-réseau* [3] (notons que, quel que soit le réseau, $\Lambda(\mathbf{0}) = M$). Lorsqu'un écart, autre que $\mathbf{d} = \mathbf{0}$, apparaît plusieurs fois dans le réseau, cet écart est dit *redondant* et le réseau est lui même qualifié de *redondant*. Dans le cas contraire, si chaque écart,

¹Nous ne considérerons dans ce chapitre que des répartitiones de ce type que nous qualifierons de régulières

autre que $\mathbf{d} = \mathbf{0}$, n'apparaît au plus qu'une seule fois, le réseau est dit *sans redondance*. Enfin, lorsqu'un écart n'apparaît pas dans le réseau, on dit que celui-ci présente un *trou* pour cet écart.

Pour illustrer ceci, la FIG.3.3 présente les fonctions co-réseau des 3 réseaux précédents, représentées uniquement dans le domaine $d \geq 0$, puisqu'il s'agit de fonctions paires.

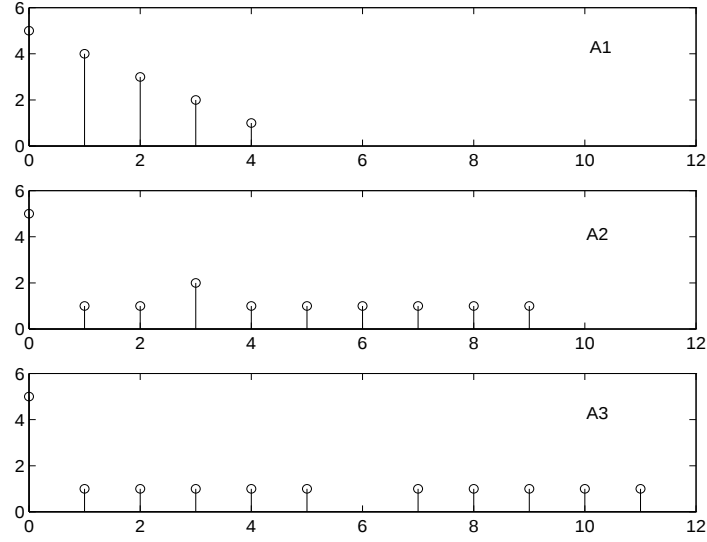


FIG. 3.3: Fonctions co-réseau des 3 réseaux de la FIG.3.1.

Le réseau A_2 est “sans trou” et ne présente (hormis pour $d = 0$) de redondance que pour $d = 3$ (et $d = -3$). Nous verrons qu’il correspond en fait à l’un des réseaux réguliers, linéaires, de 5 capteurs ne présentant pas de trou et un nombre de redondances minimum. Ce type de réseau sera qualifié de *linéaire à redondance minimum* (LMR).

Le réseau A_3 pour sa part est sans redondance et ne présente qu’un seul trou. Il s’agit cette fois d’un exemple de réseau régulier, linéaire, de 5 capteurs présentant un minimum de trous. Ce type de réseau sera qualifié de *linéaire à minimum de trous* (LMH).

Les matrices de covariance spatiale des signaux de sortie de ces réseaux sont de dimension 5×5 et permettent donc la détection et la localisation de 4 sources non corrélées. Toutefois, sous les hypothèses précédentes, la matrice de covariance spatiale associée à un réseau uniforme de M capteurs, hermitienne Toeplitz est caractérisée par M termes uniquement. D’autre part, le réseau A_2 fournit au moins une estimation de la fonction de corrélation pour $\chi = 0, 1, \dots, 9$. Avec ces 10 valeurs on peut alors envisager d’estimer “correctement” la matrice de covariance associée à un réseau virtuel linéaire uniforme de 10 capteurs et permettre ainsi le traitement d’un maximum de 9 sources. Le même principe peut être envisagé avec le réseau A_3 dont le co-réseau toutefois présente lui un trou et qui donc nécessiterait une technique d’augmentation légèrement différente.

L’intérêt des réseaux non uniformes dans le domaine de la détection-localisation de sources ne se limite toutefois pas à l’accroissement du nombre de sources que l’on peut traiter mais permet également, dans certaines situations, d’améliorer la précision de l’estimation des directions d’arrivée fournie par des méthodes du second ordre. En utilisant un critère de D-optimalité consistant à minimiser le déterminant de la matrice de covariance des estimées au sens du maximum de vraisemblance, Huang *et al.* [5] ont ainsi proposé des réseaux linéaires réguliers non-uniformes offrant, à nombre de capteurs identique, des performances, en terme de variance d’estimée, supérieures à celles de réseaux à redondance minimum elles mêmes supérieures à celles des réseaux uniformes. Abramovich *et al.* [4, 6, 7] de leur côté ont proposé différentes méthodes d’augmentation adaptées aux réseaux sans trou et aux réseaux sans redondance qu’ils qualifient respectivement de *complètement augmentables* (*fully augmentable array*) et de *partiellement augmentable* (*partially augmentable array*) et ont montré dans quelle mesure ces réseaux

pouvaient améliorer les performances des réseaux uniformes. Une synthèse de leur étude est présentée dans le paragraphe 3.3.

Un réseau quelconque de M capteurs fournit au maximum $M(M - 1)$ écarts non nuls distincts, ce nombre étant d'autant plus grand que la redondance est petite. De ce fait, pour un nombre donné de capteurs, les réseaux à redondance minimum ou éventuellement à minimum de trous permettront d'optimiser la démarche précédente.

Le problème qui se pose alors est de déterminer, pour un nombre fixé de capteurs, les différentes configurations, mono ou multidimensionnelles, présentant ces propriétés.

3.2 Panorama des réseaux d'antennes non uniformes

La majorité des études et des recherches qui ont été menées sur les réseaux non-uniformes s'est focalisée sur les réseaux linéaires et ce sont eux que nous décrirons tout d'abord. Nous verrons ensuite comment certaines notions introduites en dimension 1 peuvent s'étendre en dimension 2 ou 3.

3.2.1 Réseaux linéaires non uniformes

S'appuyant sur les travaux de Leech [8] en Théorie des nombres, Moffet [9] fut parmi les premiers à définir les notions de réseau à minimum de redondance au “sens strict” et au “sens large”. Une caractéristique de ces géométries particulières est qu'il n'existe pas de relation analytique permettant de donner la répartition optimale d'un nombre fixé de capteurs. Les études dans ce domaine se sont alors orientées d'une part vers la recherche de réseaux de plus en plus étendus [10] ou sous-optimaux [11] et d'autre part, vers l'étude des performances de tels réseaux en localisation de sources [4, 6, 7].

Pour décrire ces réseaux de façon la plus simple et la plus complète possible nous utiliserons la classification employée par Abramovich *et al.* dans [7].

Soient M capteurs répartis sur un axe de sorte que toutes les distances entre capteurs soient des multiples entiers d'une distance fondamentale d_0 . Dans le repère ayant pour origine le 1^{er} capteur, la géométrie du réseau sera caractérisée par les coordonnées entières $[x_1 = 0 < x_2 < \dots < x_M]$ de ses capteurs. Rappelons que l'ensemble $\mathcal{D} = \{\mathbf{d}_{ij} = x_i - x_j | i, j = 1, 2, \dots, M\}$ des écarts entre capteurs constitue ce que l'on appelle le co-réseau et que la fonction $\Lambda(d), d \in \mathbb{Z}$ représentant le nombre de fois ou un écart d apparaît dans le réseau est appelée fonction co-réseau.

Le réseau est alors dit *complètement augmentable* ou simplement *sans trou* (ST) si tous les écarts intermédiaires apparaissent au moins une fois dans le réseau ; en d'autres termes si $\forall d \in \mathbb{Z}$ tel que $-x_M \leq d \leq x_M$, $\Lambda(d) \geq 1$.

L'idéal serait que $\Lambda(d) = 1, \forall d = 1, \dots, x_M$ de sorte que le réseau soit complètement augmentable et strictement non-redondant ou encore *parfait*. Malheureusement ceci n'est possible, comme nous le montrerons au paragraphe suivant, que si $M \leq 4$. Sinon, pour $M \geq 5$, certains écarts, autres que $d = 0$, apparaissent plusieurs fois. Soit R , le nombre de ces redondances appelé redondance du réseau. Les réseaux complètement augmentables qui, pour un nombre fixé de capteurs, minimisent R sont dits à *minimum de redondance* (MR) (ceci correspondant à la dénomination “minimum de redondance au sens strict” utilisée par Moffet). Le réseau $A_2 : 1\ 3\ 3\ 2$ de la FIG.3.1 en est un exemple. Notons au passage que pour spécifier ces réseaux nous emploierons la codification communément utilisée qui consiste à lister les écarts entre capteurs plutôt que leurs positions absolues.

Le TAB.3.1 fournit tous les réseaux linéaires MR obtenus pour $M \leq 7$. A notre connaissance le plus grand réseau connu de ce type compte 17 capteurs (cf [11] pour une liste complète).

Comme il n'existe pas d'expression analytique de ces géométries optimales, ces différentes configurations ont été obtenues par recherche exhaustive. Toutefois, certaines configurations “presque optimales” ont été proposées par Pearson [12] et par Linebarger [11].

Nb. de capt. M	Redondance R	Ouverture A	Réseaux
3	0	3	1 2
4	0	6	1 3 2
5	2	9	1 3 3 2 1 1 4 3
6	4	13	1 1 4 4 3 1 3 1 6 2 1 5 3 2 2
7	8	17	1 1 4 4 4 3 1 1 1 5 5 4 1 1 6 4 2 3 1 3 6 2 3 2 1 7 3 2 2 2

TAB. 3.1: Réseaux linéaires à minimum de redondance (LMR).

Par opposition aux réseaux complètement augmentables, les réseaux dits *partiellement augmentables* admettent un certain nombre de “trous” dans leur co-réseau. L’optimisation de ces réseaux peut alors se faire selon 2 critères différents :

On peut par exemple chercher à maximiser N_{max} , le plus grand multiple entier de l’écart unitaire d_0 de telle sorte que tous les écarts de 0 à N_{max} inclus soient présents dans le réseau (ceci correspond à la notion de *réseau à minimum de redondance au sens large* de Moffet). En d’autres termes, le plus petit écart “manquant” dans le co-réseau est $N_{max} + 1$. Ces réseaux sont aussi appelés [7] : à *maximum d’écarts contigus*.

Le TAB.3.2 fournit tous les réseaux linéaires à minimum de redondance au sens large pour $5 \leq M \leq 7$ (pour $2 \leq M \leq 4$, ces réseaux coïncident avec les réseaux LMR).

Nb. de capt. M	N_{max}	Ouverture A	Réseaux
5	9	11	3 1 5 2
		13	4 1 2 6
6	13	16	4 1 1 7 3
		19	6 1 2 2 8
		17	1 3 6 2 5
		17	1 7 3 2 4
7	18	24	6 3 1 7 5 2
		25	8 1 3 6 5 2
		31	14 1 3 6 2 5
		31	13 1 2 5 4 6

TAB. 3.2: Réseaux linéaires à minimum de redondance au sens large.

Une autre façon d’optimiser les réseaux partiellement augmentables est de minimiser le nombre de trous sous la contrainte que la redondance R soit nulle, en d’autres termes que chaque écart non nul n’apparaisse au plus qu’une seule fois. Ce type de réseau est appelé réseau (ou règle) de Golomb ou encore réseau à *minimum de trous* (MT). Le réseau $A_3 : 1 3 5 2$ de la FIG.3.1 en fournit un exemple.

Le TAB.3.3 fournit tous les réseaux linéaires à minimum de trous pour $M \leq 7$, un article récent a étendu cette liste jusqu’à $M = 19$ [13].

Nb. de capt. M	Nb de trous H	Ouverture A	Réseaux
3	0	3	1 2
4	0	6	1 3 2
5	1	11	1 3 5 2 2 5 1 3
6	2	17	1 7 3 2 4 1 7 4 2 3 1 3 6 2 5 1 3 6 5 2
7	4	25	1 10 5 3 4 2 1 6 4 9 3 2 1 3 6 8 5 2 2 5 6 8 1 3 2 1 7 6 5 4

TAB. 3.3: Réseaux linéaires à minimum de trous (LMT).

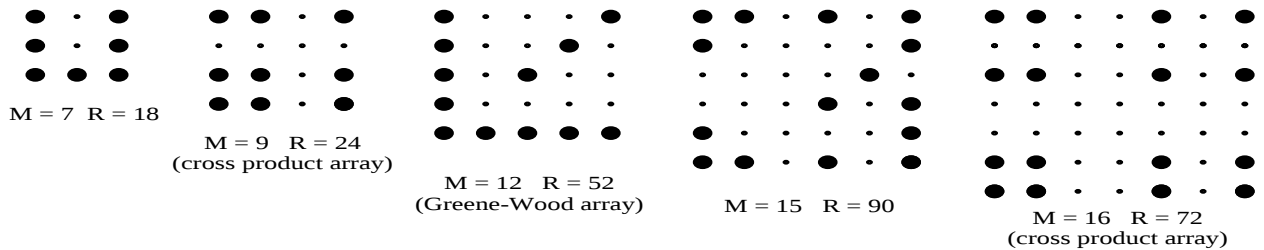
3.2.2 Réseaux non uniformes 2D et 3D

Les notions de minimum de redondance et de minimum de trous se généralisent facilement aux réseaux réguliers plans ou de volume à condition de se limiter à des formes simples de réseaux : rectangles ou parallélépipèdes, comme nous le ferons dans la suite. Pas plus que dans le cas linéaire, il n'existe de relation fournissant, à ouverture donnée, le nombre nécessaire ni la répartition des capteurs conduisant à un réseau optimal. Toutefois, dans le cas des réseaux à minimum de redondance, il existe deux type de structures particulières qui fournissent des candidats intéressants.

Les premiers, appelés “cross product” (CP) par Pumphrey [14], sont construits à partir d'un (ou plusieurs) réseau(x) linéaire(s) MR. Un réseau CP carré par exemple pourra être construit à partir d'un même réseau linéaire MR et aura un capteur en (x, y) si ce réseau linéaire MR a un capteur en x et un capteur en y (cf. FIG.3.4). On peut bien sur utiliser des réseaux linéaires différents.

Les seconds ont été proposés par Greene et Wood (GW) [15] en tant que réseaux carrés. De façon générale la position (x, y, z) d'un capteur d'un tel réseau cubique d'ouverture $A \geq 3$ vérifie : $x = 0$ ou $y = 0$ ou $z = 0$ ou $x = y = z = 2, \dots, A$ ($z = 0$ pour un réseau carré cf. FIG.3.4).

La FIG.3.4 présente des exemples de réseaux carrés à minimum de redondance des deux types précédents ou obtenus par recherche exhaustive pour des ouvertures allant de 2 à 6.

FIG. 3.4: Exemples de réseaux carrés à minimum de redondance avec le nombre de capteurs (M) et la redondance (R) associés.

La FIG.3.5 présente des exemples de réseaux carrés à minimum de trous pour des ouvertures allant de 2 à 7 obtenus par recherche exhaustive.

Il est à noter que pour les réseaux carrés ou cubiques, contrairement à ce qui se passe pour les réseaux linéaires, il n'existe pas nécessairement de solution optimale pour chaque valeur de M . La recherche

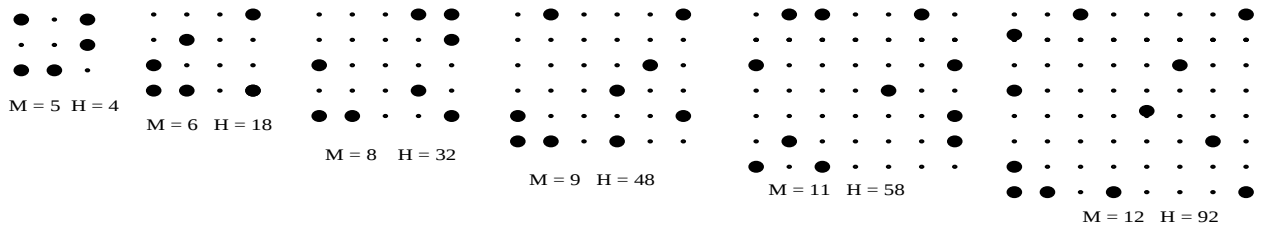


FIG. 3.5: Exemples de réseaux carrés à minimum de trous avec le nombre de capteurs (M) et le nombre de trous (H) associés.

des solutions optimales consiste donc, pour une ouverture donnée, à déterminer le nombre de capteurs nécessaires et leur(s) répartition(s). Comme nous le montrerons au paragraphe 3.4 relation (4.1), les ouvertures A_x , A_y et A_z d'un réseau parallélépipédique dans chacune des dimensions, le nombre de capteurs M , la redondance R et le nombre de trous H vérifient la relation

$$(2A_x + 1)(2A_y + 1)(2A_z + 1) = (M^2 - M + 1) + H - R.$$

Dans le cas des réseaux MR, à ouvertures fixées, minimiser R sous la contrainte $H = 0$ est donc équivalent à maximiser M sous cette même contrainte. De façon analogue, dans le cas des réseaux à minimum de trous, minimiser H sous la contrainte $R = 0$ est équivalent à minimiser M sous cette contrainte.

L'un des principaux intérêts des réseaux réguliers non-uniformes est d'améliorer les performances des algorithmes de DOA en possibilité de détection et/ou en précision d'estimation. La section suivante présente une synthèse succincte des études menées par Abramovich *et al.* [4] dans ce but.

3.3 L'augmentation des réseaux linéaires non-uniformes

3.3.1 Réseaux complètement augmentables

Dans le cas de ces réseaux sans trou, Abramovich *et al.* [4] distinguent 2 situations : le cas "conventionnel" où le nombre de sources non-corrélées K est inférieur au nombre M de capteurs du réseau lacunaire et le cas "supérieur" où $M \leq K < M_\alpha$, M_α représentant le nombre de capteurs du réseau virtuel, linéaire uniforme, ayant le même co-réseau que le réseau lacunaire.

- cas conventionnel ($K < M$)

L'étude de la borne de Cramer-Rao (BCR) permet d'envisager une bonne précision, même pour des valeurs de rapport signal à bruit (S/B) et des tailles d'observation N relativement faibles. Toutefois, et en dépit d'une optimalité asymptotique, comme le montre la FIG.3.6, les performances des méthodes du second ordre classiques sont assez éloignées de la BCR dans un domaine dit "préasymptotique" où séparation angulaire, S/B et N sont simultanément faibles. Il est alors possible de tirer parti des propriétés des réseaux complètement augmentables pour améliorer les performances de MUSIC, notamment dans ce domaine préasymptotique.

Le principe de la démarche proposée dans [4] est dans un premier temps de construire, à partir de la matrice de covariance empirique fournie par le réseau lacunaire $\mathbf{R}_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{t=1}^N \mathbf{y}_t \mathbf{y}_t^H$, la matrice Toeplitz définie positive (d.p.) $\hat{\mathbf{T}}_0$ d'ordre M_α dont le spectre du maximum d'entropie (ME) soit égal au pseudo-spectre MUSIC de $\mathbf{R}_N$ dans le domaine asymptotique. Cette première étape, dénommée "MUSIC-ME equalization" (MMEE), fournit une matrice Toeplitz d.p. dont le sous-espace signal n'est toutefois pas nécessairement de dimension K . On remédie à cet inconvénient dans un deuxième temps par une troncature de l'espace signal (SST) obtenue par une méthode itérative de projections alternées, d'une part sur l'espace des matrices Hermitiennes d.p. dont la

dimension de l'espace signal est égale à K puis sur l'espace des matrices Toeplitz. L'algorithme MUSIC est enfin appliqué à la matrice $\hat{\mathbf{T}}_K$ ainsi obtenue.

Le principe de l'augmentation peut donc être résumé par le schéma :

$$\mathbf{R}_N \xrightarrow{MSEE} \hat{\mathbf{T}}_0 \xrightarrow{SST} \hat{\mathbf{T}}_K.$$

Des simulations [6] montrent que cette méthode permet une diminution significative du seuil caractérisant la zone préasymptotique tout en préservant les performances dans la zone asymptotique.

- cas supérieur ($M \leq K < M_\alpha$)

L'étude de la BCR (cf. FIG.3.6 (b)) montre que, même si le rapport S/B est relativement bon, cette situation requière une taille d'observation importante et que, contrairement à ce qui se passe dans le cas conventionnel, une diminution de la taille de l'observation ne peut être compensée par une augmentation de S/B. De plus il est montré dans [16] que, même asymptotiquement, la technique d'augmentation directe (DAA) consistant à former la matrice de covariance Toeplitz du réseau virtuel par moyennage sur chacune des diagonales des termes fournis par la matrice de covariance directe ne permettrait d'approcher la BCR que dans le cas de signaux faibles ($S/B < 0$ dB). Dans le cas de signaux forts, la précision obtenue en utilisant l'augmentation directe est bien plus faible que ce que laisse espérer la BCR.

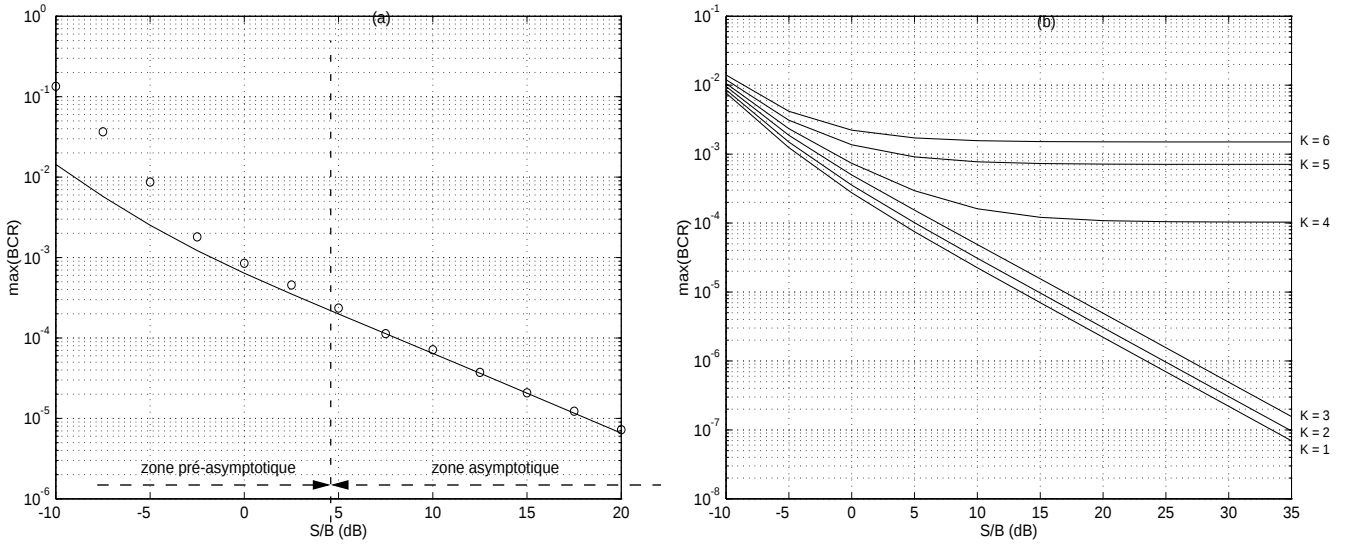


FIG. 3.6: (a) EQM (◦) sur les fréq. spat. de $K = 2$ sources, séparées de 10 deg, estimées par MUSIC avec un réseau lacunaire de $M = 4$ capteurs (1.3.2) comparée à la BCR (-) en fonction du rapport S/B ($N = 200$). (b) Maximum, sur l'ensemble des sources, de la BCR pour le même réseau et un nombre variable $K = 1, \dots, 6$ de sources.

Afin d'améliorer les performances des algorithmes du second ordre tels que MUSIC, en particulier dans le cas de signaux sources forts, il est alors proposé de les faire opérer sur une matrice augmentée obtenue en deux étapes :

$$\mathbf{R}_N \xrightarrow{DAA} \bar{\mathbf{T}}_0 \xrightarrow{SST} \bar{\mathbf{T}}_K.$$

Une augmentation directe appliquée à $\mathbf{R}_N$ fournit une matrice Toeplitz $\bar{\mathbf{T}}_0$ non nécessairement d.p. à laquelle est appliquée une troncature de l'espace signal, telle que celle décrite précédemment, fournissant ainsi une matrice $\bar{\mathbf{T}}_K$ dont les $(M_\alpha - K)$ plus petites valeurs propres sont égales.

L'algorithme MUSIC appliqué à cette matrice fournit alors des premières estimations des directions et des puissances des sources. La dernière étape consiste alors en une estimation, au sens du maximum de vraisemblance, des directions d'arrivée à partir des valeurs initiales préalablement obtenues. La encore des simulations [6] montrent que les performances de cette méthode sont beaucoup plus proches de la BCR que celles obtenues par une simple augmentation directe, en particulier dans le cas de signaux sources forts.

3.3.2 Réseaux partiellement augmentables

Si le nombre de sources est inférieur au nombre d'écarts contigus ($K < N_{max}$) les techniques évoquées dans le paragraphe précédent concernant les réseaux complètement augmentables peuvent être directement appliquées. Sinon, si $N_{max} \leq K < M_\alpha$, des techniques d'augmentation similaires à ces dernières ont été proposées dans [7].

3.4 Bornes pour les réseaux lacunaires plans ou de volume

La recherche exhaustive de réseaux lacunaires, plan ou de volume, optimaux (à minimum de redondance ou à minimum de trous) est un problème de complexité très élevé (de l'ordre de $C_{(A+1)^2}^M$ pour un réseau plan de M capteurs, d'ouverture $A \times A$), ce qui limite son emploi à des réseaux de faible dimension. Il est alors important de disposer de bornes portant sur le nombre de capteurs permettant d'évaluer la pertinence de solutions sous-optimales fournies par d'autres méthodes de recherche (recuit simulé ou algorithmes génétiques [17]). De telles bornes ont été présentées par Linebarger [11] mais uniquement pour des réseaux linéaires. Nous présentons dans cette partie les résultats d'une étude [18] consacrée à l'amélioration sensible de ces bornes et à leur extension au cas de réseaux plans ou volumiques ².

Considérons une antenne volumique formée de M capteurs situés en des sommets d'un réseau cartésien. Rappelons que le co-réseau $\mathcal{D}$ de cette antenne représente l'ensemble des écarts entre capteurs fournis par ce réseau

$$\mathcal{D} = \{\mathbf{d}_{ij} = \mathbf{s}_i - \mathbf{s}_j \mid i, j = 1, 2, \dots, M\}$$

où $\mathbf{s}_i = (x_i, y_i, z_i)^T$ désigne la position du i ème capteur et où x_i, y_i and z_i sont des entiers. Les ouvertures A_x, A_y, A_z de ce réseau sont définies par

$$(A_x, A_y, A_z) = (\max(x_i - x_j), \max(y_i - y_j), \max(z_i - z_j))$$

Rappelons enfin que R et H désignent respectivement le nombre de redondances et le nombre de trous de ce réseau. Si on désigne par E le nombre d'écarts différents dans le réseau (y compris l'écart $\mathbf{0}$) alors le nombre total M^2 d'écarts $\mathbf{d}_{ij}$ est composé de E écarts apparaissant au moins une fois et de $R + M - 1$ écarts strictement redondants et donc $M^2 = E + R + (M - 1)$. Par ailleurs, chacun des $(2A_x + 1)(2A_y + 1)(2A_z + 1)$ écarts du réseau parallélépipédique régulier uniforme associé apparaît au moins une fois ou pas du tout d'où, $E + H = (2A_x + 1)(2A_y + 1)(2A_z + 1)$. De ce fait les ouvertures et les nombres de capteurs, de redondances et de trous vérifient la relation

$$(2A_x + 1)(2A_y + 1)(2A_z + 1) = (M^2 - M + 1) + H - R. \quad (4.1)$$

Afin d'éliminer les ouvertures dans la relation précédente nous introduisons maintenant une nouvelle fonction $\chi_{\mathcal{A}} : \mathbb{Z}^3 \rightarrow \{0, 1\}$ définie de la façon suivante : $\chi_{\mathcal{A}}(\mathbf{s}) = 1$ si il y a un capteur en position $\mathbf{s}$ et $\chi_{\mathcal{A}}(\mathbf{s}) = 0$ sinon. $\chi_{\mathcal{A}}$ peut être considérée comme la "fonction indicatrice" du réseau. La fonction d'autocorrélation $\Lambda(\mathbf{d}) : \mathbb{Z}^3 \rightarrow \mathbb{N}$ de $\chi_{\mathcal{A}}$ représente alors le nombre de fois où l'écart $\mathbf{d}$ apparaît dans le réseau. Grâce à ces notions nous pouvons maintenant montrer le résultat suivant qui affine et généralise [11, rel. 19].

²Remarquons qu'un réseau plan présente une ambiguïté de "côté" qui peut être levée par un réseau volumique

Résultat 3.1 Les nombres de capteurs M , de redondances R et de trous H de tout réseau régulier linéaire, plan ou volumique vérifient l'inégalité : ³

$$R(3\pi + 2(1 + \alpha_A)) + H(3\pi - 2(1 + \alpha_A)) \geq 2(1 + \alpha_A)M^2 - (M - 1)(3\pi + 2(1 + \alpha_A)) \quad (4.2)$$

où α_A est donné en Annexe ⁴.

Preuve Soit $\chi_{\overline{\mathcal{D}}}$ la fonction indicatrice du réseau régulier uniforme d'ouvertures (A_x, A_y, A_z) et $\chi_{\mathbf{0}} : \mathcal{Z}^3 \rightarrow \{0, 1\}$ celle de l'écart $\mathbf{0}$ dont les transformées de Fourier (TF) respectives sont $\frac{\sin \pi(2A_x+1)f_x}{\sin \pi f_x} \frac{\sin \pi(2A_y+1)f_y}{\sin \pi f_y} \frac{\sin \pi(2A_z+1)f_z}{\sin \pi f_z}$ et 1.

Comme $\Lambda(\mathbf{d})$ représente le nombre de fois où l'écart $\mathbf{d}$ apparaît dans le réseau lacunaire, la différence

$$\epsilon(\mathbf{d}) \stackrel{\text{def}}{=} \Lambda(\mathbf{d}) - \chi_{\overline{\mathcal{D}}}(\mathbf{d}) - (M - 1)\chi_{\mathbf{0}}(\mathbf{d}) \quad (4.3)$$

vérifie pour tout $\mathbf{d} \in \overline{\mathcal{D}}$

$$\epsilon(\mathbf{d}) = \begin{cases} \text{nombre de redondances de } \mathbf{d} & \text{si } \mathbf{d} \in \mathcal{D} \text{ et } \mathbf{d} \neq \mathbf{0} \\ 0 & \text{si } \mathbf{d} = \mathbf{0} \\ -1 & \text{si } \mathbf{d} \notin \mathcal{D} \end{cases}$$

De ce fait $\sum_{\mathbf{d} \in \overline{\mathcal{D}}} |\epsilon(\mathbf{d})| = H + R$, et comme la TF $E(\mathbf{f})$ de la fonction réelle et paire $\epsilon(\mathbf{d})$ est réelle, $E(\mathbf{f}) \leq |E(\mathbf{f})| \leq \sum_{\mathbf{d} \in \overline{\mathcal{D}}} |\epsilon(\mathbf{d})| = H + R$. En prenant alors la TF de (4.3) et en remarquant que la TF $L(\mathbf{f})$ de la fonction d'autocorrélation $\Lambda(\mathbf{d})$ est réelle non négative, il apparaît que pour tout $\mathbf{f} = (f_x, f_y, f_z)$ on a :

$$-\frac{\sin \pi(2A_x+1)f_x}{\sin \pi f_x} \frac{\sin \pi(2A_y+1)f_y}{\sin \pi f_y} \frac{\sin \pi(2A_z+1)f_z}{\sin \pi f_z} - (M - 1) \leq H + R - L(\mathbf{f}) \leq H + R.$$

Après quelques manipulations simples détaillées en Annexe, on obtient :

$$\frac{2}{3\pi}(1 + \alpha_A)(2A_x + 1)(2A_y + 1)(2A_z + 1) - (M - 1) \leq H + R. \quad (4.4)$$

En substituant (4.1) dans cette expression, on obtient finalement (4.2). ■

Le Résultat 3.1 est identique à [11, rel. (19)] au coefficient α_A près. Il améliore la borne classique du cas mono-dimensionnel [1] et l'étend aux cas bi- et tri-dimensionnels. De façon surprenante, ce résultat ne dépend pas de la dimensionnalité du réseau.

En annulant respectivement H et R dans la relation (4.2) on obtient des bornes inférieures sur respectivement le nombre de redondances des réseaux sans trou ou sur le nombre de trous des réseaux sans redondance. Quel que soit la dimensionnalité du réseau nous obtenons

$$R \geq \frac{2(1 + \alpha_A)M^2}{3\pi + 2(1 + \alpha_A)} - (M - 1)$$

pour tout réseau sans trou et

$$H \geq \frac{2(1 + \alpha_A)M^2}{3\pi - 2(1 + \alpha_A)} - (M - 1) \frac{3\pi + 2(1 + \alpha_A)}{3\pi - 2(1 + \alpha_A)}$$

pour tout réseau sans redondance.

Ainsi, pour des réseaux relativement grands, les bornes inférieures fournies par [11, rel. (20) et (21)] sont approximativement multipliées par 1.0237.

Concernant les bornes sur M , on déduit facilement des relations (4.2) et (4.1) les résultats suivants :

³Selon notre formulation, les écarts entre capteurs $\mathbf{d}$ sont dans $\mathbb{Z}^3$, aussi nos définitions de H et R diffèrent-elles de celles de [11, rel. 19] d'un facteur 2.

⁴En particulier il est montré en Annexe que $\alpha_A \ll 1$ et que $\alpha_A \approx 0.0237$ pour des ouvertures ≥ 6 .

1. Pour les réseaux parfaits ($R = 0$ et $H = 0$) (4.2) implique $2(M^2 - M + 1) - 3\pi(M - 1) \leq 0$, et donc que $M \in \{2, 3, 4\}$. Grâce à (4.1) on en déduit que les seules solutions sont :

$$M = 3; A_x = 3; A_y = A_z = 0$$

et

$$M = 4; A_x = 6; A_y = A_z = 0.$$

Ceci montre que les seuls réseaux parfaits sont linéaires.

2. Pour les réseaux sans trou ($H = 0$) on obtient :

$$M^2 \geq (2A_x + 1)(2A_y + 1)(2A_z + 1) \left(1 + \frac{2(1 + \alpha_A)}{3\pi} \right), \quad (4.5)$$

ce qui, pour $A_y = A_z = 0$ fournit une borne inférieure sur M sensiblement plus fine que [11, rel. (17)]. Toutefois, en raison des effets d'arrondis, l'amélioration dans le cas mono-dimensionnel n'est pas régulière et n'apparaît que pour certaines valeurs de A . Par exemple, pour les réseaux LMR, $A = 69$ est la plus petite valeur de A pour laquelle les bornes inférieures sur M fournies par [11, rel. (17)] et par (4.5) sont différentes (respectivement $M \geq 13$ et $M \geq 14$).

3. Pour les réseaux sans redondance ($R = 0$) on obtient :

$$(M - 1)^2 \leq (2A_x + 1)(2A_y + 1)(2A_z + 1) \left(1 - \frac{2(1 + \alpha_A)}{3\pi} \right) - 1, \quad (4.6)$$

ce qui, pour $A_y = A_z = 0$ fournit une borne supérieure sur M sensiblement plus fine que [11, rel. (26)]. Mais, pour les mêmes raisons que précédemment, l'amélioration dans le cas mono-dimensionnel n'est pas régulière. Par exemple, pour les réseaux LMH, $A = 51$ est la plus petite valeur de A pour laquelle les bornes supérieures sur M fournies par [11, rel. (26)] et par (4.6) sont différentes (respectivement $M \leq 10$ et $M \leq 9$).

Comme nous l'avons vu, les réseaux CP et GW sont sans trou et il est intéressant de comparer le nombre de leur capteurs à la borne que l'on vient d'obtenir.

Un réseau CP carré (respectivement cubique) d'ouverture A formé de M_{ca} (respectivement M_{cu}) capteurs est construit à partir d'un réseau LMR de M_l capteurs et de même ouverture, avec $M_{ca} = M_l^2$ (respectivement $M_{cu} = M_l^3$). Or Leech [8] a montré que pour les réseaux linéaires MR on a $2.434 \leq \lim_{M_l \rightarrow \infty} \frac{M_l^2}{A} \leq 3.348$. Il s'ensuit que

$$2.434 \leq \lim_{M_{ca} \rightarrow \infty} \frac{M_{ca}}{A} \leq 3.348 \quad (4.7)$$

et

$$2.434 \leq \lim_{M_{cu} \rightarrow \infty} \frac{M_{cu}^{2/3}}{A} \leq 3.348. \quad (4.8)$$

Comparé à (4.5), qui donne respectivement pour des réseaux linéaire, carré et cubique $M^2 \geq 2.434A + 1.217$, $M \geq 2.207A + 1.103$ et $M^{2/3} \geq 2.135A + 1.068$, ces réseaux CP sont donc potentiellement efficaces.

Concernant les réseaux GW, un réseau carré de ce type nécessite $M = 3A$ capteurs et un réseau cubique $M = A(3A + 4)$ capteurs⁵. Rapprochés de (4.5) ces chiffres montrent qu'un réseau GW carré est potentiellement efficace alors qu'un réseau cubique de ce type ne l'est pas.

Les nombres de structures carrées optimales MR et MH obtenues par recherche exhaustive sont récapitulés dans les TAB.3.4 et TAB.3.5. Leur performance en terme de nombre de capteurs vis à vis des bornes précédentes sont présentées FIG.3.7.

⁵ A noter qu'une structure cubique plus performante peut être obtenue en "empilant" des réseaux GW carrés pour laquelle le nombre de capteurs est $M = 3A(A + 1)$

Ouv.	M	M_{inf}	Nb. de configurations		
			Total	GW	CP
2	7	6	2		
3	9	8	4	1	1
4	12	10	4	1	
5	15	13	83	1	
6	16	15	≥ 1		1

TAB. 3.4: Nombre de capteurs et de configurations de réseaux carrés MR obtenus par recherche exhaustive comparé à la borne inférieure (M_{inf}) en fonction de l'ouverture.

Ouv.	M	M_{sup}	Nb. de configurations
2	5	5	1
3	6	7	36
4	8	8	4
5	9	10	347
6	11	12	1
7	12	14	113

TAB. 3.5: Nombre de capteurs et de configurations de réseaux carrés MH obtenus par recherche exhaustive comparé à la borne supérieure (M_{sup}), en fonction de l'ouverture.

3.5 Conclusion

Après avoir dressé un panorama des structures régulières non uniformes de capteurs, nous avons présenté les 2 principaux types de structures optimales pour des réseaux réguliers non uniformes : les réseaux à minimum de redondance et les réseaux à minimum de trous. Une relation générale entre le nombre de capteurs, le nombre de redondances, le nombre de trous et les ouvertures d'un réseau, indépendante de la dimensionnalité de ce réseau, nous a permis de déterminer des bornes sur le nombre de capteurs pour ces 2 types de structures. Ces bornes permettent d'évaluer la pertinence de solutions sous optimales fournies par des méthodes de recherche non exhaustives qu'il reste à élaborer pour des tailles de réseaux toujours plus grandes.

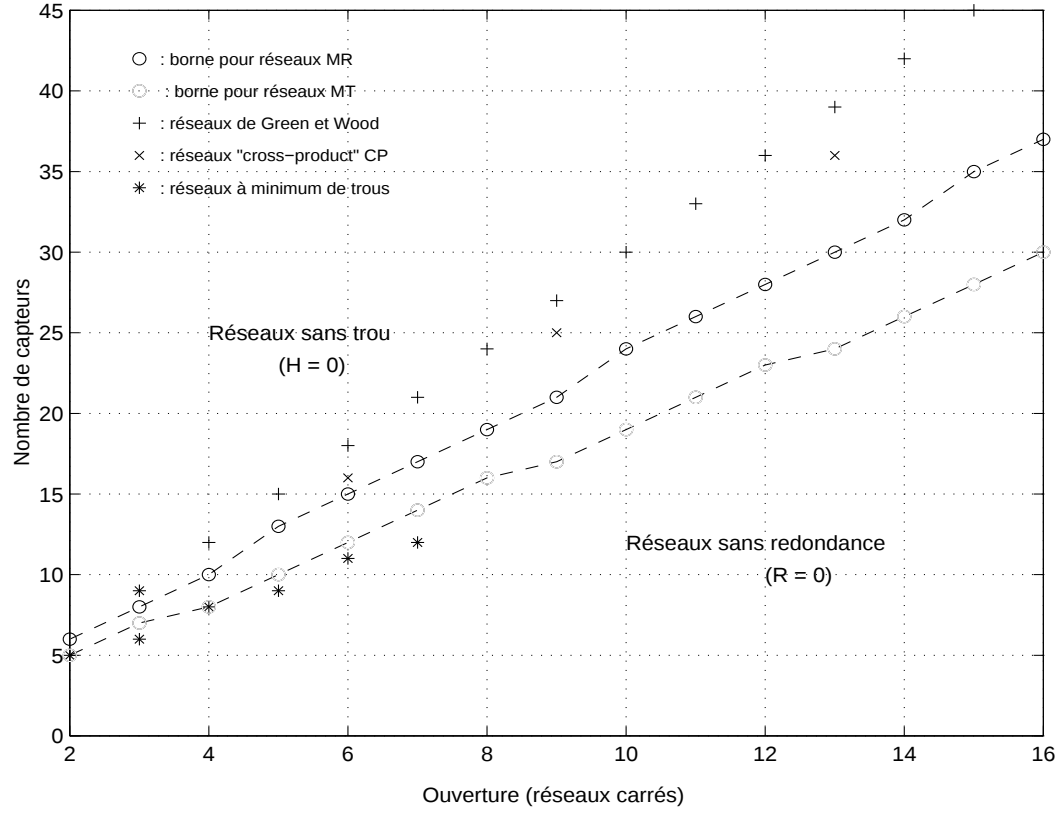


FIG. 3.7: Nombre de capteurs et bornes correspondantes des réseaux lacunaires réguliers carrés.

Annexe

Annexe 1 : Preuve de la relation (4.4)

Posons $s(\mathbf{A}, \mathbf{f}) = -\frac{\sin \pi(2A_x+1)f_x}{\sin \pi f_x} \frac{\sin \pi(2A_y+1)f_y}{\sin \pi f_y} \frac{\sin \pi(2A_z+1)f_z}{\sin \pi f_z}$. Le calcul de $\max_{\mathbf{f}} s(\mathbf{A}, \mathbf{f})$ étant compliqué, nous allons considérer deux valeurs légèrement plus petites.

Tout d'abord, comme $\max_x (-\frac{\sin x}{x}) = \frac{2}{3\pi}(1 + \alpha')$ avec $\alpha' \approx 0.0237$ et que $\sin x < x$ pour $x > 0$, un calcul direct montre que :

$$\frac{2}{3\pi}(1 + \alpha')(2A_x + 1)(2A_y + 1)(2A_z + 1) \leq \max_{\mathbf{f}} s(\mathbf{A}, \mathbf{f}).$$

En supposant ensuite que $A_x = \min(A_x, A_y, A_z)$, pour $\mathbf{f}_0 \stackrel{\text{def}}{=} (f_0, 0, 0)$, et $f_0 \stackrel{\text{def}}{=} \frac{3}{2(2A_x+1)}$ on a :

$$-\frac{\sin \pi(2A_x+1)f_0}{\sin \pi f_0}(2A_y+1)(2A_z+1) = s(\mathbf{A}, \mathbf{f}_0) \leq \max_{\mathbf{f}} s(\mathbf{A}, \mathbf{f}).$$

En remarquant alors que $\frac{\sin \pi(2A_x+1)f_0}{\sin \pi f_0}$ est la transformée de Fourier d'un signal rectangulaire échantillonné on a :

$$\begin{aligned} -\frac{\sin \pi(2A_x+1)f_0}{\sin \pi f_0} &= -\sum_{k=-\infty}^{+\infty} \frac{\sin(\pi(2A_x+1)(f_0-k))}{\pi(f_0-k)} = -\frac{\sin \pi(2A_x+1)f_0}{\pi} \sum_{k=-\infty}^{+\infty} \frac{(-1)^k}{f_0-k} \\ &= -\frac{\sin \pi(2A_x+1)f_0}{\pi} \left(\frac{1}{f_0} + 2f_0 \sum_{k=1}^{+\infty} \frac{(-1)^k}{f_0^2 - k^2} \right) \\ &= \frac{2}{3\pi}(2A_x+1) \left(1 - \frac{9}{2} \sum_{k=1}^{+\infty} \frac{(-1)^k}{k^2(2A_x+1)^2 - \frac{9}{4}} \right) \stackrel{\text{def}}{=} \frac{2}{3\pi}(2A_x+1)(1 + \alpha_A''). \end{aligned}$$

De ce fait

$$\frac{2}{3\pi}(1 + \alpha_A'')(2A_x + 1)(2A_y + 1)(2A_z + 1) \leq \max_{\mathbf{f}} s(\mathbf{A}, \mathbf{f}).$$

En posant alors $\alpha_A \stackrel{\text{def}}{=} \max(\alpha', \alpha_A'')$, on obtient la relation (4.4). On peut remarquer que α_A'' est une fonction décroissante de A_x et une étude minutieuse montre que $\alpha_A = \alpha_A''$ pour $A_x < 6$ (par ex. $\alpha_3 \approx 0.0797$) et $\alpha_A = \alpha' \approx 0.0237$ pour $A_x \geq 6$.

Bibliographie

- [1] S. De Graff, D.H. Johnson, "Optimal linear arrays for narrow-band beamforming," in *Proc. ICASSP-84*, pp. 40.8.1-4, 1984.
- [2] R. A. Haubrich, "Array design," *Bull. Seismol. Soc. Am*, vol. 58, pp. 977-991, 1968.
- [3] D.H. Johnson, D.E. Dudgeon *Array signal processing technics*, Prentice Hall, 1993.
- [4] Y.I. Abramovich, D.A. Gray, A.Y. Gorokhov and N.K. Spencer, "Positive-definite Toeplitz completion in DOA estimation for fully augmentable nonuniform linear antenna arrays," in *Proc. ICASSP*, Atlanta, pp. 2551-2554, 1996.
- [5] X. Huang, J.P. Reilly and M. Wong, "Optimal design of linear array of sensors," in *Proc. ICASSP-91*, Toronto, pp. 1405-1408, 1991.
- [6] Y.I. Abramovich, D.A. Gray, A.Y. Gorokhov and N.K. Spencer, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays - Part I : Fully augmentable arrays," *IEEE Trans. Signal Processing*, vol. 46, pp. 2458-2471, Sept 1998.
- [7] Y.I. Abramovich, D.A. Gray, A.Y. Gorokhov and N.K. Spencer, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays - Part II : Partially augmentable arrays," *IEEE Trans. Signal Processing*, vol. 47, pp. 1502-1521, June 1999.
- [8] J. Leech, "On the representation of 1, 2, ..., n by differences," *J. London Math. Soc.*, vol. 31, pp. 160-169, 1956.
- [9] A.T. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. on Antennas and Propagat.*, vol. AP-16, pp. 172-175, March 1968.
- [10] K. Blanton, J. McClellan, "New search algorithms for minimum redundancy arrays," in *Proc. ICASSP-91*, pp. 1361-1364, 1991.
- [11] D.A. Linebarger, I.H. Sudborough and I.G. Tollis, "Difference bases and sparse sensor arrays," *IEEE Trans. on Information Theory*, vol. 39, no. 2, pp. 716-721, March 1993.
- [12] D. Pearson, S.U. Pillai and Y. Lee, "An algorithm for near-optimal placement of sensors elements," in *ONR Annual Report*, Polytechnic University, June 1988.
- [13] A. Dollas, W.T. Rankin, and D. McCracken, "An new algorithm for Golomb ruler derivation and proof of the 19 mark ruler," *IEEE Trans. on Information Theory*, vol. 44, pp. 379-382, Jan. 1998.
- [14] H.C. Pumphrey, "Design of sparse arrays in one, two, and three dimensions," *J. Acoust. Soc. Am*, vol. 93, no. 3, pp. 1620-1628, March 1993.
- [15] C.R. Greene, R.C. Wood, "Sparse array performance," *J. Acoust. Soc. Am.*, Vol. 63, pp. 1866-1872, 1978.

- [16] Y.I. Abramovich and A.Y. Gorokhov, “Improved analysis of high resolution spatial spectrum estimators in minimum redundancy arrays,” in *Proc. RADAR*, Paris, France, pp. 127-132, 1994.
- [17] J.P. Robinson, “Genetic search for Golomb arrays,” *IEEE Trans. Inform. Theory*, Vol. 46, pp. 1170-1173, May 2000.
- [18] Y. Meurisse, and J.P. Delmas, “Bounds for sparse planar and volume arrays,” *IEEE Trans. Inform. Theory*, Vol. 47, pp. 464-468, Jan. 2001.

Chapitre 4

Sensibilité des algorithmes d'estimation de DOA “bande étroite” à la distribution des signaux et à la corrélation temporelle des signaux

4.1 Introduction

L'étude statistique des performances asymptotiques des algorithmes du second ordre d'estimation de directions d'arrivée (DOA) de sources bande étroite a fait l'objet de nombreux articles. Deux types de modèle ont été utilisés pour réaliser cette étude : le modèle déterministe ou le modèle stochastique. Le modèle déterministe considère que les signaux des sources sont fixés a priori et identiques pour chaque réalisations et que le bruit est un processus aléatoire gaussien temporellement non corrélé. Le modèle stochastique, quant à lui suppose en général que les signaux des sources et le bruit sont des processus aléatoires gaussiens temporellement non corrélés et indépendants entre eux. De nombreuses publications (cf. [1] [2] [3] [4]) ont été consacrées à la comparaison des performances asymptotiques des algorithmes d'estimation de DOA en utilisant ces deux modèles et ont comparé ces performances aux bornes de Cramer-Rao associées. Il s'avère que la plupart des estimateurs de DOA ont les mêmes performances asymptotiques sous chacun de ces deux modèles (cf. [3] [4]) et ceci indépendamment de la distribution statistique des signaux des sources dans le modèle stochastique (cf. [5]). Toutefois ces différentes études présupposent que les observations successives des signaux de sortie des capteurs sont statistiquement indépendantes. On peut alors légitimement se demander ce que deviennent ces performances si cette dernière hypothèse est levée.

Le but de ce chapitre est donc d'étudier les performances asymptotiques des algorithmes d'estimation de DOA en utilisant le modèle stochastique et en considérant que les signaux des sources et/ou le bruit sont des processus aléatoires éventuellement corrélés temporellement et éventuellement non gaussiens.

Sous les hypothèses classiques d'indépendance et de distribution gaussienne des observations, les analyses de performances reposent, soit sur des calculs de perturbation induits par la loi de Wishart complexe de la matrice de covariance spatiale empirique $\mathbf{R}_x(n) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{t=1}^n \mathbf{x}_t \mathbf{x}_t^H$, soit sur un théorème de continuité (voir par exemple [6, théorème p. 122]) exploitant la loi asymptotique gaussienne de $\mathbf{R}_x(n)$ issue du théorème central limite classique. Lorsque les échantillons ne sont pas indépendants aucune de ces deux approches ne peut être envisagée car, d'une part la distribution de probabilité de $\mathbf{R}_x(n)$ ne suit pas une loi de Wishart complexe même dans le cas gaussien et d'autre part le théorème central limite ne peut être appliqué dans la seconde approche. Avec nos hypothèses, nous adoptons l'approche fonctionnelle de [7] dans laquelle les lois asymptotiques gaussiennes des diverses estimées de DOA sont déduites de la distribution asymptotique gaussienne de $\mathbf{R}_x(n)$. Cela nous permet de donner des expressions analytiques

des matrices de covariance asymptotique des estimées de DOA pour la plupart des algorithmes du second ordre et pour différents modèles de dépendance statistique des échantillons.

Après avoir introduit les notations classiques de l'imagerie d'antenne bande étroite et précisé le modèle de corrélation temporelle des observations $\mathbf{x}_t$, un théorème central limit portant sur $\mathbf{R}_x(n)$ est donné dans un cadre général où signaux sources et bruit, tout en étant complexes circulaires peuvent être non gaussiens et temporellement corrélés. Nous en déduisons, par une approche fonctionnelle, la gaussianité asymptotique des estimées au second ordre de DOA ainsi que des contraintes que doivent satisfaire les différentielles associées à différentes classes d'algorithmes du second ordre. Ces contraintes nous permettront de démontrer des propriétés de non sensibilité aux distributions de probabilité et aux corrélations temporelles de signaux sources pour certains algorithmes. L'influence de la corrélation temporelle du bruit est examinée et les cas particuliers des algorithmes de "Toeplitzation", d'"augmentation" et de "covariance matching estimation" (COMET) sont considérés.

4.2 Modèle et notations

Considérons un réseau d'antennes composé de M capteurs recevant les signaux "bande-étroite" provenant de K sources discrètes. On suppose que ces signaux ont une même fréquence centrale f_0 et sont dans une même bande B . Pour la source k ($k = 1, \dots, K$) on note respectivement $u_{t,k}$, Θ_k et $\mathbf{e}(\Theta_k)$ l'enveloppe complexe du signal émis à l'instant t par cette source, les paramètres spatiaux inconnus à estimer et le vecteur directionnel de cette source qui modélise la réponse du réseau à une onde de paramètres spatiaux Θ_k . Le M -vecteur d'observation des enveloppes complexes des sorties des capteurs du réseau s'écrit alors :

$$\mathbf{x}_t = \sum_{k=1}^K u_{t,k} \mathbf{e}(\Theta_k) + \mathbf{v}_t,$$

où $\mathbf{v}_t$ représente le M -vecteur d'observation de l'enveloppe complexe du bruit additif à l'instant t . Habituellement on suppose que les sorties des capteurs sont échantillonnées à une fréquence suffisamment basse pour que les $(u_{t,k}, \mathbf{v}_t)_{t=1, \dots, n}$ puissent être considérés comme indépendants. Nous supposons pour notre part que la fréquence d'échantillonnage est supérieure ou égale à B de sorte que les observations $(\mathbf{x}_t)_{t=1, \dots, n}$ ne peuvent plus être supposées toujours indépendantes. Par ailleurs, le bruit additif $\mathbf{v}_t$ et les signaux des sources $u_{t,k}$ sont considérés comme des processus aléatoires stationnaires complexes circulaires centrés indépendants, non nécessairement gaussiens et de moments du quatrième ordre finis. La matrice de covariance spatiale $\mathbf{R}_x \stackrel{\text{def}}{=} E(\mathbf{x}_t \mathbf{x}_t^H)$ s'écrit :

$$\mathbf{R}_x = \mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) + \sigma_v^2 \mathbf{C},$$

avec $\mathbf{E}(\Theta) \stackrel{\text{def}}{=} [\mathbf{e}(\Theta_1), \dots, \mathbf{e}(\Theta_K)]$, $\mathbf{R}_u \stackrel{\text{def}}{=} E(\mathbf{u}_t \mathbf{u}_t^H)$ et $\mathbf{R}_v \stackrel{\text{def}}{=} E(\mathbf{v}_t \mathbf{v}_t^H) = \sigma_v^2 \mathbf{C}$ où $\mathbf{u}_t \stackrel{\text{def}}{=} (u_{t,1}, \dots, u_{t,K})^T$. Pour déterminer la distribution asymptotique de la matrice de covariance estimée $\mathbf{R}_x(n) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{t=1}^n \mathbf{x}_t \mathbf{x}_t^H$, nous considérons deux modèles de corrélation temporelle de sources :

1. Les $u_{t,k}$ sont des processus aléatoires harmoniques : $u_{t,k} = \sum_{l=1}^{L_k} a_{k,l} e^{i\alpha_{k,l}} e^{i2\pi f_{k,l}t}$ où $(f_{k,l})_{k=1, \dots, K, l=1, \dots, L_k}$ sont fixées et *distinctes* dans $] -1/2, +1/2[$, $a_{k,l}$ sont des réels fixés positifs et $\alpha_{k,l}$ sont des variables aléatoires uniformément distribuées dans $[0, 2\pi[$, indépendantes entre elles et de $\mathbf{v}_t$.
2. Les $u_{t,k}$ sont des processus ARMA complexes circulaires centrés, de puissance σ_k^2 , de densité spectrale de puissance $S_{u_k}(f)$ et de densité interspectrale de puissance la matrice $K \times K$ $\mathbf{S}_u(f)$ (ceci inclut le cas où les échantillons $\mathbf{x}_t$ sont indépendants pour lequel $\mathbf{S}_u(f) = \mathbf{R}_u$). Leur matrice

$K^2 \times K^2$ des cumulants d'ordre quatre $\mathbf{Q}_u$ est définie par

$$[\mathbf{Q}_u]_{i+K(j-1), k+K(l-1)} \stackrel{\text{def}}{=} \int_{-1/2}^{+1/2} \int_{-1/2}^{+1/2} \rho_{i,j,l,k}(f, -f', f') df df' \quad (2.1)$$

où $\rho_{k_1, k_2, k_3, k_4}(f, f', f'')$ désigne le polyspectre :

$$\rho_{k_1, k_2, k_3, k_4}(f, f', f'') = \sum_{\tau, \tau', \tau''} \text{Cum}(u_{0, k_1}, u_{\tau, k_2}^*, u_{\tau', k_3}, u_{\tau'', k_4}^*) e^{i2\pi(f\tau + f'\tau' + f''\tau'')}.$$

Pour ce qui concerne le bruit $\mathbf{v}_t$, nous utiliserons le même type d'hypothèse ARMA et le même type de notation.

Habituellement, dans le contexte de sources bande étroite, les observations $(\mathbf{x}_t)_{t=1, \dots, n}$ sont supposées gaussiennes complexes circulaires centrées et indépendantes. Dans ces conditions la matrice de covariance spatiale empirique $\mathbf{R}_x(n)$ est une statistique exhaustive sur laquelle se fondent tous les algorithmes d'estimation de DOA. Les analyses de performance asymptotique de ces algorithmes se fondent sur la loi de probabilité suivie par cette matrice qui, dans ce contexte, est une loi de Wishart complexe. Dans ce chapitre nous considérons toujours des algorithmes du second ordre mais, comme les observations ne sont plus supposées indépendantes, tous les résultats basés sur la loi de Wishart complexe ne sont plus utilisables. Il nous est donc nécessaire de déterminer la distribution asymptotique de $\mathbf{R}_x(n)$ sous nos hypothèses.

4.3 Distribution asymptotique de la matrice de covariance empirique $\mathbf{R}_x(n)$

Nous allons montrer, grâce à un théorème central limit, que la matrice $\mathbf{R}_x(n)$ vectorisée a une distribution asymptotique gaussienne complexe centrée dont la matrice de covariance dépend fortement du modèle de dépendance envisagé entre les observations. Plus précisément nous démontrons en Annexe 1, le théorème suivant :

Résultat 4.1 $\sqrt{n} (\text{Vec}(\mathbf{R}_x(n)) - \text{Vec}(\mathbf{R}_x))$ converge en loi vers une loi gaussienne complexe ¹ non circulaire centrée de matrice de covariance $\mathbf{C}_{R_x}$:

$$\sqrt{n} (\text{Vec}(\mathbf{R}_x(n)) - \text{Vec}(\mathbf{R}_x)) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_{R_x}, \mathbf{C}_{R_x} \mathbf{K}_{M^2}^c) \quad (3.1)$$

De plus

$$\mathbb{E}(\mathbf{R}_x(n)) = \mathbf{R}_x$$

et ²

$$\lim_{n \rightarrow \infty} n \text{Cov}(\text{Vec}(\mathbf{R}_x(n))) = \mathbf{C}_{R_x} \quad (3.2)$$

¹ Une variable aléatoire p -dimensionnelle complexe centrée $\mathbf{z}$ a une distribution complexe gaussienne non circulaire caractérisée par une matrice $p \times p$ définie positive $\mathbf{\Sigma}_1$ et une matrice complexe symétrique $\mathbf{\Sigma}_2$, distribution notée $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_1, \mathbf{\Sigma}_2)$, si le vecteur aléatoire réel de dimension $2p$ formé des parties réelles et imaginaires de $\mathbf{z}$ est gaussien centré, c'est à dire si pour tout p -vecteur complexe $\mathbf{w}$, la v.a. scalaire réelle $\mathbf{w}^H \mathbf{z} + (\mathbf{w}^H \mathbf{z})^H$ a une distribution gaussienne centrée de variance $2\mathbf{w}^H \mathbf{\Sigma}_1 \mathbf{w} + \mathbf{w}^H \mathbf{\Sigma}_2 \mathbf{w}^* + \mathbf{w}^T \mathbf{\Sigma}_2^* \mathbf{w}$ où $\mathbb{E}(\mathbf{z}\mathbf{z}^H) = \mathbf{\Sigma}_1$ et $\mathbb{E}(\mathbf{z}\mathbf{z}^T) = \mathbf{\Sigma}_2$.

² $\text{Cov}(\text{Vec}(\mathbf{R}_x(n)))$ dénote la matrice $\mathbb{E}(\text{Vec}(\mathbf{R}_x(n) - \mathbf{R}_x) \text{Vec}^H(\mathbf{R}_x(n) - \mathbf{R}_x))$ et puisque $\mathbb{E}(\text{Vec}(\mathbf{R}_x(n) - \mathbf{R}_x) \text{Vec}^T(\mathbf{R}_x(n) - \mathbf{R}_x)) = \mathbb{E}(\text{Vec}(\mathbf{R}_x(n) - \mathbf{R}_x) \text{Vec}^H(\mathbf{R}_x(n) - \mathbf{R}_x)) \mathbf{K}_{M^2}^c$ (où $\mathbf{K}_{M^2}^c$ désigne la matrice de commutation définie dans le Formulaire relation (0.8)) la loi asymptotique de $\text{Vec}(\mathbf{R}_x(n))$, gaussienne complexe non circulaire, est caractérisée par $\mathbf{C}_{R_x}$ uniquement.

avec :

$$\begin{aligned} \mathbf{C}_{R_x} &= (\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta)) \mathbf{C}_{R_u} (\mathbf{E}^H(\Theta) \otimes_c \mathbf{E}^H(\Theta)) + \mathbf{C}_{R_v} \\ &+ (\mathbf{E}(\Theta) \otimes_c \mathbf{I}_M) \mathbf{C}_{R_{u,v}} (\mathbf{E}^H(\Theta) \otimes_c \mathbf{I}_M) + (\mathbf{I}_M \otimes_c \mathbf{E}(\Theta)) \mathbf{C}_{R_{v,u}} (\mathbf{I}_M \otimes_c \mathbf{E}^H(\Theta)) \end{aligned} \quad (3.3)$$

où dans le cas d'un modèle de signaux source ARMA

$$\mathbf{C}_{R_u} = \int_{-1/2}^{+1/2} \mathbf{S}_u(f) \otimes_c \mathbf{S}_u(f) df + \mathbf{Q}_u \quad (3.4)$$

et dans le cas harmonique $\mathbf{C}_{R_u} = \mathbf{O}$ et avec

$$\mathbf{C}_{R_v} = \int_{-1/2}^{+1/2} \mathbf{S}_v(f) \otimes_c \mathbf{S}_v(f) df + \mathbf{Q}_v,$$

$$\mathbf{C}_{R_{u,v}} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{R}_u^{t-t'} \otimes_c \mathbf{R}_v^{t-t'} \quad \text{et} \quad \mathbf{C}_{R_{v,u}} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{R}_v^{t-t'} \otimes_c \mathbf{R}_u^{t-t'}$$

avec $\mathbf{R}_u^{t-t'} \stackrel{\text{def}}{=} \mathbf{E}(\mathbf{u}_t \mathbf{u}_{t'}^H)$ et $\mathbf{R}_v^{t-t'} \stackrel{\text{def}}{=} \mathbf{E}(\mathbf{v}_t \mathbf{v}_{t'}^H)$.

Remarques

1. Des formules simplifiées des matrices d'intercovariance $\mathbf{C}_{R_{u,v}}$ et $\mathbf{C}_{R_{v,u}}$ sont obtenues si les suites $(\mathbf{R}_u^t \otimes \mathbf{R}_v^t)_{t=\dots, -1, 0, +1, \dots}$ et $(\mathbf{R}_v^t \otimes \mathbf{R}_u^t)_{t=\dots, -1, 0, +1, \dots}$ sont absolument sommables³. Dans ce cas nous obtenons à partir de [8, A10, p.411] et du théorème de Parseval :

$$\mathbf{C}_{R_{u,v}} = \sum_{t=-\infty}^{+\infty} \mathbf{R}_u^t \otimes_c \mathbf{R}_v^t = \int_{-1/2}^{+1/2} \mathbf{S}_u(f) \otimes_c \mathbf{S}_v(f) df$$

et

$$\mathbf{C}_{R_{v,u}} = \sum_{t=-\infty}^{+\infty} \mathbf{R}_v^t \otimes_c \mathbf{R}_u^t = \int_{-1/2}^{+1/2} \mathbf{S}_v(f) \otimes_c \mathbf{S}_u(f) df.$$

2. Dans le cas où les échantillons de bruit $\mathbf{v}_t$ sont temporellement non corrélés, les matrices d'intercovariance $\mathbf{C}_{R_{u,v}}$ et $\mathbf{C}_{R_{v,u}}$ se réduisent respectivement à $\mathbf{R}_u \otimes \mathbf{R}_v$ et $\mathbf{R}_v \otimes \mathbf{R}_u$. Dans cette situation, l'information de corrélation temporelle des signaux sources disparaît dans les matrices d'intercovariance $\mathbf{C}_{R_{u,v}}$ et $\mathbf{C}_{R_{v,u}}$.
3. Dans le cas de signaux complexes circulaires gaussiens et d'observations $\mathbf{x}_t$ indépendantes, $\mathbf{Q}_u = \mathbf{O}$, $\mathbf{Q}_v = \mathbf{O}$, $\mathbf{S}_u(f) = \mathbf{R}_u$ et $\mathbf{S}_v(f) = \mathbf{R}_v$. La relation (3.3) s'écrit alors :

$$\mathbf{C}_{R_x} = (\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) + \mathbf{R}_v) \otimes_c (\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) + \mathbf{R}_v) = \mathbf{R}_x \otimes_c \mathbf{R}_x \quad (3.5)$$

ce qui constitue un résultat classique ([9, p. 336]).

³Cette condition est satisfaite si les suites $\mathbf{R}_u^t$ et $\mathbf{R}_v^t$ sont absolument sommables ou si l'une de ces suites est absolument sommable et l'autre bornée.

4. Dans le cas de signaux ARMA complexes circulaires gaussiens non nécessairement indépendants, la relation (3.3) peut s'écrire :

$$\mathbf{C}_{R_x} = \int_{-1/2}^{+1/2} \mathbf{R}(f) \otimes_c \mathbf{R}(f) df \quad (3.6)$$

où $\mathbf{R}(f) = \mathbf{E}(\Theta) \mathbf{S}_u(f) \mathbf{E}^H(\Theta) + \mathbf{S}_v(f)$ ce qui constitue une extension de (3.5).

5. Toujours dans le cas de signaux ARMA complexes gaussiens, on peut noter que si l'un des signaux source $u_{t,k}$ tend à devenir blanc dans la bande $[f_k - b, f_k + b]$ tout en conservant une puissance finie σ_k^2 , l'expression (3.3) n'est pas bornée par rapport à b car $S_{u_k}(f)$ tend vers $\frac{\sigma_k^2}{2b} (\text{rect}_b(f - f_k) + \text{rect}_b(f + f_k))$ et donc $\mathbf{C}_{R_x}$ contient le terme $\int_{-1/2}^{+1/2} S_{u_k}^2(f) df$ qui tend vers $\frac{\sigma_k^4}{2b}$. En conséquence, la covariance asymptotique de la matrice de covariance spatiale empirique est non bornée quand la largeur de bande des signaux devient très petite par rapport à la fréquence d'échantillonnage. En fait, dans cette situation, l'expression (3.3) correspondant aux signaux ARMA complexes ne converge pas vers l'expression correspondant aux signaux harmoniques car dans ce dernier cas l'enveloppe complexe $u_{t,k}$ n'est pas gaussienne.
6. On peut noter que si deux fréquences de deux signaux sources harmoniques provenant de directions différentes sont égales, $\lim_{n \rightarrow \infty} n \text{Cov}(\text{Vec}(\mathbf{R}_x(n))) \neq 0$. Par exemple, si on considère le cas particulier : $K = 2, L_1 = L_2 = 1, f_{1,1} = f_{2,1}$,

$$\lim_{n \rightarrow \infty} \text{Cov}(\text{Vec}(\mathbf{R}_x(n))) = a_{1,1}^2 a_{2,1}^2 (\mathbf{e}(\Theta_1) \mathbf{e}^H(\Theta_1) \otimes_c \mathbf{e}(\Theta_2) \mathbf{e}^H(\Theta_2) + \mathbf{e}(\Theta_2) \mathbf{e}^H(\Theta_2) \otimes_c \mathbf{e}(\Theta_1) \mathbf{e}^H(\Theta_1)).$$

Ce cas est exclu de notre modèle harmonique défini au paragraphe 4.2.

Alors que $\mathbf{R}_x(n)$ est très sensible à la distribution et au modèle de corrélation temporelle des sources, nous montrons dans la section suivante que les estimées $\Theta(n)$ n'y sont en général pas sensibles.

4.4 Distribution asymptotique des estimées de DOA

4.4.1 Approche fonctionnelle

Pour déterminer les performances asymptotiques des algorithmes de DOA fondés sur la matrice de covariance spatiale, nous adoptons une analyse fonctionnelle [10] qui consiste à considérer tout algorithme fournissant une estimée $\Theta(n)$ de Θ comme une relation fonctionnelle entre cette estimée $\Theta(n)$ et la statistique $\mathbf{R}_x(n)$ dont elle est déduite. Cette relation fonctionnelle est notée $\Theta(n) = \text{alg}(\mathbf{R}_x(n))$. Comme $\Theta = \text{alg}(\mathbf{R}_x)$, ces algorithmes $\text{alg}(\cdot)$ constituent diverses extensions de l'application $\mathbf{R}_x \rightarrow \Theta$ générées par les matrices non structurées $\mathbf{R}_x(n)$. Dans la suite nous allons considérer des algorithmes "réguliers". Plus précisément nous supposons que ces algorithmes remplissent 2 conditions :

1. La fonctionnelle $\text{alg}(\cdot)$ est différentiable dans un voisinage de $\mathbf{R}_x$, i.e. si $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}}$ ⁴ désigne la $K \times M^2$ matrice de la différentielle de cette fonctionnelle évaluée au point $\mathbf{R}_x$

$$\text{alg}(\mathbf{R}_x + \delta \mathbf{R}) = \Theta + \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \text{Vec}(\delta \mathbf{R}) + o(\delta \mathbf{R}) \quad (4.1)$$

2. Pour tout Θ et pour toute matrice définie positive $\mathbf{R}_u$ (condition 2a) ou pour tout Θ et pour toute matrice diagonale définie positive $\mathbf{R}_u$ (condition 2b) :

$$\text{alg}(\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) + \sigma_v^2 \mathbf{C}) = \Theta. \quad (4.2)$$

⁴Des expressions de $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}}$ sont données dans l'Annexe 2.

Remarques :

- La condition 1 est vérifiée par tous les algorithmes du second ordre. En effet, que ces algorithmes soient purement empiriques (issues de l'exploitation de la structure de $\mathbf{R}_x$) ou répondent à des critères particuliers (maximum de vraisemblance déterministe gaussien, maximum de vraisemblance stochastique gaussien ou minimum de variance asymptotique), ils sont toujours suffisamment “réguliers” pour être différentiables au point $\mathbf{R}_x$.
- La condition 2a est vérifiée par des algorithmes qui dans leur conception ne supposent aucune information sur la corrélation spatiale des sources. Nous appellerons ces algorithmes, “algorithmes standards” dans les paragraphes suivants. Ils considèrent donc que les sources sont spatialement corrélées. C’est en particulier le cas des algorithmes MUSIC, ESPRIT, . . .
- Au contraire, la condition 2b ne s’applique qu’aux algorithmes qui dans leur conception exigent des sources spatialement non corrélées. C’est le cas par exemple des algorithmes de Toeplitzation et d’augmentation appliqués à des réseaux linéaires ou plans que l’on examinera respectivement aux paragraphes 4.7.1 et 4.7.2.
- Certains algorithmes vérifient la condition 2a ou 2b selon la connaissance a priori sur la corrélation spatiale des sources qu’ils intègrent. Ainsi les algorithmes de type COMET [11] introduits par Ottersten *et al.* vérifient la condition 2a ou 2b selon le paramétrage choisi pour la matrice $\mathbf{R}_u$.

Les conditions précédentes impliquent des contraintes suivantes sur $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}}$:

Lemme 4.1 *Sous les conditions 1 et 2a [resp. les conditions 1 et 2b], on a les contraintes respectives suivantes sur $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}}$*

$$\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} (\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta)) = \mathbf{0} \quad (4.3)$$

$$[resp. \quad \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} (\mathbf{e}(\Theta_k) \otimes_c \mathbf{e}(\Theta_k)) = \mathbf{0}, \quad k = 1, \dots, K]. \quad (4.4)$$

Preuve : Pour les conditions 1 et 2a, nous avons :

$$\begin{aligned} \text{alg} (\mathbf{E}(\Theta) (\mathbf{R}_u + \delta \mathbf{R}_u) \mathbf{E}^H(\Theta) + \sigma_v^2 \mathbf{C}) &= \Theta \\ &= \Theta + \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \text{Vec} (\mathbf{E}(\Theta) \delta \mathbf{R}_u \mathbf{E}^H(\Theta)) + o(\delta \mathbf{R}_u) \\ &= \Theta + \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} (\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta)) \text{Vec}(\delta \mathbf{R}_u) + o(\delta \mathbf{R}_u), \end{aligned}$$

et pour les conditions 1 et 2b, la perturbation $\delta \mathbf{R}_u$ est diagonale, $\delta \mathbf{R}_u = \text{Diag}(\delta \sigma_1^2, \dots, \delta \sigma_K^2)$, ainsi $\text{Vec} (\mathbf{E}(\Theta) \delta \mathbf{R}_u \mathbf{E}^H(\Theta)) = \sum_{k=1}^K \delta \sigma_k^2 \text{Vec} (\mathbf{e}(\Theta_k) \mathbf{e}^H(\Theta_k)) = \sum_{k=1}^K \delta \sigma_k^2 (\mathbf{e}(\Theta_k) \otimes_c \mathbf{e}(\Theta_k))$. ■

Remarque : La condition 2a est plus forte que la condition 2b et par suite la contrainte (4.3) est aussi plus forte que la contrainte (4.4). Plus l’étendue de l’application d’un algorithme du second ordre est grande et plus la contrainte que doit suivre sa différentielle est forte.

4.4.2 Algorithmes standards

La condition de régularité (4.1) nous permet de déduire directement les propriétés asymptotiques de $\Theta(n)$ de celles de $\mathbf{R}_x(n)$. En utilisant un théorème classique de continuité (ex [6, théorème, p. 122]) nous avons en effet :

$$\sqrt{n} (\Theta(n) - \Theta) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_{\Theta}) \quad (4.5)$$

avec

$$\mathbf{C}_\Theta = \lim_{n \rightarrow \infty} n \mathbf{E}((\Theta(n) - \Theta)(\Theta(n) - \Theta)^T) = \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \mathbf{C}_{R_x} \left(\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \right)^H. \quad (4.6)$$

Selon qu'un algorithme satisfait les conditions 2a ou 2b du lemme 4.1, différentes propriétés de sensibilité aux distributions des signaux et à leurs corrélations temporelles peuvent être déduites. Ainsi les algorithmes que l'on appelle ici "algorithmes standards", satisfaisant les conditions 1 et 2a ne sont pas sensibles à la distribution de probabilité des signaux sources qu'elle que soit la corrélation temporelle du bruit ou des signaux sources. Ce résultat généralise un résultat présenté dans [5] dans le cadre d'observations indépendantes.

Résultat 4.2 *Pour des sources complexes gaussiennes ou non gaussiennes, ARMA ou harmoniques, la covariance asymptotique des estimées des algorithmes de DOA du second ordre qui n'exigent pas des sources spatialement non corrélées ne dépend pas de la distribution de probabilité des sources et a pour expression :*

$$\begin{aligned} \mathbf{C}_\Theta = \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \left((\mathbf{E}(\Theta) \otimes_c \mathbf{I}_M) \mathbf{C}_{R_{u,v}} (\mathbf{E}^H(\Theta) \otimes_c \mathbf{I}_M) + (\mathbf{I}_M \otimes_c \mathbf{E}(\Theta)) \mathbf{C}_{R_{v,u}} (\mathbf{I}_M \otimes_c \mathbf{E}^H(\Theta)) \right. \\ \left. + \mathbf{C}_{R_v} \right) \left(\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \right)^H. \end{aligned} \quad (4.7)$$

Preuve :

Les termes $\mathbf{C}_{R_{u,v}}$ et $\mathbf{C}_{R_{v,u}}$ apparaissant dans (3.3) ne contiennent que des termes du second ordre $\mathbf{E}(\mathbf{u}_t \mathbf{u}_{t'}^H)$ et $\mathbf{E}(\mathbf{v}_t \mathbf{v}_{t'}^H)$ et par suite ne dépendent pas de la distribution de probabilité des signaux.

- Dans le cas de signaux harmoniques, le résultat est immédiat compte tenu de (3.3) avec $\mathbf{C}_{R_u} = \mathbf{O}$.
- Dans le cas de signaux ARMA, même si les termes $\int_{-1/2}^{+1/2} \mathbf{S}_u(f) \otimes_c \mathbf{S}_u(f) df$ et $\mathbf{Q}_u$ ne sont pas nuls, ils disparaissent néanmoins lors de l'évaluation de (4.6) grâce à la contrainte (4.3) sur $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}}$. ■

En conséquence, les performances asymptotiques des algorithmes du second ordre "standards" (ceux qui n'intègrent pas d'information de non corrélation spatiale des signaux sources) dépendent en général de la corrélation temporelle des signaux source par l'intermédiaire des termes $\mathbf{C}_{R_{u,v}}$ et $\mathbf{C}_{R_{v,u}}$ qui contiennent $(\mathbf{R}_u^n)_{n=\dots, -1, 0, 1, \dots}$. Par contre, comme nous le verrons dans la suite, lorsque le bruit est temporellement non corrélé, l'expression précédente se simplifie et ne dépend plus, ni de la distribution, ni de la corrélation temporelle des sources. Ceci nous conduit à distinguer ces deux cas.

4.5 Bruit temporellement corrélé

Comme nous venons de le remarquer, la corrélation entre les échantillons du bruit peut influencer les performances des algorithmes de DOA. Afin d'évaluer cette influence, nous proposons deux approches. D'une part, en remarquant que la corrélation entre les échantillons d'un bruit de caractéristiques données dépend de la fréquence d'échantillonnage $f_e = \frac{1}{T_e}$, il semble intéressant d'exprimer la covariance asymptotique des estimées en fonction de cette fréquence d'échantillonnage. Nous verrons d'autre part comment cette covariance asymptotique dépend de la densité spectrale de puissance du bruit.

4.5.1 Covariance asymptotique des estimées et fréquence d'échantillonnage

Habituellement, les performances des algorithmes de DOA sont évaluées en fonction du nombre de "snapshots" observés sans tenir compte de la fréquence d'échantillonnage. En fait, selon la valeur de cette fréquence, les échantillons sont plus ou moins temporellement corrélés. Il s'agit d'échantillons de

l'enveloppe complexe du signal bande-étroite obtenue par filtrage passe-bande, dans une bande de largeur B autour de la fréquence de la porteuse, du signal reçu par les capteurs. Si le bruit reçu par les capteurs est blanc, le signal à temps continu $\mathbf{v}(t)$, correspondant à l'enveloppe complexe de ce bruit en sortie de l'ensemble des capteurs, est blanc dans la bande $[-\frac{B}{2}, +\frac{B}{2}]$, de densité spectrale de puissance N_0 dans cette bande et complexe circulaire. Nous supposons en outre que $\mathbf{v}(t)$ est gaussien, spatialement non corrélé. Le matrice de densité inter-spectrale de puissance de l'enveloppe complexe à temps continu des sources, noté $\mathbf{S}_u^c(f)$ a son support dans $[-\frac{B}{2}, +\frac{B}{2}]$. Après échantillonnage de l'enveloppe complexe des signaux à la fréquence $f_e = \frac{1}{T_e}$, les densités spectrales de puissance des signaux source et du bruit *échantillonnés* sont respectivement

$$\mathbf{S}_u(f) = \frac{1}{T_e} \sum_{k=-\infty}^{\infty} \mathbf{S}_u^c(f - \frac{k}{T_e})$$

$$\mathbf{S}_v(f) = \frac{1}{T_e} N_0 \sum_{k=-\infty}^{\infty} \text{Diag} \left(\mathbb{1}_{[-\frac{B}{2}, +\frac{B}{2}]}(f - \frac{k}{T_e}), \dots, \text{Diag}(\mathbb{1}_{[-\frac{B}{2}, +\frac{B}{2}]}(f - \frac{k}{T_e})) \right).$$

Selon que les signaux sont sur-échantillonnés ou sous-échantillonnés par rapport à la fréquence de Nyquist, les termes de covariance de (4.7) deviennent alors

- dans le cas de sur-échantillonnage ($f_e > B$) :

$$\begin{aligned} \mathbf{C}_{R_v} &= T_e \int_{-1/2T_e}^{+1/2T_e} \mathbf{S}_v^c(f) \otimes \mathbf{S}_v^c(f) df = \frac{N_0^2}{T_e} \text{Diag}(B, \dots, B) = \frac{\sigma_v^4}{BT_e} \mathbf{I}_{M^2} \\ \mathbf{C}_{R_{u,v}} &= T_e \int_{-1/2T_e}^{+1/2T_e} \mathbf{S}_u^c(f) \otimes \mathbf{S}_v^c(f) df = \frac{N_0}{T_e} \int_{-B/2}^{+B/2} \mathbf{S}_u^c(f) \otimes \mathbf{I}_M df = \frac{1}{BT_e} \mathbf{R}_u \otimes \sigma_v^2 \mathbf{I}_M \end{aligned}$$

où $\sigma_v^2 = N_0 B$ représente la puissance du bruit.

- dans le cas de sous-échantillonnage ($f_e < B$) :

Selon la valeur de f_e , les spectres de signaux continus se recouvrent plus ou moins pour former les spectres des signaux discrets et les valeurs des intégrales précédentes varient autour de leur valeur limite obtenue quand $T_e \rightarrow \infty$, c'est à dire les valeurs correspondant à des "snapshots" indépendants :

$$\mathbf{C}_{R_v} = \mathbf{R}_v \otimes_c \mathbf{R}_v = \sigma_v^4 \mathbf{I}_{M^2} \quad \text{et} \quad \mathbf{C}_{R_{u,v}} = \mathbf{R}_u \otimes_c \mathbf{R}_v = \mathbf{R}_u \otimes_c \sigma_v^2 \mathbf{I}_M.$$

En considérant maintenant la matrice de covariance du paramètre Θ comme une fonction de la durée d'observation $T = nT_e$, les expressions précédentes de $\mathbf{C}_{R_v}$ et de $\mathbf{C}_{R_{u,v}}$ montrent que :

- si les signaux sont sur-échantillonnés :

$$\mathbb{E}((\Theta(T) - \Theta)(\Theta(T) - \Theta)^T) \approx \frac{1}{BT} \mathbf{C}_\Theta > \frac{1}{n} \mathbf{C}_\Theta \quad \text{quand} \quad n \gg 1$$

et ceci quel que soit $f_e > B$.

- si les signaux sont sous-échantillonnés :

$$\mathbb{E}((\Theta(T) - \Theta)(\Theta(T) - \Theta)^T) \approx \frac{T_e}{T} \mathbf{C}_\Theta = \frac{1}{n} \mathbf{C}_\Theta \gg \frac{1}{BT} \mathbf{C}_\Theta \quad \text{quand} \quad n \gg 1 \quad \text{et} \quad BT_e \gg 1$$

- si les signaux sont échantillonnés à la fréquence de Nyquist ($f_e = B$) :

$\mathbf{C}_{R_v}$ et de $\mathbf{C}_{R_{u,v}}$ prennent les valeurs correspondantes au cas de “snapshots” indépendants et

$$\mathbb{E}((\Theta(T) - \Theta)(\Theta(T) - \Theta)^T) \approx \frac{1}{n} \mathbf{C}_\Theta = \frac{1}{BT} \mathbf{C}_\Theta \quad \text{quand } n \gg 1$$

où $\mathbf{C}_\Theta$ représente la matrice de covariance asymptotique du paramètre Θ sous l’hypothèse d’indépendance des “snapshots”.

Dans la plupart des applications, le paramètre primordial des performances asymptotiques est le temps d’observation $T = nT_e$ et non le nombre n de “snapshots”. Avec ce paramètre T , nous voyons que *les performances asymptotiques (i.e. pour $n \gg 1$) ne dépendent que de la durée d’observation $T = nT_e$ et non de la fréquence d’échantillonnage f_e dès que $f_e \geq B$* . De plus, du point de vue de la durée d’observation, nous n’avons jamais intérêt à sous échantillonner les signaux issus des capteurs.

4.5.2 Covariance asymptotique des estimées et densité spectrale de puissance du bruit

Nous étudions dans cette partie l’influence de la densité spectrale de puissance de bruit sur les performances des algorithmes de DOA. D’après la relation (4.7), ces performances dépendent des intégrales

$$\int_{-1/2}^{+1/2} \mathbf{S}_v(f) \otimes_c \mathbf{S}_v(f) df, \quad \int_{-1/2}^{+1/2} \mathbf{S}_u(f) \otimes_c \mathbf{S}_v(f) df \quad \text{et} \quad \int_{-1/2}^{+1/2} \mathbf{S}_v(f) \otimes_c \mathbf{S}_u(f) df.$$

contenues respectivement dans les termes $\mathbf{C}_{R_v}$, $\mathbf{C}_{R_{u,v}}$ et $\mathbf{C}_{R_{v,u}}$. A puissance de bruit σ_v^2 fixée, la première intégrale diverge lorsque la largeur de la bande de fréquence du bruit devient très petite vis à vis de la fréquence d’échantillonnage. Les deux autres, ou bien s’annulent si les spectres $\mathbf{S}_v(f)$ et $\mathbf{S}_u(f)$ ont des supports disjoints, ou bien divergent également si le support commun à ces spectres devient très étroit à puissances σ_v^2 et σ_u^2 fixées.

Ces remarques peuvent être illustrées par un “exemple d’école” très simple. Considérons en effet un réseau ne recevant qu’une seule source de densité spectrale de puissance $S_u(f)$, le bruit étant supposé gaussien complexe circulaire et spatialement non corrélé dont chaque composante a la même densité spectrale de puissance $S_v(f)$. La variance asymptotique de l’estimé de θ par l’algorithme MUSIC est donnée d’après (4.7) et d’après l’expression (4.10) de la matrice $\mathbf{D}_{\theta, \mathbf{R}_x}^{\text{MUSIC}}$ de la différentielle de l’algorithme MUSIC donnée au chapitre 5. Il est démontré en Annexe 3 :

$$\text{Var}(\theta) = \frac{1}{\alpha \sigma_u^4} \left(\int_{-1/2}^{+1/2} S_u(f) S_v(f) df + \frac{1}{\|\mathbf{e}(\theta)\|^2} \int_{-1/2}^{+1/2} S_v^2(f) df \right). \quad (5.1)$$

avec $\alpha \stackrel{\text{def}}{=} 2\mathbf{e}'^H(\theta) \mathbf{P} \mathbf{e}'(\theta)$ où $\mathbf{e}'(\theta) \stackrel{\text{def}}{=} \frac{d\mathbf{e}(\theta)}{d\theta}$ et $\mathbf{P}$ est la matrice de projection sur l’espace bruit. Cette expression généralise celle déduite de la borne de Cramer Rao évaluée sous l’hypothèse classique de “snapshots” indépendants (cf. par exemple [2] et [3]), qui représente la variance asymptotique pour l’algorithme MUSIC qui est efficace dans le cas d’une seule source :

$$\text{Var}(\theta) = \frac{1}{\alpha \sigma_u^4} \left(\sigma_u^2 \sigma_v^2 + \frac{1}{\|\mathbf{e}(\theta)\|^2} \sigma_v^4 \right). \quad (5.2)$$

Dans le cas particulier où les signaux source et les différentes composantes du bruit sont des signaux AR d’ordre 1 de coefficients respectifs a_u et a_v , dont les densités spectrales de puissance sont séparées de Δf_c , l’expression (5.1) devient (voir Annexe 4) :

$$\text{Var}(\theta) = \frac{1}{\alpha \sigma_u^4} \left(\sigma_u^2 \sigma_v^2 \frac{1 - a_u^2 a_v^2}{1 + a_u^2 a_v^2 - 2a_u a_v \cos(2\pi \Delta f)} + \frac{\sigma_v^4}{\|\mathbf{e}(\theta)\|^2} \frac{1 + a_v^2}{1 - a_v^2} \right).$$

Les figures FIG.4.1 et FIG.4.2 proposent une illustration numérique de ces remarques dans le cas de cet exemple. Elles représentent, respectivement pour un rapport S/B de 0 dB et de 10 dB, la variance asymptotique théorique et estimée de l'estimation $\theta(n)$ de la direction d'arrivée d'une source par l'algorithme MUSIC dans le cas d'un réseau linéaire de 5 capteurs en fonction du coefficient AR commun ($a_u = a_v$). Deux situations sont envisagées : les fréquences centrales des densités spectrales de puissance des signaux source et bruit sont les mêmes ($\Delta f_c = 0$) ou sont séparées de 1/2 ($\Delta f_c = 0.5$). Le nombre de snapshots est de 500 et l'estimation de la variance asymptotique est réalisée à partir de 1000 itérations et représentée avec un intervalle de confiance à 99%.

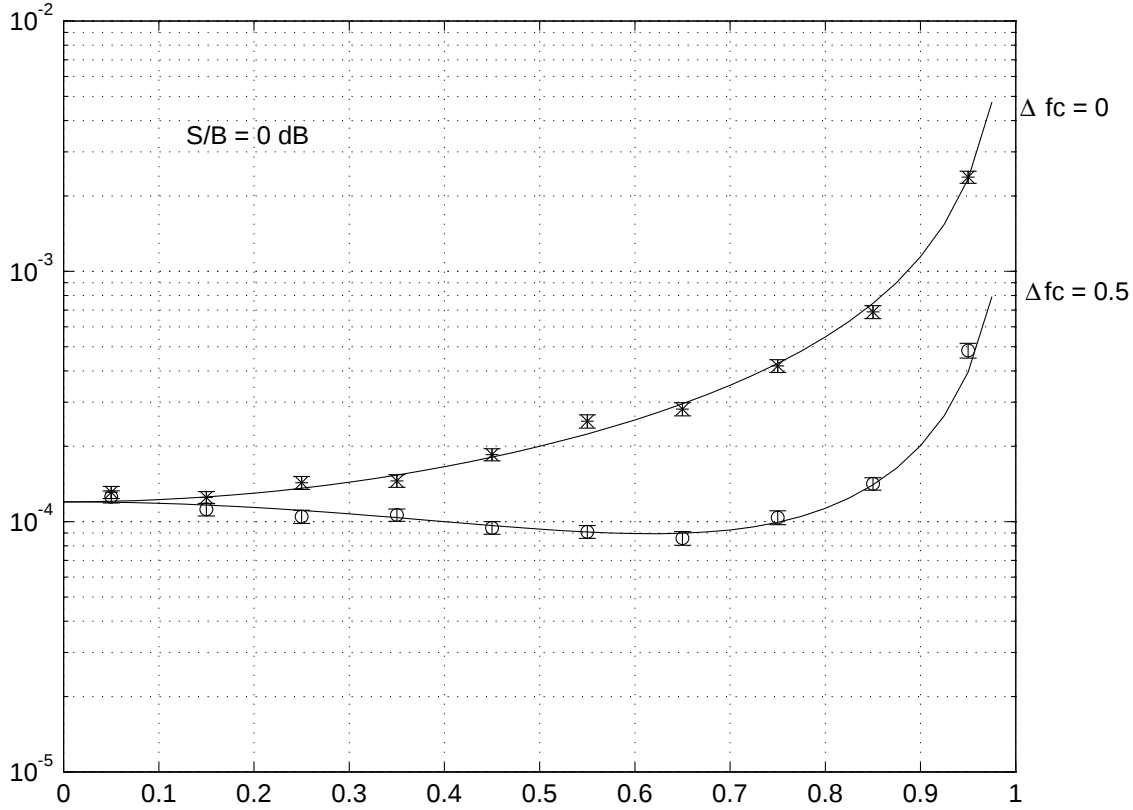


FIG. 4.1: Variance asymptotique théorique et estimée de $\theta(n)$ en fonction du coefficient AR identique du signal source et du bruit pour 2 valeurs Δf_c d'écartement des fréquences centrales de ces signaux, un rapport S/B = 0 dB et $n = 500$ snapshots.

En considérant les courbes correspondant à $\Delta f_c = 0.5$, on constate, lorsque le coefficient AR augmente progressivement que, les supports de $\mathbf{S}_u(f)$ et de $\mathbf{S}_v(f)$ tendant à se disjoindre, la variance asymptotique diminue corrélativement avec la diminution de la valeur de l'intégrale $\int_{-1/2}^{+1/2} \mathbf{S}_u(f) \otimes \mathbf{S}_v(f) df$. Ensuite, pour des valeurs importantes du coefficient AR, la variance asymptotique augmente corrélativement avec l'augmentation de la valeur de l'intégrale $\int_{-1/2}^{+1/2} \mathbf{S}_v(f)^2 df$. Au contraire, dans le cas où $\mathbf{S}_u(f)$ et $\mathbf{S}_v(f)$ ont même support ($\Delta f_c = 0$), les deux intégrales précédentes augmentent continûment avec le coefficient AR ce qui se traduit par une augmentation continue de la variance asymptotique.

Dans de nombreuses situations, même si les signaux source sont temporellement corrélés, le bruit peut néanmoins être considéré comme temporellement non corrélé. Cette situation se présente par exemple si le bruit à temps continu est blanc et si après démodulation, la fréquence d'échantillonnage est choisie égale à la largeur de bande du filtre antirepliement ($f_e = B$). Il convient donc d'étudier en détail ce cas particulier.

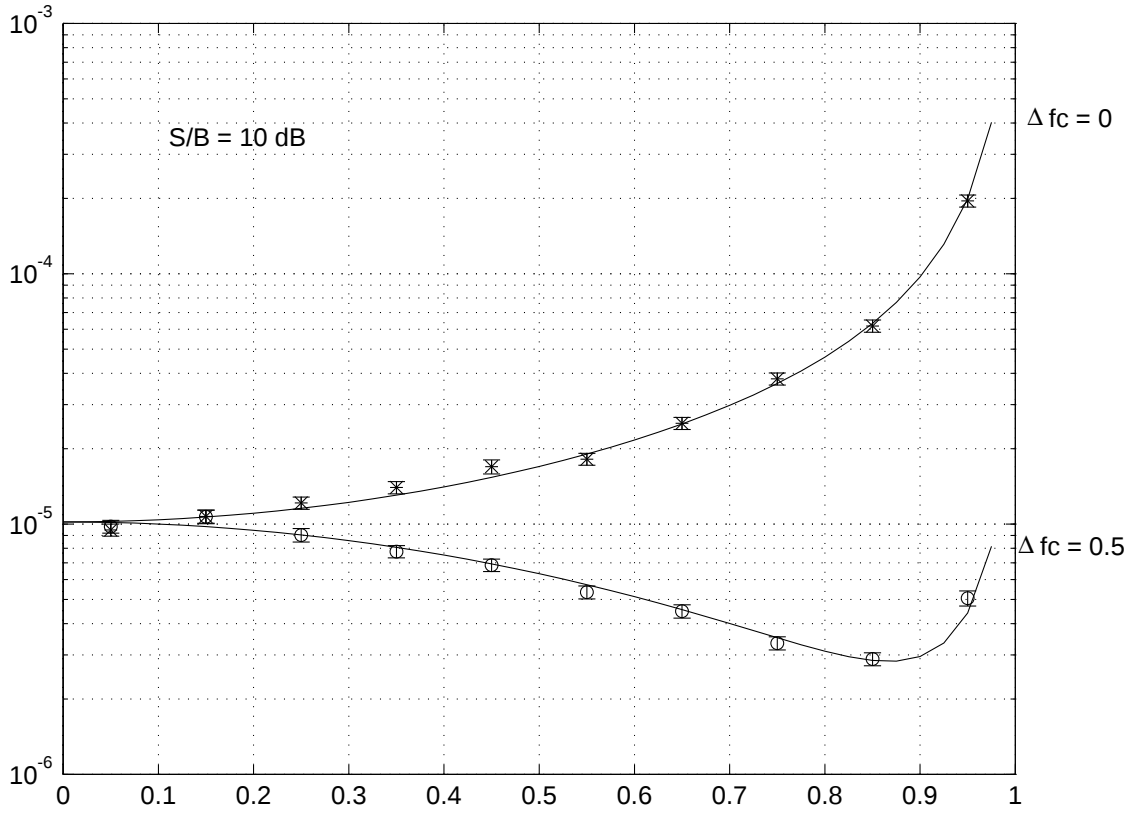


FIG. 4.2: Variance asymptotique théorique et estimée de $\theta(n)$ en fonction du coefficient AR identique du signal source et du bruit pour 2 valeurs Δf_c d'écartement des fréquences centrales de ces signaux, un rapport S/B = 10 dB et $n = 500$ snapshots.

4.6 Bruit temporellement non corrélé

4.6.1 Antenne de géométrie quelconque

Dans le cas où le bruit est temporellement non corrélé, l'expression des matrices de covariance du bruit est extrêmement simple : $\mathbf{R}_v^n = \delta_{0,n} \sigma_v^2 \mathbf{C}$. Il s'en suit une simplification de l'expression (4.7) qui nous amène au résultat suivant :

Résultat 4.3 *Dans le cas d'un bruit temporellement non corrélé et pour des sources complexes circulaires gaussiennes ou non gaussiennes, ARMA ou harmoniques, la covariance asymptotique des estimées des algorithmes de DOA du second ordre qui n'exigent pas des sources spatialement non corrélées n'est pas sensible à la distribution de probabilité et au modèle de corrélation temporelle des sources et a pour expression :*

$$\mathbf{C}_\Theta = \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} (\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) \otimes_c \sigma_v^2 \mathbf{C} + \sigma_v^2 \mathbf{C} \otimes_c \mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) + \sigma_v^4 \mathbf{C} \otimes_c \mathbf{C} + \mathbf{Q}_v) \left(\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \right)^H \quad (6.1)$$

Preuve : Si le bruit est temporellement non corrélé

$$\mathbf{C}_{R_v} = \sigma_v^4 \mathbf{C} \otimes_c \mathbf{C} + \mathbf{Q}_v, \quad \mathbf{C}_{R_{u,v}} = \mathbf{R}_u \otimes_c \sigma_v^2 \mathbf{C} \quad \text{et} \quad \mathbf{C}_{R_{v,u}} = \sigma_v^2 \mathbf{C} \otimes_c \mathbf{R}_u.$$

Le résultat s'en déduit car la propriété (0.1) du produit de Kronecker assure que :

$$\begin{aligned} (\mathbf{E}(\Theta) \otimes_c \mathbf{I}_M) \mathbf{C}_{R_{u,v}} (\mathbf{E}^H(\Theta) \otimes_c \mathbf{I}_M) &= (\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) \otimes_c \sigma_v^2 \mathbf{C}) \\ (\mathbf{E}(\Theta) \otimes_c \mathbf{I}_M) \mathbf{C}_{R_{v,u}} (\mathbf{E}^H(\Theta) \otimes_c \mathbf{I}_M) &= \sigma_v^2 \mathbf{C} \otimes_c (\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta)). \end{aligned}$$

■

Remarque : Le résultat précédent montre que, même si la covariance asymptotique de $\mathbf{R}_x(n)$ dépend de la distribution de probabilité et du modèle de corrélation temporelle des sources, les performances asymptotiques des algorithmes de DOA du second ordre qui n'exigent pas des sources spatialement non corrélées n'y sont au contraire pas sensibles.

Le résultat 4.3 nous permet d'étudier la sensibilité des algorithmes du second ordre d'estimation de DOA à diverses transformations, notées $\text{trans}(\cdot)$, de $\mathbf{R}_x(n)$ appropriées à des géométries d'antenne particulières en considérant l'application composée $\text{alg}[\text{trans}(\cdot)]$ des applications $\text{alg}(\cdot)$ et $\text{trans}(\cdot)$ comme un nouvel algorithme $\text{alg}^{\text{trans}}(\cdot)$. Ainsi dans le paragraphe suivant, nous considérerons le cas des algorithmes "standards" appliqués aux antennes centrosymétriques.

4.6.2 Antenne centrosymétrique

Les antennes centrosymétriques dont les capteurs sont symétriques par rapport à un centre géométrique (les réseaux uniformes linéaires et carrés en sont des exemples) sont souvent utilisées pour leur facilité d'implémentation. Une propriété importante et intéressante de ce type de réseau est que les vecteurs directionnels possèdent la propriété dite de centro-symétrie :

$$\mathbf{e}(\Theta_k) = \mathbf{J}\mathbf{e}^*(\Theta_k), \quad (6.2)$$

où $\mathbf{J}$ est la matrice $M \times M$ identité antidiagonale. Cette propriété entraîne que la matrice de covariance moyennée $\mathbf{R}_x^{\text{ch}} \stackrel{\text{def}}{=} \frac{1}{2}(\mathbf{R}_x + \mathbf{J}\mathbf{R}_x^*\mathbf{J})$ est centro-Hermitienne (CH). Si $\mathbf{C}$ est elle même centro-Hermitienne alors⁵ :

$$\mathbf{R}_x^{\text{ch}} = \mathbf{E}(\Theta) \left(\frac{\mathbf{R}_u + \mathbf{R}_u^*}{2} \right) \mathbf{E}^H(\Theta) + \sigma_v^2 \mathbf{C}.$$

Les algorithmes de DOA peuvent alors être appliqués à la matrice de covariance estimée et moyennée CH,

$$\mathbf{R}_x^{\text{ch}}(n) \stackrel{\text{def}}{=} \frac{1}{2}(\mathbf{R}_x(n) + \mathbf{J}\mathbf{R}_x^*(n)\mathbf{J}) = \frac{1}{2}(\mathbf{R}_x(n) + \mathbf{J}\mathbf{R}_x^T(n)\mathbf{J}).$$

La précision de l'estimation de la matrice de covariance spatiale est alors améliorée de façon significative. Par exemple, pour un réseau linéaire uniforme, un calcul simple montre que $\lim_{n \rightarrow \infty} n \|\mathbf{R}_x(n) - \mathbf{R}_x\|^2 = \text{Tr}(\mathbf{C}_{R_x})$ décroît pour une source ARMA complexe circulaire gaussienne de

$$M^2 \left(\int_{-1/2}^{+1/2} S_{u_1}^2(f) df + 2\sigma_{u_1}^2 \sigma_v^2 + \sigma_v^4 \right) \quad \dot{\rightarrow} \quad M^2 \int_{-1/2}^{+1/2} S_{u_1}^2(f) df + \frac{M(M+1)}{2} (2\sigma_{u_1}^2 \sigma_v^2 + \sigma_v^4)$$

et pour des sources harmoniques de

$$M^2 \left(2 \sum_{k=1}^K (a_k^2 \sigma_v^2 + \sigma_v^4) \right) \quad \dot{\rightarrow} \quad \frac{M(M+1)}{2} \left(2 \sum_{k=1}^K (a_k^2 \sigma_v^2 + \sigma_v^4) \right).$$

Comme $\text{Vec}(\mathbf{R}_x^{\text{ch}}(n)) = \frac{1}{2}(\mathbf{I} + (\mathbf{J} \otimes_c \mathbf{J})\mathbf{K}_M^c) \text{Vec}(\mathbf{R}_x(n))$ (grâce à la propriété (0.2) du produit de Kronecker), ce moyennage CH effectue une transformation linéaire sur $\mathbf{R}_x(n)$ grâce à l'opérateur de symétrie CH

$$\mathbf{A}_{\text{ch}} \stackrel{\text{def}}{=} \frac{1}{2}(\mathbf{I} + (\mathbf{J} \otimes_c \mathbf{J})\mathbf{K}_M^c).$$

⁵Pour des sources spatialement non corrélées $\mathbf{R}_u = \mathbf{R}_u^*$ et $\mathbf{R}_x^{\text{ch}} = \mathbf{R}_x$.

Les algorithmes de DOA $\text{alg}(\cdot)$ appliqués à $\mathbf{R}_x^{\text{ch}}(n)$ définissent une application $\text{alg}^{\text{ch}}(\cdot)$ appliquée à $\mathbf{R}_x(n)$. Comme l'opérateur de symétrie CH ne suppose aucune structure particulière pour la matrice $\mathbf{R}_u$, l'application $\text{alg}^{\text{ch}}(\cdot)$ vérifie la condition 2.a du paragraphe 4.4.1 pour tout algorithme alg "standard" et par suite le résultat 4.3 s'applique en utilisant comme matrice de la différentielle, celle de l'application $\text{alg}^{\text{ch}}(\cdot)$ évaluée au point $\mathbf{R}_x$, c'est à dire : $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}^{\text{ch}}} = \mathbf{D}_{\Theta, \mathbf{R}_x^{\text{ch}}}^{\text{alg}} \mathbf{A}_{\text{ch}}$.

Comme il est montré dans [12] que pour des sources complexes circulaires gaussiennes, temporellement non corrélées et indépendantes entre elles, la matrice de covariance $\mathbf{C}_{\Theta}$ est invariante par moyennage CH pour les algorithmes MUSIC et ESPRIT, nous avons dans ces conditions pour ces algorithmes :

$$\begin{aligned} \mathbf{C}_{\Theta} &= \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} (\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) \otimes_c \sigma_v^2 \mathbf{I}_M + \sigma_v^2 \mathbf{I}_M \otimes_c \mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) + \sigma_v^4 \mathbf{I}_{M^2}) \left(\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \right)^H \\ &= \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \mathbf{A}_{\text{ch}} (\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) \otimes_c \sigma_v^2 \mathbf{I}_M + \sigma_v^2 \mathbf{I}_M \otimes_c \mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta) + \sigma_v^4 \mathbf{I}_{M^2}) \mathbf{A}_{\text{ch}} \left(\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \right)^H. \end{aligned}$$

Par suite, comme les termes de corrélations temporelles des sources et la distribution de probabilité de ces sources n'interviennent pas dans ces expressions, l'invariante de $\mathbf{C}_{\Theta}$ par moyennage CH pour les algorithmes MUSIC et ESPRIT s'étend au cas de sources indépendantes, gaussiennes ou non gaussiennes, ARMA ou harmoniques. Ainsi dans le cas de ces algorithmes, l'amélioration de l'estimation de la matrice de covariance n'apporte aucune amélioration sur les performances asymptotiques de l'estimation des DOA sous ces conditions très générales.

4.7 Algorithmes non standards

Nous allons maintenant examiner quelques algorithmes qui utilisent explicitement l'hypothèse de non corrélation spatiale des sources. Nous porterons notre attention successivement sur des algorithmes de Toeplization, d'augmentation et de type COMET.

4.7.1 Algorithmes de Toeplization

Ces algorithmes sont utilisés dans le cadre de réseaux d'antennes réguliers M -linéaires [respectivement $M_1 \times M_2$ -rectangulaires], de sources spatialement non corrélées et d'un bruit spatialement blanc. Dans cette situation, la matrice $\mathbf{R}_x$ a une structure Toeplitz [resp., Toeplitz, bloc-Toeplitz]

$$\begin{aligned} \mathbf{R}_x &= \sum_{k=1}^K \sigma_k^2 \mathbf{e}_M(\theta_k) \mathbf{e}_M^H(\theta_k) + \sigma_v^2 \mathbf{I}_M \\ [\text{resp. } \mathbf{R}_x &= \sum_{k=1}^K \sigma_k^2 (\mathbf{e}_{M_1}(\theta_k) \otimes_r \mathbf{e}_{M_2}(\phi_k)) (\mathbf{e}_{M_1}^H(\theta_k) \otimes_r \mathbf{e}_{M_2}^H(\phi_k)) + \sigma_v^2 \mathbf{I}_M] \end{aligned}$$

avec $\mathbf{e}_M(\theta_k) \stackrel{\text{def}}{=} (1, e^{i\theta_k}, e^{2i\theta_k}, \dots, e^{(M-1)i\theta_k})^T$. Le moyennage le long des diagonales [resp. le long des blocs diagonaux] de $\mathbf{R}_x(n)$ améliore sensiblement la précision de l'estimation de la matrice de covariance spatiale. Par exemple, un calcul simple montre que $\lim_{n \rightarrow \infty} n \|\mathbf{R}_x(n) - \mathbf{R}_x\|^2 = \text{Tr}(\mathbf{C}_{R_x})$ décroît pour un source ARMA complexe circulaire gaussienne de

$$M^2 \left(\int_{-1/2}^{+1/2} S_{u_1}^2(f) df + 2\sigma_{u_1}^2 \sigma_v^2 + \sigma_v^4 \right) \quad \text{à} \quad M^2 \int_{-1/2}^{+1/2} S_{u_1}^2(f) df + (2M-1)(2\sigma_{u_1}^2 \sigma_v^2 + \sigma_v^4)$$

et pour des sources harmoniques de

$$M^2 \left(2 \sum_{k=1}^K (a_k^2 \sigma_v^2 + \sigma_v^4) \right) \quad \text{à} \quad (2M-1) \left(2 \sum_{k=1}^K (a_k^2 \sigma_v^2 + \sigma_v^4) \right).$$

Soit $\mathbf{R}_x^{\text{to}}(n)$ la matrice “Toeplitzée” ainsi obtenue. La “Toeplitzation” effectue une transformation linéaire sur $\mathbf{R}_x(n)$ par l’intermédiaire d’une matrice de projection $\mathbf{A}_{\text{to}}$ telle que $\text{Vec}(\mathbf{R}_x^{\text{to}}(n)) = \mathbf{A}_{\text{to}} \text{Vec}(\mathbf{R}_x(n))$. Cette transformation linéaire permet d’appliquer le résultat 4.1 à $\mathbf{R}_x^{\text{to}}(n)$ avec comme matrice de covariance asymptotique $\mathbf{C}_{R_x}^{\text{to}} = \mathbf{A}_{\text{to}} \mathbf{C}_{R_x} \mathbf{A}_{\text{to}}$. Compte tenu de la condition de régularité (4.1), les estimées de DOA sont asymptotiquement normales, de matrice de covariance asymptotique

$$\mathbf{C}_{\Theta}^{\text{to}} = \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \mathbf{A}_{\text{to}} \mathbf{C}_{R_x} \mathbf{A}_{\text{to}} \left(\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \right)^H. \quad (7.1)$$

Les algorithmes de DOA $\text{alg}(\cdot)$ appliqués à $\mathbf{R}_x^{\text{to}}(n)$ définissent comme précédemment une application $\text{alg}^{\text{to}}(\cdot)$ appliquée à $\mathbf{R}_x(n)$. Cette application vérifie bien la condition 2b du lemme 4.1, par contre elle ne vérifie pas en général la condition 2a et ceci quel que soit l’algorithme de DOA $\text{alg}(\cdot)$ appliqué à $\mathbf{R}_x^{\text{to}}(n)$ qu’il soit “standard” ou non. Le résultat 4.3 ne peut donc plus s’appliquer et les algorithmes de DOA basés sur la matrice de covariance “Toeplitzée” sont a priori sensibles à la distribution de probabilité et à la forme du spectre des sources dans le cas de bruit temporellement non corrélé. Plus précisément, nous allons démontrer les résultats suivants :

Résultat 4.4 *La “Toeplitzation” n’est pas sensible à la distribution de probabilité des sources ARMA si ces sources sont non seulement spatialement décorrélées mais aussi indépendantes.*

Preuve : Si les sources sont indépendantes, les seuls termes non nuls de la matrice des cumulants du 4ème ordre $\mathbf{Q}_u$ (2.1) sont les termes

$$[\mathbf{Q}_u]_{K(k-1)+k, K(k-1)+k} \stackrel{\text{def}}{=} c_k = \int_{-1/2}^{+1/2} \int_{-1/2}^{+1/2} \rho_{k,k,k,k}(f, f', -f') df df'.$$

Ainsi $\mathbf{Q}$ devient :

$$\mathbf{Q}_u = \sum_{k=1}^K c_k (\mathbf{e}_{K,k} \mathbf{e}_{K,k}^T) \otimes_c (\mathbf{e}_{K,k} \mathbf{e}_{K,k}^T) = \sum_{k=1}^K c_k (\mathbf{e}_{K,k} \otimes_c \mathbf{e}_{K,k}) (\mathbf{e}_{K,k}^T \otimes_c \mathbf{e}_{K,k}^T)$$

où $\mathbf{e}_{K,k}$ est le k ième vecteur unité de $\mathbb{C}^K$, et le terme correspondant aux cumulants d’ordre quatre de (3.3) s’écrit alors sous la forme :

$$\sum_{k=1}^K c_k (\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta)) (\mathbf{e}_{K,k} \otimes_c \mathbf{e}_{K,k}) (\mathbf{e}_{K,k}^T \otimes_c \mathbf{e}_{K,k}^T) (\mathbf{E}^H(\Theta) \otimes_c \mathbf{E}^H(\Theta))$$

qui, en utilisant la propriété (0.1) du produit de Kronecker, se réduit à $\sum_{k=1}^K c_k (\mathbf{e}_M(\theta_k) \otimes_c \mathbf{e}_M(\theta_k)) (\mathbf{e}_M^H(\theta_k) \otimes_c \mathbf{e}_M^H(\theta_k))$. De plus :

$$\begin{aligned} \mathbf{A}_{\text{to}} (\mathbf{e}_M(\theta_k) \otimes_c \mathbf{e}_M(\theta_k)) &= \mathbf{A}_{\text{to}} \text{Vec} (\mathbf{e}_M(\theta_k) \mathbf{e}_M^H(\theta_k)) = \text{Vec} (\mathbf{e}_M(\theta_k) \mathbf{e}_M^H(\theta_k)) \\ &= \mathbf{e}_M(\theta_k) \otimes_c \mathbf{e}_M(\theta_k) \end{aligned}$$

où la structure de Toeplitz est utilisée dans la deuxième égalité. Les algorithmes de DOA $\text{alg}(\cdot)$ appliqués à $\mathbf{R}_x^{\text{to}}(n)$ définissent alors de nouveaux algorithmes $\text{alg}^{\text{to}}(\cdot)$ qui satisfont la condition 2b du lemme 4.1. La contribution de $\mathbf{Q}_u$ dans $\mathbf{C}_{\Theta}^{\text{to}}$ disparaît alors grâce à la contrainte (4.4), ce qui établit le résultat. ■

Résultat 4.5 *Dans le cas d’une seule source, la “Toeplitzation” n’est pas sensible à la corrélation temporelle de la source.*

Preuve : Dans le cas d’une seule source en effet, la condition 2a du lemme 4.1 se ramène à la condition 2b. L’expression de $\mathbf{C}_{\Theta}^{\text{to}}$ dans le cas d’une source ARMA ou harmonique est alors :

$$\mathbf{C}_{\Theta}^{\text{to}} = \mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \mathbf{A}_{\text{to}} (\sigma_1^2 \mathbf{e}_M(\theta_1) \mathbf{e}_M^H(\theta_1) \otimes_c \sigma_v^2 \mathbf{I}_M + \sigma_v^2 \mathbf{I}_M \otimes_c \sigma_1^2 \mathbf{e}_M(\theta_1) \mathbf{e}_M^H(\theta_1) + \sigma_v^4 \mathbf{I}_{M^2}) \mathbf{A}_{\text{to}} \left(\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}} \right)^H.$$

■

Résultat 4.6 *Dans le cas de plusieurs sources, la “Toeplitzation” est sensible à la corrélation temporelle des sources.*

Preuve : Comme $\mathbf{e}_M(\theta_k)\mathbf{e}_M^H(\theta_l)$, pour $k \neq l$, n’a pas une structure Toeplitz, l’espace engendré par les vecteurs colonnes de $\mathbf{A}_{\text{to}}(\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta))$ n’est pas inclus dans celui engendré par les colonnes de $\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta)$. La condition 2a du lemme 4.1 n’est en général pas remplie et les termes correspondants dans (7.1) ne s’annulent pas. De ce fait, le terme

$$\mathbf{A}_{\text{to}} \left(\sum_{1 \leq l \neq k \leq K} \left(\int_{-1/2}^{+1/2} S_{u_k}(f) S_{u_l}(f) df \right) (\mathbf{e}_M(\theta_k) \otimes_c \mathbf{e}_M(\theta_l)) (\mathbf{e}_M^H(\theta_k) \otimes_c \mathbf{e}_M^H(\theta_l)) \right) \mathbf{A}_{\text{to}} \quad (7.2)$$

qui apparaît dans le cas de sources ARMA, ne s’annule pas dans l’expression de $\mathbf{C}_{\Theta}^{\text{to}}$. ■

Remarques :

- Lorsque les spectres des sources ARMA tendent à se disjoindre, les performances des algorithmes de Toeplitzation coïncident avec celles obtenues dans le cas de sources harmoniques et sont donc insensibles à la corrélation temporelle des sources. Par ailleurs, pour de faibles rapports S/B, $\mathbf{C}_{R_x}$ peut être approximé par $\sigma_v^4 \mathbf{I}$ et cette fois l’opération de “Toeplitzation” elle même devient insensible à la corrélation temporelle des sources.
- De plus, les expressions analytiques données pour 2 sources proches et pour de faibles rapports S/B dans [13, rel. 9.118 et 9.119] demeurent valides pour des sources colorées. Pour des rapports S/B élevés, le terme (7.2) devient dominant dans $\mathbf{C}_{R_x}$ et la “Toeplitzation” devient très sensible à la corrélation temporelle des sources, ce qui sera confirmé par les simulations du paragraphe 4.8.
- Les expressions analytiques de la covariance asymptotique $\text{Cov}(\theta_k(n), \theta_l(n))$ des DOA estimées par les algorithmes Min Norm et MUSIC données dans le cas de sources gaussiennes, blanches et spatialement décorréliées dans [13, rel. 9.101 et 9.103] restent valables pour des sources ARMA, non gaussiennes et indépendantes à condition de remplacer $\text{Tr}(\mathbf{R}_x \tilde{\mathbf{B}}_k \mathbf{R}_x \tilde{\mathbf{B}}_l)$ par $\int_{-1/2}^{+1/2} \text{Tr}(\mathbf{R}(f) \tilde{\mathbf{B}}_k \mathbf{R}(f) \tilde{\mathbf{B}}_l) df$ car $\mathbf{R}_x \otimes_c \mathbf{R}_x$ est maintenant remplacé par $\int_{-1/2}^{+1/2} \mathbf{R}(f) \otimes_c \mathbf{R}(f) df$ d’après la remarque 4 (3.6).

Nous venons de montrer que les algorithmes de DOA fondés sur la matrice de covariance “Toeplitzée” peuvent être sensibles à la distribution de probabilité et à la forme du spectre des sources dans le cas de bruit temporellement non corrélé. Signalons que dans le cas de sources gaussiennes complexes circulaires non corrélées spatialement et d’observations indépendantes, l’amélioration de l’estimation de la matrice de covariance apportée par l’opérateur de Toeplitzation n’améliore pas la variance des DOA estimées. Cette propriété a été mise en évidence pour les algorithmes MUSIC et MIN NORM dans [13, chap. 9] et [14]. Par contre il est démontré également dans [13, chap. 9] et [14] que, dans le cas de deux sources, sous ces hypothèses, cet opérateur de Toeplitzation améliore la résolution de ces algorithmes. Ces résultats en apparence contradictoires peuvent s’interpréter à l’aide des termes d’intercovariance de $\mathbf{C}_{\Theta}$. En effet, considérons par exemple le cas de deux sources de même puissance et de paramètres spatiaux θ_1 et θ_2 . Si on compare les coefficients de corrélation asymptotiques des estimées $\theta_1(n)$ et $\theta_2(n)$ de θ_1 et θ_2 dans un domaine de valeurs du rapport S/B, du nombre d’observations ou de l’écart spatial entre sources, où les deux directions ne sont pas toujours résolues, on constate que, sans Toeplitzation, ce coefficient est voisin de 0 alors qu’il est voisin de 1 avec Toeplitzation. En d’autres termes, les variables aléatoires $d_1 \stackrel{\text{def}}{=} \theta_1(n) - \theta_1$ et $d_2 \stackrel{\text{def}}{=} \theta_2(n) - \theta_2$, asymptotiquement gaussiennes centrées, sont “pratiquement” indépendantes sans Toeplitzation ou “pratiquement” égales avec Toeplitzation dans le domaine considéré. En conséquence, si $d \stackrel{\text{def}}{=} \theta_2(n) - \theta_1(n)$ est la variable aléatoire représentant l’écart entre les estimées,

$d = d_2 - d_1 + \theta_2 - \theta_1$ et $E(d) = \theta_2 - \theta_1$ dans les deux cas. Cependant, dans le cas “sans Toeplization”, d est asymptotiquement gaussienne, de moyenne $\theta_2 - \theta_1$, de variance $2\text{Var}(\theta_1(n))$ et la probabilité que d prenne, dans le domaine considéré, de très petites valeurs n’est pas négligeable. Par contre, dans le cas “avec Toeplization”, d est “pratiquement” constante, égale à $\theta_2 - \theta_1$, ce qui améliore la résolution de l’algorithme dans ce cas.

Les FIG.4.3 et FIG.4.4 illustrent cette interprétation dans le cas de l’estimation, par l’algorithme MUSIC, des directions d’arrivée de deux sources de même puissance à l’aide d’un réseau linéaire uniforme de 10 capteurs et de 500 observations avec un rapport S/B = 15 dB.

La FIG.4.3 présente la probabilité de résolution estimée des deux sources, en utilisant un critère de type Cox ⁶ [15], en fonction de l’écart entre les deux directions exprimé en degré. Pour ce critère, le domaine dans lequel la Toeplization améliore la résolution est l’intervalle $[2^\circ, 5^\circ]$. La FIG.4.4 présente, dans ce domaine, les valeurs des coefficients de corrélation des estimées des directions avec ou sans Toeplization.

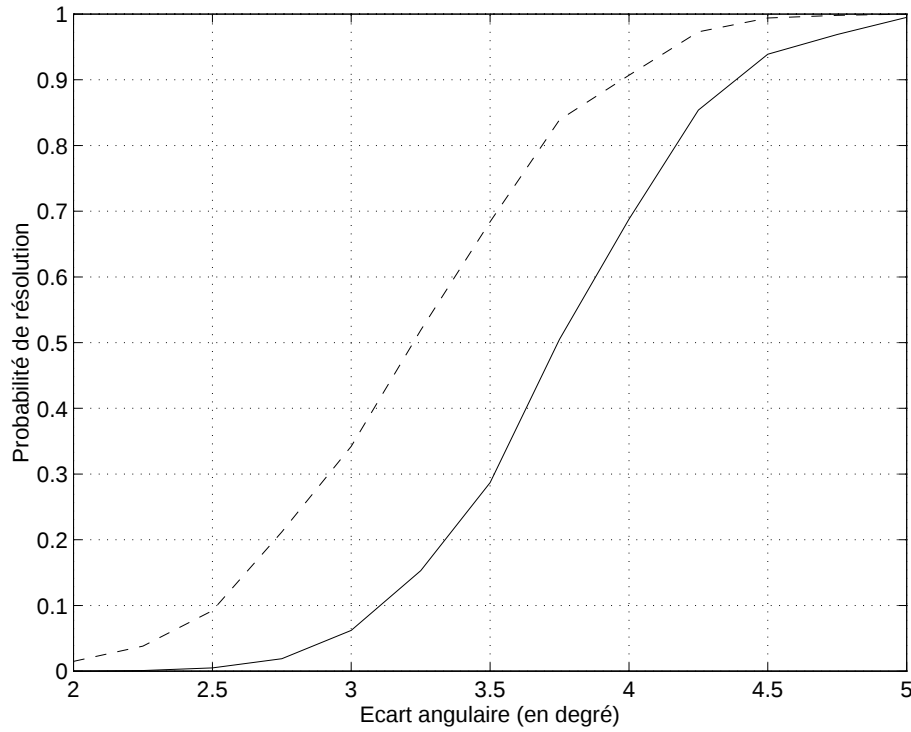


FIG. 4.3: Probabilité de résolution estimée des deux sources (critère de Cox) en fonction de l’écart en degré des DOA, sans (-) et avec (- -) Toeplization.

⁶Si $g(\theta)$ désigne la fonction de localisation de l’algorithme, dont les minima fournissent les estimées des paramètres, les deux sources seront réputées résolues si $g(\frac{\theta_1 + \theta_2}{2}) > \frac{g(\theta_1) + g(\theta_2)}{2}$.

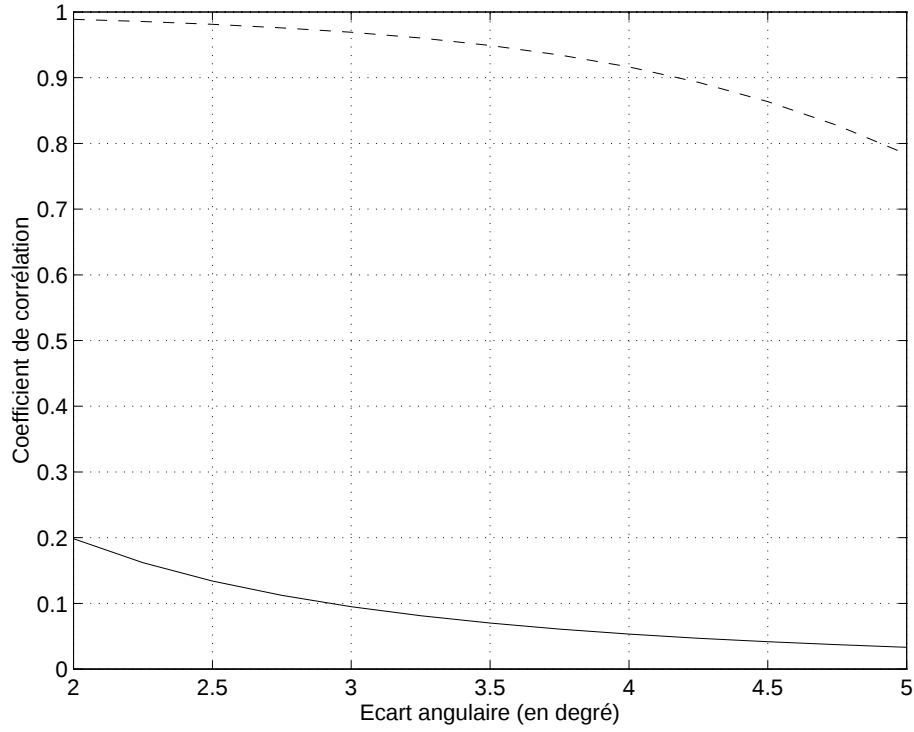


FIG. 4.4: Coefficients de corrélation des estimées des DOA sans (-) et avec (- -) Toeplitzation en fonction de l'écart en degré des DOA.

4.7.2 Algorithmes d'augmentation

Ces algorithmes sont utilisés dans le cadre de réseaux d'antennes réguliers M -linéaires [respectivement $M_1 \times M_2$ -rectangulaires] dans lesquels certains capteurs sont absents, de sources spatialement non corrélées et d'un bruit spatialement blanc. Ces réseaux lacunaires "augmentables", linéaires ou plans, ont suscité un intérêt considérable car ils permettent d'améliorer significativement les performances des algorithmes de DOA dans le cas de sources blanches, spatialement décorréées en utilisant des techniques dites "d'augmentation". Pour étudier la sensibilité de ces techniques à la distribution de probabilité et à la corrélation temporelle des sources nous ne considérerons que la méthode classique d'augmentation directe [16]. Pour fixer la notation nous considérerons les capteurs répartis sur les sommets d'un réseau cartésien plan, d'unité de longueur égale à la demi longueur d'onde des signaux incidents. La position de chaque capteur sera repérée par le couple de ses coordonnées entières (x, y) . Désignons par Γ la fonction indicatrice du réseau d'antennes :

$$\Gamma(x, y) = \begin{cases} 1 & \text{si il y a un capteur en } (x, y) \\ 0 & \text{sinon.} \end{cases}$$

Soit $\Lambda(d_x, d_y)$ la fonction d'autocorrélation de la fonction indicatrice du réseau Γ ($\Lambda(d_x, d_y)$ représente le nombre de fois où l'écart (d_x, d_y) apparaît dans le réseau). Soient $\mathbf{R}_x(n)$ et $\mathbf{R}_x^{\text{au}}(n)$ les matrices de covariance spatiale respectives associées au réseau uniforme fictif de dimension $M'_1 \times M'_2$ et au réseau augmenté de dimension $M_1 \times M_2$ ⁷. La technique d'augmentation directe [16] effectue une transformation linéaire sur $\mathbf{R}_x(n)$ ⁸ par l'intermédiaire d'un opérateur d'augmentation auquel est associé une matrice

⁷ $M_1 = M'_1$ et $M_2 = M'_2$ pour une redondance restreinte et $M_1 \leq M'_1$ et $M_2 \leq M'_2$ pour une redondance non restreinte [16].

⁸ Notons que cette transformation linéaire n'est définie que pour les besoins de l'analyse puisque la matrice $\mathbf{R}_x(n)$ n'est pas observée.

$\mathbf{A}_{\text{au}}$ de dimension $M_1^2 M_2^2 \times M_1'^2 M_2'^2$: $(\text{Vec}(\mathbf{R}_x^{\text{au}}(n)) = \mathbf{A}_{\text{au}} \text{Vec}(\mathbf{R}_x(n)))$.

$$[\mathbf{A}_{\text{au}}]_{(h-1)M_1^2 M_2^2 + (j-1)M_1 M_2 + (g-1)M_1 + i, (f-1)M_1'^2 M_2'^2 + (l-1)M_1' M_2' + (e-1)M_1' + k} = \begin{cases} \frac{\Gamma(k-1, e-1) \Gamma(l-1, f-1)}{\Lambda(l-k, f-e)} & \text{si } \begin{cases} l-k = j-i \\ f-e = h-g \end{cases} \\ 0 & \text{sinon.} \end{cases} \quad (7.3)$$

Remarque : Dans le cas des réseaux linéaires, $e = f = g = h = 1$ et (7.3) s'écrit :

$$[\mathbf{A}_{\text{au}}]_{(j-1)M+i, (l-1)M'+k} = \begin{cases} \frac{\Gamma(k-1) \Gamma(l-1)}{\Lambda(l-k)} & \text{si } l-k = j-i \\ 0 & \text{sinon} \end{cases}$$

Cette transformation linéaire permet d'appliquer le résultat 4.1 à $\mathbf{R}_x^{\text{au}}(n)$ avec la matrice de covariance asymptotique $\mathbf{C}_{R_x}^{\text{au}} = \mathbf{A}_{\text{au}} \mathbf{C}_{R_x} \mathbf{A}_{\text{au}}^T$. De par la condition de régularité (4.1) du paragraphe 4.4.1 les DOA estimées sont asymptotiquement normales, de covariance asymptotique:

$$\mathbf{C}_{\Theta}^{\text{au}} = \mathbf{D}_{\Theta, R_x}^{\text{alg}} \mathbf{A}_{\text{au}} \mathbf{C}_{R_x} \mathbf{A}_{\text{au}}^T \left(\mathbf{D}_{\Theta, R_x}^{\text{alg}} \right)^H. \quad (7.4)$$

Les algorithmes de DOA $\text{alg}(\cdot)$ appliqués à $\mathbf{R}_x^{\text{au}}(n)$ définissent comme pour l'opérateur de Toeplization, une application $\text{alg}^{\text{au}}(\cdot)$ appliquée à $\mathbf{R}_x(n)$. Cette application vérifie aussi la condition 2b du lemme 4.1, par contre elle ne vérifie pas aussi en général la condition 2a et ceci quel que soit l'algorithme de DOA $\text{alg}(\cdot)$ appliqué à $\mathbf{R}_x^{\text{au}}(n)$ qu'il soit "standard" ou non. Le résultat 4.3 ne peut donc plus s'appliquer et les algorithmes de DOA basés sur la matrice de covariance augmentée sont aussi a priori sensibles à la distribution et à la corrélation temporelle des sources. Mais puisque la propriété clé

$$\mathbf{A}_{\text{au}} (\mathbf{e}_M(\theta_k) \otimes_c \mathbf{e}_M(\theta_k)) = \mathbf{e}_M(\theta_k) \otimes_c \mathbf{e}_M(\theta_k)$$

est aussi satisfaite pour l'opérateur d'augmentation, les résultats 4.4, 4.5 et 4.6 sont aussi satisfaits ici.

4.7.3 L'algorithme COMET en DOA bande étroite

Comme nous l'avons signalé dans la dernière remarque du paragraphe 4.4.1, les algorithmes de type COMET bande étroite introduits par Ottersten *et al.* [11] vérifient la condition 2a ou 2b selon le paramétrage choisi pour la matrice $\mathbf{R}_u$. Nous ne considérerons ici et dans le Chapitre 6, où cet algorithme sera étudié en détail, que le cas où les sources sont spatialement non corrélées. Dans cette situation, seule la condition 2b du lemme 4.1 est satisfaite. Cette condition entraîne des résultats de sensibilité à la distribution de probabilité et à la corrélation temporelle des sources, quand le bruit est temporellement non corrélé, très proches de celles des algorithmes de Toeplization. Nous démontrons en particulier les résultats suivants :

Résultat 4.7 *Dans le cas de sources complexes circulaires indépendantes, l'algorithme COMET bande étroite qui suppose les sources spatialement non corrélées n'est pas sensible à la distribution des sources.*

Preuve : Si les sources sont indépendantes, avec le même raisonnement utilisé pour les algorithmes de Toeplization, le terme des cumulants du 4ème ordre de la matrice $\mathbf{C}_{\Theta}$ (3.3) se réduit à :

$$\sum_{k=1}^K c_k (\mathbf{e}_M(\theta_k) \otimes_c \mathbf{e}_M(\theta_k)) (\mathbf{e}_M^H(\theta_k) \otimes_c \mathbf{e}_M^H(\theta_k)).$$

Et grâce à (4.4), la contribution de $\mathbf{Q}_u$ dans $\mathbf{C}_{\Theta}$ disparaît alors, ce qui établit le résultat. ■

En ce qui concerne la sensibilité de COMET vis à vis de la corrélation temporelle des sources nous avons la propriété suivante :

Résultat 4.8 *L'algorithme COMET bande étroite qui suppose les sources spatialement non corrélées n'est pas sensible à la corrélation temporelle des sources dans le cas d'une seule source et dans le cas de plusieurs sources à spectres disjoints. Cette propriété d'insensibilité est perdue dans le cas de spectres non disjoints.*

Preuve : Dans le cas d'une seule source, la condition 2a du lemme 4.1 se ramène à la condition 2b, ce qui prouve le résultat dans ce cas. Dans le cas de plusieurs sources spatialement non corrélées, la matrice de densité interspectrale de puissance $\mathbf{S}_u(f)$ est diagonale et en conséquence le terme correspondant de $\mathbf{C}_{R_x}$ (3.3) s'écrit sous la forme :

$$\sum_{1 \leq l \neq k \leq K} \left(\int_{-1/2}^{+1/2} S_{u_k}(f) S_{u_l}(f) df \right) (\mathbf{e}_M(\theta_k) \otimes_c \mathbf{e}_M(\theta_l)) (\mathbf{e}_M^H(\theta_k) \otimes_c \mathbf{e}_M^H(\theta_l)).$$

Grâce à (4.4), cette contribution disparaît dans le cas de sources de spectres disjoints. ■

4.8 Simulations

Nous considérons, dans tout ce paragraphe de simulations, le cas de deux sources de même puissance σ^2 . Le bruit est supposé gaussien, spatialement et temporellement non corrélé et le rapport S/B est défini par $\frac{2\sigma^2}{\sigma_v^2}$. Les DOA sont estimées par l'algorithme MUSIC et le nombre des réalisations de Monte Carlo est de 500. Ces sources proviennent de directions $\theta_k = \pi \sin \theta'_k$ avec $\theta'_1 = 30^\circ$, $\theta'_2 = 20^\circ$ pour les réseaux linéaires et de directions $\theta_k = \pi \sin \theta'_k \sin \phi'_k$, $\phi_k = \pi \cos \theta'_k \sin \phi'_k$ avec $\theta'_1 = 30^\circ$, $\theta'_2 = 20^\circ$ et $\phi'_1 = 10^\circ$, $\phi'_2 = 40^\circ$ pour les réseaux plans. Les sources ARMA sont générées par un filtre de Butterworth (10,10) excité par un signal blanc gaussien.

La première simulation présente le cas de 2 signaux harmoniques $u_{t,k} = a_{k,1} e^{i\alpha_{k,1}} e^{i2\pi f_{k,1}t}$, $k = 1, 2$, avec $f_{1,1} = 0.25$ fixé et $f_{2,1}$ variable, observés par un réseau linéaire uniforme de 10 capteurs. La FIG.4.5 représente l'EQM théorique sur $\theta_k : \frac{1}{n}[\mathbf{C}_\Theta]_{k,k}$ ainsi que les EQM estimées : $E\|\theta_k(n) - \theta_k\|_{\text{Fro}}^2$ en fonction du nombre d'observations, sans Toeplitzation, pour différentes valeurs de l'écart entre les fréquences temporelles.

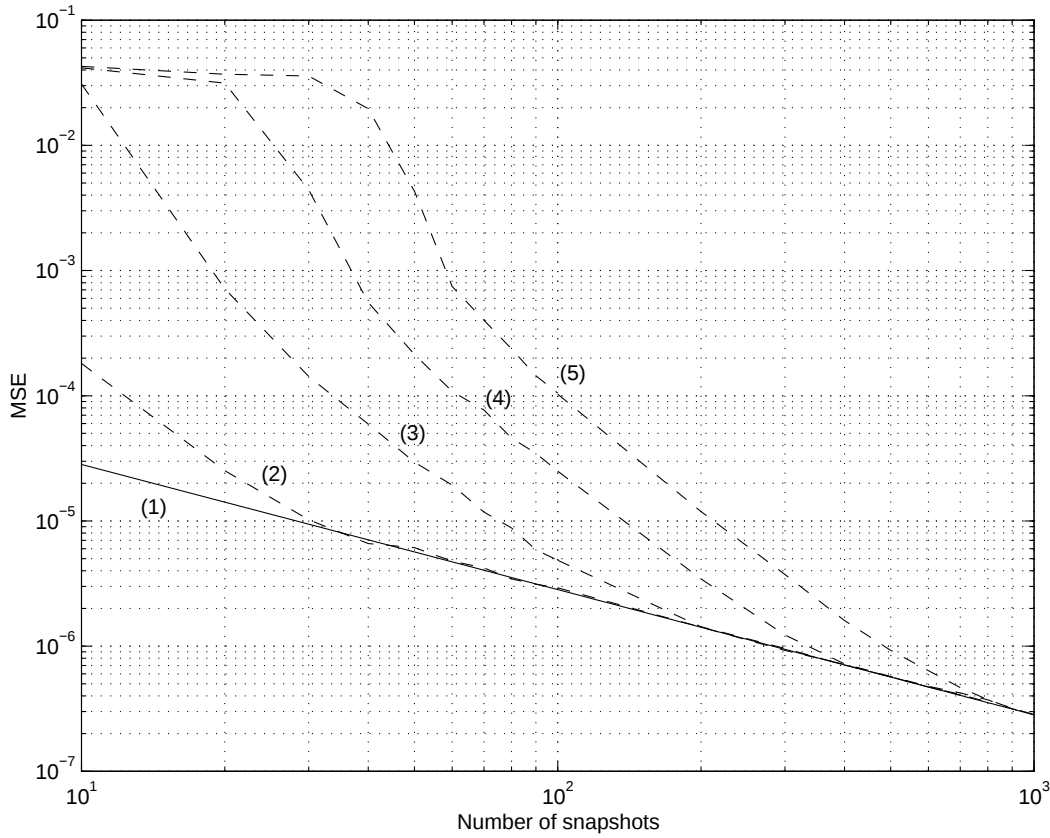


FIG. 4.5: EQM de $\theta_k(n)$, théorique (1) et estimée en fonction du nombre d'observations n pour $SNR = 20dB$, pour 2 signaux harmoniques, avec respectivement $f_{2,1} - f_{1,1} = 0.025$ (2), 0.005 (3), 0.002 (4) et 0.001 (5) sans Toeplitzation.

Nous pouvons observer que les EQM estimées correspondent bien aux EQM théoriques asymptotiques, mais avec des domaines de validité qui se réduisent lorsque les 2 fréquences temporelles se rapprochent ($n \geq 30$ pour $f_{2,1} - f_{1,1} = 0.025$ mais $n \geq 1000$ pour $f_{2,1} - f_{1,1} = 0.001$). Ceci est en accord avec la condition de validité de (3.3) qui impose $f_{2,1} \neq f_{1,1}$ comme cela a été remarqué à la fin du paragraphe 4.3.

La seconde simulation présente le cas de 2 signaux qui sont, soit temporellement blancs gaussiens spatialement décorrélés, soit ARMA gaussiens, soit harmoniques. Dans ces 2 derniers cas, les fréquences centrales des signaux ARMA ou les fréquences des sinusôides sont égales à -0.25 et 0.25. La FIG.4.6 représente les EQM théoriques sur θ_k : $\frac{1}{n}[\mathbf{C}_\Theta]_{k,k}$ ainsi que les EQM estimées : $E\|\theta_k(n) - \theta_k\|_{\text{Fro}}^2$ en fonction du rapport S/B pour un réseau linéaire uniforme de 10 capteurs après Toeplitzation (la bande passante des signaux ARMA est de 0.5).

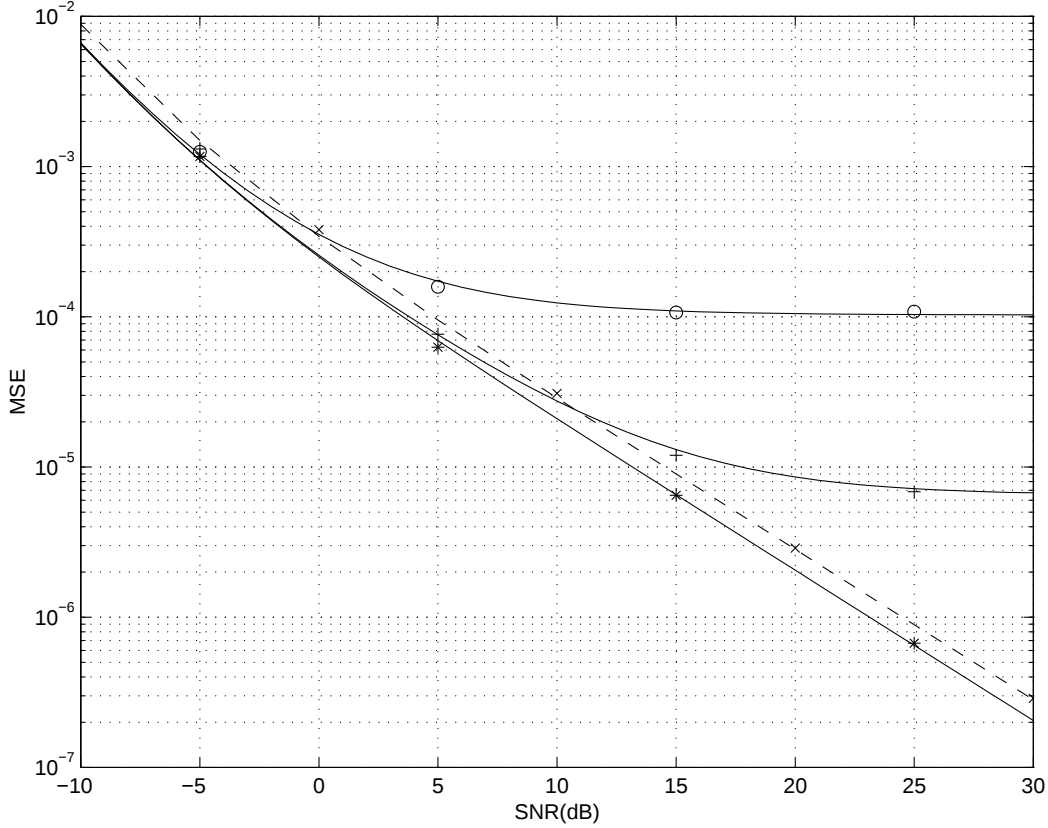


FIG. 4.6: MSE théorique et estimée de $\theta_k(n)$ en fonction du rapport S/B, pour respectivement des signaux blancs ($\circ$), colorés ARMA (+) et harmoniques (*) pour un réseau linéaire régulier de 10 capteurs, $n = 100$ avec Toeplitzation (—) et sans Toeplitzation (- - -).

Nous pouvons observer que les EQM estimées coïncident bien avec les valeurs théoriques. Par ailleurs, et conformément à nos conclusions, les performances de l'algorithme MUSIC standard opérant sur la matrice de covariance spatiale non toeplitzée et vérifiant donc dans ce cas la contrainte (4.3), sont identiques quelque soit le type de corrélation temporelle des sources. Par contre, dans le cas où l'algorithme MUSIC opère sur la matrice toeplitzée, l'algorithme résultant ne vérifiant pas la contrainte (4.3) mais la contrainte (4.4), les performances dépendent de la corrélation temporelle des sources et se dégradent d'autant plus que les spectres des signaux se recouvrent.

La FIG.4.7 représente l'EQM théorique sur θ_k : $\frac{1}{n}[\mathbf{C}_\Theta]_{k,k}$ pour des signaux ARMA gaussiens, décorréls spatialement, en fonction de la bande occupée par les sources dans le cas d'un réseau linéaire uniforme de 10 capteurs.

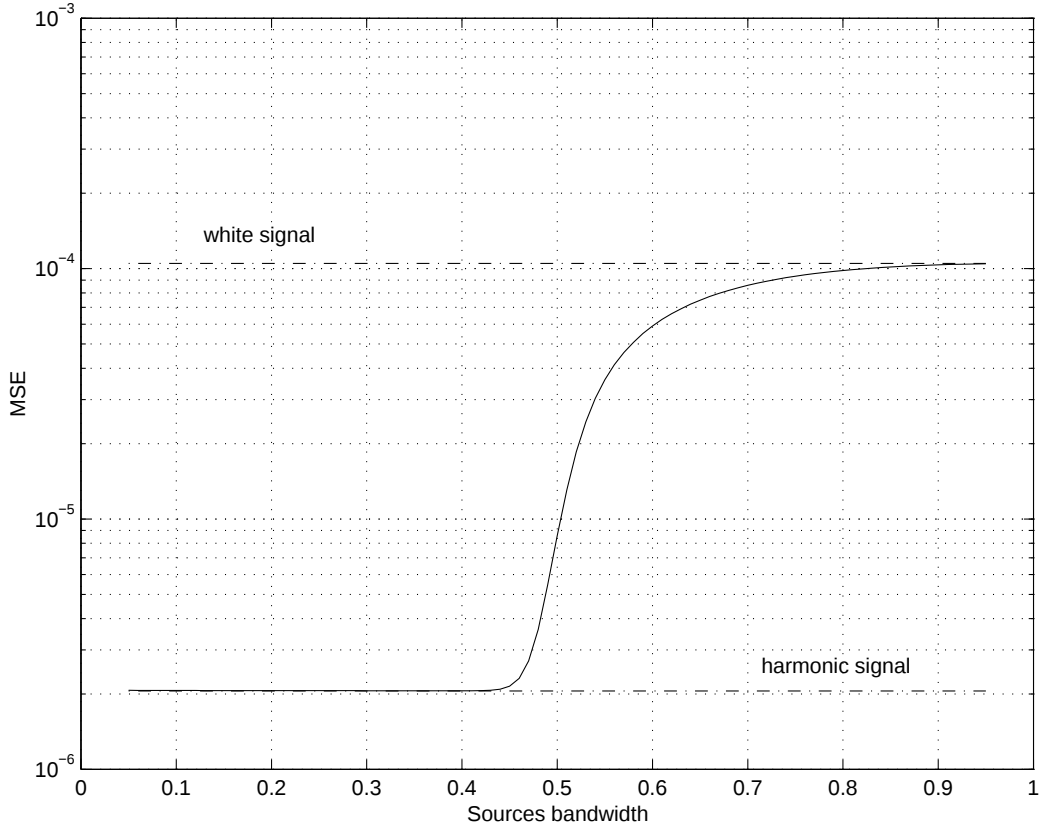


FIG. 4.7: MSE théorique et estimée de $\theta_k(n)$ en fonction de la largeur de bande de signaux ARMA centrés sur $\pm 0,25$ pour un réseau linéaire régulier de 10 capteurs, SNR = 20 dB après Toeplitzation.

On observe que l'EQM théorique augmente avec la bande occupée et que cet accroissement commence pour une largeur de bande de 0.45 qui correspond à un début de recouvrement des spectres des 2 signaux. Les valeurs minimum et maximum de l'EQM correspondent respectivement au cas de 2 sources harmoniques et au cas de 2 sources non corrélées temporellement. Le phénomène de “saturation” observé dans [12] disparaît quand les spectres des 2 sources se disjointent.

La FIG.4.8 représente les EQM théoriques sur $\theta_k : \frac{1}{n}[\mathbf{C}_\Theta]_{k,k}$ ainsi que les EQM estimées : $E\|\theta_k(n) - \theta_k\|_{\text{Fro}}^2$ en fonction du nombre de capteurs d'un réseau linéaire uniforme pour un rapport S/B de 20 dB.

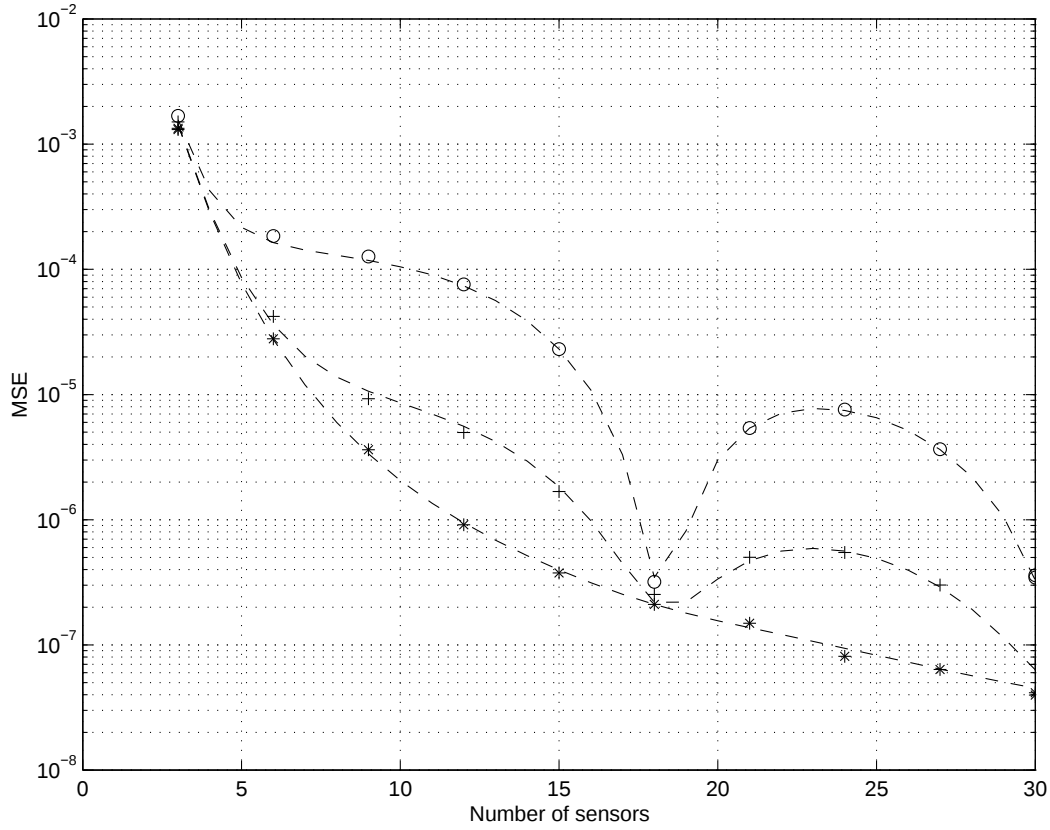


FIG. 4.8: MSE théorique et estimée de $\theta_k(n)$ en fonction du nombre de capteurs pour un réseau linéaire régulier, SNR = 20dB, $n = 100$ pour des signaux blancs (o), colorés ARMA (+) et harmoniques (*) après Toeplitzation.

Nous pouvons observer que les EQM estimées correspondent bien aux valeurs théoriques et que ces EQM sont très sensibles à la corrélation temporelle des sources (la bande de fréquence occupée par les signaux ARMA a été fixée à 0.5). Les EQM décroissent lorsque le nombre de capteurs augmente sauf dans le cas où les spectres des signaux se recouvrent et où le phénomène de “saturation” peut entraîner une dégradation de l’EQM.

Enfin la dernière simulation présente le cas de 2 signaux de mêmes paramètres temporels (pour les signaux sources non corrélés temporellement, la distribution est soit gaussienne soit discrète $\{-1, +1\}$ comme pour la deuxième simulation) mais sont observés par un réseau de 12 antennes de type Greene et Wood [17] ($\Gamma(x, y) = 1$ pour $x = y = 2, 3, 4$; $x = 0, y = 0, \dots, 4$ et $y = 0, x = 0, \dots, 4$) en utilisant une technique d'augmentation directe [16]. La FIG.4.9 représente les EQM théoriques et estimées des angles θ_k et ϕ_k en fonction du SNR. Le comportement de ces EQM est semblable à celui des EQM obtenues dans le cas de réseaux linéaires uniformes avec Toeplitzation.

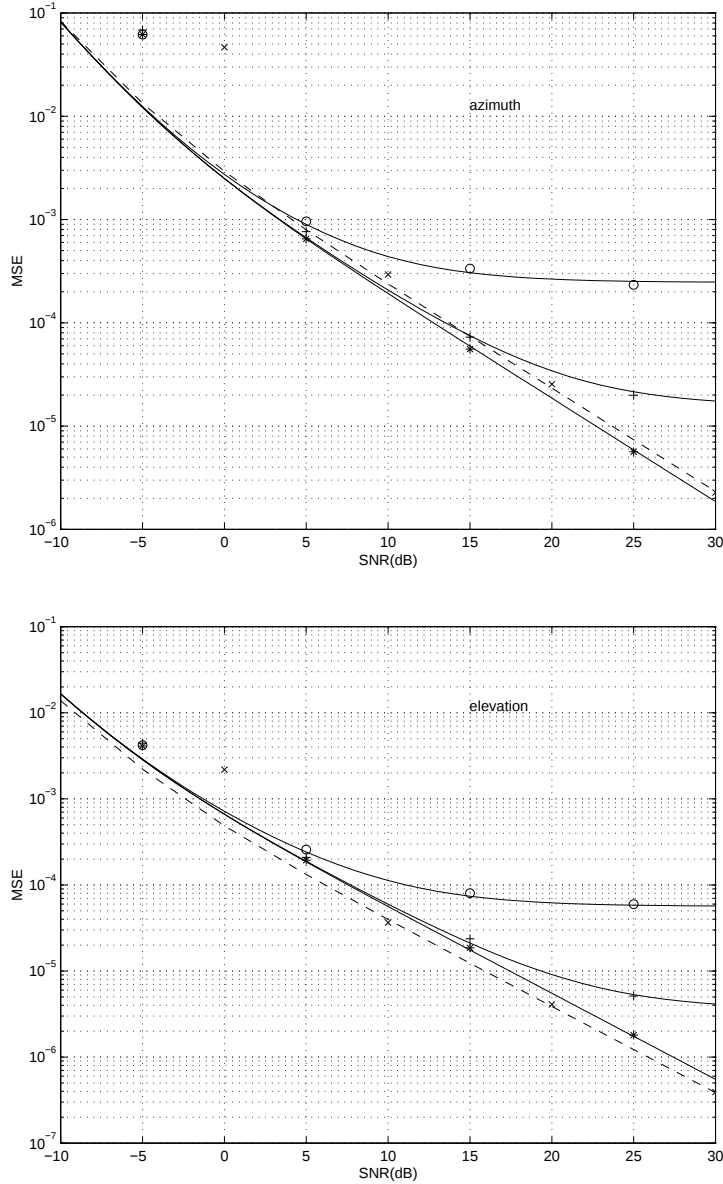


FIG. 4.9: EQM théorique et estimée des azimuth et élévation $\theta_k(n)$ et $\phi_k(n)$ en fonction du rapport S/B, pour des signaux respectivement non corrélés temporellement (de distribution, soit gaussienne, soit de Bernoulli) (o), colorés ARMA (+) et harmoniques (*) pour un réseau Greene et Wood de 12 capteurs [17], pour $n = 100$, utilisant la technique d'augmentation (—) ou non (- -).

4.9 Conclusion

Dans ce chapitre nous avons étudié les performances asymptotiques des algorithmes du second ordre de DOA lorsque les observations successives des signaux ne vérifient pas les hypothèses classiques d'indépendance et de distribution gaussienne. Nous avons supposé que les signaux sources et/ou le bruit étaient des processus aléatoires complexes circulaires éventuellement corrélés temporellement et éventuellement non gaussiens.

Un théorème central limite nous a permis tout d'abord de déterminer, sous ces hypothèses très générales, la distribution asymptotique gaussienne de la matrice de covariance spatiale empirique $\mathbf{R}_x(n)$ de l'observation et en particulier de donner l'expression générale de la matrice de covariance asymptotique $\mathbf{C}_{R_x}$ de $\mathbf{R}_x(n)$. Une approche fonctionnelle des algorithmes nous a ensuite permis de déduire la distribution de probabilité asymptotique, gaussienne également, des estimées $\Theta(n)$ des DOA fournies par ces algorithmes et en particulier d'exprimer la matrice de covariance asymptotique $\mathbf{C}_\Theta$ de ces estimées en fonction de $\mathbf{C}_{R_x}$.

Cette approche fonctionnelle nous a conduit à distinguer deux types d'algorithmes du second ordre de DOA :

1. Les algorithmes qui n'utilisent aucune information a priori sur la corrélation spatiale des sources et que nous qualifions de "standards". C'est le cas par exemple des algorithmes MUSIC et ESPRIT.
2. Les algorithmes qui supposent a priori les sources non corrélées spatialement et qui utilisent explicitement cette hypothèse dans leur conception. Nous qualifions ces algorithmes de "non standards". C'est le cas par exemple des algorithmes utilisant un pré-traitement de la matrice de covariance spatiale empirique comme la "Toeplitzation" ou "l'augmentation" ou encore de l'algorithme COMET selon le type de paramétrage envisagé.

Compte tenu de cette distinction, nous avons alors montré les résultats suivants :

- Dans le cas des algorithmes standards :
 - Ces algorithmes ne sont pas sensibles à la distribution de probabilité des sources et ceci quelle que soit la corrélation temporelle du bruit et des signaux sources.
 - Leurs performances asymptotiques dépendent toutefois a priori de la corrélation temporelle des signaux sources si le bruit est temporellement corrélé.
 - Par contre, si le bruit est temporellement non corrélé, ces algorithmes ne sont pas non plus sensibles à la corrélation temporelle des signaux sources.
 - Dans le cas particulier des algorithmes MUSIC et ESPRIT appliquées à des antennes centrosymétriques et utilisant, comme pré-traitement, un moyennage centro-hermitien de la matrice de covariance spatiale empirique, les performances sont invariantes par ce moyennage quelle que soit la distribution des sources ou leur corrélation temporelle lorsque le bruit est temporellement non corrélé.
- Dans le cas des algorithmes non standards :
 - Ces algorithmes ne sont pas sensibles à la distribution de probabilité des sources si ces sources sont non seulement spatialement non corrélées mais également indépendantes.
 - Ces algorithmes ne sont pas sensibles à la corrélation temporelle des sources dans le cas d'une seule source ou de sources à spectres disjoints. Cette propriété est perdue dans le cas de spectres non disjoints.

Au passage, nous avons également proposé, en utilisant les termes d'intercorrélation de $\mathbf{C}_\Theta$, une explication du fait que, dans le cas de sources gaussiennes complexes circulaires non corrélées spatialement

et d'observations indépendantes, bien que la Toeplitzation détériore la variance des DOA estimées par l'algorithme MUSIC, cette opération en augmente malgré tout la résolution.

Ces différents résultats ont été illustrés et corroborés par une série de simulations.

Annexe

Annexe 1 : Preuve du résultat (4.1)

Le résultat (4.1) comprend deux parties : la convergence en loi vers une loi limite gaussienne complexe non circulaire (3.1) (qui est un résultat très “technique”) et la limite (3.2) de la suite numérique $n\text{Cov}(\text{Vec}(\mathbf{R}_x(n)))$ (qui est un résultat obtenu par un simple calcul de limite). Nous commencerons par établir la première partie.

Preuve de (3.1)

Dans le cas de signaux ARMA complexes circulaires, la convergence en loi de $\sqrt{n} (\text{Vec}(\mathbf{R}_x(n)) - \text{Vec}(\mathbf{R}_x))$ vers une loi gaussienne complexe non circulaire centrée peut être considérée comme un cas particulier, adapté au cas complexe, d’un résultat général [18, théorème 3, chap. 3] présenté dans le cas de processus réels ergodiques stationnaires⁹. Comme dans ce cadre général, la covariance de la distribution asymptotique gaussienne est égale à la limite de la suite $n\text{Cov}(\text{Vec}(\mathbf{R}_x(n)))$, nous nous limiterons ici à évaluer cette limite.

Dans le cas de signaux harmoniques, là aussi la convergence en loi peut être considérée comme un cas particulier, adapté au cas complexe multidimensionnel, d’un résultat général Ta-Hsin Li *et al* [20] concernant les processus somme d’un processus linéaire $\sum_{i=-\infty}^{+\infty} h_i \epsilon_{t-i}$ (où ϵ_t est i.i.d. d’ordre quatre et où $\sum_{i=-\infty}^{+\infty} |h_i| < \infty$) et d’une somme finie de sinusôides. Ce dernier résultat utilise une démonstration basée sur des théorèmes de convergence de suite de différences de martingales et ne donnent que des résultats sous forme de somme infinie et de plus sous forme scalaire réel. Nous allons ici démontrer (3.1) dans le cadre harmonique avec des outils mathématiques beaucoup plus simples.

Par soucis de simplicité de notation, nous ne considérerons que le cas d’une seule harmonique par direction d’arrivée. Par suite $u_{t,k} = a_k e^{i\alpha_k} e^{i2\pi f_k t}$. D’après notre modèle :

$$\text{Vec}(\mathbf{R}_x(n) - \mathbf{R}) = \text{Vec} \left(\frac{1}{n} \sum_{t=1}^n \mathbf{E}(\Theta) (\mathbf{u}_t \mathbf{u}_t^H - \mathbf{R}_u) \mathbf{E}^H(\Theta) + (\mathbf{v}_t \mathbf{v}_t^H - \mathbf{R}_v) + \mathbf{E}(\Theta) \mathbf{u}_t \mathbf{v}_t^H + \mathbf{v}_t \mathbf{u}_t^H \mathbf{E}^H(\Theta) \right),$$

avec :

$$\text{Vec}(\mathbf{E}(\Theta) (\mathbf{u}_t \mathbf{u}_t^H) \mathbf{E}^H(\Theta)) = \sum_{k=1}^K \sum_{k'=1}^K a_k a_{k'} e^{i(\alpha_k - \alpha_{k'})} e^{i2\pi(f_k - f_{k'})t} (\mathbf{e}(\Theta_k) \otimes_c \mathbf{e}^H(\Theta_{k'})),$$

et

$$\text{Vec}(\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta)) = \sum_{k=1}^K a_k^2 (\mathbf{e}(\Theta_k) \otimes_c \mathbf{e}^H(\Theta_{k'}))$$

⁹Nous signalons un résultat de forme voisine au notre présenté dans la thèse de Gorokhov [19, Théorème D.1]. Ce dernier résultat s’applique à un système MIMO FIR complexe non circulaire, additivement bruité par du bruit complexe circulaire gaussien.

d'où :

$$\text{Vec}(\mathbf{E}(\Theta) (\mathbf{u}_t \mathbf{u}_t^H - \mathbf{R}_u) \mathbf{E}^H(\Theta)) = \sum_{1 \leq k \neq k' \leq K} a_k a_{k'} e^{i(\alpha_k - \alpha_{k'})} e^{i2\pi(f_k - f_{k'})t} (\mathbf{e}(\Theta_k) \otimes_c \mathbf{e}^H(\Theta_{k'})).$$

Par suite, grâce à la propriété (0.4) du produit de Kronecker, $\text{Vec}(\mathbf{R}_x(n) - \mathbf{R}_x)$ peut s'écrire sous la forme :

$$\text{Vec}(\mathbf{R}_x(n) - \mathbf{R}_x) = \frac{1}{n} \sum_{t=1}^n (\mathbf{y}'_t + \mathbf{y}''_t)$$

avec

$$\mathbf{y}'_t \stackrel{\text{def}}{=} (\mathbf{E}(\Theta) \mathbf{u}_t \otimes_c \mathbf{v}_t) + (\mathbf{v}_t \otimes_c \mathbf{E}(\Theta) \mathbf{u}_t) + (\mathbf{v}_t \otimes_c \mathbf{v}_t - \text{Vec}(\mathbf{R}_v))$$

et

$$\mathbf{y}''_t \stackrel{\text{def}}{=} \sum_{1 \leq k \neq k' \leq K} a_k a_{k'} e^{i(\alpha_k - \alpha_{k'})} e^{i2\pi(f_k - f_{k'})t} (\mathbf{e}(\Theta_k) \otimes_c \mathbf{e}(\Theta_{k'})).$$

En remarquant que :

$$\|\sqrt{n} \frac{1}{n} \sum_{t=1}^n \mathbf{y}''_t\|_{\text{Fro}} \leq \frac{1}{\sqrt{n}} \sum_{1 \leq k \neq k' \leq K} a_k a_{k'} \left| \frac{\sin(\pi n(f_k - f_{k'}))}{\sin(\pi(f_k - f_{k'}))} \right| \|\mathbf{e}(\Theta_k) \otimes_c \mathbf{e}(\Theta_{k'})\|_{\text{Fro}},$$

nous avons $\lim_{n \rightarrow +\infty} \sqrt{n} \frac{1}{n} \sum_{t=1}^n \mathbf{y}''_t = \mathbf{0}$ presque sûrement. Par suite on peut se limiter à ne considérer que le terme $\frac{1}{n} \sum_{t=1}^n \mathbf{y}'_t$ dans l'étude de la convergence en loi de $\sqrt{n} (\text{Vec}(\mathbf{R}_x(n) - \mathbf{R}_x))$.

La suite de vecteurs aléatoires $(\mathbf{y}'_t)_{t=1, \dots, n}$ n'étant pas une suite de vecteurs aléatoires indépendantes équidistribuées, le théorème central limite classique ne peut s'appliquer. Mais en conditionnant la suite $(\mathbf{y}'_t)_{t=1, \dots, n}$ par la connaissance des phases $(\alpha_k)_{k=1, \dots, K}$, nous obtenons une loi conditionnelle d'une suite de vecteurs aléatoires cette fois-ci indépendants (mais toujours non-identiquement distribués). Pour étudier sa convergence en loi, nous utiliserons le théorème central limite basé sur la condition de Lyapounov [21, 21, p. 371]. Comme cette condition concerne des suites de variables aléatoires réelles scalaires, et que nous avons une suite de vecteurs aléatoires complexes, nous l'appliquerons à la suite de variable aléatoires réelles scalaires $z_t \stackrel{\text{def}}{=} \mathbf{w}^H \mathbf{y}'_t + (\mathbf{w}^H \mathbf{y}'_t)^H$ pour tout vecteur $\mathbf{w} \in \mathbf{C}^{M^2}$ (voir la note de bas de page du paragraphe 4.3).

Pour cela, considérons $\mathbf{E}(\mathbf{y}'_t \mathbf{y}'_t^H / \alpha)$ et $\mathbf{E}(\mathbf{y}'_t \mathbf{y}'_t^T / \alpha)$. Utilisant le fait que $\mathbf{v}_t$ est un vecteur aléatoire gaussien complexe circulaire, nous obtenons immédiatement :

$$\mathbf{E}(\mathbf{y}'_t \mathbf{y}'_t^H / \alpha) = (\mathbf{E}(\Theta) \mathbf{E}(\mathbf{u}_t \mathbf{u}_t^H / \alpha) \mathbf{E}^H(\Theta)) \otimes_c \mathbf{R}_v + \mathbf{R}_v \otimes_c (\mathbf{E}(\Theta) \mathbf{E}(\mathbf{u}_t \mathbf{u}_t^H / \alpha) \mathbf{E}^H(\Theta)) + \mathbf{R}_v \otimes_c \mathbf{R}_v,$$

et en utilisant la propriété (0.9) qui implique que $\mathbf{y}'_t{}^T = \mathbf{y}'_t{}^H \mathbf{K}_M^c$:

$$\mathbf{E}(\mathbf{y}'_t \mathbf{y}'_t^T / \alpha) = ((\mathbf{E}(\Theta) \mathbf{E}(\mathbf{u}_t \mathbf{u}_t^H / \alpha) \mathbf{E}^H(\Theta)) \otimes_c \mathbf{R}_v + \mathbf{R}_v \otimes_c (\mathbf{E}(\Theta) \mathbf{E}(\mathbf{u}_t \mathbf{u}_t^H / \alpha) \mathbf{E}^H(\Theta)) + \mathbf{R}_v \otimes_c \mathbf{R}_v) \mathbf{K}_M^c.$$

Parce que :

$$\frac{1}{n} \sum_{t=1}^n \mathbf{E}(\mathbf{u}_t \mathbf{u}_t^H / \alpha) = \sum_{1 \leq k, k' \leq K} a_k a_{k'} e^{i(\alpha_k - \alpha_{k'})} \left(\frac{1}{n} \sum_{t=1}^n e^{i2\pi(f_k - f_{k'})t} \right) \mathbf{e}(\Theta_k) \mathbf{e}^H(\Theta_{k'}),$$

qui a pour limite $\mathbf{R} \stackrel{\text{def}}{=} \sum_{k=1}^K a_k^2 \mathbf{e}(\Theta_k) \mathbf{e}^H(\Theta_k) = \mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta)$ quand $n \rightarrow \infty$, nous avons :

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \mathbf{E}(\mathbf{y}'_t \mathbf{y}_t^H / \alpha) &= \mathbf{R} \otimes_c \mathbf{R}_v + \mathbf{R}_v \otimes_c \mathbf{R} + \mathbf{R}_v \otimes_c \mathbf{R}_v \\ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \mathbf{E}(\mathbf{y}'_t \mathbf{y}_t'^T / \alpha) &= (\mathbf{R} \otimes_c \mathbf{R}_v + \mathbf{R}_v \otimes_c \mathbf{R} + \mathbf{R}_v \otimes_c \mathbf{R}_v) \mathbf{K}_M^c. \end{aligned}$$

Et en regroupant les termes constituant $z_t \stackrel{\text{def}}{=} \mathbf{w}^H \mathbf{y}'_t + (\mathbf{w}^H \mathbf{y}'_t)^H$, nous obtenons :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \mathbf{E}(z_t^2 / \alpha) = 2 \mathbf{w}^H \mathbf{C}_{R_x} \mathbf{w} + \mathbf{w}^H \mathbf{C}_{R_x} \mathbf{K}_M^c \mathbf{w}^* + \mathbf{w}^T \mathbf{C}_{R_x}^* \mathbf{K}_M^c \mathbf{w}, \quad (9.1)$$

avec $\mathbf{C}_{R_x} \stackrel{\text{def}}{=} \mathbf{R} \otimes_c \mathbf{R}_v + \mathbf{R}_v \otimes_c \mathbf{R} + \mathbf{R}_v \otimes_c \mathbf{R}_v$ qui ne dépend pas de α . Par suite, quand $n \rightarrow \infty$, $\sum_{t=1}^n \mathbf{E}(z_t^2 / \alpha) \propto n$. D'autre part, le terme $\frac{1}{n} \sum_{t=1}^n \mathbf{E}(z_t^4 / \alpha)$ peut être décomposé en la somme de termes de la forme

$$c_{i,j,k,l}(\mathbf{w}) \frac{1}{n} \sum_{t=1}^n \mathbf{E} \left((\mathbf{y}'_t)_i^{s_i} (\mathbf{y}'_t)_j^{s_j} (\mathbf{y}'_t)_k^{s_k} (\mathbf{y}'_t)_l^{s_l} / \alpha \right), \quad \text{pour } 0 \leq i, j, k, l \leq M^2, \quad (9.2)$$

où $c_{i,j,k,l}(\mathbf{w})$ est une fonction appropriée de $\mathbf{w}$ et où $(\mathbf{y}'_t)_i^{s_i}$ dénote la i -ème composante de $\mathbf{y}'_t$ et cette composante est conjuguée ou non. Parce que l'examen des termes de (9.2) montre que l'expression (9.2) admet aussi une limite lorsque $n \rightarrow \infty$, $\sum_{t=1}^n \mathbf{E}(z_t^4 / \alpha) \propto n$ quand $n \rightarrow \infty$. Cela implique donc que la condition de Lyapounov [21, 21, p. 371] suivante est satisfaite :

$$\lim_{n \rightarrow \infty} \frac{\sum_{t=1}^n \mathbf{E}(z_t^{2+2} / \alpha)}{\left(\sqrt{\sum_{t=1}^n \mathbf{E}(z_t^2 / \alpha)} \right)^{2+2}} = 0. \quad (9.3)$$

Par suite, selon les conclusions de cette condition de Lyapounov, la suite de variables aléatoires scalaires réelles $\frac{\sum_{t=1}^n z_t}{\sqrt{\sum_{t=1}^n \mathbf{E}(z_t^2 / \alpha)}}$ converge donc en loi vers une variable aléatoire réelle gaussienne centrée réduite.

Puis grâce à (9.1), la suite de variables aléatoires réelles scalaires $\sqrt{n} \frac{1}{n} \sum_{t=1}^n z_t$ converge en loi vers une loi gaussienne centrée de variance $2 \mathbf{w}^H \mathbf{C}_{R_x} \mathbf{w} + \mathbf{w}^H \mathbf{C}_{R_x} \mathbf{K}_M^c \mathbf{w}^* + \mathbf{w}^T \mathbf{C}_{R_x}^* \mathbf{K}_M^c \mathbf{w}$.

D'après le théorème de Cramer-Wold [21, theorem 29.4], le vecteur aléatoire complexe $\sqrt{n} \frac{1}{n} \sum_{t=1}^n \mathbf{y}'_t$ converge donc en loi vers une loi gaussienne complexe centrée non-circulaire $\mathcal{N}(\mathbf{0}, \mathbf{C}_{R_x}, \mathbf{C}_{R_x} \mathbf{K}_M^c)$, ce qui prouve le résultat (3.1) pour le modèle harmonique des sources.

Notons que ce théorème exclut des fréquences égales $f_{k,l}$ dans deux directions différentes car dans ce dernier cas la suite $\sqrt{n} \frac{1}{n} \sum_{t=1}^n \mathbf{y}_t''$ ne convergerait plus vers $\mathbf{0}$ presque sûrement.

Preuve de (3.2)

En posant $\mathbf{R}_u(n) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{t=1}^n \mathbf{u}_t \mathbf{u}_t^H$ et $\mathbf{R}_v(n) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{t=1}^n \mathbf{v}_t \mathbf{v}_t^H$, $\text{Vec}(\mathbf{R}_x(n))$ s'écrit, en utilisant la propriété (0.1) du produit de Kronecker,

$$\text{Vec}(\mathbf{R}_x(n)) = (\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta)) \text{Vec}(\mathbf{R}_u(n)) + \text{Vec}(\mathbf{R}_v(n)) + \frac{1}{n} \sum_{t=1}^n (\mathbf{E}(\Theta) \mathbf{u}_t) \otimes_c \mathbf{v}_t + \frac{1}{n} \sum_{t=1}^n \mathbf{v}_t \otimes_c (\mathbf{E}(\Theta) \mathbf{u}_t)$$

Par suite $n \text{Cov}(\text{Vec}(\mathbf{R}_x(n)))$ est constitué de 16 termes dont 10 sont nuls de par l'indépendance et le caractère centré de $\mathbf{u}_t$ et $\mathbf{v}_t$. Après élimination de ces termes, il reste :

$$n \text{Cov}(\text{Vec}(\mathbf{R}_x(n))) = \mathbf{A}_1(n) + \mathbf{A}_2(n) + \mathbf{A}_3(n) + \mathbf{A}_4(n) + \mathbf{A}_5(n) + \mathbf{A}_6(n) \quad (9.4)$$

avec

$$\begin{aligned}
\mathbf{A}_1(n) &\stackrel{\text{def}}{=} (\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta)) n \text{Cov}(\text{Vec}(\mathbf{R}_u(n))) (\mathbf{E}^H(\Theta) \otimes_c \mathbf{E}^H(\Theta)) \\
\mathbf{A}_2(n) &\stackrel{\text{def}}{=} n \text{Cov}(\text{Vec}(\mathbf{R}_v(n))) \\
\mathbf{A}_3(n) &\stackrel{\text{def}}{=} \frac{1}{n} \mathbf{E} \left[\left(\sum_{t=1}^n (\mathbf{E}(\Theta) \mathbf{u}_t) \otimes_c \mathbf{v}_t \right) \left(\sum_{t'=1}^n (\mathbf{E}(\Theta) \mathbf{u}_{t'})^H \otimes_c \mathbf{v}_{t'}^H \right) \right] \\
\mathbf{A}_4(n) &\stackrel{\text{def}}{=} \frac{1}{n} \mathbf{E} \left[\left(\sum_{t=1}^n (\mathbf{E}(\Theta) \mathbf{u}_t) \otimes_c \mathbf{v}_t \right) \left(\sum_{t'=1}^n \mathbf{v}_{t'}^H \otimes_c (\mathbf{E}(\Theta) \mathbf{u}_{t'})^H \right) \right] \\
\mathbf{A}_5(n) &\stackrel{\text{def}}{=} \frac{1}{n} \mathbf{E} \left[\left(\sum_{t=1}^n \mathbf{v}_t \otimes_c (\mathbf{E}(\Theta) \mathbf{u}_t) \right) \left(\sum_{t'=1}^n (\mathbf{E}(\Theta) \mathbf{u}_{t'})^H \otimes_c \mathbf{v}_{t'}^H \right) \right] \\
\mathbf{A}_6(n) &\stackrel{\text{def}}{=} \frac{1}{n} \mathbf{E} \left[\left(\sum_{t=1}^n \mathbf{v}_t \otimes_c (\mathbf{E}(\Theta) \mathbf{u}_t) \right) \left(\sum_{t'=1}^n \mathbf{v}_{t'}^H \otimes_c (\mathbf{E}(\Theta) \mathbf{u}_{t'})^H \right) \right].
\end{aligned}$$

Evaluons la limite de chacun de ces termes quand $n \rightarrow \infty$. Dans le premier terme, $n \text{Cov}(\text{Vec}(\mathbf{R}_u(n))) = n \mathbf{E}[\text{Vec}(\mathbf{R}_u(n)) \text{Vec}^H(\mathbf{R}_u(n))] - n \text{Vec}(\mathbf{R}_u) \text{Vec}^H(\mathbf{R}_u)$ dont les composantes s'écrivent :

$$\begin{aligned}
[n \text{Cov}(\text{Vec}(\mathbf{R}_u(n)))]_{K(j-1)+i, K(l-1)+k} &= \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E}(u_{t,i} u_{t,j}^* u_{t',l} u_{t',k}^*) \\
&\quad - \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E}(u_{t,i} u_{t,j}^*) \mathbf{E}(u_{t',l} u_{t',k}^*). \tag{9.5}
\end{aligned}$$

Pour ce terme considérons successivement le cas de signaux sources harmoniques et ARMA.

- **Signaux sources harmoniques** : Afin d'alléger la notation, nous nous limiterons là aussi au cas où chaque source n'est constituée que d'une seule harmonique et par suite $u_{t,k} = a_k e^{i\alpha_k} e^{i2\pi f_k t}$.

$$\frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E}(u_{t,i} u_{t,j}^* u_{t',l} u_{t',k}^*) = \begin{cases} n a_i^2 a_k^2 & \text{si } i = j \text{ et } k = l \\ \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n a_i^2 e^{i2\pi(f_i-f_l)t} a_l^2 e^{-i2\pi(f_i-f_l)t'} & \text{si } i = k \neq j = l \\ 0 & \text{sinon} \end{cases}$$

Or

$$\frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n a_i^2 e^{i2\pi(f_i-f_l)t} a_l^2 e^{-i2\pi(f_i-f_l)t'} = \frac{1}{n} a_i^2 a_l^2 \frac{\sin^2 \pi(f_i - f_l)n}{\sin^2 \pi(f_i - f_l)}$$

et, sous la condition qu'aucune fréquence issue de deux directions ne soient égales ($f_i \neq f_l \ \forall i \neq l$), la limite de ces termes quand $n \rightarrow \infty$ est nulle. Comme par ailleurs $\mathbf{R}_u = \text{Diag}(a_1^2, \dots, a_K^2)$, on a

$$\lim_{n \rightarrow \infty} n \text{Cov}(\text{Vec}(\mathbf{R}_u(n))) \stackrel{\text{def}}{=} \mathbf{C}_{R_u} = \mathbf{O}.$$

- **Signaux sources ARMA complexes circulaires** : Dans ce cas les composantes de (9.5) s'écrivent :

$$\begin{aligned}
[n \text{Cov}(\text{Vec}(\mathbf{R}_u(n)))]_{K(j-1)+i, K(l-1)+k} &= \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E}(u_{t,i} u_{t',k}^*) \mathbf{E}(u_{t',l} u_{t,j}^*) \\
&\quad + \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \text{Cum}(u_{t,i}, u_{t,j}^*, u_{t',l}, u_{t',k}^*). \tag{9.6}
\end{aligned}$$

Le premier terme est celui de la matrice qui a pour limite quand n tend vers $+\infty$ d'après [8, lemme 4.1 p.101] : $\sum_{-\infty}^{+\infty} \mathbf{R}_u^t \otimes_c \mathbf{R}_u^t$, qui grâce au Théorème de Parseval s'écrit $\int_{-1/2}^{+1/2} \mathbf{S}_u(f) \otimes_c \mathbf{S}_u(f) df$. Le deuxième terme s'écrit sous l'hypothèse que $\sum_{-\infty}^{+\infty} |\text{Cum}(u_{0,i}, u_{0,j}^*, u_{t,l}, u_{t,k}^*)| < \infty$, d'après [8, lemme 4.1 p.101] étendu au cas de signaux complexes $u_{t,i}$, $u_{t,j}^*$, $u_{t,l}$ et $u_{t,k}^*$ sous la forme suivante :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \text{Cum}(u_{t,i}, u_{t,j}^*, u_{t',l}, u_{t',k}^*) = \sum_{t=-\infty}^{\infty} \text{Cum}(u_{0,i}, u_{0,j}^*, u_{t,l}, u_{t,k}^*)$$

Si $\rho_{i,j,k,l}(f_1, f_2, f_3)$ désigne le trispectre des signaux complexes $u_{t,i}$, $u_{t,j}^*$, $u_{t',l}$ et $u_{t',k}^*$ et grâce à $\int_{-1/2}^{+1/2} e^{i2\pi f t} df = \delta_{t,0}$ ($\delta_{t,0} = 0$ si $t = 0$, $\delta_{t,0} = 1$ si $t \neq 0$), ce terme s'écrit sous la forme :

$$\begin{aligned} & \sum_{t=-\infty}^{\infty} \text{Cum}(u_{0,i}, u_{0,j}^*, u_{t,l}, u_{t,k}^*) \\ &= \sum_{t=-\infty}^{\infty} \sum_{t_1=-\infty}^{\infty} \sum_{t_2=-\infty}^{\infty} \int_{-1/2}^{+1/2} \int_{-1/2}^{+1/2} \text{Cum}(u_{0,i}, u_{t_1,j}^*, u_{t,l}, u_{t_2,k}^*) e^{i2\pi(f t_1 + f'(t_2 - t))} df df' \\ &= \int_{-1/2}^{+1/2} \int_{-1/2}^{+1/2} \left(\sum_{t=-\infty}^{\infty} \sum_{t_1=-\infty}^{\infty} \sum_{t_2=-\infty}^{\infty} \text{Cum}(u_{0,i}, u_{t_1,j}^*, u_{t,l}, u_{t_2,k}^*) e^{i2\pi(f t_1 + f'(t_2 - t))} \right) df df' \\ &= \int_{-1/2}^{+1/2} \int_{-1/2}^{+1/2} \rho_{i,j,k,l}(f, -f', f') df df'. \end{aligned} \quad (9.7)$$

Par suite le deuxième terme de (9.6) est égal à $\mathbf{Q}_u$ défini par (2.1). Par suite :

$$\lim_{n \rightarrow \infty} n \text{Cov}(\text{Vec}(\mathbf{R}_u(n))) \stackrel{\text{def}}{=} \mathbf{C}_{R_u} = \int_{-1/2}^{+1/2} \mathbf{S}_u(f) \otimes_c \mathbf{S}_u(f) df + \mathbf{Q}_u.$$

D'où $\lim_{n \rightarrow \infty} \mathbf{A}_1(n) = (\mathbf{E}(\Theta) \otimes_c \mathbf{E}(\Theta)) \mathbf{C}_{R_u} (\mathbf{E}^H(\Theta) \otimes_c \mathbf{E}^H(\Theta))$ avec respectivement dans le cas de sources harmoniques $\mathbf{C}_{R_u} = \mathbf{O}$ et dans le cas de sources ARMA complexes circulaires $\mathbf{C}_{R_u} = \int_{-1/2}^{+1/2} \mathbf{S}_u(f) \otimes_c \mathbf{S}_u(f) df + \mathbf{Q}_u$. Le deuxième terme de (9.4) se traite de façon similaire au premier terme et, avec des notations identiques :

$$\lim_{n \rightarrow \infty} \mathbf{A}_2(n) \stackrel{\text{def}}{=} \mathbf{C}_{R_v} = \int_{-1/2}^{+1/2} \mathbf{S}_v(f) \otimes_c \mathbf{S}_v(f) df + \mathbf{Q}_v.$$

Considérons $\lim_{n \rightarrow \infty} \mathbf{A}_3(n)$. Grâce à la propriété (0.1)

$$\begin{aligned} \mathbf{A}_3(n) &= \frac{1}{n} \mathbf{E} \left[\left(\sum_{t=1}^n (\mathbf{E}(\Theta) \mathbf{u}_t) \otimes_c \mathbf{I}_M \mathbf{v}_t \right) \left(\sum_{t'=1}^n (\mathbf{E}(\Theta) \mathbf{u}_{t'}) \otimes_c \mathbf{I}_M \mathbf{v}_{t'} \right)^H \right] \\ &= \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E} \left((\mathbf{E}(\Theta) \otimes_c \mathbf{I}_M) (\mathbf{u}_t \otimes_c \mathbf{v}_t) (\mathbf{u}_{t'} \otimes_c \mathbf{v}_{t'})^H (\mathbf{E}(\Theta) \otimes_c \mathbf{I}_M)^H \right) \\ &= (\mathbf{E}(\Theta) \otimes_c \mathbf{I}_M) \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E}(\mathbf{u}_t \mathbf{u}_{t'}^H) \otimes_c \mathbf{E}(\mathbf{v}_t \mathbf{v}_{t'}^H) (\mathbf{E}^H(\Theta) \otimes_c \mathbf{I}_M). \end{aligned}$$

Par suite

$$\lim_{n \rightarrow \infty} \mathbf{A}_3(n) = (\mathbf{E}(\Theta) \otimes_c \mathbf{I}_M) \mathbf{C}_{R_{u,v}} (\mathbf{E}^H(\Theta) \otimes_c \mathbf{I}_M) \quad \text{avec} \quad \mathbf{C}_{R_{u,v}} \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E}(\mathbf{u}_t \mathbf{u}_{t'}^H) \otimes_c \mathbf{E}(\mathbf{v}_t \mathbf{v}_{t'}^H).$$

De manière analogue

$$\lim_{n \rightarrow \infty} \mathbf{A}_6(n) = (\mathbf{I}_M \otimes_c \mathbf{E}(\Theta)) \mathbf{C}_{R_v, u} (\mathbf{I}_M \otimes_c \mathbf{E}^H(\Theta)) \quad \text{avec} \quad \mathbf{C}_{R_v, u} \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E}(\mathbf{v}_t \mathbf{v}_t^H) \otimes_c \mathbf{E}(\mathbf{u}_t \mathbf{u}_t^H).$$

Enfin, en utilisant les propriétés (0.5) et (0.6) du produit de Kronecker,

$$\mathbf{A}_4(n) = \frac{1}{n} \sum_{t=1}^n \sum_{t'=1}^n \mathbf{E} [(\mathbf{E}(\Theta) \otimes_c \mathbf{v}_t) \mathbf{u}_t \mathbf{u}_{t'}^T (\mathbf{v}_{t'}^H \otimes_c \mathbf{E}^H(\Theta))],$$

et à cause de la circularité de $\mathbf{u}_t$, $\mathbf{A}_4(n) = \mathbf{0}$. De la même manière on montrerait que $\mathbf{A}_5(n) = \mathbf{0}$. En rassemblant les six termes de (9.4), la relation (3.2) est prouvée.

Annexe 2 : Expression de $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{MUSIC-2D}}$

Les expressions de $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}}$ peuvent être obtenues par un calcul de perturbation comme par exemple dans [7] pour l'algorithme MUSIC-1D conventionnel. Ces expressions sont classiques pour la plupart des algorithmes 1D à haute résolution basés sur des statistiques du second ordre. La même approche peut être utilisée pour les algorithmes 2D. Considérons par exemple l'algorithme MUSIC-2D standard pour lequel la détermination des estimés se fait en deux étapes : estimation du projecteur sur l'espace bruit puis recherche d'extrema de la fonction discriminante. L'expression de la $2K \times M_1^2 M_2^2$ matrice $\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{alg}}$ associée à la différentielle de cet algorithme, évaluée au point $\mathbf{R}_x$, peut donc s'écrire sous la forme :

$$\mathbf{D}_{\Theta, \mathbf{R}_x}^{\text{music}} = \mathbf{D}_{\Theta, \mathbf{P}_x}^{\text{music}} \mathbf{D}_{\mathbf{P}_x, \mathbf{R}_x}$$

où $\mathbf{P}_x$ désigne la matrice de projection sur l'espace bruit de $\mathbf{R}_x$ et où $\mathbf{D}_{\Theta, \mathbf{P}_x}^{\text{music}}$ [resp., $\mathbf{D}_{\mathbf{P}_x, \mathbf{R}_x}$] désigne la $2K \times M_1^2 M_2^2$ [resp., $M_1^2 M_2^2 \times M_1^2 M_2^2$] matrice associée à la différentielle de l'application $\mathbf{P}_x \rightarrow \Theta$ [resp. $\mathbf{R}_x \rightarrow \mathbf{P}_x$] au point Θ [resp., $\mathbf{R}_x$]. Il est montré dans [7] que :

$$\mathbf{D}_{\mathbf{P}_x, \mathbf{R}_x} = -\mathbf{P}_x \otimes_c \mathbf{\Gamma}_x - \mathbf{\Gamma}_x \otimes_c \mathbf{P}_x,$$

avec $\mathbf{\Gamma}_x \stackrel{\text{def}}{=} (\mathbf{E}(\Theta) \mathbf{R}_u \mathbf{E}^H(\Theta))^\#$ (où $\#$ désigne la matrice pseudo-inverse de Moore-Penrose). En utilisant la même méthode que dans [7], on obtient l'expression suivante pour $\mathbf{D}_{\Theta, \mathbf{P}_x}^{\text{music}}$:

$$\mathbf{D}_{\Theta, \mathbf{P}_x}^{\text{music}} = - \begin{bmatrix} \mathbf{A}_1^{-1} & \begin{pmatrix} \mathbf{e}'_{\theta_1}{}^T(\Theta_1) \otimes_c \mathbf{e}^T(\Theta_1) + \mathbf{e}^T(\Theta_1) \otimes_c \mathbf{e}'_{\theta_1}{}^T(\Theta_1) \\ \mathbf{e}'_{\phi_1}{}^T(\Theta_1) \otimes_c \mathbf{e}^T(\Theta_1) + \mathbf{e}^T(\Theta_1) \otimes_c \mathbf{e}'_{\phi_1}{}^T(\Theta_1) \end{pmatrix} \\ \vdots & \vdots \\ \mathbf{A}_K^{-1} & \begin{pmatrix} \mathbf{e}'_{\theta_K}{}^T(\Theta_K) \otimes_c \mathbf{e}^T(\Theta_K) + \mathbf{e}^T(\Theta_K) \otimes_c \mathbf{e}'_{\theta_K}{}^T(\Theta_K) \\ \mathbf{e}'_{\phi_K}{}^T(\Theta_K) \otimes_c \mathbf{e}^T(\Theta_K) + \mathbf{e}^T(\Theta_K) \otimes_c \mathbf{e}'_{\phi_K}{}^T(\Theta_K) \end{pmatrix} \end{bmatrix},$$

avec $\Theta \stackrel{\text{def}}{=} (\Theta_1^T, \dots, \Theta_K^T)^T$, $\Theta_k \stackrel{\text{def}}{=} (\theta_k, \phi_k)^T$, $\mathbf{e}(\Theta_k) \stackrel{\text{def}}{=} \mathbf{e}_{M_1}(\theta_k) \otimes_r \mathbf{e}_{M_2}(\phi_k)$, $\mathbf{e}'_{\theta_k}(\Theta_k) \stackrel{\text{def}}{=} \frac{d}{d\theta_k} \mathbf{e}_{M_1}(\theta_k) \otimes_r \mathbf{e}_{M_2}(\phi_k)$, $\mathbf{e}'_{\phi_k}(\Theta_k) \stackrel{\text{def}}{=} \mathbf{e}_{M_1}(\theta_k) \otimes_r \frac{d}{d\phi_k} \mathbf{e}_{M_2}(\phi_k)$ et

$$\mathbf{A}_k \stackrel{\text{def}}{=} 2 \begin{bmatrix} \text{Re} \left(\text{Tr} \left(\mathbf{e}'_{\theta_k}{}^H(\Theta_k) \mathbf{P}_x \mathbf{e}'_{\theta_k}(\Theta_k) \right) \right) & \text{Re} \left(\text{Tr} \left(\mathbf{e}'_{\theta_k}{}^H(\Theta_k) \mathbf{P}_x \mathbf{e}'_{\phi_k}(\Theta_k) \right) \right) \\ \text{Re} \left(\text{Tr} \left(\mathbf{e}'_{\phi_k}{}^H(\Theta_k) \mathbf{P}_x \mathbf{e}'_{\theta_k}(\Theta_k) \right) \right) & \text{Re} \left(\text{Tr} \left(\mathbf{e}'_{\phi_k}{}^H(\Theta_k) \mathbf{P}_x \mathbf{e}'_{\phi_k}(\Theta_k) \right) \right) \end{bmatrix}, k = 1, \dots, K.$$

Annexe 3 : Preuve de l'expression (5.1)

Les matrices $\mathbf{C}_{R_{u,v}}$, $\mathbf{C}_{R_{v,u}}$ et $\mathbf{C}_{R_v}$ qui figurent dans l'expression (4.7) de la variance donnée dans le cas général de bruit temporellement corrélé, s'écrivent ici sous les formes suivantes :

- $\mathbf{C}_{R_{u,v}} = \int_{-1/2}^{+1/2} S_u(f) \mathbf{S}_v(f) df = (\int_{-1/2}^{+1/2} S_u(f) S_v(f) df) \mathbf{I}_M = \sigma_{u,v}^4 \mathbf{I}_M$ avec $\sigma_{u,v}^4 \stackrel{\text{def}}{=} \int_{-1/2}^{+1/2} S_u(f) S_v(f) df$
- $\mathbf{C}_{R_{v,u}} = \int_{-1/2}^{+1/2} \mathbf{S}_v(f) S_u(f) df = (\int_{-1/2}^{+1/2} S_u(f) S_v(f) df) \mathbf{I}_M = \sigma_{u,v}^4 \mathbf{I}_M$
- $\mathbf{C}_{R_v} = \int_{-1/2}^{+1/2} \mathbf{S}_v(f) \mathbf{S}_v^*(f) df = (\int_{-1/2}^{+1/2} S_v^2(f) df) \mathbf{I}_{M^2} = \sigma_{v,v}^4 \mathbf{I}_{M^2}$ avec $\sigma_{v,v}^4 \stackrel{\text{def}}{=} \int_{-1/2}^{+1/2} S_v^2(f) df$

D'autre part, la matrice de la différentielle de l'algorithme MUSIC a pour expression (cf. Chapitre 3, relation (4.10)) :

$$\mathbf{D}_{\theta, \mathbf{R}_x}^{\text{MUSIC}} = \frac{1}{\alpha} \left(\mathbf{e}^H(\theta) \mathbf{\Gamma} \otimes_c \mathbf{e}'^H(\theta) \mathbf{P} + \mathbf{e}'^H(\theta) \mathbf{P} \otimes_c \mathbf{e}^H(\theta) \mathbf{\Gamma} \right),$$

où $\mathbf{P} \stackrel{\text{def}}{=} \mathbf{I}_M - \frac{\mathbf{e}(\theta) \mathbf{e}^H(\theta)}{\|\mathbf{e}(\theta)\|^2}$ est la matrice de projection sur l'espace bruit, $\mathbf{\Gamma} \stackrel{\text{def}}{=} (\sigma_u^2 \mathbf{e}(\theta) \mathbf{e}^H(\theta))^{\#} = \frac{\mathbf{e}(\theta) \mathbf{e}^H(\theta)}{\sigma_u^2 \|\mathbf{e}(\theta)\|^4}$ et $\alpha \stackrel{\text{def}}{=} 2 \mathbf{e}'^H(\theta) \mathbf{P} \mathbf{e}'(\theta)$ (où $\#$ désigne la matrice pseudo-inverse de Moore-Penrose). Par suite

$$\mathbf{D}_{\theta, \mathbf{R}_x}^{\text{MUSIC}} = \frac{1}{\alpha \sigma_u^2 \|\mathbf{e}(\theta)\|^2} \left(\mathbf{e}^H(\theta) \otimes_c \mathbf{e}'^H(\theta) \mathbf{P} + \mathbf{e}'^H(\theta) \mathbf{P} \otimes_c \mathbf{e}^H(\theta) \right).$$

D'après la relation (4.7) qui donne la variance asymptotique de θ dans le cas général de bruit temporellement corrélé, on obtient à l'aide des expressions précédentes de $\mathbf{C}_{R_{u,v}}$, $\mathbf{C}_{R_{v,u}}$ et $\mathbf{C}_{R_v}$:

$$\begin{aligned} \text{Var}(\theta) &= \frac{1}{\alpha^2 \sigma_u^4 \|\mathbf{e}(\theta)\|^4} \left(\mathbf{e}^H(\theta) \otimes_c \mathbf{e}'^H(\theta) \mathbf{P} + \mathbf{e}'^H(\theta) \mathbf{P} \otimes_c \mathbf{e}^H(\theta) \right) \\ &\quad \left(\sigma_{u,v}^4 (\mathbf{e}(\theta) \otimes_c \mathbf{I}_M) (\mathbf{e}^H(\theta) \otimes_c \mathbf{I}_M) + \sigma_{u,v}^4 (\mathbf{I}_M \otimes_c \mathbf{e}(\theta)) (\mathbf{I}_M \otimes_c \mathbf{e}^H(\theta)) + \sigma_{v,v}^4 \mathbf{I}_{M^2} \right) \\ &\quad \left(\mathbf{e}(\theta) \otimes_c \mathbf{P} \mathbf{e}'(\theta) + \mathbf{P} \mathbf{e}'(\theta) \otimes_c \mathbf{e}(\theta) \right). \end{aligned}$$

En développant cette expression en utilisant $\mathbf{P} \mathbf{e}(\theta) = \mathbf{0}$ et la propriété (0.1) du produit de Kronecker (voir formulaire), il vient :

$$\begin{aligned} \text{Var}(\theta) &= \frac{1}{\alpha^2 \sigma_u^4 \|\mathbf{e}(\theta)\|^4} \left(\mathbf{e}^H(\theta) \otimes_c \mathbf{e}'^H(\theta) \mathbf{P} + \mathbf{e}'^H(\theta) \mathbf{P} \otimes_c \mathbf{e}^H(\theta) \right) \\ &\quad \left(\sigma_{u,v}^4 (\mathbf{e}(\theta) \mathbf{e}^H(\theta) \otimes_c \mathbf{I}_M + \mathbf{I}_M \otimes_c \mathbf{e}(\theta) \mathbf{e}^H(\theta)) + \sigma_{v,v}^4 \mathbf{I}_{M^2} \right) \left(\mathbf{e}(\theta) \otimes_c \mathbf{P} \mathbf{e}'(\theta) + \mathbf{P} \mathbf{e}'(\theta) \otimes_c \mathbf{e}(\theta) \right) \\ &= \frac{1}{\alpha \sigma_u^4} \left(\sigma_{u,v}^4 + \frac{1}{\|\mathbf{e}(\theta)\|^2} \sigma_{v,v}^4 \right) \\ &= \frac{1}{\alpha} \left(\int_{-1/2}^{+1/2} S_u(f) S_v(f) df + \frac{1}{\|\mathbf{e}(\theta)\|^2} \int_{-1/2}^{+1/2} S_v^2(f) df \right). \end{aligned}$$

Annexe 4 : Preuve de l'expression (5.3)

Dans le cas particulier où les signaux source et les différentes composantes du bruit sont des signaux AR d'ordre 1 de coefficients respectifs a_u et a_v , dont les densités spectrales de puissance sont centrées respectivement sur les fréquences f_u et f_v nous avons,

$$R_u^n \stackrel{\text{def}}{=} \text{E}(u_t u_{t-n}^*) = (-1)^n a_u^{|n|} \sigma_u^2 e^{i2\pi f_u n}$$

et, le bruit étant spatialement non corrélé,

$$\mathbf{R}_v^n \stackrel{\text{def}}{=} \mathbb{E}(\mathbf{v}_t \mathbf{v}_{t-n}^H) = (-1)^n a_v^{|n|} \sigma_v^2 e^{i2\pi f_v n} \mathbf{I}_M.$$

Donc,

$$\mathbf{C}_{R_{u,v}} \stackrel{\text{def}}{=} \sum_{n=-\infty}^{+\infty} R_u^n \otimes_c \mathbf{R}_v^n = \sigma_u^2 \sigma_v^2 \frac{1 - a_u^2 a_v^2}{1 + a_u^2 a_v^2 - 2a_u a_v \cos(2\pi \Delta f)} \mathbf{I}_M \quad \text{avec} \quad \Delta f \stackrel{\text{def}}{=} f_u - f_v$$

et

$$\mathbf{C}_{R_{v,v}} \stackrel{\text{def}}{=} \sum_{n=-\infty}^{+\infty} \mathbf{R}_v^n \otimes_c \mathbf{R}_v^n = \sigma_v^4 \frac{1 + a_v^2}{1 - a_v^2} \mathbf{I}_{M^2}.$$

Par suite

$$\int_{-1/2}^{+1/2} S_u(f) S_v(f) df = \sigma_u^2 \sigma_v^2 \frac{1 - a_u^2 a_v^2}{1 + a_u^2 a_v^2 - 2a_u a_v \cos(2\pi \Delta f)} \quad \text{et} \quad \int_{-1/2}^{+1/2} S_v^2(f) df = \sigma_v^4 \frac{1 + a_v^2}{1 - a_v^2}$$

et en reportant ces expressions dans la relation (5.1) il vient :

$$\text{Var}(\theta) = \frac{1}{\alpha \sigma_u^4} \left(\sigma_u^2 \sigma_v^2 \frac{1 - a_u^2 a_v^2}{1 + a_u^2 a_v^2 - 2a_u a_v \cos(2\pi \Delta f)} + \frac{\sigma_v^4}{\|\mathbf{e}(\theta)\|^2} \frac{1 + a_v^2}{1 - a_v^2} \right).$$

Bibliographie

- [1] B. Ottersten, Ljung, "Asymptotic results for sensor array processing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (Glasgow, Scotland)*, pp. 2266-2269, March 1989.
- [2] P. Stoica, A. Nehorai, "MUSIC, Maximum likelihood, and Cramer-Rao Bound," *IEEE Trans. Acoustics Speech Signal Processing*, vol. 37, no. 5, pp. 720-741, May 1989.
- [3] P. Stoica, A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoustics Speech Signal Processing*, vol. 38, no. 10, pp. 1783-1795, Oct. 1990.
- [4] B. Ottersten, M. Viberg and T. Kailath, "Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data," *IEEE Trans. Signal Processing*, vol. 40, no. 3, pp. 590-599, March 1992.
- [5] J.F. Cardoso, E. Moulines, "A robustness property of DOA estimators based on covariance," *IEEE Trans. Signal Processing*, vol. 42, no. 11, pp. 3285-3287, Nov. 1994.
- [6] R.J. Serfling, *Approximation theorems of mathematical statistics*, John Wiley and Sons, 1980.
- [7] J.F. Cardoso, E. Moulines, "Asymptotic performance analysis of direction-finding algorithms based on fourth-order cumulants," *IEEE Trans. Signal Processing*, vol. 43, no. 1, pp. 214-224, Jan. 1995.
- [8] B. Porat, *Digital Processing of Random Signals : Theory and Methods*, Prentice-Hall, 1993.
- [9] D.R. Brillinger, *Times series, data analysis and theory*, Expanded Edition, Holden-Day, Inc., 1980.
- [10] W. Xu, M. Kaveh, "The optimum weight of angle-dependent weighted MUSIC and its approximations," in *Proc. Asilomar conf. on Signal, System and Computers*, Nov. 1993.
- [11] B. Ottersten, P. Stoica and R. Roy, "Covariance matching estimation techniques for array signal processing," *Digital Signal Processing*, vol. 8, pp. 185-210, 1998.
- [12] A. Gorokhov, Y. Abramovich and J.F. Böhme, "Unified analysis of DOA estimation algorithms for covariance matrix transforms," *Signal Processing*, vol. 55, pp. 107-115, 1996.
- [13] P. Forster, E. Villier, "Performances asymptotiques des méthodes HR," in *S. Marcos editor, les méthodes à hautes résolution*, Hermes 1998.
- [14] P. Forster, E. Villier, "Simplified formulas for performance analysis of MUSIC and Min Norm," *Proceedings of Ocean conf.*, Sept. 1998.
- [15] H. Cox, "Resolving power and sensivity to mismatch of optimum array processors," *Journal of the Acoustical Society of America*, vol. 54, no. 3, pp. 771-785, 1973.

- [16] S.U. Pillai, F. Haber, "Statistical analysis of a high resolution spatial spectrum estimator utilizing an augmented covariance matrix," *IEEE Trans. Acoustics Speech Signal Processing*, vol. 35, no. 11, pp. 1517-1523, Nov. 1987.
- [17] C.R. Greene, R.C. Wood, "Sparse array performance," *J. Acoust. Soc. Am.*, vol. 63, pp. 1866-1872, 1978.
- [18] M. Rosenblatt, *Stationnary sequences and random fields*, Birkhäuser, 1985.
- [19] A. Gorokhov, "Séparation autodidacte de mélanges convolutifs : méthodes du second ordre," *Rapport de thèse*, ENST Paris, 1997.
- [20] T.H. Li, B. Kedem and S. Yakowitz, "Asymptotic normality of sample autocovariances with an application in frequency estimation," *Stochastic Processes and their Applications*, vol. 52, pp. 329-349, 1994.
- [21] P. Billingsley, *Probability and measure*, Second Edition, John Wiley and Sons, 1986.

Chapitre 5

Robustesse des algorithmes du second ordre d'estimation de DOA face à la largeur de bande

5.1 Introduction

En traitement d'antenne, dans le domaine de l'estimation de direction d'arrivée (DOA) d'ondes planes, on distingue habituellement deux types de situations et par la même deux types de traitement à mettre en œuvre selon que l'enveloppe complexe des signaux reçus peut être ou non considérée comme constante durant la traversée du réseau. Les traitements correspondants, à enveloppe complexe constante, et qualifiés à ce titre de “bande étroite” seront, en règle générale, plus simples à mettre en œuvre que ceux adaptés à des signaux “large bande”, dont la complexité de calcul est beaucoup plus importante (voir par exemple [2] et ses références). Il est alors intéressant d'examiner où se situe la frontière entre ces deux domaines et d'étudier dans quelle mesure, en terme de dégradation de performance, des méthodes bande-étroite peuvent être appliquées à des signaux à bande étendue.

Ce problème ne semble pas avoir été étudié de façon complète précédemment. Schell et Gardner [1] ont mentionné l'approximation bande étroite comme une source possible d'erreur de modèle. Quelques auteurs ont proposé des définitions de la situation bande étroite sans toutefois les relier aux performances des algorithmes de DOA. En se fondant sur un résultat montré dans [3] selon lequel plus de 99.99% de la puissance d'un seul signal reçu est caractérisé par les $r = \lceil 2b\tau_\theta + 1 \rceil$ plus grandes valeurs propres de la matrice de covariance spatio-temporelle (où b est la largeur de bande du signal reçu et τ_θ le temps de propagation total à travers le réseau, déphaseurs compris), Buckley définit comme étant à bande-étroite le cas où $2b\tau_\theta + 1$ est suffisamment voisin de un. Pour des signaux de bande plus large et des temps de propagation à travers le réseau plus importants, cette notion de “rang effectif” permet à Buckley et Griffiths [4] de proposer une méthode “sous-espace” dans un contexte de signaux à bande étendue. Avec une approche similaire, Zatman [5] considère comme bande-étroite un signal pour lequel la seconde valeur propre de la matrice de covariance spatiale non bruitée est inférieure à la puissance du bruit. Afin de traduire cette définition en une condition sur la largeur de bande b du signal (plus précisément sur la largeur de bande réduite $\frac{b}{f_0}$ où f_0 représente la fréquence centrale de la source), il propose d'assimiler une source non bande-étroite à deux sources bande-étroite de même puissance, décorrélées et avec un écartement angulaire fonction de $\frac{b}{f_0}$. Ceci permet d'obtenir une expression de λ_2 , la deuxième valeur propre de la matrice de covariance spatiale non bruitée de la source non bande étroite en fonction de la bande réduite, de la direction d'arrivée et de la puissance de la source ainsi que de la géométrie du réseau. Le critère $\lambda_2 < P_b$ (où P_b désigne la puissance de bruit) permet alors d'obtenir une condition portant sur $\frac{b}{f_0}$ en fonction de la direction d'arrivée, du rapport S/B et de la géométrie du réseau. Plus récemment, l'effet d'une largeur de bande non nulle sur l'estimation de la direction d'arrivée de sources

à spectre symétrique par rapport à la fréquence de démodulation f_0 a été étudié par Sorelius [6].

L’objet de ce chapitre est double. D’une part nous nous proposons d’étendre l’étude faite par Sorelius au cas de spectres non symétriques et/ou non centrés sur la fréquence de démodulation. D’autre part, en considérant la détection d’ordre, la borne de Cramer-Rao et la comparaison des performances asymptotiques entre des algorithmes bande étroite et large bande, nous montrons que la définition empirique de l’hypothèse bande étroite souvent utilisée dans la littérature selon laquelle l’ouverture du réseau est beaucoup plus petite que l’inverse de la bande relative du signal (i.e., $\frac{Mb}{f_0} \ll 1$ où M représente le nombre de capteurs) est bien trop sévère. Pour étudier la robustesse des algorithmes bande-étroite de DOA vis à vis de la largeur de bande du signal, nous ne considérerons que les situations à une ou deux sources décorréliées, de même puissance et de même largeur de bande.

Ce chapitre est organisé de la manière suivante. La modélisation du problème et les notations sont introduites au paragraphe 2, les performances des méthodes classiques de détection de l’ordre, critère d’Akaike (AIC) et critère de “minimum length description” (MDL), en fonction de la largeur de bande sont ensuite présentées au paragraphe 3 et comparées à celles d’un nouveau critère [7] issu de l’analyse numérique. Les biais et covariance asymptotiques (en nombre d’observations et largeur de bande du signal) des estimées de DOA sont étudiés et illustrés au paragraphe 4 dans le cas de l’algorithme MUSIC standard. Au paragraphe 5, le problème de la détermination de la borne de Cramer-Rao du paramètre directionnel seul est abordé. En particulier, une expression analytique commune de la borne de Cramer-Rao pour le paramètre directionnel d’une source unique à bande étroite ou large bande est donnée dans le cas d’une densité spectrale de puissance symétrique quelconque et d’un réseau quelconque. Enfin, le paragraphe 6 est consacré à la comparaison des performances de deux algorithmes classiques de DOA, l’un bande-étroite, l’autre large-bande, appliqués à une ou deux sources large bande.

5.2 Modélisation du problème

Les signaux provenant de K sources sont reçus par un réseau de M capteurs, sont traités par un filtre passe-bande de largeur B autour d’une fréquence f_0 ($B < 2f_0$)¹ puis sont ramenés en bande de base afin d’en extraire l’enveloppe complexe. Si le bruit reçu par le réseau est supposé blanc, l’enveloppe complexe du bruit filtré l’est également, dans la bande $[-\frac{B}{2}, +\frac{B}{2}]$. L’enveloppe complexe de ces signaux est échantillonnée à la fréquence $\frac{1}{T_s} \ll B$ et les M -vecteurs ainsi obtenus à la sortie du réseau aux instants $t = 1, \dots, T$ constituent des observations indépendantes. Les M -vecteurs $\mathbf{n}_t$ correspondants à l’enveloppe complexe du bruit seront constamment supposés temporellement et spatialement non corrélés avec $E(\mathbf{n}_t \mathbf{n}_t^H) = \sigma_n^2 \mathbf{I}_M$ et indépendants des signaux des sources. Si s_t^k représente l’enveloppe complexe de la k ème source par rapport à la fréquence f_0 et $\mu_k(f)$ sa mesure spectrale [8]², l’enveloppe complexe de cette source observée au m ème capteur a pour expression $s_{t-\tau_{k,m}}^k e^{-i2\pi f_0 \tau_{k,m}}$, où $\tau_{k,m}$ est le retard du à la propagation du signal de la k ème source jusqu’au m ème capteur. Les paramètres $\tau_{k,m}$, $m = 1, \dots, M$ contiennent l’information relative à la position Θ_k de cette source par rapport au réseau. Le M -vecteur de l’enveloppe complexe de la sortie du réseau peut alors s’écrire :

$$\mathbf{y}_t = \sum_{k=1}^K \int_{-B/2}^{+B/2} e^{i2\pi f t} \mathbf{a}(\Theta_k, f_0 + f) d\mu_k(f) + \mathbf{n}_t,$$

¹ La densité spectrale de puissance de ces signaux est supposée nulle au voisinage de la fréquence 0.

² Cette représentation spectrale $s_t^k = \int_{-B/2}^{+B/2} e^{i2\pi f t} d\mu_k(f)$ nous permet de générer facilement des signaux stationnaires au sens large et à bande limitée grâce à l’approximation : $s_t^k \approx \sum_{l=0}^{L-1} a_l^k e^{i2\pi f_l t}$ avec f_l uniformément répartie dans $[-B/2, +B/2]$ et $L \gg 1$ (typiquement $L = 200$ dans nos simulations) et $(a_l^k)_{l=0, \dots, L-1, k=1, \dots, K}$ sont des variables aléatoires non corrélées telles que $E|a_l^k|^2 = \frac{B}{L} S_k(f_l)$ où $S_k(f)$ représente la densité spectrale de puissance de s_t^k .

où $\mathbf{a}(\Theta_k, \nu) \stackrel{\text{def}}{=} [e^{-i2\pi\nu\tau_{k,1}}, \dots, e^{-i2\pi\nu\tau_{k,M}}]^T$ est le vecteur directionnel de la k ième source à la fréquence ν . Si les sources sont spatialement non-corrélées, la matrice de covariance spatiale peut s'écrire :

$$\mathbf{R}_b = \mathbb{E}(\mathbf{y}_t \mathbf{y}_t^H) = \sum_{k=1}^K \int_{-B/2}^{+B/2} S_k(f) \mathbf{a}(\Theta_k, f_0 + f) \mathbf{a}^H(\Theta_k, f_0 + f) df + \sigma_n^2 \mathbf{I}_M, \quad (2.1)$$

où $S_k(f)$ représente la densité spectrale de puissance de la k ième source. Dans le cas de signaux bande-étroite, cette matrice devient :

$$\mathbf{R}_0 = \sum_{k=1}^K \sigma_k^2 \mathbf{a}(\Theta_k, f_0) \mathbf{a}^H(\Theta_k, f_0) + \sigma_n^2 \mathbf{I}_M, \quad (2.2)$$

où σ_k^2 représente la puissance de la k ième source. De façon à mettre en lumière les paramètres dont dépend la matrice $\mathbf{R}_b$, nous considérerons que la densité spectrale de puissance $S_k(f)$ de l'enveloppe complexe de la k ième source est paramétrée par sa fréquence centrale $f_m \stackrel{\text{def}}{=} \frac{\int f S_k(f) df}{\int S_k(f) df}$, son écart-type $f_\sigma \stackrel{\text{def}}{=} \left(\frac{\int (f - f_m)^2 S_k(f) df}{\int S_k(f) df} \right)^{\frac{1}{2}}$ ³ et sa puissance σ_k^2 .

Cette densité spectrale de puissance peut en effet s'écrire :

$$S_k(f) = \frac{\sigma_k^2}{f_\sigma} S\left(\frac{f - f_m}{f_\sigma}\right),$$

où $S(f)$ est une fonction normalisée, i.e. vérifiant :

$$\int S(f) df = 1 \quad \int f S(f) df = 0 \quad \int f^2 S(f) df = 1. \quad (2.3)$$

De ce fait la fonction de corrélation de la k ième source peut s'écrire : $R_k(t) = \sigma_k^2 R(f_\sigma t) e^{i2\pi f_m t}$ où $R(t)$ désigne la fonction de corrélation associée à $S(f)$ qui, compte tenu de (2.3), est telle que :

$$R(0) = 1 \quad , \quad R'(0) = 0 \quad , \quad R''(0) = -4\pi^2. \quad (2.4)$$

Comme $\mathbf{a}(\Theta_k, f_0 + f) \mathbf{a}^H(\Theta_k, f_0 + f) = \mathbf{a}(\Theta_k, f_0) \mathbf{a}^H(\Theta_k, f_0) \odot \mathbf{a}(\Theta_k, f) \mathbf{a}^H(\Theta_k, f)$, la matrice de covariance $\mathbf{R}_b$ peut s'écrire :

$$\mathbf{R}_b = \sum_{k=1}^K \mathbf{a}(\Theta_k, f_0) \mathbf{a}^H(\Theta_k, f_0) \odot \mathbf{R}_{s_k} + \sigma_n^2 \mathbf{I}_M \quad (2.5)$$

où $\mathbf{R}_{s_k}$ est la matrice $M \times M$ de corrélation de la k ième source de terme général :

$$\begin{aligned} [\mathbf{R}_{s_k}]_{m,n} &= \mathbb{E}(s_{t-\tau_{k,m}}^k s_{t-\tau_{k,n}}^{k*}) = \int_{-B/2}^{+B/2} S_k(f) e^{i2\pi f(\tau_{k,n} - \tau_{k,m})} df = \sigma_k^2 \int S(\nu) e^{i2\pi(\nu f_\sigma + f_m)(\tau_{k,n} - \tau_{k,m})} d\nu \\ &= \sigma_k^2 R(f_\sigma(\tau_{k,n} - \tau_{k,m})) e^{i2\pi f_m(\tau_{k,n} - \tau_{k,m})}. \end{aligned} \quad (2.6)$$

Si $\mathbf{z}_m$ est le vecteur des coordonnées du m ième capteur par rapport à un capteur choisi comme origine et $\mathbf{p}_k$, le vecteur d'onde unitaire de la k ième source, le retard différentiel de cette source sur le m ième est $\tau_{k,m} = \frac{1}{f_0} \frac{\mathbf{z}_m^T \mathbf{p}_k}{\lambda_0}$. Par exemple, pour un réseau linéaire, $\tau_{k,m} = \frac{d_m}{c} \sin \theta_k = \frac{1}{f_0} \frac{d_m}{\lambda_0} \sin \theta_k$ où θ_k est l'angle d'arrivée de la source par rapport au réseau, c est la vitesse de propagation de l'onde, d_m est la distance

³Dans le cas particulier de sources à spectre uniforme sur $[f_0 - \frac{b}{2}, f_0 + \frac{b}{2}]$, $f_m = 0$ et $f_\sigma = \frac{b}{2\sqrt{3}}$.

du m ème capteur au capteur origine et $\lambda_0 \stackrel{\text{def}}{=} \frac{c}{f_0}$.

Nous montrons en Annexe 1 que $\mathbf{R}_{s_k} = \sigma_k^2 \mathbf{1}\mathbf{1}^T + \delta\mathbf{R}_{s_k}$ avec

$$[\delta\mathbf{R}_{s_k}]_{m,n} = \sigma_k^2 \alpha_{m,n,k} \left[i2\pi \left(\frac{f_m}{f_0} \right) - 2\pi^2 \alpha_{m,n,k} \left(\frac{f_\sigma^2}{f_0^2} \right) \right] + O \left(\frac{f_m^2}{f_0^2} \right) + O \left(\frac{f_\sigma^3}{f_0^3} \right) + O \left(\frac{f_m f_\sigma^2}{f_0^3} \right) \quad (2.7)$$

pour un spectre quelconque

$$= -2\pi^2 \sigma_k^2 \alpha_{m,n,k}^2 \left(\frac{f_\sigma^2}{f_0^2} \right) + O \left(\frac{f_\sigma^3}{f_0^3} \right) \quad (2.8)$$

pour un spectre quelconque centré sur f_0 ,

$$= -2\pi^2 \sigma_k^2 \alpha_{m,n,k}^2 \left(\frac{f_\sigma^2}{f_0^2} \right) + O \left(\frac{f_\sigma^4}{f_0^4} \right) \quad (2.9)$$

pour un spectre symétrique centré sur f_0 ,

avec $\alpha_{m,n,k} \stackrel{\text{def}}{=} \frac{(\mathbf{z}_n - \mathbf{z}_m)^T \mathbf{p}_k}{\lambda_0}$.

On peut donc considérer $\mathbf{R}_b$ comme résultant d'une perturbation de $\mathbf{R}_0$:

$$\mathbf{R}_b = \mathbf{R}_0 + \delta\mathbf{R}_b \quad \text{avec} \quad \delta\mathbf{R}_b = \sum_{k=1}^K \mathbf{a}(\Theta_k, f_0) \mathbf{a}^H(\Theta_k, f_0) \odot \delta\mathbf{R}_{s_k} \quad (2.10)$$

5.3 Performance des méthodes de détection de l'ordre

Considérons tout d'abord les performances des critères de détection de l'ordre fondées sur la décomposition en éléments propres. Plusieurs critères de ce type ont été proposées. Les plus usitées utilisent les valeurs propres de la matrice de covariance spatiale $\mathbf{R}_b(T)$, estimée à partir de T observations indépendantes $\mathbf{y}_t$. Les critères d'Akaike (AIC) et du "minimum description length" (MDL) en constituent deux exemples typiques. Ces critères s'implémentent facilement et, à condition que les signaux soient effectivement bande-étroite, MDL fournit une estimée consistante du nombre de sources, indépendamment du rapport S/B. Malheureusement, si la largeur de bande des signaux augmente, les critères AIC et MDL ont rapidement tendance à surestimer le nombre de sources et, dans la pratique, le nombre estimé de sources est assez éloigné de la vraie valeur [9].

En fait, la détection du nombre de sources "non bande-étroite" à partir de la matrice de covariance estimée est similaire à la détection de l'ordre "effectif" de la réponse impulsionnelle d'un canal à partir de la matrice de covariance spatio-temporelle. La matrice de covariance spatiale estimée $\mathbf{R}_b(T)$ obtenue dans une situation "non bande-étroite" peut être considérée comme la somme d'une matrice $\mathbf{R}$, "idéale" de rang K et d'une matrice de "perturbation" $\delta\mathbf{R}_{b,T}$,

$$\mathbf{R}_b(T) = \mathbf{R} + \delta\mathbf{R}_{b,T}$$

où K est le nombre de sources et $\mathbf{R}$ est la matrice de covariance spatiale non bruitée correspondant à K sources bande-étroite. La matrice de perturbation $\delta\mathbf{R}_{b,T}$ rend compte du fait que les sources ne sont pas bande-étroite, de l'influence du bruit additif et des erreurs d'estimation. En utilisant le concept d'angles canoniques entre sous espaces et des résultats de perturbation de sous espaces invariants, Liavas *et al.* [7] ont proposé une décomposition "maximalement stable" de l'espace engendré par la matrice de covariance spatiale en sous espace signal et sous espace bruit dans le cadre de la détection de l'ordre d'un canal. Cette approche a abouti au critère suivant ; l'ordre estimé $\hat{K}$ est la valeur de k qui minimise

$$r(k) \stackrel{\text{def}}{=} \begin{cases} \frac{\lambda_{k+1}}{\lambda_k - 2\lambda_{k+1}}, & \text{if } \lambda_{k+1} \leq \frac{\lambda_k}{3} \\ 1, & \text{sinon} \end{cases}$$

où $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$ représentent les valeurs propres de la matrice estimée $\mathbf{R}_b(T)$. Nous proposons d'utiliser ce critère dans le cas du traitement d'antenne de signaux non bande-étroite. A la différence des

critères AIC et MDL qui sont fondées sur la similitude des plus petites valeurs propres, le critère proposé repose sur l'existence d'un saut entre les valeurs propres. Les simulations (cf. FIG.5.1 et [5]) montrent que, au fur et à mesure que la largeur de bande du signal augmente, les valeurs propres émergent du niveau de bruit les unes après les autres. Le critère proposé semble de ce fait prometteur.

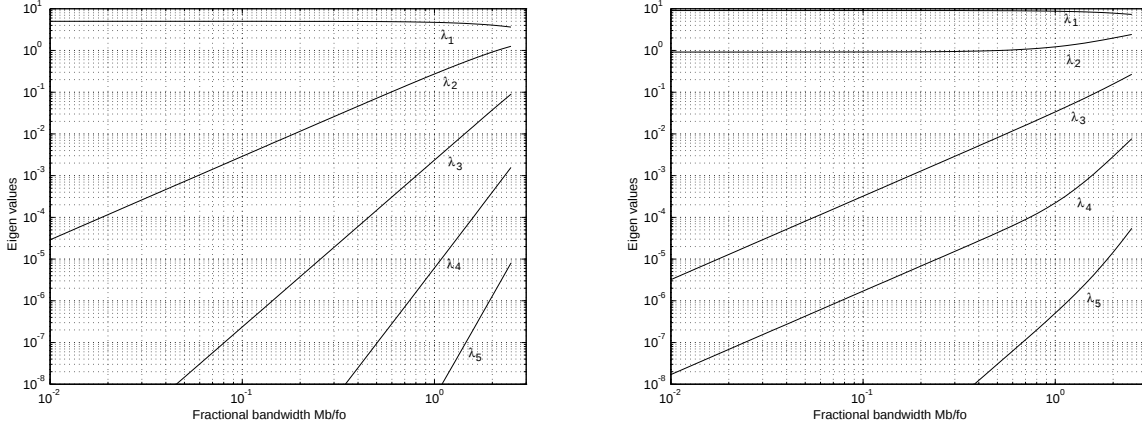


FIG. 5.1: Valeurs propres de la matrice de covariance spatiale non bruitée correspondant à une ou deux sources de même puissance, de spectre uniforme et séparées de 15° et un réseau linéaire de $M = 5$ capteurs en fonction de la largeur de bande réduite $\frac{Mb}{f_0}$.

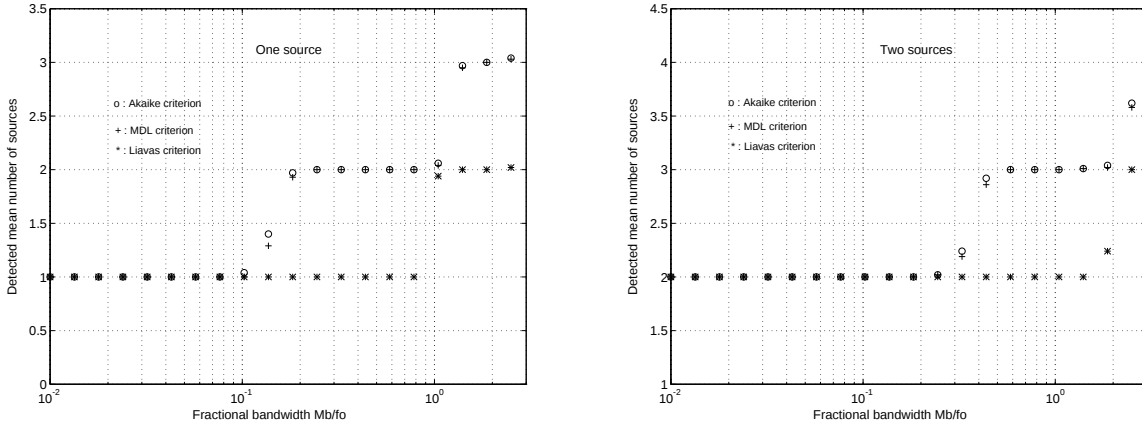


FIG. 5.2: Nombre moyen de sources détectées (100 réalisations) par les critères d'Akaike, MDL et de Liavas en fonction de la largeur de bande réduite $\frac{Mb}{f_0}$ pour un réseau linéaire de $M = 5$ capteurs, une ou deux sources (de même puissance, de spectre uniforme et séparées de 15° , un rapport S/B = 20dB et $T = 160$ observations).

La FIG.5.2 montre que le critère de Liavas est beaucoup moins sensible à l'augmentation de la largeur de bande du signal que les critères AIC et MDL pour un nombre fini d'observations. Considérons le comportement asymptotique de ces différents estimateurs lorsque la taille de l'observation tend vers l'infini.

Les FIG.5.3⁴ et FIG.5.4 montrent, respectivement pour une et deux sources, les domaines de la bande réduite $\frac{Mb}{f_0}$ dans lesquels les critères MDL et Liavas estiment correctement le nombre de sources pour des observations de grande taille (i.e. en statistiques exactes).

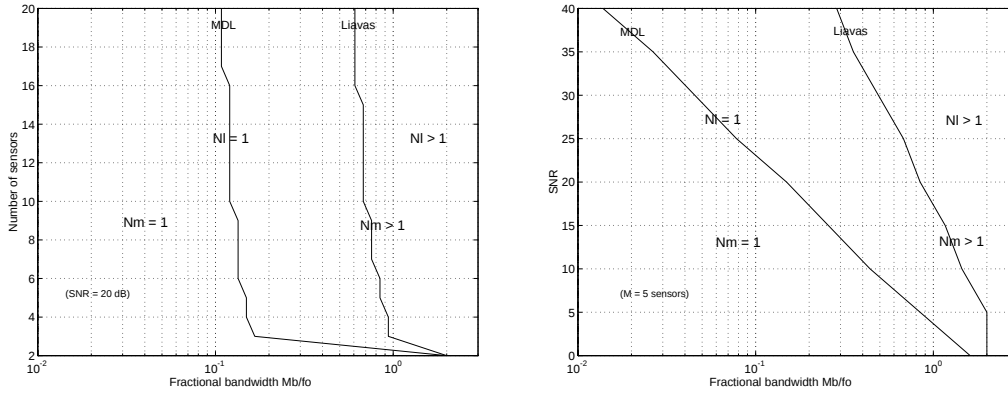


FIG. 5.3: Nombre asymptotique (par rapport au nombre d’observations) de sources détectées par les critères MDL (N_m) et Liavas (N_l) en fonction de la largeur de bande réduite $\frac{Mb}{f_0}$ (en abscisse) et du nombre de capteurs ou du rapport S/B (en ordonnée) dans le cas d’un réseau linéaire uniforme et d’une source à spectre uniforme.

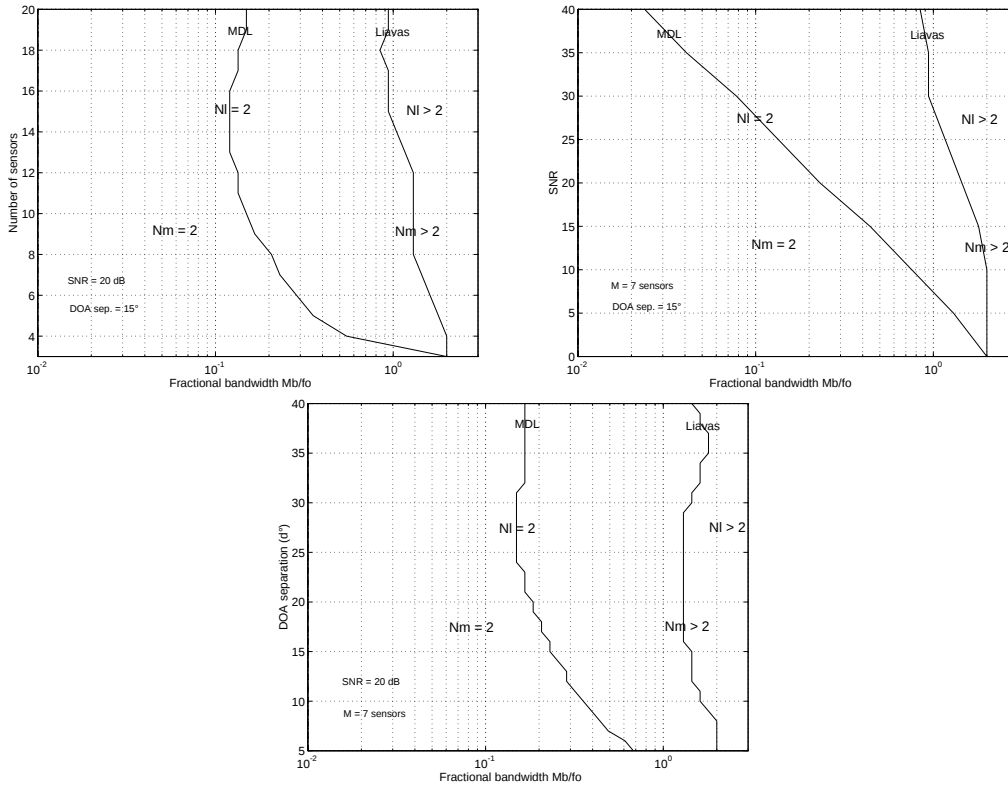


FIG. 5.4: Nombre asymptotique (par rapport au nombre d’observations) de sources détectées par les critères MDL (N_m) et Liavas (N_l) en fonction de la largeur de bande réduite $\frac{Mb}{f_0}$ (en abscisse) et du nombre de capteurs, du rapport S/B ou de la séparation angulaire (en ordonnée) dans le cas d’un réseau linéaire uniforme et de deux sources de même puissance et à spectre uniforme.

⁴ La FIG.5.3 par exemple doit être interprétée de la façon suivante : le critère MDL détecte une source ($N_m=1$) dans la zone à gauche de la courbe “MDL” soit pour $\frac{Mb}{f_0}$ approximativement plus petit que 1.5×10^{-1} et sinon plus d’une source alors que le critère Liavas détecte une source ($N_l=1$) dans la zone à gauche de la courbe “Liavas” soit pour $\frac{Mb}{f_0}$ approximativement plus petite que 7×10^{-1} et sinon plus d’une source.

Dans les situations décrites par ces figures le critère de Liavas surclasse toujours le critère MDL. Par ailleurs nous pouvons remarquer que les performances des critères MDL et de Liavas se dégradent lorsque le rapport S/B augmente (bien que le critère de Liavas soit plus stable) et reste relativement constant lorsque le nombre de capteurs augmente. Si nous considérons maintenant l'influence de la séparation angulaire entre deux sources de même puissance, nous constatons que les performances du critère de Liavas restent constantes alors que celles du critère MDL se dégradent lorsque l'écart angulaire augmente.

Une comparaison exhaustive de ces deux critères nécessiterait de considérer un grand nombre de situations différentes, pour différentes valeurs des paramètres en question : nombre de capteurs, écarts angulaire, spectres et rapport S/B. Par exemple, une étude spécifique du cas de deux sources de même puissance et même spectre uniforme et d'un réseau linéaire uniforme montre que parfois le critère de Liavas ne parvient pas à détecter l'une des sources pour un faible nombre de capteurs, un bon rapport S/B et un faible écart angulaire alors que le critère MDL détecte correctement les deux. Il en est de même dans le cas de sources de puissances différentes.

5.4 Biais et covariance asymptotiques

Pour étudier le biais et la covariance asymptotiques (en nombre d'observations et en largeur de bande du signal) des estimées des directions d'arrivées par des méthodes bande-étroite du second ordre, nous adoptons une approche fonctionnelle consistant à considérer l'ensemble du processus de calcul d'une estimation $\Theta(T)$ de Θ comme une fonctionnelle qui, appliquée à la matrice estimée $\mathbf{R}_b(T) = \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^H$, fournit cette estimation $\Theta(T)$. Cette relation fonctionnelle est notée $\Theta(T) = \text{alg}(\mathbf{R}_b(T))$. Les différentes méthodes bande-étroite du second ordre sont telles que $\Theta = \text{alg}(\mathbf{R}_0)$ et de ce fait peuvent être considérées comme autant d'extensions de l'application $\mathbf{R}_0 \rightarrow \Theta$ à n'importe quelle matrice Hermitienne non structurée $\mathbf{R}_b(T)$. On peut considérer que $\mathbf{R}_b(T)$ résulte d'une perturbation de $\mathbf{R}_b$:

$$\mathbf{R}_b(T) = \mathbf{R}_b + \delta \mathbf{R}_T, \quad (4.1)$$

où $\delta \mathbf{R}_T$ est l'erreur d'estimation, vérifiant $E(\delta \mathbf{R}_T) = \mathbf{O}$ et $\text{Cov}(\delta \mathbf{R}_T) = O(\frac{1}{T})$. Pour la plupart des algorithmes du second ordre, l'application $\text{alg}(\cdot)$ est suffisamment régulière pour que l'on puisse écrire, à partir de (4.1),

$$\Theta(T) = \text{alg}(\mathbf{R}_b) + (D_{\mathbf{R}_b}^{\text{alg}}, \delta \mathbf{R}_T) + O(\|\delta \mathbf{R}_T\|^2), \quad (4.2)$$

où $(D_{\mathbf{R}_b}^{\text{alg}}, \delta \mathbf{R}_T)$ représente la différentielle de l'application $\text{alg}(\cdot)$, évaluée au point $\mathbf{R}_b$ et appliquée à $\delta \mathbf{R}_T$. On en déduit, en calculant l'espérance mathématique :

$$E(\Theta(T)) = \text{alg}(\mathbf{R}_b) + O\left(\frac{1}{T}\right). \quad (4.3)$$

En considérant $\mathbf{R}_b$ comme résultant d'une perturbation de $\mathbf{R}_0$, une analyse au premier ordre du résultat de l'application d'un algorithme du second ordre à $\mathbf{R}_b$ évalué au point $\mathbf{R}_0$ donne :

$$\begin{aligned} \text{alg}(\mathbf{R}_b) &= \text{alg}(\mathbf{R}_0) + (D_{\mathbf{R}_0}^{\text{alg}}, \delta \mathbf{R}_b) + O(\|\delta \mathbf{R}_b\|^2) \\ &= \Theta + \mathbf{D}_{\mathbf{R}_0}^{\text{alg}} \text{Vec}(\delta \mathbf{R}_b) + O(\|\delta \mathbf{R}_b\|^2), \end{aligned} \quad (4.4)$$

où $\mathbf{D}_{\mathbf{R}_0}^{\text{alg}}$ ⁵ représente la matrice de la différentielle de l'algorithme bande étroite du second ordre $\text{alg}(\cdot)$ au point $\mathbf{R}_0$. Grâce à (4.3) et (4.4), on obtient alors le résultat suivant :

⁵ Les expressions de $\mathbf{D}_{\mathbf{R}_0}^{\text{alg}}$ et de $\mathbf{D}_{\mathbf{R}_b}^{\text{alg}}$ peuvent être obtenus grâce à un calcul de perturbation.

Résultat 5.1 *Le biais asymptotique (par rapport au nombre d'observations et à la largeur de bande du signal) de l'estimation $\Theta(T)$ fournie par un algorithme bande étroite du second ordre est donné par :*

$$E(\Theta(T)) - \Theta = \mathbf{D}_{\mathbf{R}_0}^{\text{alg}} \text{Vec}(\delta \mathbf{R}_b) + O(\|\delta \mathbf{R}_b\|^2) + O\left(\frac{1}{T}\right) \quad \text{pour un spectre quelconque} \quad (4.5)$$

$$= \sum_{k=1}^K \mathbf{b}_k^{\text{alg}} \left(\frac{f_\sigma^2}{f_0^2} \right) + O\left(\frac{f_\sigma^3}{f_0^3}\right) + O\left(\frac{1}{T}\right) \quad \text{pour un spectre quelconque centré sur } f_0 \quad (4.6)$$

$$= \sum_{k=1}^K \mathbf{b}_k^{\text{alg}} \left(\frac{f_\sigma^2}{f_0^2} \right) + O\left(\frac{f_\sigma^4}{f_0^4}\right) + O\left(\frac{1}{T}\right) \quad \text{pour un spectre symétrique centré sur } f_0 \quad (4.7)$$

avec $\mathbf{b}_k^{\text{alg}} \stackrel{\text{def}}{=} \mathbf{D}_{\mathbf{R}_0}^{\text{alg}} \text{Vec}(\mathbf{a}(\Theta_k, f_0) \mathbf{a}^H(\Theta_k, f_0) \odot \sigma_k^2 \mathbf{U}_k)$ où $[\mathbf{U}_k]_{m,n} \stackrel{\text{def}}{=} -2\pi^2 \alpha_{m,n,k}^2$.

Les relations (4.2) et (4.3) nous fournissent une expression asymptotique de la différence entre l'estimation et son espérance mathématique :

$$\Theta(T) - E(\Theta(T)) = \mathbf{D}_{\mathbf{R}_b}^{\text{alg}} \text{Vec}(\delta \mathbf{R}_T) + O\left(\frac{1}{T}\right) + O(\|\delta \mathbf{R}_b\|^2). \quad (4.8)$$

On en déduit alors une expression asymptotique de la covariance de l'estimateur :

Résultat 5.2 *La covariance asymptotique (par rapport au nombre d'observations et à la largeur de bande du signal) de l'estimation $\Theta(T)$ fournie par un algorithme bande étroite du second ordre est donné par :*

$$\begin{aligned} \text{Cov}(\Theta(T)) &= \frac{1}{T} \mathbf{D}_{\mathbf{R}_b}^{\text{alg}} \mathbf{C}_{R_b} \left(\mathbf{D}_{\mathbf{R}_b}^{\text{alg}} \right)^H + O\left(\frac{1}{T^2}\right) + O\left(\frac{1}{T}\right) O(\|\delta \mathbf{R}_b\|^2) + O(\|\delta \mathbf{R}_b\|^4) \\ &= \frac{1}{T} \mathbf{D}_{\mathbf{R}_b}^{\text{alg}} (\mathbf{R}_0 \otimes_c \mathbf{R}_0) \left(\mathbf{D}_{\mathbf{R}_b}^{\text{alg}} \right)^H + \frac{1}{T} \mathbf{D}_{\mathbf{R}_b}^{\text{alg}} (\mathbf{R}_0 \otimes_c \delta \mathbf{R}_b) \left(\mathbf{D}_{\mathbf{R}_b}^{\text{alg}} \right)^H + \frac{1}{T} \mathbf{D}_{\mathbf{R}_b}^{\text{alg}} (\delta \mathbf{R}_b \otimes_c \mathbf{R}_0) \left(\mathbf{D}_{\mathbf{R}_b}^{\text{alg}} \right)^H \\ &+ O\left(\frac{1}{T^2}\right) + O\left(\frac{1}{T}\right) O(\|\delta \mathbf{R}_b\|^2) + O(\|\delta \mathbf{R}_b\|^4) \end{aligned} \quad (4.9)$$

avec $\mathbf{C}_{R_b} = \lim_{T \rightarrow \infty} T \text{Cov}(\text{Vec}(\mathbf{R}_b(T))) = \lim_{T \rightarrow \infty} T E(\text{Vec}(\delta \mathbf{R}_T) \text{Vec}^H(\delta \mathbf{R}_T)) = \mathbf{R}_b \otimes_c \mathbf{R}_b$ pour des observations complexes circulaires indépendantes $\mathbf{y}_t$, [10, p. 336].

Ces résultats généraux étendent, au cas de spectres non symétriques et/ou non centrés, les résultats présentés dans [6] dans le cadre de spectres symétriques.

Le paragraphe suivant propose une illustration des ces résultats généraux dans le cas de l'algorithme MUSIC classique et dans le cas de deux situations particulières : une seule source ou deux sources de même puissance.

Rappelons tout d'abord que la matrice associée à la différentielle $\mathbf{D}_{\mathbf{R}_0}^{\text{alg}}$ de l'algorithme MUSIC standard, obtenue par un calcul de perturbation (cf. [11] par exemple) est donnée par :

$$\mathbf{D}_{\mathbf{R}_0}^{\text{music}}(l, :) = \frac{1}{\alpha_{\theta_l}} \left(\mathbf{a}_{\theta_l}^H \mathbf{\Gamma}_s \otimes_c \mathbf{a}_{\theta_l}'^H \mathbf{\Pi}_n + \mathbf{a}_{\theta_l}'^H \mathbf{\Pi}_n \otimes_c \mathbf{a}_{\theta_l}^H \mathbf{\Gamma}_s \right), \quad l = 1, \dots, K \quad (4.10)$$

où $\mathbf{\Pi}_n$ et $\mathbf{\Gamma}_s$ représentent respectivement la matrice de projection orthogonale sur le sous espace bruit associé à $\mathbf{R}_0$ et la matrice pseudo-inverse de Moore-Penrose $(\mathbf{R}_0 - \sigma_n^2 \mathbf{I}_M)^\#$. Dans ce paragraphe nous

poserons pour simplifier l'écriture $\mathbf{a}_{\theta_l} \stackrel{\text{def}}{=} \mathbf{a}(\Theta_l, f_0)$, $l = 1, \dots, K$ et $\alpha_{\theta_l} = 2\mathbf{a}_{\theta_l}^H \mathbf{\Pi}_n \mathbf{a}_{\theta_l}'$ avec $\mathbf{a}_{\theta_l}' \stackrel{\text{def}}{=} \frac{d\mathbf{a}_{\theta_l}}{d\theta_l}$. Avec ces notations, le biais asymptotique devient :

$$\begin{aligned} [\mathbf{D}_{\mathbf{R}_0}^{\text{music}} \text{Vec}(\delta \mathbf{R}_b)]_l &= \frac{2}{\alpha_{\theta_l}} \Re \left(\mathbf{a}_{\theta_l}^H \mathbf{\Gamma}_s \delta \mathbf{R}_b \mathbf{\Pi}_n \mathbf{a}_{\theta_l}' \right) = \frac{2}{\alpha_{\theta_l}} \Re \left(\mathbf{a}_{\theta_l}^H \mathbf{\Gamma}_s (\mathbf{R}_b - \sigma_n^2 \mathbf{I}_M) \mathbf{\Pi}_n \mathbf{a}_{\theta_l}' \right) \\ &= \frac{2}{\alpha_{\theta_l}} \sum_{k=1}^K \int_{-B/2}^{+B/2} S_k(f) \Re \left(\mathbf{a}_{\theta_k}^H \mathbf{\Gamma}_s \Delta_{\theta_k, f} \mathbf{a}_{\theta_k} \mathbf{a}_{\theta_k}^H \Delta_{\theta_k, f}^H \mathbf{\Pi}_n \mathbf{a}_{\theta_l}' \right) df \end{aligned} \quad (4.11)$$

où $\Re(\cdot)$ représente “la partie réelle de” et $\Delta_{\theta_k, f} \stackrel{\text{def}}{=} \text{Diag}(e^{-i2\pi f \tau_{k,1}}, \dots, e^{-i2\pi f \tau_{k,M}})$.

5.4.1 Cas d'une seule source

Dans le cas d'une seule source, nous avons les résultats suivants, démontrés dans l'Annexe 2:

Résultat 5.3 *Le biais asymptotique (par rapport au nombre d'observations et à la largeur de bande du signal) est donné par :*

$$\begin{aligned} E(\Theta_1(T)) - \Theta_1 &= \frac{8\pi^2}{M\alpha_{\theta_1}} \left(\frac{f_m}{f_0} \right) \sum_{n,m=1}^M \frac{\mathbf{z}_n^T \mathbf{p}_1'}{\lambda_0} \frac{(\mathbf{z}_n - \mathbf{z}_m)^T \mathbf{p}_1}{\lambda_0} + O\left(\frac{f_\sigma^3}{f_0^3}\right) + O\left(\frac{f_m^2}{f_0^2}\right) \\ &+ O\left(\frac{f_\sigma^2 f_m}{f_0^3}\right) + O\left(\frac{1}{T}\right) \quad \text{pour un spectre quelconque} \end{aligned} \quad (4.12)$$

$$= O\left(\frac{f_\sigma^3}{f_0^3}\right) + O\left(\frac{1}{T}\right) \quad \text{pour un spectre centré sur } f_0 \quad (4.13)$$

$$= O\left(\frac{f_\sigma^4}{f_0^4}\right) + O\left(\frac{1}{T}\right) \quad \text{pour un spectre symétrique par rapport à } f_0 \quad (4.14)$$

avec $\mathbf{p}_1' \stackrel{\text{def}}{=} \frac{d\mathbf{p}_1}{d\Theta_1}$.

Dans le cas d'un réseau linéaire quelconque et d'un signal à spectre symétrique par rapport à $f_0 + f_m$

$$E(\phi_1(T)) - \phi_1 = \phi_1 \frac{f_m}{f_0} + O\left(\frac{f_\sigma^4}{f_0^4}\right) + O\left(\frac{1}{T}\right) \quad (4.15)$$

où $\phi_1 = \pi \sin \theta_1$ est la pulsation spatiale et θ_1 l'angle d'arrivée par rapport à la normale du réseau linéaire.

Notons que l'expression (4.14) du biais asymptotique dans la cas d'un réseau quelconque et d'un spectre symétrique est plus précise que l'expression $\beta\left(\frac{f_m^2}{f_0^2}\right) + O\left(\frac{1}{T}\right)$ donnée dans [6]. De plus le calcul numérique du biais asymptotique $\text{music}(\mathbf{R}_b) - \theta_1$ (par rapport au nombre d'échantillons) montre que ce biais asymptotique vérifie $|\text{music}(\mathbf{R}_b) - \theta_1| < 10^{-6}$ pour une source de spectre uniforme et pour des réseaux linéaires ou circulaires uniformes comptant jusqu'à 20 capteurs et ceci indépendamment de la valeur de b^6 . Le résultat 5.3 montre en outre que le comportement de l'estimateur est fortement influencé par le fait que le spectre du signal soit centré ou non sur f_0 et, dans le cas où le spectre est centré sur f_0 , par le fait que le spectre soit symétrique ou non par rapport à f_0 .

Concernant maintenant la variance asymptotique (en nombre d'observations et en largeur de bande du signal), nous montrons en Annexe 3 le résultat suivant concernant les signaux à spectre centré sur la fréquence f_0 .

⁶Nous notons qu'une évaluation précise de ce biais asymptotique à partir de calculs numériques n'a pu être obtenue en raison du caractère très “plat” de la fonction de localisation de MUSIC au voisinage de l'extremum et de la précision limitée de l'outil de calcul Matlab.

Résultat 5.4

$$\text{Var}(\theta_1^b(T)) = \text{Var}(\theta_1^0(T)) \left(1 + c \left(\frac{f_\sigma^2}{f_0^2} \right) + O \left(\frac{f_\sigma^4}{f_0^4} \right) \right) + O \left(\frac{1}{T^2} \right) \quad (4.16)$$

où $\text{Var}(\theta_1^0(T))$ est la variance asymptotique de l'algorithme MUSIC dans le cas bande étroite (cf. [12, rel. (3.12)] par exemple) et où c est donné en Annexe 3.

Le rôle crucial, dans les performances, de la position du spectre du signal par rapport à la fréquence f_0 est illustré par les FIG.5.5 et FIG.5.6. La FIG.5.5 présente les courbes des iso-valeurs de l'erreur quadratique moyenne sur l'estimation de la pulsation spatiale ϕ_1 par l'algorithme MUSIC standard en fonction, d'une part de $\frac{f_m}{f_0}$, l'écart réduit de la fréquence centrale à la fréquence f_0 et d'autre part de $\frac{b}{f_0}$, la largeur de bande réduite du signal pour une source de spectre uniforme. La FIG.5.6 représente l'erreur quadratique moyenne en fonction de la largeur de bande réduite du signal pour différentes valeurs de l'écart réduit de la fréquence centrale à la fréquence f_0 .

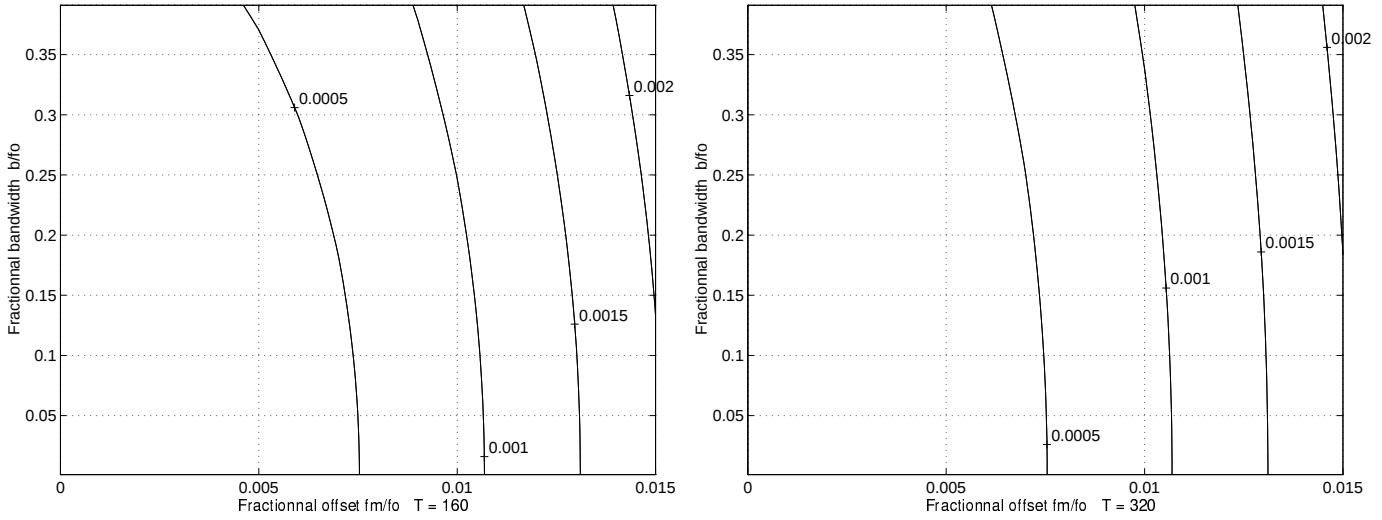


FIG. 5.5: Courbes des iso-valeurs de l'erreur quadratique moyenne en fonction la largeur de bande réduite du signal $\frac{b}{f_0}$ et du décentrage réduit $\frac{f_m}{f_0}$ dans le cas d'une source de spectre uniforme, d'un réseau linéaire uniforme de 5 capteurs, d'un S/B = 20dB et de $T = 160$ ou 320 observations.

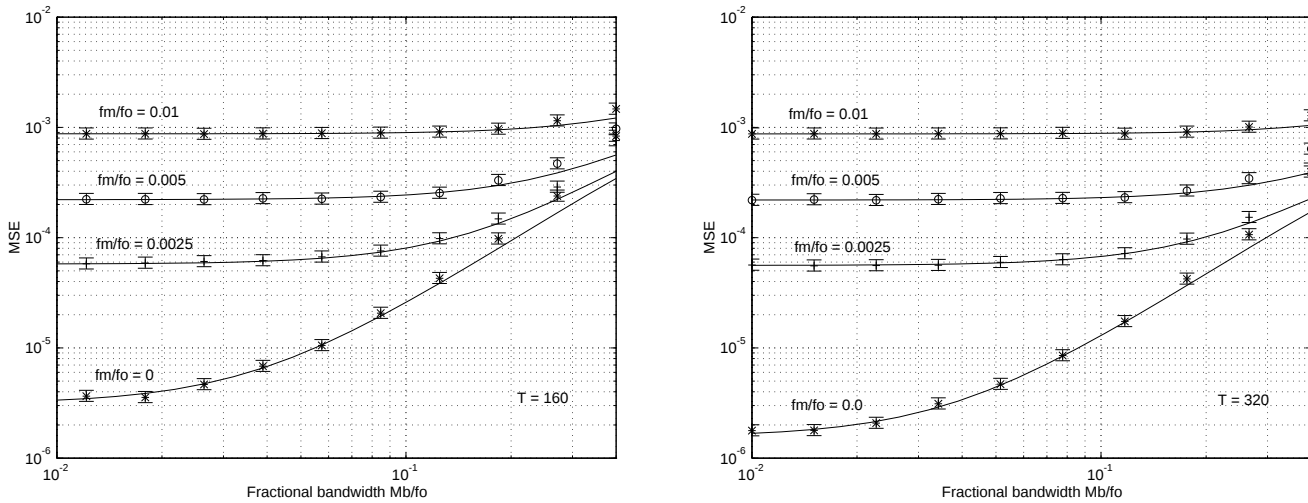


FIG. 5.6: Erreur quadratique moyenne en fonction de $\frac{b}{f_0}$ pour différentes valeurs de $\frac{f_m}{f_0}$ avec un S/B = 20dB et $T = 160$ ou 320 observations.

L'erreur quadratique moyenne est donnée par :

$$\text{EQM}(\phi_1(T)) = \text{biais}^2(\phi_1(T)) + \text{Var}(\phi_1(T)) \quad (4.17)$$

où biais et variance sont donnés respectivement par $\phi_1 \frac{f_m}{f_0}$ (cf. (4.15)) et par $\frac{1}{T} \mathbf{D}_{\mathbf{R}_b}^{\text{music}}(\mathbf{R}_0 \otimes_c \mathbf{R}_0 + \mathbf{R}_0 \otimes_c \delta \mathbf{R}_b + \delta \mathbf{R}_b \otimes_c \mathbf{R}_0) \left(\mathbf{D}_{\mathbf{R}_b}^{\text{music}} \right)^H$ (cf. (4.9)) avec $\delta \mathbf{R}_b$ donné par (2.9) et (2.10). Le terme principal de cette EQM peut s'écrire :

$$\text{EQM}(\phi_1(T)) = \phi_1^2 \left(\frac{f_m^2}{f_0^2} \right) + \left[\frac{c_0}{T} + \frac{c_1}{T} \left(\frac{f_m}{f_0} \right) + \frac{c_2}{T} \left(\frac{f_\sigma^2}{f_0^2} \right) \right], \quad (4.18)$$

où c_0/T représente la variance asymptotique (en nombre d'observations) de l'algorithme MUSIC appliqué à une source bande étroite. Nous voyons le rôle clef joué par l'offset en fréquence f_m entre la fréquence centrale du spectre et la fréquence de démodulation f_0 : les algorithmes bande étroite du second ordre sont beaucoup plus sensibles à cet offset en fréquence qu'à la largeur de bande du signal. De plus, le biais étant indépendant de T , cette sensibilité à f_m s'accroît avec le nombre d'observations comme le montre la FIG.5.6.

5.4.2 Cas de deux sources

Dans le cas de deux sources, même si les spectres de ces sources sont symétriques par rapport à f_0 , les termes $\mathbf{D}_{\mathbf{R}_0}^{\text{music}}(l, :) \text{Vec}(\delta \mathbf{R}_b)$, $l = 1, 2$ ne s'annulent pas et de ce fait les estimations des directions d'arrivée sont asymptotiquement (par rapport au nombre d'observations et à la largeur de bande) biaisées. Nous montrons en Annexe 4 que, dans le cas particulier de deux sources à spectre centré sur f_0 , les termes précédents peuvent s'écrire

$$\mathbf{D}_{\mathbf{R}_0}^{\text{music}}(l, :) \text{Vec}(\delta \mathbf{R}_b) = c_l \left(\frac{f_\sigma^2}{f_0^2} \right) + O \left(\frac{f_\sigma^4}{f_0^4} \right), \quad l = 1, 2, \quad (4.19)$$

où l'expression de c_l est donnée en Annexe 4.

Grâce à cela, trois expressions asymptotiques (par rapport au nombre d'observations et à la largeur de bande) du biais peuvent être obtenues ; les deux premières analytiques et la dernière, numérique :

$$(a) \mathbf{D}_{\mathbf{R}_0}^{\text{music}}(l, :) \text{Vec}(\delta \mathbf{R}_b), \quad (b) c_l \left(\frac{f_\sigma^2}{f_0^2} \right) \quad \text{et} \quad (c) \text{music}_l(\mathbf{R}_b) - \theta_l, \quad l = 1, 2.$$

Avec les valeurs des paramètres de la FIG.5.7, on constate que ces trois valeurs sont très voisines jusqu'à $\frac{Mb}{f_0} = 1$.

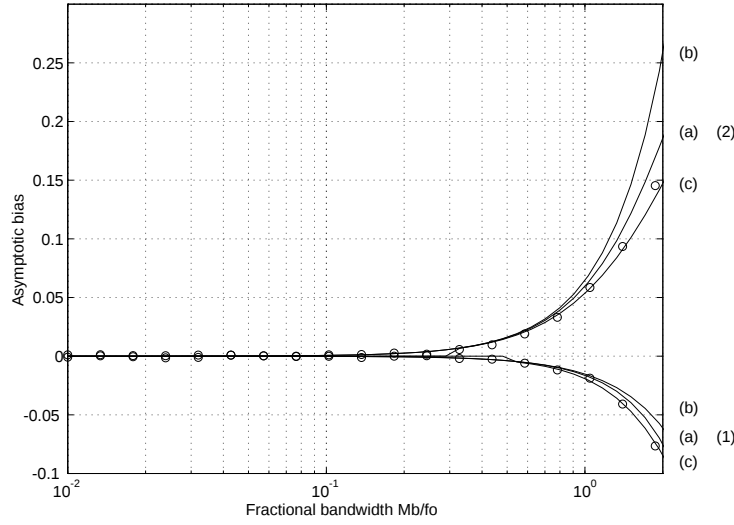


FIG. 5.7: Biases asymptotiques (en b et T) (a) (b) et (c) et estimés par méthode de Monte Carlo (100 réalisations) (o) des pulsations spatiales $\phi_k = \pi \sin \theta_k$ de deux sources de même puissance, de spectre uniforme centré sur f_0 et d'azimuts respectifs $\theta_1 = 25^\circ$ (1), $\theta_2 = 40^\circ$ (2) pour un réseau linéaire uniforme de 5 capteurs, $S/B = 20\text{dB}$, $T = 160$ en fonction de $\frac{Mb}{f_0}$.

De plus, comme les biais sont de signes opposés, les directions d'arrivée estimées s'écartent lorsque la largeur de la bande de fréquence augmente, ce qui pourrait se traduire par une augmentation du pouvoir séparateur de MUSIC dans ce cas. En fait, quand le nombre de capteurs augmente, les directions d'arrivée estimées continuent à s'écarter, mais les biais ne sont plus nécessairement de signes opposés comme le montre par exemple la FIG.5.8 qui correspond à la situation précédente sauf pour ce qui concerne le nombre de capteurs qui a été doublé. De ce fait le critère habituel de résolubilité des sources (critère de Cox ⁷) fondé sur un seuil évalué à mi-angle $\frac{1}{2}(\theta_1 + \theta_2)$ (cf. [13]) ne convient pas dans cette situation.

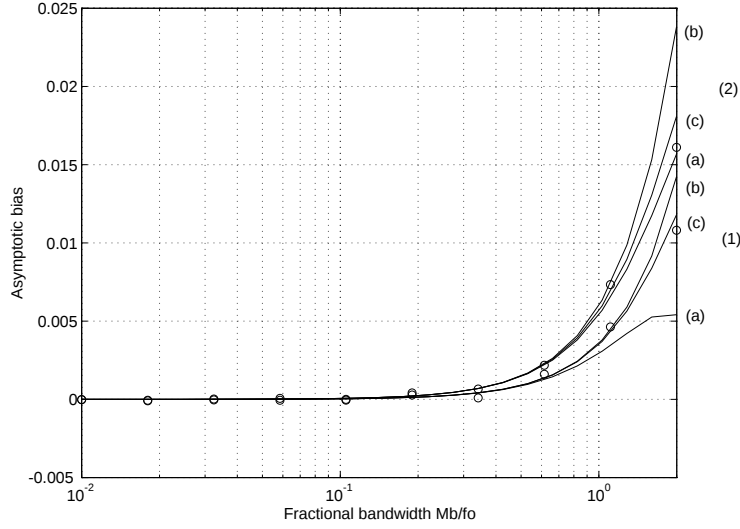


FIG. 5.8: Biases asymptotiques (en b et T) (a) (b) et (c) et estimés par méthode de Monte Carlo (100 réalisations) (o) des pulsations spatiales $\phi_k = \pi \sin \theta_k$ de deux sources de même puissance, de spectre uniforme centré sur f_0 et d'azimuts respectifs $\theta_1 = 25^\circ$ (1), $\theta_2 = 40^\circ$ (2) pour un réseau linéaire uniforme de 10 capteurs, $S/B = 20\text{dB}$, $T = 160$ en fonction de $\frac{Mb}{f_0}$.

⁷ Si $g(\theta)$ désigne la fonction de localisation de l'algorithme, dont les minima fournissent les estimées des paramètres, les deux sources seront réputées résolues si $g(\frac{\theta_1 + \theta_2}{2}) > \frac{g(\theta_1) + g(\theta_2)}{2}$.

5.5 Borne de Cramer-Rao

Nous avons montré au paragraphe 5.4 que, dans le cas de sources large bande à spectre centré sur la fréquence de démodulation f_0 , l'estimée du paramètre DOA donnée par tout algorithme du second ordre est asymptotiquement (par rapport au nombre d'observations et à la largeur de bande) non biaisée (c.f. rel. (4.6) et (4.7))⁸. Il est donc pertinent dans ce cas de considérer le problème de la détermination de la borne de Cramer-Rao associée à l'estimation de ce paramètre.

Dans le cas de T observations complexes circulaires gaussiennes centrées $\mathbf{y}_t$, la BCR pour l'ensemble des paramètres inconnus Ψ contenus dans la matrice de covariance spatiale $\mathbf{R}_b$ est obtenue en inversant la matrice d'information de Fisher (MIF) $\mathbf{I}(\Psi)$ (cf. par ex. [14, rel. (15.52)]) :

$$[\mathbf{I}(\Psi)]_{k,l} = T \operatorname{Tr} \left(\mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \psi_k} \mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \psi_l} \right). \quad (5.1)$$

Du fait de la nécessité d'inverser la MIF, il est très difficile d'obtenir une expression analytique de la BCR, même pour le seul paramètre Θ . Pour contourner cette difficulté, de nombreux auteurs (cf. [15] par exemple) ont utilisé l'efficacité asymptotique de l'estimateur au sens du maximum de vraisemblance (ML). La BCR du vecteur complet des paramètres Ψ est alors donnée par la covariance asymptotique de cet estimateur. Mais, obtenir directement la matrice de covariance asymptotique de l'estimateur ML du seul paramètre Θ suppose que la fonction de vraisemblance soit séparable. Malheureusement, la propriété de séparabilité de la fonction de vraisemblance utilisée dans le cas de sources bande-étroite (cf. [15] par exemple) ne peut s'étendre en dehors de cette hypothèse.

Nous allons montrer dans la suite que, dans le cas particulier d'une source large bande à spectre symétrique par rapport à f_0 et d'un réseau quelconque, on peut obtenir une expression analytique de la BCR pour l'angle d'arrivée θ_1 seul. Cela nous permettra de relier les cas bande étroite et large bande. Nous considérerons ensuite le cas de plusieurs sources où nous montrerons que l'approche utilisée dans le cas d'une seule source échoue.

5.5.1 Cas d'une seule source

Rappelons tout d'abord que dans le cas d'une seule source la matrice $\mathbf{R}_b$ s'écrit :

$$\mathbf{R}_b = \int_{-B/2}^{+B/2} S_1(f) \mathbf{a}(\Theta_1, f_0 + f) \mathbf{a}^H(\Theta_1, f_0 + f) df + \sigma_n^2 \mathbf{I}_M, \quad (5.2)$$

et que par ailleurs

$$\mathbf{a}(\theta_1, f + f_0) \mathbf{a}^H(\theta_1, f + f_0) = \Delta_{\theta_1} (\mathbf{a}(\theta_1, f) \mathbf{a}^H(\theta_1, f)) \Delta_{\theta_1}^H$$

avec $\Delta_{\theta_1} \stackrel{\text{def}}{=} \operatorname{Diag}(e^{-i2\pi f_0 \tau_{1,1}}, \dots, e^{-i2\pi f_0 \tau_{1,M}})$. Donc (5.2) peut s'écrire :

$$\mathbf{R}_b = \Delta_{\theta_1} \mathbf{R}_\Phi \Delta_{\theta_1}^H, \quad (5.3)$$

avec $\mathbf{R}_\Phi = \mathbf{R}_{s_1} + \sigma_n^2 \mathbf{I}_M$ où $\mathbf{R}_{s_1}$ est une matrice réelle symétrique dont tous les termes de la diagonale principale sont égaux à σ_1^2 . La matrice symétrique réelle $\mathbf{R}_\Phi$ ⁹ peut donc être paramétrée par le vecteur Φ dont les $L = \frac{M(M-1)}{2} + 1$ composantes sont $\sigma_1^2 + \sigma_n^2$ et les $\frac{M(M-1)}{2}$ termes d'un bloc triangulaire. La matrice $\mathbf{R}_b$ est donc entièrement déterminée par le paramètre $\Psi \stackrel{\text{def}}{=} (\theta_1, \Phi)$.

⁸Rappelons également que dans le cas d'une seule source à spectre symétrique par rapport à f_0 nous avons constaté que le biais est négligeable pour tout b dans le cas de réseaux linéaires ou circulaires réguliers.

⁹Dans le cas d'un réseau linéaire régulier $\mathbf{R}_\Phi$ est de plus de structure Toeplitz.

En utilisant le fait que $\mathbf{R}_b$ dépend linéairement de Φ et en suivant la même démarche que [16, Appendix]¹⁰ on montre en Annexe 5 que :

$$[\mathbf{I}(\Psi)]_{1,k} = 0, \quad k = 2, \dots, L \quad (5.4)$$

$$[\mathbf{I}(\Psi)]_{k,l} = T \operatorname{Tr} \left(\mathbf{R}_\Phi^{-1} \frac{\partial \mathbf{R}_\Phi}{\partial \phi_k} \mathbf{R}_\Phi^{-1} \frac{\partial \mathbf{R}_\Phi}{\partial \phi_l} \right), \quad k, l = 2, \dots, L \quad (5.5)$$

$$[\mathbf{I}(\Psi)]_{1,1} = T \left(-2\operatorname{Tr}(\Delta'_{\theta_1})^2 + 2\operatorname{Tr}(\Delta'_{\theta_1} \mathbf{R}_\Phi^{-1} \Delta'_{\theta_1} \mathbf{R}_\Phi) \right), \quad (5.6)$$

où $\Delta'_{\theta_1} \stackrel{\text{def}}{=} -2\pi f_0 \operatorname{Diag}(\frac{d\tau_{1,1}}{d\theta_1}, \dots, \frac{d\tau_{1,M}}{d\theta_1})$.

En conséquence la MIF du paramètre Ψ a une structure bloc-diagonale :

$$\begin{bmatrix} I(\theta_1) & \mathbf{0}^T \\ \mathbf{0} & \mathbf{I}(\Phi) \end{bmatrix}$$

où $I(\theta_1)$ représente la matrice d'information de Fisher du paramètre θ_1 considéré comme le seul paramètre inconnu et $\mathbf{I}(\Phi)$, la MIF du paramètre Φ déduit de T observations complexes circulaires gaussiennes centrées indépendantes de matrice de covariance $\mathbf{R}_\Phi$. La BCR du paramètre θ_1 est donc :

$$BCR_{\theta_1} = I^{-1}(\theta_1) = \frac{1}{2T} \left(-\operatorname{Tr}(\Delta'_{\theta_1})^2 + \operatorname{Tr}(\Delta'_{\theta_1} \mathbf{R}_\Phi^{-1} \Delta'_{\theta_1} \mathbf{R}_\Phi) \right)^{-1}, \quad (5.7)$$

et comme $[\mathbf{I}^{-1}(\Psi)]_{1,1} = [\mathbf{I}(\Psi)]_{1,1}^{-1}$, BCR_{θ_1} n'est pas diminuée par la connaissance des autres paramètres inconnus.

Nous proposons maintenant une relation entre $BCR_{\theta_1^b}$ et $BCR_{\theta_1^0}$, les BCR respectives du paramètre θ_1 dans les cas large bande et bande étroite. En écrivant, dans (5.7), $\mathbf{R}_\Phi$ sous la forme :

$$\mathbf{R}_\Phi = \sigma_1^2 \mathbf{1}\mathbf{1}^T + (\sigma_n^2 \mathbf{I}_M + \delta \mathbf{R}_{s_1}),$$

le résultat suivant est démontré en Annexe 6.

Résultat 5.5 *La BCR de l'angle d'arrivée θ_1 d'une source large bande est donnée par*

$$BCR_{\theta_1^b} = BCR_{\theta_1^0} \left(1 + c' \left(\frac{f_\sigma^2}{f_0^2} \right) + O \left(\frac{f_\sigma^4}{f_0^4} \right) \right) \quad (5.8)$$

où $BCR_{\theta_1^0}$ est la BCR classique du scénario bande-étroite (cf. [15, rel. (17)] par ex.) et c' est donné en Annexe 6.

A l'aide d'évaluations numériques des coefficients c et c' intervenant dans les expressions de la variance asymptotique (4.16) (donnée par l'algorithme MUSIC) et de la borne de Cramer-Rao asymptotique (5.8) nous pouvons remarquer (cf. FIG.5.9) que ces coefficients c et c' croissent très rapidement avec le rapport S/B et deviennent extrêmement importants.

¹⁰ Nous remarquons qu'en utilisant la propriété de linéarité de la matrice $\mathbf{R}_b$ par rapport au paramètre Φ , une autre expression analytique plus compliquée (exigeant une matrice de projection) peut être obtenue (voir par exemple [17] à l'aide d'une démonstration assez longue, [18, Appendix 4.C] avec une démonstration très courte et (2.18) au chapitre 6).

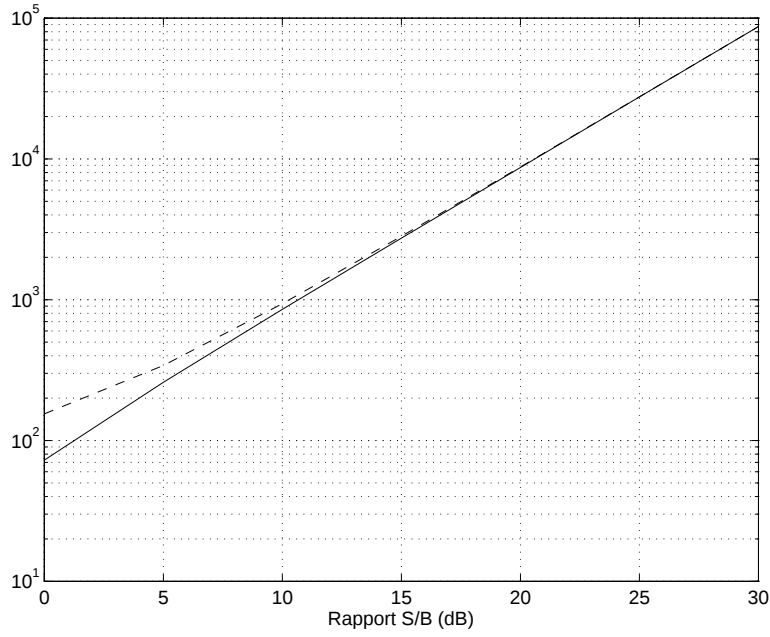


FIG. 5.9: Coefficients c (-) et c' (- -) en fonction du rapport S/B pour une source large bande ($\frac{Mb}{f_0} = 1$) et un réseau linéaire uniforme de 5 capteurs.

Ceci a comme conséquence d'atténuer l'effet qu'a l'augmentation du rapport S/B sur l'amélioration de la précision de l'estimation dans le cas d'une source à bande étendue comme le confirme la FIG.5.10. On constate en outre sur cette figure que l'approximation fournie par les deux premiers termes du développement de (4.16) est très bonne (y compris pour $\frac{Mb}{f_0} = 1$).

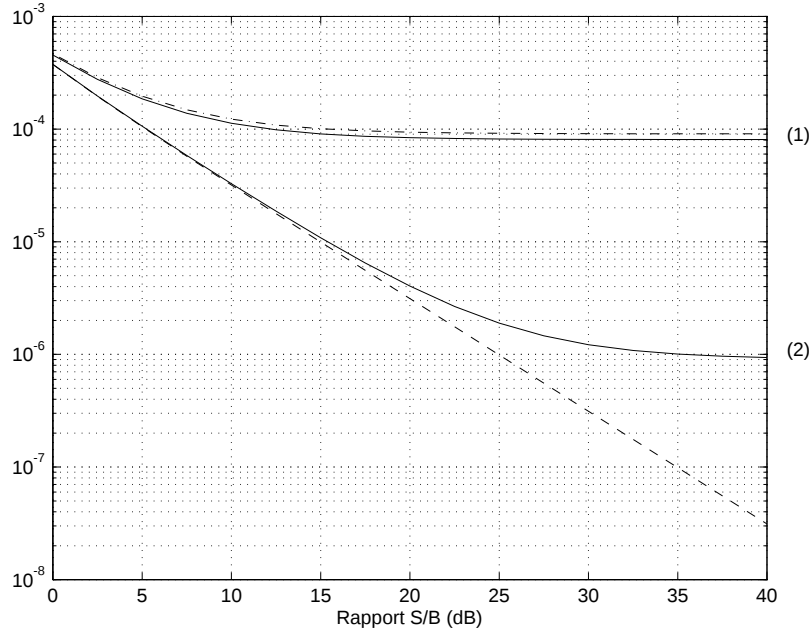


FIG. 5.10: Comparaison de $\text{Var}(\theta_1^b(T))$ (-) et de $\text{Var}(\theta_1^0(T)) \left(1 + c \left(\frac{f^2}{f_0^2}\right)\right)$ (- . -) à $\text{Var}(\theta_1^0(T))$ (- -) dans la cas d'une source large bande pour $\frac{Mb}{f_0} = 1$ (1) ou $\frac{Mb}{f_0} = 0.1$ (2), un réseau linéaire uniforme de 5 capteurs et $T = 160$ (les courbes (-) et (- . -) sont pratiquement confondues dans le cas (1) et confondues dans le cas (2)).

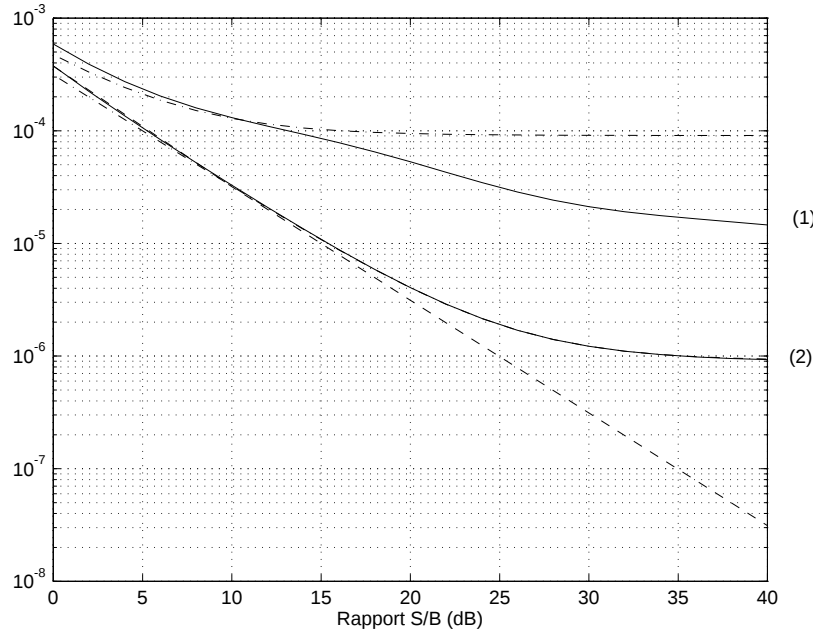


FIG. 5.11: Comparaison de $BCR_{\theta_1^b}$ (-) et de $BCR_{\theta_1^0} \left(1 + c' \left(\frac{f_\sigma^2}{f_0^2}\right)\right)$ (- . -) à $BCR_{\theta_1^0}$ (- -) dans la cas d'une source large bande pour $\frac{Mb}{f_0} = 1$ (1) ou $\frac{Mb}{f_0} = 0.1$ (2), un réseau linéaire uniforme de 5 capteurs (dans le cas (2) les courbes (-) et (- . -) sont confondues sauf au voisinage de $S/B = 0$).

Nous constatons également sur la FIG.5.11 que le comportement de la borne de Cramer-Rao en fonction de l'augmentation du rapport S/B est profondément différent selon la largeur de bande du signal. Pour une bande strictement étroite, $\mathbf{R}_\Phi$ devient singulière de rang 1 lorsque σ_n^2 tend vers 0 ; par suite $BCR_{\theta_1^0}$ tend vers 0 d'après (5.7). Nous avons dans ce cas, un modèle d'observation déterministe et le paramètre θ_1 est identifiable à partir d'un nombre fini d'observations. Au contraire, dans le cas d'une bande non strictement étroite, le modèle d'observation reste stochastique, $\mathbf{R}_\Phi$ reste de rang plein lorsque σ_n^2 tend vers 0 et $BCR_{\theta_1^b}$ tend vers la limite :

$$\lim_{\sigma_n^2 \rightarrow 0} BCR_{\theta_1^b} = \frac{1}{2T} \left(-\text{Tr}(\Delta'_{\theta_1})^2 + \text{Tr}(\Delta'_{\theta_1} \mathbf{R}_{s_1}^{-1} \Delta'_{\theta_1} \mathbf{R}_{s_1}) \right)^{-1}.$$

De plus, nous constatons sur la FIG.5.11 que l'approximation fournie par les deux premiers termes du développement de la borne de Cramer-Rao (5.8) a un domaine de validité plus restreint que pour la variance donnée par l'algorithme MUSIC. Si d'adéquation est parfaite entre $BCR_{\theta_1^b}$ (-) et $BCR_{\theta_1^0} \left(1 + c' \left(\frac{f_\sigma^2}{f_0^2}\right)\right)$ pour $\frac{Mb}{f_0} = 0.1$ et ce pour tout rapport S/B , cette adéquation est perdue pour $\frac{Mb}{f_0} = 1$ contrairement au cas de la variance donnée par l'algorithme MUSIC.

Remarque : Dans le cas particulier d'une source large bande à spectre symétrique par rapport à f_0 et d'un réseau *linéaire* quelconque, nous montrons en Annexe 7 que si la forme normalisée $S(f)$ du spectre de la source est connue, $\mathbf{R}_b$ devenant paramétrisée par (θ_1, Ψ) avec $\Psi = (\theta_1, f_\sigma, \sigma_1^2, \sigma_n^2)$, la démarche précédente conduit à la même expression (5.7) de la borne de Cramer-Rao de θ_1 . Mais curieusement cette démarche n'est plus valable pour un réseau non linéaire. On constate d'ailleurs, à l'issue de la démonstration que les paramètres θ_1 et (σ_1^2, σ_n^2) sont découplés seulement si la largeur de la bande f_σ du signal est connue.

La Fig. 5.12 illustre les résultats précédents dans le cas d'une source de spectre uniforme et centré sur f_0 , de largeur b et d'un réseau linéaire uniforme de 5 capteurs.

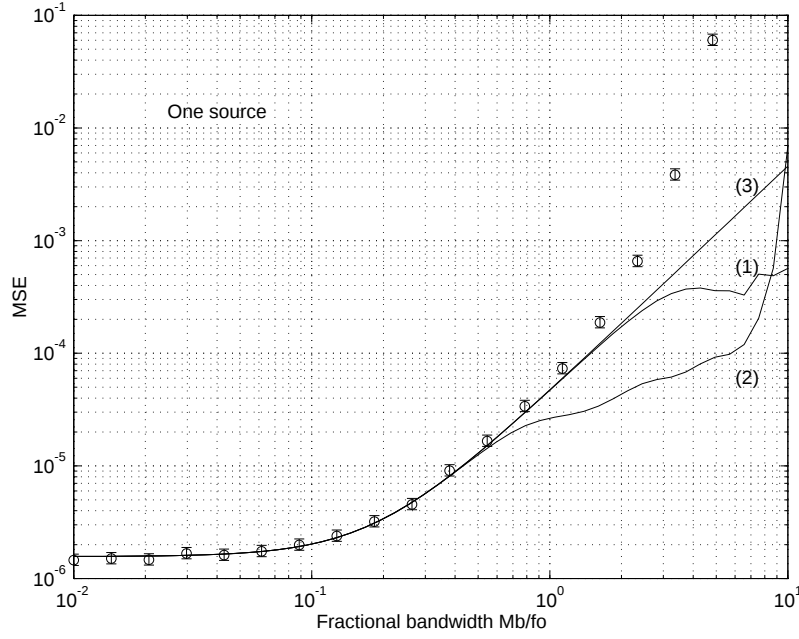


FIG. 5.12: EQM et BCR pour la pulsation spatiale $\phi_1 = \pi \sin \theta_1$ en fonction de $\frac{Mb}{f_0}$.

EQM asymptotique (en b et T) (1) et estimée (1000 réalisations) par MUSIC standard (o et intervalle de confiance à 99%), CRB_{θ_1} exacte (2) et asymptotique (en b) (3) pour une source à spectre uniforme et centré et un réseau linéaire de 5 capteurs, $S/B = 20\text{dB}$, $T = 320$.

Elle présente, en fonction de la largeur de bande réduite $\frac{Mb}{f_0}$ pour un rapport S/B de 20 dB et un nombre d'observations de $T = 320$:

- (1) l'EQM asymptotique (par rapport à la largeur de bande et au nombre d'observations) analytique de l'algorithme MUSIC (4.17)
- (2) la BCR exacte (5.7)
- (3) la BCR asymptotique (par rapport à la largeur de bande) $BCR_{\theta_1^b}$ (5.8) et l'EQM asymptotique (par rapport à la largeur de bande) de l'algorithme MUSIC (4.16) (ces 2 courbes sont en fait confondues)
- (o) l'EQM asymptotique (par rapport à la largeur de bande) estimée de l'algorithme MUSIC (1000 réalisations et intervalle de confiance à 99%)

Concernant les résultats présentés par la Fig. 5.12, nous pouvons faire les remarques suivantes :

1. les valeurs de la BCR exacte (2) et de l'expression asymptotique de la BCR (3) coïncident jusqu'à $\frac{Mb}{f_0} = 0.6$,
2. les valeurs estimées de l'EQM de l'algorithme MUSIC concordent bien avec les valeurs théoriques de l'EQM asymptotique jusqu'à $\frac{Mb}{f_0} = 1.8$,
3. les valeurs de la BCR asymptotique coïncident avec les valeurs de l'EQM asymptotique pour toutes les valeurs de $\frac{Mb}{f_0}$ envisagées.

En conséquence, l'algorithme MUSIC conventionnel, efficace asymptotiquement (par rapport au nombre d'observations) dans le cas d'une largeur de bande nulle, le reste lorsque la largeur de bande augmente jusqu'à $\frac{Mb}{f_0} = 0.6$.

5.5.2 Cas de plusieurs sources

Dans le cas de plusieurs sources de spectres symétriques par rapport à f_0 , la matrice de covariance spatiale $\mathbf{R}_b$ est paramétrée par $(\theta_k)_{k=1,\dots,K}$, σ_n^2 et les termes des blocs triangulaires des matrices symétriques réelles $(\mathbf{R}_{s_k})_{k=1,\dots,K}$ (cf. eq. (2.5)). Toutefois le paramétrage n'est maintenant plus unique car ces paramètres ne sont pas identifiables de façon unique à partir de $\mathbf{R}_b$. Pour pouvoir estimer les paramètres directionnels à partir de la seule connaissance de la matrice de covariance spatiale $\mathbf{R}_b$, il est nécessaire de disposer d'une connaissance a priori supplémentaire. Par exemple, si la forme normalisée $S(f)$ des spectres des sources est identique et connue, $\mathbf{R}_b$ peut être paramétrée de façon unique par $\Psi = (\theta_1, \dots, \theta_K, f_\sigma, \sigma_1^2, \dots, \sigma_K^2, \sigma_n^2)$. Mais, même dans ce cas, la démarche utilisée dans le cas d'une source n'est pas applicable et nous devons nous contenter d'inverser la matrice d'information de Fisher (5.1) et d'extraire le bloc supérieur gauche de $\mathbf{I}(\Psi)^{-1}$.

Notons enfin que si la forme normalisée $S(f)$ des spectres des sources est connue ainsi que leur écart-type f_σ , $\mathbf{R}_b$ peut être paramétrée de façon unique par $\Psi = (\Theta, \Phi)$ avec $\Theta \stackrel{\text{def}}{=} (\theta_1, \dots, \theta_K)$ et $\Phi \stackrel{\text{def}}{=} (\sigma_1^2, \dots, \sigma_K^2, \sigma_n^2)$ et en conséquence, $\mathbf{R}_b$ dépend ici linéairement de Φ . Le lemme d'inversion matricielle utilisé dans [18, Appendix 4C] permet d'obtenir directement une expression analytique de la BCR de Θ .

La FIG.5.13 illustre ces résultats dans le cas de deux sources de même puissance, de spectre uniforme et centré de largeur b et d'un réseau linéaire uniforme de 5 capteurs. Les directions d'arrivée sont respectivement $\theta_1 = 65^\circ$ et $\theta_2 = 50^\circ$, le rapport S/B = 20 dB et le nombre d'observations $T = 320$.

Elle présente

- (1) l'EQM asymptotique (par rapport à la largeur de bande et au nombre d'observations) analytique de l'algorithme MUSIC (4.17)
- (2) la BCR exacte (5.7)
- (o) l'EQM asymptotique (par rapport à la largeur de bande) estimée de l'algorithme MUSIC (1000 réalisations et intervalle de confiance à 99%)

Nous pouvons constater que les valeurs estimées de l'EQM de l'algorithme MUSIC concordent bien avec les valeurs théoriques de l'EQM asymptotique jusqu'à $\frac{Mb}{f_0} = 1$. D'autre part, bien que cet algorithme ne soit pas efficace, la FIG.5.13 montre qu'il demeure pratiquement efficace jusqu'à $\frac{Mb}{f_0} = 0.4$.

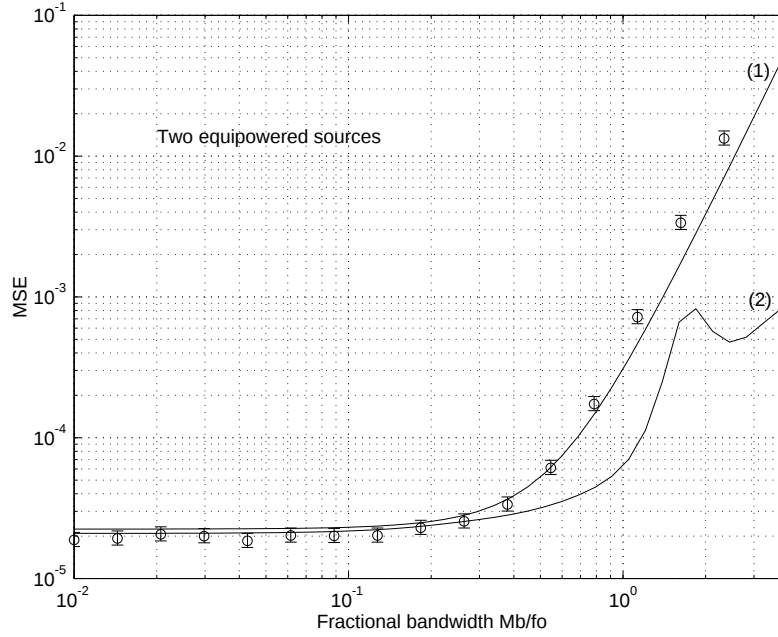


FIG. 5.13: EQM et BCR pour la pulsation spatiale $\phi_1 = \pi \sin \theta_1$ en fonction de $\frac{Mb}{f_0}$. EQM asymptotique (en b et T) (1) et estimée (1000 réalisation) par MUSIC standard (o et intervalle de confiance à 99%), CRB_{θ_1} exacte (2) pour deux sources à spectre uniforme et centré et un réseau linéaire de 5 capteurs, $S/B = 20\text{dB}$, $T = 320$.

5.6 Performances comparées d'algorithmes bande étroite et large bande

Dans ce paragraphe, nous comparons les performances de deux algorithmes classiques d'estimation de DOA, l'un bande étroite, l'autre large bande, en fonction de la largeur de bande du signal. Une comparaison complète nécessiterait l'étude de différentes situations, en type de réseaux, formes de spectres, directions d'arrivée ou valeurs de rapport S/B , ce qui sortirait du cadre de ce chapitre. Nous nous contenterons de considérer les deux scénarios à une et deux sources de la section précédente. Nous utiliserons respectivement l'algorithme MUSIC standard opérant sur la matrice de covariance spatiale comme méthode bande étroite et l'algorithme proposé par Friedlander et Weiss [19] utilisant les matrices de densité inter-spectrale comme algorithme large bande. Il paraît a priori naturel d'effectuer cette comparaison avec un nombre identique d'observations. Remarquons cependant que, dans le cas des algorithmes bande étroite utilisant la matrice de covariance spatiale, ces observations consistent en échantillons temporels généralement supposés indépendants alors que dans le cas des algorithmes large bande utilisant les matrices de densité inter-spectrale, les sorties des capteurs sont subdivisées en tranches temporelles et les observations correspondent en fait aux transformées de Fourier de ces différentes tranches. Pour pouvoir comparer ces deux méthodes dans les mêmes conditions il convient donc de le faire sur une même durée d'observation c'est à dire, si la fréquence d'échantillonnage est la même, avec le même nombre d'échantillons temporels consécutifs. Nous fournissons donc aux deux algorithmes les mêmes jeux de données constitués d'échantillons de l'enveloppe complexe des signaux prélevés à la fréquence de Nyquist $f_s = \frac{1}{T_s} = b$. Remarquons que dans cette situation les échantillons temporels ne peuvent plus être considérés comme indépendants. Par suite, les expressions de la BCR et de l'EQM de MUSIC données dans le paragraphe 5.5 ne sont plus valables.

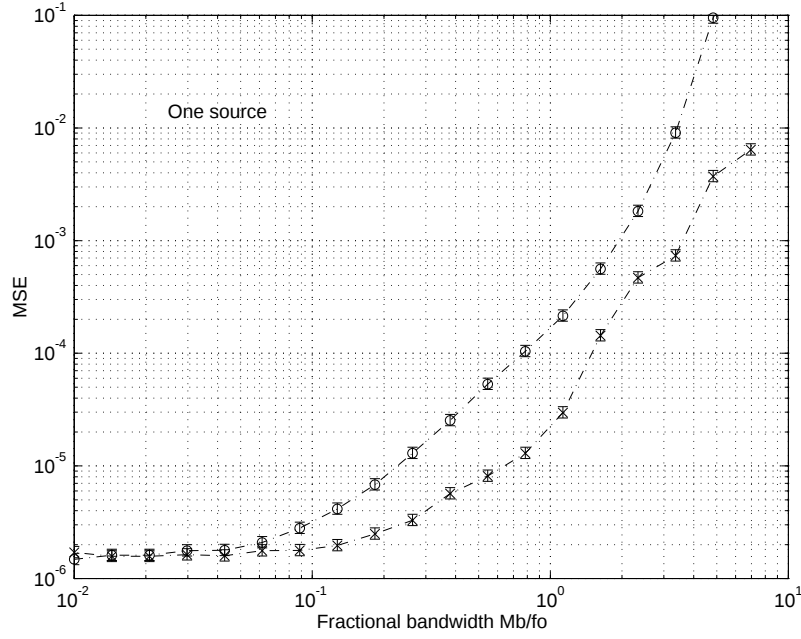


FIG. 5.14: EQM estimée sur la pulsation spatiale $\phi_1 = \pi \sin \theta_1$ pour l'algorithme MUSIC standard (o) et pour l'algorithme large bande (*) en fonction de $\frac{Mb}{f_0}$ (1000 réalisations et intervalles de confiance à 99%).

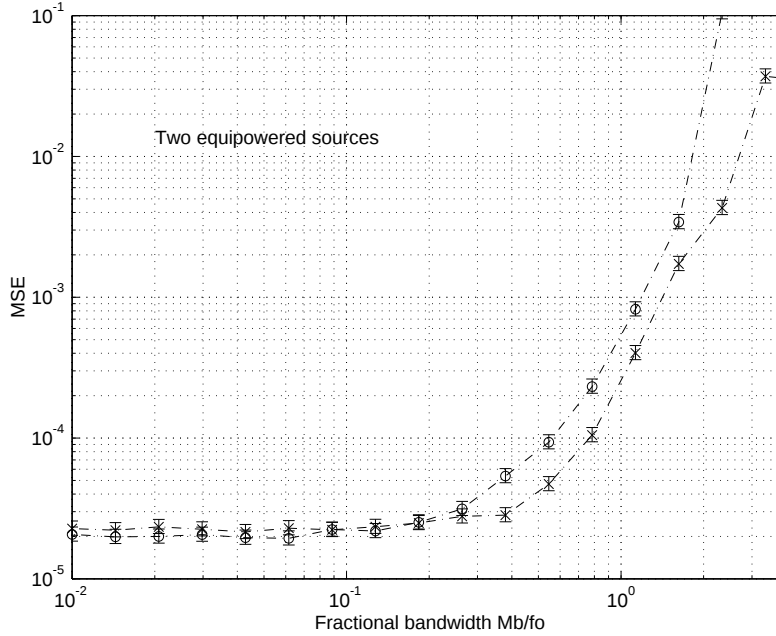


FIG. 5.15: EQM estimée sur la pulsation spatiale $\phi_1 = \pi \sin \theta_1$ pour l'algorithme MUSIC standard (o) et pour l'algorithme large bande (*) en fonction de $\frac{Mb}{f_0}$ (1000 réalisations et intervalles de confiance à 99%).

Cependant, il a été démontré dans [20] et dans le paragraphe 4.5.1 du chapitre 4 que dans le cas de sources à bande strictement étroite, les performances asymptotiques de tout algorithme du second ordre d'estimation de DOA ne dépend pas de la fréquence d'échantillonnage si $\frac{1}{T_s} \geq b$, mais seulement de la durée d'observation. Les Fig.5.14 et 5.15 montrent que les performances de l'algorithme MUSIC (qui sont légèrement dégradées par rapport à celles des Fig.5.12 et 5.13) sont équivalentes à celles de l'algorithme

large bande jusqu'à $Mb/f_0 < 0.1$. L'algorithme large bande surclasse ensuite l'algorithme MUSIC pour $0.1 < Mb/f_0 < 2$.

5.7 Conclusion

Dans ce chapitre nous avons étudié les performances des algorithmes de DOA du second ordre "bande étroite" lorsqu'ils sont appliqués à des sources à bande étendue.

En considérant tout d'abord le problème de la détection de l'ordre, nous avons comparé les performances des méthodes classiques AIC et MDL à celles d'une méthode proposée récemment par Liavas *et al.* [7] et montré que cette dernière semblait plus adaptée au problème posé par les signaux à bande étendue. Nous avons ensuite étendu les résultats obtenus dans [6] dans le cadre de sources à spectre symétrique par rapport à la fréquence de démodulation f_0 au cas de spectres non symétriques et/ou non centrés sur cette fréquence. Nous avons proposé des expressions générales du biais et de la covariance asymptotiques (en nombre d'observations et en largeur de bande du signal) des estimées des directions d'arrivées par des méthodes bande étroite du second ordre en fonction des paramètres de centrage et de dispersion du spectre des sources. Ces expressions nous ont permis de montrer, notamment dans le cas d'une seule source, que ces algorithmes sont beaucoup sensibles au caractère non centré du spectre qu'à sa largeur de bande.

En considérant ensuite plus particulièrement le cas de sources à spectre symétrique par rapport à la fréquence de démodulation, nous avons donné une expression générale de la borne de Cramer Rao dans le cas d'une source large bande et exprimé cette borne en fonction de la BCR du cas bande étroite. Ceci nous a permis notamment de montrer que le comportement de la BCR lorsque le rapport S/B augmente est profondément différent selon la largeur de bande du signal.

Enfin, nous avons comparé les performances asymptotiques de deux algorithmes classiques de DOA, l'un bande étroite, l'autre large bande. Cette comparaison ainsi que les résultats des études précédentes concernant la détection de l'ordre et la BCR nous ont montré que la définition empirique de l'hypothèse bande étroite souvent utilisée dans la littérature selon laquelle l'ouverture du réseau doit être beaucoup plus petite que l'inverse de la bande relative du signal (i.e., $\frac{Mb}{f_0} \ll 1$) est bien trop sévère dans le cas d'une source ou de deux sources de même puissance à spectre symétrique par rapport à la fréquence de démodulation. En effet, dans ces situations, notre étude théorique, confirmée par des simulations, a montré que, dans le cas de réseaux linéaires uniformes ¹¹

- pour une source, l'algorithme MUSIC standard reste asymptotiquement efficace jusqu'à $\frac{Mb}{f_0} = 0.6$,
- pour deux sources, cet algorithme reste "pratiquement" efficace jusqu'à $\frac{Mb}{f_0} = 0.4$.

En conclusion, on peut affirmer que les algorithmes bande-étroite sont robustes vis à vis de la largeur de bande du signal. Toutefois, la prise en considération de sources spatialement corrélées ou de puissances différentes mériteraient d'autres investigations.

¹¹ Notons que, mis à part les réseaux linéaires lacunaires, les réseaux linéaires uniformes représentent le cas le plus défavorable de répartition de M capteurs du point de vue du temps de traversée du réseau.

Annexe

Annexe 1 : Preuve des relations (2.7), (2.8) et (2.9)

A partir de (2.6), un développement limité de $[\mathbf{R}_{s_k}]_{m,n}$ au voisinage de 0 nous permet d'écrire, grâce à (2.4)

$$\begin{aligned} [\mathbf{R}_{s_k}]_{m,n} &= \sigma_k^2 R\left(\frac{f_\sigma}{f_0} \alpha_{m,n,k}\right) e^{i2\pi \frac{f_m}{f_0} \alpha_{m,n,k}} \\ &= \sigma_k^2 \left[1 - 2\pi^2 \left(\frac{f_\sigma^2}{f_0^2}\right) \alpha_{m,n,k}^2 + \frac{1}{6} \left(\frac{f_\sigma^3}{f_0^3}\right) \alpha_{m,n,k}^3 R'''(0) + O\left(\frac{f_\sigma^4}{f_0^4}\right) \right] \left[1 + i2\pi \frac{f_m}{f_0} \alpha_{m,n,k} + O\left(\frac{f_m^2}{f_0^2}\right) \right], \end{aligned}$$

qui donne directement (2.7) pour un spectre quelconque. Pour un spectre quelconque centré sur f_0 , $f_m = 0$ ce qui donne (2.8). Pour un spectre symétrique centré sur f_0 , $f_m = 0$ et $R'''(0) = 0$ ce qui donne (2.9).

Annexe 2 : Preuve des relations (4.12), (4.13), (4.14) et (4.15)

En remplaçant les expressions $\mathbf{\Gamma}_s = \frac{1}{M^2 \sigma_1^2} \mathbf{a}_{\theta_1} \mathbf{a}_{\theta_1}^H$, $\mathbf{\Pi}_n = \mathbf{I}_M - \frac{\mathbf{a}_{\theta_1} \mathbf{a}_{\theta_1}^H}{M}$ et $\mathbf{a}'_{\theta_1} = i \Delta'_{\theta_1} \mathbf{a}_{\theta_1}$ avec $\Delta'_{\theta_1} \stackrel{\text{def}}{=} \text{Diag}(-2\pi f_0 \tau'_{1,1}, \dots, -2\pi f_0 \tau'_{1,M})$ dans (4.11) on obtient :

$$[\mathbf{D}_{\mathbf{R}_0}^{\text{music}} \text{Vec}(\delta \mathbf{R}_b)]_1 = -\frac{2}{M \alpha_{\theta_1} \sigma_1^2} \int_{-B/2}^{+B/2} S_1(f) \Im \left(\mathbf{a}_{\theta_1}^H \Delta_{\theta_1, f} \mathbf{a}_{\theta_1} \mathbf{a}_{\theta_1}^H \Delta_{\theta_1, -f} (\mathbf{I}_M - \frac{\mathbf{a}_{\theta_1} \mathbf{a}_{\theta_1}^H}{M}) \Delta'_{\theta_1} \mathbf{a}_{\theta_1} \right) df.$$

Comme $\mathbf{a}_{\theta_1}^H \Delta_{\theta_1, f} \mathbf{a}_{\theta_1} = \sum_{m=1}^M e^{-i2\pi f \tau_{1,m}}$ et $\mathbf{a}_{\theta_1}^H \Delta_{\theta_1, -f} \Delta'_{\theta_1} \mathbf{a}_{\theta_1} = -\sum_{n=1}^M (2\pi f_0 \tau'_{1,n}) e^{i2\pi f \tau_{1,n}}$, après quelques calculs le biais asymptotique s'écrit :

$$\begin{aligned} [\mathbf{D}_{\mathbf{R}_0}^{\text{music}} \text{Vec}(\delta \mathbf{R}_b)]_1 &= \frac{2}{M \alpha_{\theta_1} \sigma_1^2} \sum_{n=1}^M 2\pi f_0 \tau'_{1,n} \Im \left(\sum_{m=1}^M \int_{-B/2}^{+B/2} S_1(f) e^{i2\pi f (\tau_{1,n} - \tau_{1,m})} df \right) \\ &- \frac{2}{M^2 \alpha_{\theta_1} \sigma_1^2} \sum_{n=1}^M 2\pi f_0 \tau'_{1,n} \Im \left(\sum_{m,l=1}^M \int_{-B/2}^{+B/2} S_1(f) e^{i2\pi f (\tau_{1,l} - \tau_{1,m})} df \right) \\ &= \frac{2}{M \alpha_{\theta_1}} \sum_{n=1}^M 2\pi f_0 \tau'_{1,n} \Im \left(\sum_{m=1}^M R(f_\sigma(\tau_{1,n} - \tau_{1,m})) e^{i2\pi f_m (\tau_{1,n} - \tau_{1,m})} \right) \\ &- \frac{2}{M^2 \alpha_{\theta_1}} \sum_{n=1}^M 2\pi f_0 \tau'_{1,n} \Im \left(\sum_{m,l=1}^M R(f_\sigma(\tau_{1,l} - \tau_{1,m})) e^{i2\pi f_m (\tau_{1,l} - \tau_{1,m})} \right). \end{aligned} \tag{7.1}$$

La fonction de corrélation normalisée $R(t)$ vérifie $R(t) = 1 - 2\pi^2 t^2 + O(t^3)$ ce qui justifie la relation (4.12) dans le cas d'un spectre quelconque et la relation (4.13) dans le cas d'un spectre quelconque centré sur f_0

($f_m = 0$). Pour un spectre symétrique par rapport à f_0 , $R(t)$ est réelle symétrique. Le biais asymptotique est alors nul et grâce à l'expression (2.10) de $\delta \mathbf{R}_b$, la relation (4.14) est démontrée.

Dans le cas particulier d'un réseau linéaire quelconque et d'un spectre symétrique par rapport à $f_0 + f_m$, $\mathbf{R}_b$ peut s'écrire $\mathbf{R}_b = \sigma_1^2 \mathbf{a}(\theta_1, f_0 + f_m) \mathbf{a}^H(\theta_1, f_0 + f_m) \odot \mathbf{R}'_{s_1} + \sigma_n^2 \mathbf{I}_M$ où $\mathbf{R}'_{s_1}$ est réelle symétrique. De ce fait, on peut considérer que $\mathbf{R}_b$ résulte d'une perturbation de $\mathbf{R}'_0 \stackrel{\text{def}}{=} \sigma_1^2 \mathbf{a}(\theta_1, f_0 + f_m) \mathbf{a}^H(\theta_1, f_0 + f_m) + \sigma_n^2 \mathbf{I}_M$:

$$\mathbf{R}_b = \mathbf{R}'_0 + \delta \mathbf{R}'_b$$

et une analyse au premier ordre de la perturbation d'un algorithme du second ordre estimant ϕ_1 et opérant sur $\mathbf{R}_b$, évaluée au point $\mathbf{R}'_0$ donne :

$$\begin{aligned} \text{alg}(\mathbf{R}_b) &= \text{alg}(\mathbf{R}'_0) + (D_{\mathbf{R}'_0}^{\text{alg}}, \delta \mathbf{R}'_b) + O(\|\delta \mathbf{R}'_b\|^2) \\ &= \phi_1 \frac{f_0 + f_m}{f_0} + \mathbf{D}_{\mathbf{R}'_0}^{\text{alg}} \text{Vec}(\delta \mathbf{R}'_b) + O(\|\delta \mathbf{R}'_b\|^2). \end{aligned} \quad (7.2)$$

La preuve de (4.15) est alors obtenue en utilisant la même démarche que dans la démonstration de (4.14). Remarquons que (4.15) peut se déduire de (4.12) appliqué à ϕ_1 par un calcul direct mais laborieux.

Annexe 3 : Preuve de la relation (4.16)

Partant du résultat 5.2 où $\mathbf{D}_{\mathbf{R}_b}^{\text{music}}$ est remplacé par $\mathbf{D}_{\mathbf{R}'_0}^{\text{music}}$, comme $\mathbf{D}_{\mathbf{R}_b}^{\text{music}} = \mathbf{D}_{\mathbf{R}'_0}^{\text{music}} + O(\|\delta \mathbf{R}_b\|^2)$, $\text{Var}(\theta_1^b(T))$ est donné par :

$$\begin{aligned} \text{Var}(\theta_1^b(T)) &= \text{Var}(\theta_1^0(T)) + \frac{1}{T} \left(\frac{f_\sigma^2}{f_0^2} \right) \mathbf{D}_{\mathbf{R}'_0}^{\text{music}} (\mathbf{R}_0 \otimes_c (\sigma_1^2 \mathbf{a}_{\theta_1}^H \mathbf{a}_{\theta_1} \odot \mathbf{U}_1) + (\sigma_1^2 \mathbf{a}_{\theta_1}^H \mathbf{a}_{\theta_1} \odot \mathbf{U}_1) \otimes_c \mathbf{R}_0) (\mathbf{D}_{\mathbf{R}'_0}^{\text{music}})^H \\ &+ O\left(\frac{1}{T^2}\right) + O\left(\frac{f_\sigma^4}{f_0^4}\right). \end{aligned}$$

Or, d'après (4.10), $\mathbf{D}_{\mathbf{R}'_0}^{\text{music}} = \frac{1}{\alpha_{\theta_1}} \left(\mathbf{a}_{\theta_1}^H \mathbf{\Gamma}_s \otimes_c \mathbf{a}_{\theta_1}'^H \mathbf{\Pi}_n + \mathbf{a}_{\theta_1}'^H \mathbf{\Pi}_n \otimes_c \mathbf{a}_{\theta_1}^H \mathbf{\Gamma}_s \right)$ avec $\alpha_{\theta_1} = 2 \left(\|\mathbf{a}_{\theta_1}'\|^2 - \frac{|\mathbf{a}_{\theta_1}^H \mathbf{a}_{\theta_1}'|^2}{M} \right)$.

La relation (4.16) peut alors être établie après un calcul direct mais laborieux. Durant ce calcul, en utilisant l'identité $\mathbf{x}^T (\mathbf{A} \odot \mathbf{B}) \mathbf{y} = \text{Tr}[\text{Diag}(\mathbf{x}) \mathbf{A} \text{Tr}[\text{Diag}(\mathbf{y}) \mathbf{B}^T]]$ avec $\text{Diag}(\mathbf{a}_{\theta_1}) = \mathbf{\Delta}_{\theta_1}$, $\text{Diag}(\mathbf{a}_{\theta_1}') = i \mathbf{\Delta}_{\theta_1}' \mathbf{\Delta}_{\theta_1}$, $\mathbf{\Delta}_{\theta_1}^H \mathbf{a}_{\theta_1} = \mathbf{1}$ et $\mathbf{d}_{\theta_1} \stackrel{\text{def}}{=} \mathbf{\Delta}_{\theta_1}' \mathbf{1}$, il apparaît que la constante c peut s'écrire :

$$c = 2 \frac{\sigma_1^2}{\sigma_n^2} \left(\mathbf{d}_{\theta_1}^H \mathbf{U}_1 \mathbf{d}_{\theta_1} + \frac{2}{M} \Im[(\mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}) (\mathbf{d}_{\theta_1}^H \mathbf{U}_1 \mathbf{1})] + \frac{1}{M} \frac{\sigma_1^2 |\mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}'|^2 + \sigma_n^2 \|\mathbf{a}_{\theta_1}'\|^2}{\sigma_n^2 + M \sigma_1^2} (\mathbf{1}^T \mathbf{U}_1 \mathbf{1}) \right). \quad (7.1)$$

Annexe 4 : Preuve de la relation (4.19)

En utilisant l'expression de $\delta \mathbf{R}_b$ déduite de (2.9)

$$\delta \mathbf{R}_b = \left(\frac{f_\sigma^2}{f_0^2} \right) \sum_{k=1}^2 (\mathbf{a}_{\theta_k} \mathbf{a}_{\theta_k}^H \odot \sigma_k^2 \mathbf{U}_k) + O\left(\frac{f_\sigma^4}{f_0^4}\right),$$

on obtient l'expression (4.19) où les coefficients c_l , $l = 1, 2$ sont donnés par :

$$c_l = \frac{1}{\alpha_{\theta_l}} \left(\mathbf{a}_{\theta_l}^H \mathbf{\Gamma}_s \otimes_c \mathbf{a}_{\theta_l}'^H \mathbf{\Pi}_n + \mathbf{a}_{\theta_l}'^H \mathbf{\Pi}_n \otimes_c \mathbf{a}_{\theta_l}^H \mathbf{\Gamma}_s \right) \sum_{k=1}^2 (\mathbf{a}_{\theta_k} \otimes_c \mathbf{a}_{\theta_k}) \odot \sigma^2 \text{Vec}(\tilde{\mathbf{R}}_{s_k})$$

où $\sigma \stackrel{\text{def}}{=} \sigma_1 = \sigma_2$ et où $\mathbf{\Gamma}_s$ et $\mathbf{\Pi}_n$ sont respectivement donnés par

$$\begin{aligned}\mathbf{\Gamma}_s &= \frac{1}{M^2 \sigma^2 (1 - |\beta|^2)^2} ((1 + |\beta|^2) (\mathbf{a}_{\theta_1} \mathbf{a}_{\theta_1}^H + \mathbf{a}_{\theta_2} \mathbf{a}_{\theta_2}^H) - 2\beta^* \mathbf{a}_{\theta_2} \mathbf{a}_{\theta_1}^H - 2\beta \mathbf{a}_{\theta_1} \mathbf{a}_{\theta_2}^H) \\ \mathbf{\Pi}_n &= \mathbf{I}_M - \frac{1}{M(1 - |\beta|^2)} (\mathbf{a}_{\theta_1} \mathbf{a}_{\theta_1}^H + \mathbf{a}_{\theta_2} \mathbf{a}_{\theta_2}^H - \beta^* \mathbf{a}_{\theta_2} \mathbf{a}_{\theta_1}^H - \beta \mathbf{a}_{\theta_1} \mathbf{a}_{\theta_2}^H)\end{aligned}$$

avec $\beta \stackrel{\text{def}}{=} \frac{\mathbf{a}_{\theta_1}^H \mathbf{a}_{\theta_2}}{\|\mathbf{a}_{\theta}\|^2}$.

Annexe 5 : Preuve des relations (5.4), (5.5) et (5.6)

Les expressions de termes de la matrice d'information de Fisher $\mathbf{I}(\Psi')$ peuvent être obtenues en plusieurs étapes. Tout d'abord, partant de l'expression (5.3) de $\mathbf{R}_b$ faisant intervenir la matrice diagonale Δ_{θ_1} , on peut écrire :

$$\begin{aligned}\frac{\partial \mathbf{R}_b}{\partial \theta_1} &= i \Delta'_{\theta_1} \Delta_{\theta_1} \mathbf{R}_{\Phi} \Delta_{\theta_1}^H - i \Delta_{\theta_1} \mathbf{R}_{\Phi} \Delta_{\theta_1}^H \Delta'_{\theta_1} \\ \mathbf{R}_b^{-1} &= \Delta_{\theta_1} \mathbf{R}_{\Phi}^{-1} \Delta_{\theta_1}^H\end{aligned}$$

où $\Delta'_{\theta_1} \stackrel{\text{def}}{=} -2\pi f_0 \text{Diag}(\frac{d\tau_{1,1}}{d\theta_1}, \dots, \frac{d\tau_{1,M}}{d\theta_1})$. De ce fait

$$\begin{aligned}\mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \theta_1} &= \Delta_{\theta_1} \mathbf{R}_{\Phi}^{-1} \Delta_{\theta_1}^H \left(i \Delta'_{\theta_1} \Delta_{\theta_1} \mathbf{R}_{\Phi} \Delta_{\theta_1}^H - i \Delta_{\theta_1} \mathbf{R}_{\Phi} \Delta_{\theta_1}^H \Delta'_{\theta_1} \right) \\ &= i \left(\Delta_{\theta_1} \mathbf{R}_{\Phi}^{-1} \Delta'_{\theta_1} \mathbf{R}_{\Phi} \Delta_{\theta_1}^H - \Delta'_{\theta_1} \right)\end{aligned}$$

et par suite

$$\text{Tr} \left(\mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \theta_1} \mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \theta_1} \right) = 2 \text{Tr} \left(\Delta'_{\theta_1} \mathbf{R}_{\Phi}^{-1} \Delta'_{\theta_1} \mathbf{R}_{\Phi} \right) - 2 \left(\text{Tr}(\Delta'_{\theta_1})^2 \right).$$

Ensuite, comme

$$\frac{\partial \mathbf{R}_b}{\partial \phi_k} = \Delta_{\theta_1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \Delta_{\theta_1}^H, \quad k = 2, \dots, L$$

$$\begin{aligned}\text{Tr} \left(\mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \theta_1} \mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \phi_k} \right) &= i \text{Tr} \left(\Delta_{\theta_1} \mathbf{R}_{\Phi}^{-1} \Delta'_{\theta_1} \mathbf{R}_{\Phi} \Delta_{\theta_1}^H \Delta_{\theta_1} \mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \Delta_{\theta_1}^H \right) - i \text{Tr} \left(\Delta'_{\theta_1} \Delta_{\theta_1} \mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \Delta_{\theta_1}^H \right) \\ &= i \left(\text{Tr} \left(\mathbf{R}_{\Phi}^{-1} \Delta'_{\theta_1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \right) - \text{Tr} \left(\Delta'_{\theta_1} \mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \right) \right).\end{aligned}$$

Or, puisque le spectre de la source est symétrique par rapport à sa fréquence centrale f_0 , $\mathbf{R}_{s_1}$ est réelle symétrique. Comme il en est de même de $\mathbf{R}_{\Phi}^{-1}$ et de $\frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k}$, nous avons donc :

$$\text{Tr} \left(\mathbf{R}_{\Phi}^{-1} \Delta'_{\theta_1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \right) - \text{Tr} \left(\Delta'_{\theta_1} \mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \right) = \text{Tr} \left(\frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \Delta'_{\theta_1} \mathbf{R}_{\Phi}^{-1} \right) - \text{Tr} \left(\Delta'_{\theta_1} \mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \right) = 0.$$

De ce fait

$$\text{Tr} \left(\mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \theta_1} \mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \theta_k} \right) = 0 \text{ for } k = 2, \dots, L.$$

Enfin, comme

$$\begin{aligned}\text{Tr} \left(\mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \phi_k} \mathbf{R}_b^{-1} \frac{\partial \mathbf{R}_b}{\partial \phi_l} \right) &= \text{Tr} \left(\Delta_{\theta_1} \mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \Delta_{\theta_1}^H \Delta_{\theta_1} \mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_l} \Delta_{\theta_1}^H \right) \text{ pour } k = 2, \dots, L \\ &= \text{Tr} \left(\mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_k} \mathbf{R}_{\Phi}^{-1} \frac{\partial \mathbf{R}_{\Phi}}{\partial \phi_l} \right)\end{aligned} \tag{7.1}$$

ceci conclut la démonstration de (5.4), (5.5) et (5.6).

Annexe 6 : Preuve de la relation (5.8)

Grâce à $\mathbf{R}_\Phi = \sigma_1^2 \mathbf{1}\mathbf{1}^T + (\sigma_n^2 \mathbf{I}_M + \delta \mathbf{R}_{s_1})$ et à

$$(\sigma_n^2 \mathbf{I}_M + \delta \mathbf{R}_{s_1})^{-1} = \frac{1}{\sigma_n^2} \mathbf{I}_M - \frac{1}{\sigma_n^4} \delta \mathbf{R}_{s_1} + O(\delta \mathbf{R}_{s_1}^2),$$

le lemme d'inversion matricielle (cf. [14, p.571] par ex.) donne, après quelques manipulations algébriques,

$$\begin{aligned} \mathbf{R}_\Phi^{-1} &= \left(\frac{1}{\sigma_n^2} \mathbf{I}_M - \frac{1}{\sigma_n^4} \delta \mathbf{R}_{s_1} \right) - \left(\frac{1}{\sigma_n^2} \mathbf{I}_M - \frac{1}{\sigma_n^4} \delta \mathbf{R}_{s_1} \right) \\ &\quad \left(\frac{\sigma_n^2 \sigma_1^2}{M \sigma_1^2 + \sigma_n^2} \mathbf{1}\mathbf{1}^T + \frac{\sigma_1^4}{(M \sigma_1^2 + \sigma_n^2)^2} \mathbf{1}\mathbf{1}^T \delta \mathbf{R}_{s_1} \mathbf{1}\mathbf{1}^T + O(\delta \mathbf{R}_{s_1}^2) \right) \left(\frac{1}{\sigma_n^2} \mathbf{I}_M - \frac{1}{\sigma_n^4} \delta \mathbf{R}_{s_1} \right) + O(\delta \mathbf{R}_{s_1}^2) \\ &= \frac{1}{\sigma_n^2} \mathbf{I}_M - \frac{\sigma_1^2}{\sigma_n^2 (M \sigma_1^2 + \sigma_n^2)} \mathbf{1}\mathbf{1}^T + a(\delta \mathbf{R}_{s_1}) + O(\delta \mathbf{R}_{s_1}^2) \end{aligned}$$

où $a(\delta \mathbf{R}_{s_1})$ est une expression linéaire de $\delta \mathbf{R}_{s_1}$ définie par :

$$\begin{aligned} a(\delta \mathbf{R}_{s_1}) &\stackrel{\text{def}}{=} -\frac{1}{\sigma_n^4} \delta \mathbf{R}_{s_1} + \frac{\sigma_1^2}{\sigma_n^4 (M \sigma_1^2 + \sigma_n^2)} (\delta \mathbf{R}_{s_1} \mathbf{1}\mathbf{1}^T + \mathbf{1}\mathbf{1}^T \delta \mathbf{R}_{s_1}) - \frac{\sigma_1^4}{\sigma_n^4 (M \sigma_1^2 + \sigma_n^2)^2} \mathbf{1}\mathbf{1}^T \delta \mathbf{R}_{s_1} \mathbf{1}\mathbf{1}^T \\ &= \sigma_1^2 \left(\frac{f_\sigma^2}{f_0^2} \right) \left(-\frac{1}{\sigma_n^4} \mathbf{U}_1 + \frac{\sigma_1^2}{\sigma_n^4 (M \sigma_1^2 + \sigma_n^2)} (\mathbf{U}_1 \mathbf{1}\mathbf{1}^T + \mathbf{1}\mathbf{1}^T \mathbf{U}_1) - \frac{\sigma_1^4}{\sigma_n^4 (M \sigma_1^2 + \sigma_n^2)^2} (\mathbf{1}\mathbf{1}^T \mathbf{U}_1 \mathbf{1}\mathbf{1}^T) \right) \\ &+ O\left(\frac{f_\sigma^4}{f_0^4}\right) = \sigma_1^2 \left(\frac{f_\sigma^2}{f_0^2} \right) \bar{\mathbf{R}}_{s_1} + O\left(\frac{f_\sigma^4}{f_0^4}\right) \end{aligned}$$

avec $\bar{\mathbf{R}}_{s_1} \stackrel{\text{def}}{=} -\frac{1}{\sigma_n^4} \mathbf{U}_1 + \frac{\sigma_1^2}{\sigma_n^4 (M \sigma_1^2 + \sigma_n^2)} (\mathbf{U}_1 \mathbf{1}\mathbf{1}^T + \mathbf{1}\mathbf{1}^T \mathbf{U}_1) - \frac{\sigma_1^4}{\sigma_n^4 (M \sigma_1^2 + \sigma_n^2)^2} (\mathbf{1}\mathbf{1}^T \mathbf{U}_1 \mathbf{1}\mathbf{1}^T)$ où $\mathbf{U}_1$ est définie dans le Résultat 5.1. Grâce aux différentes expressions précédentes nous pouvons écrire

$$\text{Tr}(\Delta'_{\theta_1} \mathbf{R}_\Phi^{-1} \Delta'_{\theta_1} \mathbf{R}_\Phi) - \text{Tr}(\Delta'_{\theta_1})^2 = \text{Tr}(\mathbf{A}) + \text{Tr}(\mathbf{B}) + \text{Tr}(\mathbf{C}) - \text{Tr}(\Delta'_{\theta_1})^2 + O(\delta \mathbf{R}_{s_1}^2)$$

avec

$$\begin{aligned} \mathbf{A} &= \Delta'_{\theta_1} \left(\frac{1}{\sigma_n^2} \mathbf{I}_M - \alpha \mathbf{1}\mathbf{1}^T \right) \Delta'_{\theta_1} (\sigma_n^2 \mathbf{I}_M + \sigma_1^2 \mathbf{1}\mathbf{1}^T) \\ \mathbf{B} &= \Delta'_{\theta_1} \left(\frac{1}{\sigma_n^2} \mathbf{I}_M - \alpha \mathbf{1}\mathbf{1}^T \right) \Delta'_{\theta_1} \delta \mathbf{R}_{s_1} \\ \mathbf{C} &= \Delta'_{\theta_1} a(\delta \mathbf{R}_{s_1}) \Delta'_{\theta_1} (\sigma_n^2 \mathbf{I}_M + \sigma_1^2 \mathbf{1}\mathbf{1}^T) \end{aligned}$$

en ayant posé $\alpha \stackrel{\text{def}}{=} \frac{\sigma_1^2}{\sigma_n^2 (M \sigma_1^2 + \sigma_n^2)}$ et avec, rappelons le, $\Delta_{\theta_1} \stackrel{\text{def}}{=} \text{Diag}(\mathbf{a}_{\theta_1})$ et $\Delta'_{\theta_1} \stackrel{\text{def}}{=} -2\pi f_0 \text{Diag}(\tau'_{1,1}, \dots, \tau'_{1,M})$ de sorte que $\mathbf{a}'_{\theta_1} = i \Delta'_{\theta_1} \mathbf{a}_{\theta_1} = i \Delta'_{\theta_1} \Delta_{\theta_1} \mathbf{1}$. Notons qu'avec ces notations $\text{Tr}(\Delta'_{\theta_1})^2 = \|\mathbf{a}'_{\theta_1}\|^2$ et que $\mathbf{1}^T \Delta'_{\theta_1} \mathbf{1} \mathbf{1}^T \Delta'_{\theta_1} \mathbf{1} = |\mathbf{a}'_{\theta_1}{}^H \mathbf{a}_{\theta_1}|^2$. Après quelques manipulations algébriques on obtient tout d'abord :

$$\text{Tr}(\mathbf{A}) - \text{Tr}(\Delta'_{\theta_1})^2 = \frac{\sigma_1^4 \left(M \|\mathbf{a}'_{\theta_1}\|^2 - |\mathbf{a}'_{\theta_1}{}^H \mathbf{a}_{\theta_1}|^2 \right)}{\sigma_n^2 (M \sigma_1^2 + \sigma_n^2)} = \frac{\frac{\sigma_1^4}{\sigma_n^4} \left(M \|\mathbf{a}'_{\theta_1}\|^2 - |\mathbf{a}'_{\theta_1}{}^H \mathbf{a}_{\theta_1}|^2 \right)}{1 + \frac{\sigma_1^2}{\sigma_n^2} M}$$

Notons au passage que

$$\frac{1}{2T} \left(\text{Tr}(\mathbf{A}) - \text{Tr}(\Delta'_{\theta_1})^2 \right)^{-1} = \frac{1}{2T} \frac{1 + \frac{\sigma_1^2}{\sigma_n^2} M}{\frac{\sigma_1^4}{\sigma_n^4} \left(\|\mathbf{a}_{\theta_1}\|^2 \|\mathbf{a}'_{\theta_1}\|^2 - |\mathbf{a}'_{\theta_1}{}^H \mathbf{a}_{\theta_1}|^2 \right)}$$

n'est autre que la borne de Cramer-Rao $BCR_{\theta_1^0}$ dans le cas bande-étroite (cf. par exemple [15, rel. (17)]). On obtient ensuite

$$\mathbf{B} = \frac{1}{\sigma_n^2} \Delta'_{\theta_1}{}^2 \delta \mathbf{R}_{s_1} - \alpha \Delta'_{\theta_1} \mathbf{1} \mathbf{1}^T \Delta'_{\theta_1} \delta \mathbf{R}_{s_1}.$$

On en déduit

$$\begin{aligned} \text{Tr}(\mathbf{B}) &= \frac{1}{\sigma_n^2} \text{Tr}(\Delta'_{\theta_1}{}^2 \delta \mathbf{R}_{s_1}) - \alpha \mathbf{a}_{\theta_1}^H \Delta'_{\theta_1} \delta \mathbf{R}_{s_1} \Delta'_{\theta_1} \mathbf{a}_{\theta_1} \\ &= \sigma_1^2 \frac{f_\sigma^2}{f_0^2} \left(\frac{1}{\sigma_n^2} \text{Tr}(\Delta'_{\theta_1}{}^2 \mathbf{U}_1) - \alpha \mathbf{d}_{\theta_1}' \mathbf{U}_1 \mathbf{d}_{\theta_1} \right) + O\left(\frac{f_\sigma^4}{f_0^4}\right), \end{aligned}$$

avec, rappelons le $\mathbf{d}_{\theta_1} \stackrel{\text{def}}{=} \Delta'_{\theta_1} \mathbf{1}$. On en déduit enfin

$$\mathbf{C} = \mathbf{C}_1 + \mathbf{C}_2$$

avec

$$\mathbf{C}_1 \stackrel{\text{def}}{=} \sigma_1^2 \sigma_n^2 \frac{f_\sigma^2}{f_0^2} \Delta'_{\theta_1} \bar{\mathbf{R}}_{s_1} \Delta'_{\theta_1} \quad \text{et} \quad \mathbf{C}_2 \stackrel{\text{def}}{=} \sigma_1^4 \frac{f_\sigma^2}{f_0^2} \Delta'_{\theta_1} \bar{\mathbf{R}}_{s_1} \Delta'_{\theta_1} \mathbf{1} \mathbf{1}^T.$$

D'où

$$\begin{aligned} \text{Tr}(\mathbf{C}_1) &= \sigma_1^2 \frac{f_\sigma^2}{f_0^2} \left(-\frac{1}{\sigma_n^2} \text{Tr}(\Delta'_{\theta_1}{}^2 \mathbf{U}_1) + \alpha \mathbf{d}_{\theta_1}'^T \Delta'_{\theta_1} \mathbf{U}_1 \mathbf{1} + \alpha \mathbf{1}^T \mathbf{U}_1 \Delta'_{\theta_1} \mathbf{d}_{\theta_1} - \alpha^2 \sigma_n^2 \|\mathbf{a}_{\theta_1}'\|^2 \mathbf{1}^T \mathbf{U}_1 \mathbf{1} \right) + O\left(\frac{f_\sigma^4}{f_0^4}\right) \\ \text{Tr}(\mathbf{C}_2) &= \sigma_1^2 \frac{f_\sigma^2}{f_0^2} \left(-\frac{\sigma_1^2}{\sigma_n^4} \mathbf{d}_{\theta_1}'^T \mathbf{U}_1 \mathbf{d}_{\theta_1} + \alpha \frac{\sigma_1^2}{\sigma_n^2} \left(i \mathbf{d}_{\theta_1}'^T \mathbf{U}_1 \mathbf{1} \mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1} - i \mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1} \mathbf{1}^T \mathbf{U}_1 \mathbf{d}_{\theta_1} \right) \right. \\ &\quad \left. - \alpha^2 |\mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}|^2 \mathbf{1}^T \mathbf{U}_1 \mathbf{1} \right) + O\left(\frac{f_\sigma^4}{f_0^4}\right) \end{aligned}$$

En définitive

$$\begin{aligned} \text{Tr}(\Delta'_{\theta_1} \mathbf{R}_\Phi^{-1} \Delta'_{\theta_1} \mathbf{R}_\Phi) - \text{Tr}(\Delta'_{\theta_1}{}^2) &= \frac{\frac{\sigma_1^4}{\sigma_n^4} \left(\|\mathbf{a}_{\theta_1}\|^2 \|\mathbf{a}_{\theta_1}'\|^2 - |\mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}|^2 \right)}{1 + \frac{\sigma_1^2}{\sigma_n^2} M} - \frac{f_\sigma^2}{f_0^2} \left[\left(\sigma_1^2 \alpha + \frac{\sigma_1^4}{\sigma_n^4} \right) \mathbf{d}_{\theta_1}'^T \mathbf{U}_1 \mathbf{d}_{\theta_1} \right. \\ &\quad \left. + \sigma_1^2 \alpha^2 \left(\sigma_n^2 \|\mathbf{a}_{\theta_1}'\|^2 + \sigma_1^2 |\mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}|^2 \right) \mathbf{1}^T \mathbf{U}_1 \mathbf{1} - 2 \sigma_1^2 \alpha \mathbf{d}_{\theta_1}'^T \Delta'_{\theta_1} \mathbf{U}_1 \mathbf{1} \right. \\ &\quad \left. + 2 \alpha \frac{\sigma_1^4}{\sigma_n^2} \Im(\mathbf{d}_{\theta_1}'^T \mathbf{U}_1 \mathbf{1} \mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}) \right] + O\left(\frac{f_\sigma^4}{f_0^4}\right) \end{aligned}$$

que l'on peut également écrire sous la forme

$$\text{Tr}(\Delta'_{\theta_1} \mathbf{R}_\Phi^{-1} \Delta'_{\theta_1} \mathbf{R}_\Phi) - \text{Tr}(\Delta'_{\theta_1}{}^2) = \frac{\frac{\sigma_1^4}{\sigma_n^4} \left(\|\mathbf{a}_{\theta_1}\|^2 \|\mathbf{a}_{\theta_1}'\|^2 - |\mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}|^2 \right)}{1 + \frac{\sigma_1^2}{\sigma_n^2} M} \left(1 - c' \left(\frac{f_\sigma^2}{f_0^2} \right) + O\left(\frac{f_\sigma^4}{f_0^4}\right) \right)$$

avec

$$c' = \frac{1}{\|\mathbf{a}_{\theta_1}\|^2 \|\mathbf{a}_{\theta_1}'\|^2 - |\mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}|^2} \left[\beta_1 \mathbf{d}_{\theta_1}'^T \mathbf{U}_1 \mathbf{d}_{\theta_1} + \beta_2 \mathbf{1}^T \mathbf{U}_1 \mathbf{1} - 2 \mathbf{d}_{\theta_1}'^T \Delta'_{\theta_1} \mathbf{U}_1 \mathbf{1} + \beta_3 \Im(\mathbf{d}_{\theta_1}'^T \mathbf{U}_1 \mathbf{1} \mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}) \right]$$

où

$$\beta_1 \stackrel{\text{def}}{=} \frac{M \sigma_1^2 + 2 \sigma_n^2}{\sigma_n^2}, \quad \beta_2 \stackrel{\text{def}}{=} \frac{\sigma_1^2 \sigma_n^2 \|\mathbf{a}_{\theta_1}'\|^2 + \sigma_1^4 |\mathbf{a}_{\theta_1}'^H \mathbf{a}_{\theta_1}|^2}{\sigma_n^2 (M \sigma_1^2 + \sigma_n^2)} \quad \text{et} \quad \beta_3 \stackrel{\text{def}}{=} 2 \frac{\sigma_1^2}{\sigma_n^2}.$$

Le résultat 5.5 est alors établi en utilisant (5.7) avec

$$BCR_{\theta_1^0} = \frac{1}{2T} \frac{1 + \frac{\sigma_1^2}{\sigma_n^2} M}{\frac{\sigma_1^4}{\sigma_n^4} \left(\|\mathbf{a}_{\theta_1}\|^2 \|\mathbf{a}'_{\theta_1}\|^2 - |\mathbf{a}_{\theta_1}^H \mathbf{a}_{\theta_1}|^2 \right)}$$

ou encore, d'après (5.2) du chapitre 4

$$BCR_{\theta_1^0} = \frac{1}{2T \sigma_1^4 \mathbf{a}_{\theta_1}^H \mathbf{P} \mathbf{a}_{\theta_1}} \left(\sigma_1^2 \sigma_n^2 + \frac{\sigma_n^4}{\|\mathbf{a}_{\theta_1}\|^2} \right).$$

où $\mathbf{P}$ est la matrice de projection sur l'espace bruit.

Annexe 7 : Autre justification de (5.7)

Partant de l'expression (5.2) de la matrice $\mathbf{R}_b$ et en considérant que la densité spectrale de puissance $S_1(f)$ de la source large bande est paramétrée par son écart type f_σ et la puissance de la source σ_1^2 nous constatons que le problème est paramétré par $\Psi = (\theta_1, f_\sigma, \sigma_1^2, \sigma_n^2)^T$. Nous avons toujours

$$\mathbf{R}_b = \Delta_{\theta_1} \mathbf{R}_\Phi \Delta_{\theta_1}^H$$

avec $\mathbf{R}_\Phi = \mathbf{R}_{s_1} + \sigma_n^2 \mathbf{I}_M$. La matrice $\mathbf{R}_{s_1}$ est paramétrée par $(f_\sigma \tau_{1,m})_{m=1,\dots,M}$ et σ_1^2 . Dans le cas particulier d'un réseau *linéaire*, comme $\tau_{1,m} = \frac{d_m}{c} \sin \theta_1$, $\mathbf{R}_\Phi$ est entièrement déterminée par le paramètre $\Phi = (\phi_2, \phi_3, \phi_4)$ avec $\phi_2 \stackrel{\text{def}}{=} f_\sigma \sin \theta_1$, $\phi_3 \stackrel{\text{def}}{=} \sigma_1^2$ et $\phi_4 \stackrel{\text{def}}{=} \sigma_n^2$ donc $\mathbf{R}_b$ est entièrement déterminée par le paramètre $\Psi' \stackrel{\text{def}}{=} (\theta_1, \Phi)$. Notons que dans le cas d'un réseau quelconque, la paramétrisation de $\mathbf{R}_\Phi$ nécessite plus que les 3 paramètres ϕ_k et qu'alors (θ_1, Φ) n'est plus un paramétrage suffisant de $\mathbf{R}_b$.

Grâce à la bijection $\Psi \leftrightarrow \Psi'$, la BCR du paramètre Ψ est liée à la MIF du paramètre Ψ' par la relation [14, rel. (3.30)] :

$$CRB_\Psi = \frac{\partial \Psi}{\partial \Psi'^T} \mathbf{I}^{-1}(\Psi') \frac{\partial \Psi^T}{\partial \Psi'} \quad (7.1)$$

En suivant la même démarche que celle empruntée en Annexe D on montre que :

$$[\mathbf{I}(\Psi')]_{1,k} = 0, \quad k = 2, 3, 4 \quad (7.2)$$

$$[\mathbf{I}(\Psi')]_{k,l} = T \operatorname{Tr} \left(\mathbf{R}_\Phi^{-1} \frac{\partial \mathbf{R}_\Phi}{\partial \phi_k} \mathbf{R}_\Phi^{-1} \frac{\partial \mathbf{R}_\Phi}{\partial \phi_l} \right), \quad k, l = 2, 3, 4 \quad (7.3)$$

$$[\mathbf{I}(\Psi')]_{1,1} = T \left(2 \operatorname{Tr} \left(\Delta_{\theta_1}' \mathbf{R}_\Phi^{-1} \Delta_{\theta_1}' \mathbf{R}_\Phi \right) - 2 \operatorname{Tr}(\Delta_{\theta_1}'^2) \right), \quad (7.4)$$

où $\Delta_{\theta_1}' \stackrel{\text{def}}{=} \operatorname{Diag}(-2\pi f_0 \frac{d_1}{c} \cos \theta_1, \dots, -2\pi f_0 \frac{d_M}{c} \cos \theta_1)$. En conséquence la MIF du paramètre Ψ' a une structure bloc-diagonale :

$$\begin{bmatrix} I(\theta_1) & \mathbf{0}^T \\ \mathbf{0} & \mathbf{I}(\Phi) \end{bmatrix}$$

où $I(\theta_1)$ représente la matrice d'information de Fisher du paramètre θ_1 considéré comme le seul paramètre inconnu et $\mathbf{I}(\Phi)$ la MIF du paramètre Φ déduit de T observations complexes circulaires gaussiennes centrées indépendantes de matrice de covariance $\mathbf{R}_\Phi$. Comme $\frac{\partial \Psi}{\partial \Psi'^T}$ a la structure particulière suivante :

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{f_\sigma}{\tan \theta_1} & \frac{1}{\sin \theta_1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

La BCR du paramètre Ψ est donnée par :

$$\begin{bmatrix} I^{-1}(\theta_1) & aI^{-1}(\theta_1) & 0 & 0 \\ aI^{-1}(\theta_1) & a^2I^{-1}(\theta_1) + b^2[\mathbf{I}^{-1}(\Phi)]_{1,1} & b[\mathbf{I}^{-1}(\Phi)]_{1,2} & b[\mathbf{I}^{-1}(\Phi)]_{1,3} \\ 0 & b[\mathbf{I}^{-1}(\Phi)]_{2,1} & [\mathbf{I}^{-1}(\Phi)]_{2,2} & [\mathbf{I}^{-1}(\Phi)]_{2,3} \\ 0 & b[\mathbf{I}^{-1}(\Phi)]_{3,1} & [\mathbf{I}^{-1}(\Phi)]_{3,2} & [\mathbf{I}^{-1}(\Phi)]_{3,3} \end{bmatrix}$$

avec $a \stackrel{\text{def}}{=} -\frac{f_\sigma}{\tan \theta_1}$ et $b \stackrel{\text{def}}{=} \frac{1}{\sin \theta_1}$. La BCR du paramètre θ_1 est donc :

$$BCR_{\theta_1} = I^{-1}(\theta_1) = \frac{1}{2T} \left(\text{Tr}(\Delta'_{\theta_1} \mathbf{R}_\Phi^{-1} \Delta'_{\theta_1} \mathbf{R}_\Phi) - \text{Tr}(\Delta'^2_{\theta_1}) \right)^{-1}. \quad (7.5)$$

Nous pouvons remarquer que si la largeur de la bande du signal est connue, et donc si f_σ est connu, la matrice $\mathbf{I}(\Psi)$ est bloc-diagonale et par suite les paramètres θ_1 et (σ_1^2, σ_n^2) sont découplés.

Bibliographie

- [1] S.V. Schell, W.A. Gardner, "High-resolution direction finding," in *Handbook of Statistics 10, Signal Processing and its Applications*, edited by N.K. Bose and C.R. Rao, Elsevier Science, 1994.
- [2] J. Krolick, *Focused wideband array processing for spatial estimation* in Advances in spectrum analysis and array processing, vol. II, Simon Haykin, Editor, Prentice Hall, 1991.
- [3] K. Buckley, "Spatial/spectral filtering with linearly constrained minimum variance beamformers," *IEEE Trans. Acoust. Speech and Signal Process.*, vol. 35, no. 3, pp. 249-266, March 1987.
- [4] K. Buckley, L. Griffiths, "Broadband signal-subspace spatial-spectrum (BASS-ALE) estimation for sensor array processing," *IEEE Trans. Acoust. Speech and Signal Process.*, vol. 36, no. 7, pp. 953-964, July 1988.
- [5] M. Zatman, "How narrow is narrowband," *IEE Proc. Radar, Sonar Navig.*, vol. 145, no. 2, pp. 85-91, April 1998.
- [6] J. Sorelius, R.L. Moses, T. Söderström and A.L. Swindlehurst, "Effects of nonzero bandwidth on direction of arrival estimators in array processing," *IEE Proc. Radar, Sonar Navig.*, vol. 145, no. 6, pp. 317-324, December 1998.
- [7] A.P. Liavas, P.A. Regalia and J.P. Delmas, "Blind channel approximation: Effective channel order determination," *IEEE Trans. Signal Processing*, vol. 47, no. 12, pp. 3336-3344, Dec. 1999.
- [8] P.E. Caines, *Linear stochastic systems*, Wiley series in probability and mathematical statistics, 1988.
- [9] W. Xu, M. Kaveh, "Analysis de the performance and sensitivity de eigendecomposition-based detectors," *IEEE Trans. on Signal Processing*, vol. 43, no. 6, pp. 1413-1426, June 1995.
- [10] D.R. Brillinger, *Times series, data analysis and theory*. Expanded Edition, Holden-Day, Inc., 1980.
- [11] J.F. Cardoso, E. Moulines, "Asymptotic performance analysis de direction-finding algorithms based on fourth-order cumulants," *IEEE Trans. Signal Processing*, vol. 43, no. 1, pp. 214-224, Jan. 1995.
- [12] P. Stoica, A. Nehorai, "MUSIC, Maximum likelihood, and Cramer-Rao Bound," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, no. 5, pp. 720-741, May 1989.
- [13] M. Kaveh, A.J. Barabell, "The statistical performance de MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 34, no. 2, pp. 331-341, April 1986.
- [14] S.M. Kay, *Fundamentals de statistical signal processing*, New York: Prentice Hall, 1993.
- [15] B. Ottersten, M. Viberg and T. Kailath, "Analysis de subspace fitting and ML techniques for parameter estimation from sensor array data," *IEEE Trans. on Signal Processing*, vol. 40, no. 3, pp. 590-599, March 1992.

- [16] O. Besson, P. Stoica, "Decoupled estimation de DOA and angular spread for spatially distributed source," *IEEE Trans. on Signal Processing*, vol. 48, no. 7, pp. 1872-1882, June 2000.
- [17] B. Ottersten, P. Stoica and R. Roy, "Covariance matching estimation techniques for array signal processing," *Digital Signal Processing*, vol. 8, pp. 185-210, 1998.
- [18] M. Bengtsson, *Antenna array signal processing for high rank data models*, PhD thesis, Royal Institute of technology, Stockholm, Sweden, December 1999. TRITA-S3-SB-9938.
- [19] B. Friedlander, A.J. Weiss, "Direction finding for wide-band signals using an interpolated array," *IEEE Trans. on Signal Processing*, vol. 41, no. 4, pp. 1618-1634, April 1993.
- [20] J.P. Delmas, "Asymptotic performance of second-order algorithms," *IEEE Trans. Signal Processing*, vol. 50, no. 1, pp. 49-57, January 2002.

Chapitre 6

Application des techniques de “Covariance Matching Estimation” (COMET) à l’estimation de direction d’arrivée ou de fréquence de sinusoides

6.1 Introduction

Les techniques dites de “Covariance Matching Estimation” (COMET) sont des méthodes d’estimation de type “moindres carrés pondérés non linéaires du deuxième ordre” qui tirent parti du fait que l’ensemble des paramètres inconnus peut être séparé en paramètres *linéaires* et *non-linéaires* selon que la matrice de covariance de l’observation dépend linéairement ou non de chacun d’eux. L’exploitation de cette propriété permet une séparation du problème d’estimation : les paramètres linéaires peuvent être obtenus analytiquement en fonction des paramètres non-linéaires ce qui permet de réduire la dimension du problème d’optimisation à résoudre. Ces techniques sont présentées notamment dans [1] dans le cadre d’observations complexes circulaires gaussiennes et indépendantes et il est montré, sous cette hypothèse, que les estimateurs obtenus sont asymptotiquement efficaces.

L’objet de ce chapitre est l’étude des performances de l’algorithme COMET, appliqué à l’estimation d’un paramètre réel Θ lorsque *les observations ne sont plus nécessairement gaussiennes et indépendantes*. Les résultats généraux de cette étude sont illustrés en considérant l’application de la méthode COMET à deux problèmes classiques en traitement du signal : l’estimation de direction d’arrivée de sources (DOA) et l’estimation de fréquences de sinusoides.

Le plan de ce chapitre est le suivant :

- Dans une première partie le modèle général des signaux et les notations sont présentés. Le principe général de l’estimateur COMET sous l’hypothèse d’observations complexes circulaires gaussiennes indépendantes est rappelé au paragraphe 6.2 en s’appuyant tout d’abord, comme dans [1], sur le principe d’invariance étendue (EXIP), puis étendu au cas d’observations réelles gaussiennes indépendantes. Les performances asymptotiques de cette méthode sont également reformulées mais d’une façon plus simple que dans [1]. Dans le paragraphe 6.3 nous reformulons, par à une approche différente, l’algorithme COMET dans le cadre d’observations gaussiennes réelles puis complexes circulaires indépendantes en le considérant comme un cas particulier d’une classe plus générale d’estimateurs *asymptotiquement de variance minimale* (AMV). L’étude des estimateurs AMV est ensuite étendu au cas d’observations non gaussiennes et/ou non indépendantes dans le paragraphe 6.4 et leurs performances asymptotiques étudiées dans le paragraphe 6.5.

- L’application de l’algorithme COMET classique à l’estimation de DOA est étudiée dans une deuxième partie au paragraphe 6.6, dans le cadre d’observations non nécessairement gaussiennes et indépendantes,

de sources spatialement non corrélées et d'un bruit additif spatialement blanc non uniforme sur les capteurs. Il est démontré que dans le cas de sources bande étroite, si les sources sont spatialement indépendantes, les performances de l'algorithme COMET sont insensibles la distribution statistique des signaux source. Nous montrons également que l'algorithme COMET peut être appliqué à une source large bande à condition que son spectre soit symétrique par rapport à la fréquence porteuse de démodulation et que dans ce cas la propriété précédente n'est plus vraie. Nous étudions également les conditions sous lesquelles l'algorithme COMET bande étroite est insensible à la corrélation temporelle des sources. Nous donnons enfin une expression analytique de la borne de Cramer-Rao de l'estimée de la DOA seule dans le cas d'une source bande étroite ou large bande et d'observations gaussiennes indépendantes. Cette expression nous permet de démontrer que les performances de COMET obtenues quand les puissances de bruit sur quelques capteurs tendent vers l'infini sont équivalentes à celles obtenues à l'aide du réseau privé de ces capteurs.

- Dans une troisième partie, paragraphe 6.7, un algorithme de type COMET modifié, asymptotiquement de variance minimale, est appliqué à l'estimation de fréquence de sinusôides observées dans un bruit additif MA. Les performances de cet algorithme sont étudiées et comparées à celles d'un estimateur de type "Pisarenko étendu" dans le cas d'un bruit MA et à celles de MUSIC dans le cas particulier d'un bruit blanc. Enfin, la robustesse de cet algorithme, confronté à une erreur sur l'ordre du modèle est étudiée.

6.2 Principe de l'algorithme COMET

La méthode COMET fait appel au *Principe d'Invariance Etendue* (EXIP) qui a été introduit dans [2]. Nous commencerons donc par rappeler brièvement ce Principe d'Invariance Etendue avant de rappeler les grandes lignes de l'algorithme COMET. Initialement décrit dans le cadre d'observations gaussiennes complexes circulaires indépendantes [1], nous le présenterons également dans le cadre d'observations réelles indépendantes.

6.2.1 Principe d'Invariance Etendue

Commençons par rappeler le principe d'Invariance Etendue. Soit Θ un paramètre vectoriel de dimension M_Θ , prenant ses valeurs dans un ensemble connexe $D_\Theta \subset \mathbb{R}^{M_\Theta}$, qui régit une observation $\mathbf{x}$ de densité de probabilité notée $p(\mathbf{x}; \Theta)$. L'estimateur au sens du maximum de vraisemblance de Θ est

$$\hat{\Theta} = \arg \max_{\Theta \in D_\Theta} p(\mathbf{x}; \Theta)$$

ou encore, de façon équivalente,

$$\hat{\Theta} = \arg \min_{\Theta \in D_\Theta} L_\Theta(\Theta) \quad (2.1)$$

avec, comme fonction de coût, $L_\Theta(\Theta) \stackrel{\text{def}}{=} -\ln(p(\mathbf{x}; \Theta))$, l'opposée de la fonction de log-vraisemblance.

Soit maintenant un paramètre Γ , de dimension M_Γ , déduit de Θ par une application $f : D_\Theta \rightarrow D_\Gamma \subset \mathbb{R}^{M_\Gamma}$. Définissons la fonction de vraisemblance $p(\mathbf{x}; \Gamma)$ induite de $\Gamma = f(\Theta)$

$$p(\mathbf{x}; \Gamma) \stackrel{\text{def}}{=} \max_{\Theta \in f^{-1}(\Gamma)} p(\mathbf{x}; \Theta)$$

et la fonction de coût associée

$$L_\Gamma(\Gamma) \stackrel{\text{def}}{=} -\ln(p(\mathbf{x}; \Gamma)).$$

Il est immédiat de démontrer (cf. [3, p.74]) que si $\hat{\Theta}$ est l'estimateur de maximum de vraisemblance de Θ , alors

$$\hat{\Gamma} = f(\hat{\Theta}) \quad (2.2)$$

est la valeur qui maximise la fonction de vraisemblance induite $p(\mathbf{x}; \Gamma)$,

$$\hat{\Gamma} = \arg \min_{\Gamma \in D_{\Gamma}} L_{\Gamma}(\Gamma). \quad (2.3)$$

C'est le *principe d'invariance classique de l'estimateur de maximum de vraisemblance*.

Remarque 1 : Si la fonction f est bijective alors $L_{\Gamma}(\Gamma) = L_{\Theta}(f^{-1}(\Gamma))$ et le principe d'invariance classique n'est autre que la propriété d'invariance de l'estimateur de maximum de vraisemblance par reparamétrisation bijective.

Remarque 2 : Ce principe d'invariance classique est en général présenté pour une sous-paramétrisation du problème ($M_{\Gamma} \leq M_{\Theta}$). Il peut l'être aussi dans le cas d'une sur-paramétrisation ($M_{\Gamma} \geq M_{\Theta}$), en particulier pour traiter le cas où le modèle $p(\mathbf{x}; \Theta)$ est plus "structuré" que le modèle $p(\mathbf{x}; \Gamma)$. Dans ce cas [2], la fonction de perte $L_{\Gamma}(\Gamma)$ est arbitrairement définie pour "déstructurer" le problème de minimisation (2.1) et vérifie

$$L_{\Gamma}(f(\Theta)) = L_{\Theta}(\Theta) \quad \forall \Theta \in D_{\Theta}$$

Dans la pratique, le problème de minimisation (2.3) peut être difficile en raison de la contrainte $\Gamma \in D_{\Gamma}$.

Pour simplifier ce problème de minimisation, l'idée est d'élargir le domaine de contrainte D_{Γ} à un ensemble $\tilde{D}_{\Gamma}$ ($D_{\Gamma} \subset \tilde{D}_{\Gamma}$) et le problème (2.3) devient

$$\tilde{\Gamma} = \arg \min_{\gamma \in \tilde{D}_{\Gamma}} L_{\Gamma}(\gamma). \quad (2.4)$$

Remarquons qu'en général D_{Γ} est un sous-ensemble de mesure nulle de $\tilde{D}_{\Gamma}$ et qu'alors le principe d'invariance classique ne peut être utilisé comme dans (2.2) pour obtenir un estimateur de Θ à partir de $\tilde{\Gamma}$ car $\nexists$ p.s. $\Theta \in D_{\Theta}$ t.q. $\tilde{\Gamma} = f(\hat{\Theta})$.

Par contre, si les différentes estimées $\hat{\Theta}$, $\hat{\Gamma}$ et $\tilde{\Gamma}$ sont indexées par un entier N (le nombre d'échantillons par exemple) le résultat suivant, qui constitue la formulation du principe d'invariance étendue (EXIP), a été démontré dans [2] :

Si

- $\hat{\Gamma}_N$ et $\tilde{\Gamma}_N$ convergent en probabilité vers la même limite,
- $L_{\Gamma}(\Gamma)$ est indéfiniment dérivable,
- $L_{\Gamma}''(\tilde{\Gamma})$ est définie positive, où $L_{\Gamma}''(\Gamma) \stackrel{\text{def}}{=} \frac{\delta^2 L_{\Gamma}(\Gamma)}{\delta \Gamma \delta \Gamma^T}$,

alors

$$\tilde{\Theta} = \arg \min_{\Theta \in D_{\Theta}} \left[\tilde{\Gamma} - f(\Theta) \right]^T L_{\Gamma}''(\tilde{\Gamma}) \left[\tilde{\Gamma} - f(\Theta) \right] \quad (2.5)$$

est un estimateur asymptotiquement équivalent¹ à $\hat{\Theta}$ quand $N \rightarrow \infty$.

¹ au sens où $\sqrt{N}(\tilde{\Theta} - \Theta)$ et $\sqrt{N}(\hat{\Theta} - \Theta)$ ont même distribution asymptotique.

6.2.2 Principe de l'algorithme COMET original

De façon très générale nous supposons que le M -vecteur représentant l'observation à l'instant t peut s'écrire :

$$\mathbf{y}_t = \mathbf{x}_t + \mathbf{n}_t$$

où $\mathbf{x}_t$ représente le signal contenant le paramètre d'intérêt Θ_1 et éventuellement un paramètre de nuisance Υ et $\mathbf{n}_t$ représente le bruit dépendant d'un paramètre de nuisance Σ et indépendant de $\mathbf{x}_t$. La matrice de covariance de l'observation peut alors s'écrire :

$$\mathbf{R}_y(\Theta_1, \Upsilon, \Sigma) = E(\mathbf{y}_t \mathbf{y}_t^H) = \mathbf{R}_x(\Theta_1, \Upsilon) + \mathbf{R}_n(\Sigma)$$

Nous supposons de plus, et ceci est fondamental pour la mise en œuvre de l'algorithme COMET, que les matrices de covariance du signal $\mathbf{R}_x(\Theta_1, \Upsilon) = E(\mathbf{x}_t \mathbf{x}_t^H)$ et du bruit $\mathbf{R}_n(\Sigma) = E(\mathbf{n}_t \mathbf{n}_t^H)$ dépendent linéairement des paramètres respectifs Υ et Σ (en général la matrice de covariance signal dépend de façon fortement non-linéaire du paramètre à estimer Θ_1). La matrice de covariance $\mathbf{R}_y$ dépend alors linéairement du paramètre $\Theta_2 \stackrel{\text{def}}{=} (\Upsilon^T, \Sigma^T)^T$ et sa vectorisation s'écrit sous la forme :

$$\mathbf{r}_y \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_y) = \Psi(\Theta_1) \Theta_2.$$

Nous supposons enfin que $\Theta \stackrel{\text{def}}{=} (\Theta_1^T, \Theta_2^T)^T$ est strictement identifiable à partir de la matrice $\mathbf{R}_y$, en d'autres termes que :

$$\Psi(\Theta_1) \Theta_2 = \Psi(\Theta'_1) \Theta'_2 \quad \Rightarrow \quad \Theta_1 = \Theta'_1 \text{ et } \Theta_2 = \Theta'_2.$$

Remarquons que cette condition et donc le choix du paramètre Θ_2 qui lui est lié dépend de l'application considérée. Par exemple, en estimation de DOA, le choix du paramètre Υ dépend, comme nous le verrons plus en détail par la suite, de la géométrie du réseau d'antenne.

Les observations sont supposées maintenant gaussiennes et indépendantes. Nous distinguerons successivement le cas d'observations complexes circulaires et le cas d'observations réelles.

Dans le cas d'observations $(\mathbf{y}_t)_{t=1, \dots, T}$ *gaussiennes complexes circulaires centrées et indépendantes*, l'estimateur au sens du maximum de vraisemblance de Θ est :

$$\begin{aligned} \hat{\Theta}_T &= \arg \max_{\Theta \in D_\Theta} \left(C - T \ln(\text{Det}(\mathbf{R}_y(\Theta))) - \sum_{t=1}^T \mathbf{y}_t^H \mathbf{R}_y^{-1}(\Theta) \mathbf{y}_t \right) \\ &= \arg \max_{\Theta \in D_\Theta} (C - T (\ln(\text{Det}(\mathbf{R}_y(\Theta))) + \text{Tr}(\mathbf{R}_y^{-1}(\Theta) \mathbf{R}_T))) \end{aligned}$$

avec $\mathbf{R}_T \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^H$, ce qui peut s'écrire

$$\hat{\Theta}_T = \arg \min_{\Theta \in D_\Theta} L_\Theta(\Theta) \quad \text{avec} \quad L_\Theta(\Theta) = \ln(\text{Det}(\mathbf{R}_y(\Theta))) + \text{Tr}(\mathbf{R}_y^{-1}(\Theta) \mathbf{R}_T). \quad (2.6)$$

Considérons un "reparamétrage" du problème par l'intermédiaire des parties réelles et imaginaires des termes de la matrice hermitienne $\mathbf{R}_y(\Theta)$. Ceci définit un nouveau paramètre inconnu

$$\Gamma = \mathbf{J} \text{Vec}(\mathbf{R}_y(\Theta)),$$

où $\mathbf{J}$ est une matrice inversible, ainsi qu'une application

$$f : D_\Theta \rightarrow D_\Gamma \quad \text{t.q.} \quad \Gamma = f(\Theta).$$

Le paramètre Θ étant supposé strictement identifiable à partir de $\mathbf{R}_y$, l'application f est bijective de D_Θ dans D_Γ . L'estimateur au sens du maximum de vraisemblance de Γ est alors, d'après le principe d'invariance de cet estimateur,

$$\hat{\Gamma}_T = \arg \min_{\Gamma \in D_\Gamma} L_\Gamma(\Gamma) \quad \text{avec} \quad L_\Gamma(\Gamma) \stackrel{\text{def}}{=} L_\Theta(f^{-1}(\Gamma))$$

Si maintenant on élargit le domaine de contrainte D_Γ en considérant un modèle $\mathbf{R}_y(\Gamma)$ de $\mathbf{R}_y$ non structuré en Θ , la matrice $\mathbf{R}_y(\Gamma)$ étant simplement supposée hermitienne, ceci définit un nouveau domaine de contrainte $\tilde{D}_\Gamma$ ($D_\Gamma \subset \tilde{D}_\Gamma$).

L'estimateur au sens du maximum de vraisemblance de Γ sur ce domaine est :

$$\tilde{\Gamma}_T = \arg \min_{\Gamma \in \tilde{D}_\Gamma} L_\Gamma(\Gamma).$$

Or

$$dL_\Gamma(\Gamma) = \text{Tr} (\mathbf{R}_y^{-1}(\Gamma)(\mathbf{R}_y(\Gamma) - \mathbf{R}_T)\mathbf{R}_y^{-1}(\Gamma)d\mathbf{R}_y(\Gamma))$$

et

$$dL_\Gamma(\Gamma) = 0 \Leftrightarrow \mathbf{R}_y(\Gamma) = \mathbf{R}_T \Leftrightarrow \Gamma = \mathbf{J}\text{Vec}(\mathbf{R}_T).$$

Par ailleurs,

$$d^2L_\Gamma(\Gamma) = -\text{Tr} (\mathbf{R}_y^{-1}(\Gamma)d\mathbf{R}_y(\Gamma)\mathbf{R}_y^{-1}(\Gamma)d\mathbf{R}_y(\Gamma)\mathbf{R}_y^{-1}(\Gamma)(\mathbf{R}_y(\Gamma) - 2\mathbf{R}_T))$$

d'où

$$d^2L_\Gamma(\Gamma)|_{\mathbf{R}_y(\Gamma)=\mathbf{R}_T} = \text{Tr} (\mathbf{R}_T^{-1}d\mathbf{R}_y(\Gamma)\mathbf{R}_T^{-1}d\mathbf{R}_y(\Gamma)) > 0$$

donc $L_\Gamma(\Gamma)$ admet un minimum unique en $\Gamma = \mathbf{J}\text{Vec}(\mathbf{R}_T)$ et par suite $\tilde{\Gamma}_T = \mathbf{J}\text{Vec}(\mathbf{R}_T)$.

Par ailleurs, $\mathbf{R}_T$ est un estimateur consistant de $\mathbf{R}_y$, et l'estimateur du maximum de vraisemblance de Γ étant consistant, on a la convergence en probabilité suivante :

$$\lim_{T \rightarrow \infty} \tilde{\Gamma}_T = \mathbf{J}\text{Vec}(\mathbf{R}_y) = \lim_{T \rightarrow \infty} \hat{\Gamma}_T = f(\Theta).$$

D'autre part, grâce aux propriétés (0.7) et (0.2) de l'opérateur de vectorisation et du produit de Kronecker, nous avons

$$\begin{aligned} d^2L_\Gamma(\Gamma)|_{\mathbf{R}_y(\Gamma)=\mathbf{R}_T} &= \text{Tr} (d\mathbf{R}_y(\Gamma)\mathbf{R}_T^{-1}d\mathbf{R}_y(\Gamma)\mathbf{R}_T^{-1}) \\ &= \text{Vec}^T(d\mathbf{R}_y^T(\Gamma))\text{Vec}(\mathbf{R}_T^{-1}d\mathbf{R}_y(\Gamma)\mathbf{R}_T^{-1}) \\ &= \text{Vec}^T(d\mathbf{R}_y^T(\Gamma))(\mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1})\text{Vec}(d\mathbf{R}_y(\Gamma)). \end{aligned}$$

La matrice $\mathbf{J}$ étant inversible et la matrice $\mathbf{R}_y(\Gamma)$ étant hermitienne, de $\Gamma = \mathbf{J}\text{Vec}(\mathbf{R}_y(\Gamma))$ on déduit que

$$\text{Vec}(d\mathbf{R}_y(\Gamma)) = \mathbf{J}^{-1}d\Gamma \quad \text{et} \quad \text{Vec}^T(d\mathbf{R}_y^T(\Gamma)) = d\Gamma^T\mathbf{J}^{-H}.$$

Par suite,

$$d^2L_\Gamma(\Gamma)|_{\mathbf{R}_y(\Gamma)=\mathbf{R}_T} = d\Gamma^T\mathbf{J}^{-H}(\mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1})\mathbf{J}^{-1}d\Gamma,$$

et, d'après la propriété d'identification [4, Théorème 6, p. 87], nous avons :

$$\frac{\partial^2 L_\Gamma(\Gamma)}{\partial \Gamma \partial \Gamma^T} \Big|_{\tilde{\Gamma}_T} = \mathbf{J}^{-H}(\mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1})\mathbf{J}^{-1}. \quad (2.7)$$

Pour être dans les conditions d'application d'EXIP, ce principe étant défini dans le cas réel, il faut que la matrice $\frac{\partial^2 L_T(\mathbf{\Gamma})}{\partial \mathbf{\Gamma} \partial \mathbf{\Gamma}^T} \big|_{\hat{\mathbf{\Gamma}}_T}$ soit réelle. Pour montrer que tel est bien le cas, considérons le raisonnement suivant.

Soit une suite de variables aléatoires $(\mathbf{z}_t)_{t \in \mathbb{N}}$ indépendantes, complexes circulaires gaussiennes centrées qui aurait comme matrice de covariance $\mathbf{R}_T$. Puisque $\mathbf{R}_T \otimes_c \mathbf{R}_T$ est la matrice de covariance asymptotique de la covariance empirique $\mathbf{R}_{z,N} = \frac{1}{N} \sum_{t=1}^N \mathbf{z}_t \mathbf{z}_t^H$ vectorisée,

$$\mathbf{R}_T \otimes_c \mathbf{R}_T = \lim_{N \rightarrow \infty} \mathbb{E} \left[(\text{Vec}(\mathbf{R}_{z,N}) - \text{Vec}(\mathbf{R}_T)) (\text{Vec}(\mathbf{R}_{z,N}) - \text{Vec}(\mathbf{R}_T))^H \right]$$

et que $\mathbf{J} \text{Vec}(\mathbf{R}_T)$ et $\mathbf{J} \text{Vec}(\mathbf{R}_{z,N})$ sont réels (les matrices $\mathbf{R}_T$ et $\mathbf{R}_{z,N}$ sont structurées, du point de vue parties réelles et parties imaginaires des termes, comme $\mathbf{R}_y(\Theta)$), $\mathbf{J}(\mathbf{R}_T \otimes_c \mathbf{R}_T) \mathbf{J}^H$ est réelle. Ce raisonnement, appliqué à (2.7), montre que $\frac{\partial^2 L_T(\mathbf{\Gamma})}{\partial \mathbf{\Gamma} \partial \mathbf{\Gamma}^T} \big|_{\hat{\mathbf{\Gamma}}_T}$ est bien réelle.

Nous sommes bien dans les conditions d'application d'EXIP.

En posant $\mathbf{s}_T \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_T)$ et en remarquant que $\mathbf{J} \mathbf{s}_T - \mathbf{J} \Psi(\Theta_1) \Theta_2$ est réel, l'estimateur

$$\begin{aligned} \tilde{\Theta}_T = (\tilde{\Theta}_{1,T}^T, \tilde{\Theta}_{2,T}^T)^T &= \arg \min_{\Theta \in D_\Theta} ((\mathbf{J} \mathbf{s}_T - \mathbf{J} \Psi(\Theta_1) \Theta_2)^T \mathbf{J}^{-H} (\mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1}) \mathbf{J}^{-1} (\mathbf{J} \mathbf{s}_T - \mathbf{J} \Psi(\Theta_1) \Theta_2)) \\ &= \arg \min_{\Theta \in D_\Theta} ((\mathbf{s}_T - \Psi(\Theta_1) \Theta_2)^H (\mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1}) (\mathbf{s}_T - \Psi(\Theta_1) \Theta_2)) \\ &= \arg \min_{\Theta \in D_\Theta} V(\Theta) \end{aligned} \quad (2.8)$$

avec

$$V(\Theta) = V(\Theta_1, \Theta_2) \stackrel{\text{def}}{=} [\mathbf{s}_T - \Psi(\Theta_1) \Theta_2]^H \mathbf{W}_T [\mathbf{s}_T - \Psi(\Theta_1) \Theta_2]$$

et

$$\mathbf{W}_T \stackrel{\text{def}}{=} \mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1},$$

est, d'après EXIP, un estimateur asymptotiquement équivalent à l'estimateur au sens du maximum de vraisemblance $\hat{\Theta}_T$.

Nous constatons que l'estimateur ainsi obtenu n'est autre que l'estimateur COMET de matrice de pondération $\mathbf{W}_T = \mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1}$ et que, dans cette situation, cet estimateur, asymptotiquement équivalent à l'estimateur du maximum de vraisemblance, est asymptotiquement efficace.

Il reste maintenant à voir comment la "séparation" du paramètre Θ en termes linéaires et non linéaires permet de simplifier l'estimation du paramètre d'intérêt Θ_1 .

Remarque : Dans le développement qui suit, nous supposons uniquement que la matrice de pondération $\mathbf{W}_T$ de l'algorithme COMET est telle que $\mathbf{W}_T^{-1}$ est la matrice de covariance asymptotique d'une matrice de covariance empirique vectorisée d'une suite de variables aléatoires $\mathbf{z}_t \in \mathbb{C}^M$ de distribution quelconque.

La minimisation (2.8) peut être réalisée en 2 étapes en considérant tout d'abord le paramètre linéaire Θ_2 ,

$$\begin{aligned} \tilde{\Theta}_{2,T} &= \arg \left(\min_{\Theta_2} [\mathbf{s}_T - \Psi(\Theta_1) \Theta_2]^H \mathbf{W}_T [\mathbf{s}_T - \Psi(\Theta_1) \Theta_2] \right) \\ &= \arg \left(\min_{\Theta_2} \|\mathbf{s}_T - \Psi(\Theta_1) \Theta_2\|_{\mathbf{W}_T}^2 \right) \\ &= \arg \left(\min_{\Theta_2} \|\mathbf{W}_T^{1/2} \mathbf{s}_T - \mathbf{W}_T^{1/2} \Psi(\Theta_1) \Theta_2\|^2 \right). \end{aligned}$$

Par une approche géométrique élémentaire, ce minimum est donné dans $\mathbb{C}^{N_{\Theta_2}}$ par

$$\tilde{\Theta}_{2,T} = [\Psi^H(\Theta_1) \mathbf{W}_T \Psi(\Theta_1)]^{-1} \Psi^H(\Theta_1) \mathbf{W}_T \mathbf{s}_T \quad (2.9)$$

mais, par définition le paramètre Θ_2 est réel, aussi il convient de montrer que $\tilde{\Theta}_{2,T}$ est bien réel dans ce cas. Ceci est justifié en Annexe 1.

Pour terminer, calculons $\tilde{\Theta}_{1,T} = \arg \min_{\Theta_1} V(\Theta_1, \tilde{\Theta}_{2,T})$.

$$V(\Theta_1, \tilde{\Theta}_{2,T}) = \mathbf{s}_T^H \mathbf{W}_T \mathbf{s}_T - \mathbf{s}_T^H \mathbf{W}_T \Psi(\Theta_1) (\Psi^H(\Theta_1) \mathbf{W}_T \Psi(\Theta_1))^{-1} \Psi^H(\Theta_1) \mathbf{W}_T \mathbf{s}_T$$

donc

$$\begin{aligned} \tilde{\Theta}_{1,T} &= \arg \max_{\Theta_1} \mathbf{s}_T^H \mathbf{W}_T \Psi(\Theta_1) (\Psi^H(\Theta_1) \mathbf{W}_T \Psi(\Theta_1))^{-1} \Psi^H(\Theta_1) \mathbf{W}_T \mathbf{s}_T \\ &= \arg \max_{\Theta_1} \mathbf{g}^H(\Theta_1, \mathbf{R}_T) \mathbf{G}^{-1}(\Theta_1, \mathbf{R}_T) \mathbf{g}(\Theta_1, \mathbf{R}_T) \end{aligned} \quad (2.10)$$

avec

$$\mathbf{g}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^H(\Theta_1) \mathbf{W}_T \mathbf{s}_T \quad \text{et} \quad \mathbf{G}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^H(\Theta_1) \mathbf{W}_T \Psi(\Theta_1).$$

En conclusion, par rapport à l'estimateur du maximum de vraisemblance, l'algorithme COMET simplifie notablement le problème de minimisation et, dans le cas d'observations complexes gaussiennes circulaires indépendantes, cet algorithme est asymptotiquement efficace.

Voyons maintenant quelles sont les adaptations à faire dans le cas d'observations réelles.

6.2.3 Formulation de l'algorithme COMET dans le cas réel

Supposons maintenant les observations $(\mathbf{y}_t)_{t=1,\dots,T}$ *réelles gaussiennes centrées et indépendantes*. L'estimateur au sens du maximum de vraisemblance de Θ est donné comme précédemment par

$$\hat{\Theta}_T = \arg \min_{\Theta \in D_\Theta} L_\Theta(\Theta) \quad \text{où} \quad L_\Theta(\Theta) = \ln(\text{Det}(\mathbf{R}_y(\Theta))) + \text{Tr}(\mathbf{R}_y^{-1}(\Theta) \mathbf{R}_T) \quad (2.11)$$

avec ici $\mathbf{R}_T \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^T$. Nous considérons à nouveau un "reparamétrage" du problème par les termes $\mathbf{\Gamma}$ de la matrice réelle symétrique $\mathbf{R}_y(\Theta)$. Ainsi

$$\mathbf{\Gamma} = \text{Vec}(\mathbf{R}_y(\Theta))$$

et l'application $f : D_\Theta \rightarrow D_\Gamma$ t.q. $\mathbf{\Gamma} = f(\Theta)$ reste bijective.

L'estimateur au sens du maximum de vraisemblance de ce paramètre est comme précédemment

$$\hat{\mathbf{\Gamma}}_T = \arg \min_{\mathbf{\Gamma} \in D_\Gamma} L_\Gamma(\mathbf{\Gamma}) \quad \text{avec} \quad L_\Gamma(\mathbf{\Gamma}) = L_\Theta(f^{-1}(\mathbf{\Gamma})).$$

Le domaine de contrainte D_Γ est ensuite élargi à un domaine $\tilde{D}_\Gamma$ en considérant un modèle $\mathbf{R}_y(\mathbf{\Gamma})$ de $\mathbf{R}_y$ simplement réel symétrique et l'estimateur au sens du maximum de vraisemblance de $\mathbf{\Gamma}$ sur ce domaine est $\tilde{\mathbf{\Gamma}}_T = \text{Vec}(\mathbf{R}_T)$ de sorte que, comme dans le cas d'observations complexes :

$$\lim_{T \rightarrow \infty} \tilde{\mathbf{\Gamma}}_T = \lim_{T \rightarrow \infty} \hat{\mathbf{\Gamma}}_T = \text{Vec}(\mathbf{R}_y).$$

L'équivalent de l'expression (2.7) est

$$\frac{\partial^2 L_\Gamma(\mathbf{\Gamma})}{\partial \mathbf{\Gamma} \partial \mathbf{\Gamma}^T} \Big|_{\tilde{\mathbf{\Gamma}}_T} = \mathbf{R}_T^{-1} \otimes_r \mathbf{R}_T^{-1}. \quad (2.12)$$

En posant $\mathbf{s}_T \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_T)$, d'après EXIP, l'estimateur

$$\tilde{\Theta}_T = (\tilde{\Theta}_{1,T}^T, \tilde{\Theta}_{2,T}^T)^T = \arg \min_{\Theta \in D_\Theta} ((\mathbf{s}_T - \Psi(\Theta_1) \Theta_2)^T (\mathbf{R}_T^{-1} \otimes_r \mathbf{R}_T^{-1}) (\mathbf{s}_T - \Psi(\Theta_1) \Theta_2)) \quad (2.13)$$

est asymptotiquement équivalent à l'estimateur au sens du maximum de vraisemblance $\hat{\Theta}_T$.

Nous reconnaissons à nouveau l'estimateur COMET de matrice de pondération $\mathbf{W}_T = \mathbf{R}_T^{-1} \otimes_r \mathbf{R}_T^{-1}$ et, dans cette situation, puisque l'estimateur du maximum de vraisemblance est asymptotiquement efficace, cet estimateur est asymptotiquement efficace.

La minimisation de (2.13) s'effectue de façon analogue à celle de (2.8) et conduit à

$$\tilde{\Theta}_{2,T} = [\Psi^T(\Theta_1) \mathbf{W}_T \Psi(\Theta_1)]^{-1} \Psi^T(\Theta_1) \mathbf{W}_T \mathbf{s}_T \quad (2.14)$$

et

$$\tilde{\Theta}_{1,T} = \arg \max_{\Theta_1} \mathbf{g}^T(\Theta_1, \mathbf{R}_T) \mathbf{G}^{-1}(\Theta_1, \mathbf{R}_T) \mathbf{g}(\Theta_1, \mathbf{R}_T) \quad (2.15)$$

avec

$$\mathbf{g}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^T(\Theta_1) \mathbf{W}_T \mathbf{s}_T \quad \text{et} \quad \mathbf{G}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^T(\Theta_1) \mathbf{W}_T \Psi(\Theta_1)$$

6.2.4 Performances asymptotiques de COMET

Dans le cas particulier où les observations $(\mathbf{y}_t)_{t=1,\dots,T}$ sont complexes circulaires gaussiennes et indépendantes, l'estimateur COMET (2.10) $\Theta_{1,T}$ est asymptotiquement gaussien et efficace puisqu'il est asymptotiquement équivalent à l'estimateur du Maximum de Vraisemblance qui est lui même asymptotiquement gaussien et efficace. Par suite :

$$\sqrt{T} (\Theta_{1,T} - \Theta_1) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}; \mathbf{C}_{\Theta_1}) \quad (2.16)$$

avec

$$\frac{1}{T} \mathbf{C}_{\Theta_1} = BCR_{\Theta_1}$$

où BCR_{Θ_1} est la borne de Cramer-Rao de Θ_1 dont l'expression est donnée dans [1] mais que l'on peut obtenir plus simplement de la façon suivante.

Le terme général de la matrice d'information de Fisher du paramètre Θ est (cf. [7, p.565]) :

$$[\mathbf{I}(\Theta)]_{k,l} = T \operatorname{Tr} \left(\frac{\partial \mathbf{R}_y}{\partial \theta_k} \mathbf{R}_y^{-1} \frac{\partial \mathbf{R}_y}{\partial \theta_l} \mathbf{R}_y^{-1} \right) \quad (2.17)$$

Grâce à (0.7) et (0.2), ce terme peut également s'écrire

$$\begin{aligned} [\mathbf{I}(\Theta)]_{k,l} &= T \operatorname{Vec}^H \left(\frac{\partial \mathbf{R}_y}{\partial \theta_k} \right) \operatorname{Vec}(\mathbf{R}_y^{-1} \frac{\partial \mathbf{R}_y}{\partial \theta_l} \mathbf{R}_y^{-1}) \\ &= T \operatorname{Vec}^H \left(\frac{\partial \mathbf{R}_y}{\partial \theta_k} \right) (\mathbf{R}_y^{-1} \otimes_c \mathbf{R}_y^{-1}) \operatorname{Vec} \left(\frac{\partial \mathbf{R}_y}{\partial \theta_l} \right) \end{aligned}$$

Donc

$$\mathbf{I}(\Theta) = T \left(\frac{\partial \operatorname{Vec}(\mathbf{R}_y)}{\partial \Theta} \right)^H \mathbf{W} \frac{\partial \operatorname{Vec}(\mathbf{R}_y)}{\partial \Theta} \quad \text{avec} \quad \mathbf{W} \stackrel{\text{def}}{=} \mathbf{R}_y^{-1} \otimes_c \mathbf{R}_y^{-1}.$$

D'autre part

$$\frac{\partial \operatorname{Vec}(\mathbf{R}_y)}{\partial \Theta} = [\mathbf{S}_1, \Psi(\Theta_1)] \quad \text{avec} \quad \mathbf{S}_1 = \frac{\partial \operatorname{Vec}(\mathbf{R}_y)}{\partial \Theta_1}$$

Donc

$$\mathbf{I}(\Theta) = T \begin{bmatrix} \mathbf{S}_1^H \mathbf{W} \mathbf{S}_1 & \mathbf{S}_1^H \mathbf{W} \Psi(\Theta_1) \\ \Psi^H(\Theta_1) \mathbf{W} \mathbf{S}_1 & \Psi^H(\Theta_1) \mathbf{W} \Psi(\Theta_1) \end{bmatrix}$$

et

$$BCR_{\Theta_1} = (\mathbf{I}^{-1}(\Theta))_{[1,1]}$$

en désignant par $(\mathbf{I}^{-1}(\Theta))_{[1,1]}$ le bloc de la matrice $\mathbf{I}^{-1}(\Theta)$ relatif au paramètre Θ_1 . Grâce au lemme d'inversion matricielle partitionnée :

$$\begin{aligned} (\mathbf{I}^{-1}(\Theta))_{[1,1]} &= \frac{1}{T} \left(\mathbf{S}_1^H \mathbf{W} \mathbf{S}_1 - \mathbf{S}_1^H \mathbf{W} \Psi(\Theta_1) (\Psi^H(\Theta_1) \mathbf{W} \Psi(\Theta_1))^{-1} \Psi^H(\Theta_1) \mathbf{W} \mathbf{S}_1 \right)^{-1} \\ &= \frac{1}{T} \left(\mathbf{S}_1^H \mathbf{W}^{1/2} \left(\mathbf{I} - \mathbf{W}^{1/2} \Psi(\Theta_1) (\Psi^H(\Theta_1) \mathbf{W} \Psi(\Theta_1))^{-1} \Psi^H(\Theta_1) \mathbf{W}^{1/2} \right) \mathbf{W}^{1/2} \mathbf{S}_1 \right)^{-1} \\ &= \frac{1}{T} \left(\mathbf{S}_1^H \mathbf{W}^{1/2} \left(\mathbf{I} - \Pi_{\mathbf{W}^{1/2} \Psi(\Theta_1)} \right) \mathbf{W}^{1/2} \mathbf{S}_1 \right)^{-1} \\ &= \frac{1}{T} \left(\mathbf{S}_1^H \mathbf{W}^{1/2} \Pi_{\mathbf{W}^{1/2} \Psi(\Theta_1)}^\perp \mathbf{W}^{1/2} \mathbf{S}_1 \right)^{-1} \end{aligned} \quad (2.18)$$

avec $\Pi_{\mathbf{W}^{1/2} \Psi(\Theta_1)}^\perp \stackrel{\text{def}}{=} \mathbf{I} - \Pi_{\mathbf{W}^{1/2} \Psi(\Theta_1)}$ où $\Pi_{\mathbf{W}^{1/2} \Psi(\Theta_1)}$ est la matrice de projection orthogonale sur les colonnes de $\mathbf{W}^{1/2} \Psi(\Theta_1)$.

On en déduit donc que

$$\mathbf{C}_{\Theta_1} = \left(\left(\frac{\partial \text{Vec}(\mathbf{R}_y)}{\partial \Theta_1} \right)^H \mathbf{W}^{1/2} \Pi_{\mathbf{W}^{1/2} \Psi(\Theta_1)}^\perp \mathbf{W}^{1/2} \frac{\partial \text{Vec}(\mathbf{R}_y)}{\partial \Theta_1} \right)^{-1} \quad (2.19)$$

avec $\mathbf{W} \stackrel{\text{def}}{=} \mathbf{R}_y^{-1} \otimes_c \mathbf{R}_y^{-1}$.

Remarque : La même approche s'étend au cas d'observations réelles gaussiennes indépendantes en remplaçant (2.17) par (cf. [7, p.76]) :

$$[\mathbf{I}(\Theta)]_{k,l} = \frac{1}{2} T \text{Tr} \left(\frac{\partial \mathbf{R}_y}{\partial \theta_k} \mathbf{R}_y^{-1} \frac{\partial \mathbf{R}_y}{\partial \theta_l} \mathbf{R}_y^{-1} \right)$$

et l'expression de la covariance asymptotique devient :

$$\mathbf{C}_{\Theta_1} = \frac{1}{2} \left(\left(\frac{\partial \text{Vec}(\mathbf{R}_y)}{\partial \Theta_1} \right)^T \mathbf{W}^{1/2} \Pi_{\mathbf{W}^{1/2} \Psi(\Theta_1)}^\perp \mathbf{W}^{1/2} \frac{\partial \text{Vec}(\mathbf{R}_y)}{\partial \Theta_1} \right)^{-1}$$

avec $\mathbf{W} \stackrel{\text{def}}{=} \mathbf{R}_y^{-1} \otimes_r \mathbf{R}_y^{-1}$.

6.3 Reformulation de l'algorithme COMET original

L'estimateur COMET peut être considéré comme un cas particulier, dans le cas d'observations gaussiennes indépendantes, d'un résultat classique concernant les estimateurs asymptotiquement de variance minimale (AMV) opérant sur une statistique donnée (e.g. [3, pp.81-84]) et étendu au cas d'observations complexes circulaires. Avant de justifier ceci et afin de préciser nos notations, rappelons que si g est une application de $\mathbb{C}^M$ dans $\mathbb{R}^K$, continûment différentiable par rapport à $\mathbf{x} = \Re(\mathbf{z})$ et $\mathbf{y} = \Im(\mathbf{z})$, alors

$$g(\mathbf{z} + \delta \mathbf{z}) = f(\mathbf{x} + \delta \mathbf{x}, \mathbf{y} + \delta \mathbf{y}) = f(\mathbf{x}, \mathbf{y}) + \frac{\partial f}{\partial \mathbf{x}} \delta \mathbf{x} + \frac{\partial f}{\partial \mathbf{y}} \delta \mathbf{y} + o((\delta \mathbf{x}, \delta \mathbf{y}))$$

ce que nous noterons

$$g(\mathbf{z} + \delta \mathbf{z}) = g(\mathbf{z}) + [\mathbf{D}(\mathbf{z}), \mathbf{D}^*(\mathbf{z})] \begin{bmatrix} \delta \mathbf{z} \\ \delta \mathbf{z}^* \end{bmatrix} + o(\delta \mathbf{z}) \quad (3.1)$$

avec

$$[\mathbf{D}(\mathbf{z})]_{k,m} = \frac{1}{2} \left(\frac{\partial f_k}{\partial \mathbf{x}_m} - i \frac{\partial f_k}{\partial \mathbf{y}_m} \right) \stackrel{\text{def}}{=} \frac{\partial g_k}{\partial \mathbf{z}_m}$$

où f_k et g_k représentent respectivement les k -ième composantes des applications f et g .

Avec cette notation

$$[\mathbf{D}^*(\mathbf{z})]_{k,m} = \frac{1}{2} \left(\frac{\partial f_k}{\partial \mathbf{x}_m} + i \frac{\partial f_k}{\partial \mathbf{y}_m} \right) \stackrel{\text{def}}{=} \frac{\partial g_k}{\partial \mathbf{z}_m^*}.$$

Considérons maintenant une suite d'observations $\{\mathbf{y}_t\}_{t \in \mathbb{N}}$ constituée de vecteurs aléatoires complexes (resp. réels) de dimension fixée P dont la distribution de probabilité est régie par un paramètre réel fixé inconnu $\Theta \in D_\Theta \subset \mathbb{R}^L$, supposé identifiable à partir de la matrice de covariance $\mathbf{R}_y(\Theta)$ de l'observation. En DOA par exemple $\mathbf{y}_t = [y_{t,1}, y_{t,2}, \dots, y_{t,P}]^T$ représentera le signal de sortie d'un réseau de P capteurs à l'instant t alors que dans le cas de l'observation temporelle d'un processus aléatoire scalaire y_t , le vecteur d'observation pourra être de la forme $\mathbf{y}_t = [y_t, y_{t-1}, \dots, y_{t-P+1}]^T$.

Nous noterons respectivement dans le cas d'observations complexes et réelles

$$\mathbf{R}_y(\Theta) \stackrel{\text{def}}{=} \mathbb{E}[\mathbf{y}_t \mathbf{y}_t^H] \quad (\text{resp. } \mathbf{R}_y(\Theta) \stackrel{\text{def}}{=} \mathbb{E}[\mathbf{y}_t \mathbf{y}_t^T])$$

la matrice de covariance de l'observation et $(\mathbf{R}_T)_T$: la suite des matrices de covariance empiriques où pour tout $T \in \mathbb{N}^*$,

$$\mathbf{R}_T \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^H \quad (\text{resp. } \mathbf{R}_T \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^T).$$

Précisons que dans le cas d'observations complexes, nous considérerons les cas où ces observations sont circulaires d'ordre 2 ($\mathbb{E}[\mathbf{y}_t \mathbf{y}_t^T] = \mathbf{O}$) ou non circulaires d'ordre 2 ($\mathbb{E}[\mathbf{y}_t \mathbf{y}_t^T] \neq \mathbf{O}$) mais dans ce dernier cas, sans tenir compte de cette seconde matrice de covariance. Dans ce dernier cas, nous continuerons, dans ce document, d'appeler "algorithme du second ordre", un algorithme qui est une fonction de $\mathbf{R}_T \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^H$ seul.

Notre propos est de nous intéresser aux estimateurs de Θ fondés sur une suite de statistiques $(\mathbf{s}_T)_T$ où pour tout $T \in \mathbb{N}^*$, $\mathbf{s}_T$ est un vecteur constitué de M termes de $\mathbf{R}_T$, cette suite convergeant en un certain sens vers une constante dépendant de Θ notée $\mathbf{s}(\Theta)$, constituée par les termes associés de $\mathbf{R}_y(\Theta)$. Nous considérerons de tels estimateurs comme des applications $\mathbf{s} \rightarrow g(\mathbf{s})$ de $\mathbb{C}^M$ (resp. $\mathbb{R}^M$) dans $\mathbb{R}^L$ telles que

$$g(\mathbf{s}(\Theta)) = \Theta \quad \text{pour tout } \Theta \in D_\Theta.$$

Par suite, pour tout algorithme $\mathbf{s}_T \rightarrow g(\mathbf{s}_T)$ de ce type, nous avons alors la propriété suivante

Résultat 6.1

- Dans le cas d'observations complexes circulaires, si $\sqrt{T} (\mathbf{s}_T - \mathbf{s}(\Theta)) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}; \mathbf{C}_s, \mathbf{C}_s')$ où $\mathbf{C}_s \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} T \mathbb{E}[(\mathbf{s}_T - \mathbf{s}(\Theta))(\mathbf{s}_T - \mathbf{s}(\Theta))^H]$ et $\mathbf{C}_s' \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} T \mathbb{E}[(\mathbf{s}_T - \mathbf{s}(\Theta))(\mathbf{s}_T - \mathbf{s}(\Theta))^T]$ et si $g(\mathbf{s})$ est continûment différentiable au sens précédent par rapport à $\Re(\mathbf{s})$ et $\Im(\mathbf{s})$ alors

$$\sqrt{T} (g(\mathbf{s}_T) - g(\mathbf{s}(\Theta))) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_\Theta)$$

avec

$$\mathbf{C}_\Theta = [\mathbf{D}_s, \mathbf{D}_s^*] \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_s' \\ \mathbf{C}_s'^* & \mathbf{C}_s^* \end{bmatrix} [\mathbf{D}_s, \mathbf{D}_s^*]^H$$

où $\mathbf{D}_s$ est la matrice des dérivées partielles de $g(\mathbf{s})$ au sens précédent.

- Dans le cas d'observations réelles, si $\sqrt{T}(\mathbf{s}_T - \mathbf{s}(\Theta)) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_s)$
où $\mathbf{C}_s \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} T \mathbb{E} [(\mathbf{s}_T - \mathbf{s}(\Theta))(\mathbf{s}_T - \mathbf{s}(\Theta))^T]$ et si $g(\mathbf{s})$ est continûment différentiable alors

$$\sqrt{T} (g(\mathbf{s}_T) - g(\mathbf{s}(\Theta))) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_\Theta)$$

avec

$$\mathbf{C}_\Theta = \mathbf{D}_s \mathbf{C}_s \mathbf{D}_s^T$$

où $\mathbf{D}_s$ est la matrice Jacobienne de $g(\mathbf{s})$

Preuve : Ce résultat découle d'un théorème de Continuité (e.g. [5, théorème p. 122]) adapté au besoin au cas de statistiques complexes. ■

Notre premier objectif sera de mettre en évidence l'existence d'une borne inférieure pour $\mathbf{C}_\Theta$. Nous chercherons ensuite s'il existe des estimateurs dont la matrice de covariance asymptotique est égale à cette borne. Pour déterminer la borne inférieure de $\mathbf{C}_\Theta$ nous aurons besoin du lemme suivant, démontré par Porat dans le cas de matrices réelles [3, Lemme 3.1]) et dont l'adaptation au cas de matrices complexes est immédiate en remplaçant les opérations de transposition par des transpositions hermitiennes.

Lemme 6.1 Si $\mathbf{D}$ est une matrice $L \times M$ de rang plein avec $M \geq L$, $\mathbf{S}$ une inverse à droite de $\mathbf{D}$ (i.e. t.q. $\mathbf{D}\mathbf{S} = \mathbf{I}_L$), et Σ une matrice $M \times M$ symétrique réelle définie positive, alors

$$\mathbf{D}\Sigma\mathbf{D}^T \geq (\mathbf{S}^T \Sigma^{-1} \mathbf{S})^{-1}$$

et dans le cas complexe avec Σ hermitienne définie positive,

$$\mathbf{D}\Sigma\mathbf{D}^H \geq (\mathbf{S}^H \Sigma^{-1} \mathbf{S})^{-1}.$$

Ce lemme sera en particulier appliqué à la matrice $\Sigma = \mathbf{C}_s$ ce qui nécessitera donc que cette matrice soit régulière.

Le choix des termes de $\mathbf{R}_T$ constituant la statistique $\mathbf{s}_T$ est donc assujetti à 3 conditions

1. le paramètre Θ doit être identifiable à partir de $\mathbf{s}_T$,
2. la matrice de covariance asymptotique $\mathbf{C}_s$ de $\mathbf{s}_T$ définie dans le résultat 6.1 doit être régulière. Comme le rang de $\mathbf{C}_s$ est égal au rang des variables aléatoires scalaires composant le vecteur $\mathbf{s}_T$, $\mathbf{s}_T$ doit être constituée par des termes de $\mathbf{R}_T$ qui soient des variables aléatoires linéairement indépendantes.
3. l'algorithme doit être continûment différentiable.

Ces conditions couplées à la nature réelle ou complexe des observations nous conduisent alors à distinguer différents cas :

- Dans le cas d'observations complexes, la matrice de covariance asymptotique $\mathbf{C}_R$ de $\mathbf{R}_T$ peut être ou non régulière

- si $\mathbf{C}_R$ est régulière, ce qui sera par exemple le cas en DOA, la statistique sera constituée de l'ensemble des termes de $\mathbf{R}_T$ de sorte que $\mathbf{s}_T = \text{Vec}(\mathbf{R}_T)$
- si $\mathbf{C}_R$ n'est pas régulière, par exemple si $\mathbf{y}_t = [y_t, y_{t-1}, \dots, y_{t-P+1}]^T$ où y_t est un processus stationnaire, $\mathbf{s}_T$ sera constitué des termes de la 1ère colonne et de la 1ère ligne de $\mathbf{R}_T$ pour que l'algorithme soit différentiable.
- Dans le cas d'observations réelles, la matrice de covariance asymptotique $\mathbf{C}_R$ de $\mathbf{R}_T$ n'est pas régulière
 - $\mathbf{s}_T$ sera alors constitué des termes d'un bloc triangulaire (inférieur par exemple) de $\mathbf{R}_T$ si la matrice de covariance asymptotique correspondante est régulière comme en DOA par exemple
 - sinon $\mathbf{s}_T$ sera constitué par la 1ère colonne de $\mathbf{R}_T$ comme dans le 2ème sous-cas précédent.

Etudions tout d'abord le cas où les observations sont indépendantes en distinguant le cas complexe du cas réel.

6.3.1 Cas d'observations gaussiennes réelles indépendantes

Supposons tout d'abord uniquement les observations *réelles centrées et indépendantes*, nous ne ferons intervenir l'hypothèse gaussienne qu'en tout dernier lieu. La matrice de covariance empirique $\mathbf{R}_T$ est symétrique réelle et, en posant $\mathbf{r}_T = \text{Vec}(\mathbf{R}_T)$, si

$$\sqrt{T} (\mathbf{r}_T - \mathbf{r}(\Theta)) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_r)$$

la matrice de covariance asymptotique $\mathbf{C}_r$ est singulière, de rang $\frac{P(P+1)}{2}$. Ceci nous conduit à n'utiliser que les termes d'un bloc triangulaire de $\mathbf{R}_T$ pour définir la statistique $\mathbf{s}_T$ à utiliser. Par exemple $\mathbf{s}_T$ sera constituée des termes du bloc triangulaire inférieur

$$\mathbf{s}_T = \text{Vet}(\mathbf{R}_T).$$

En désignant par $\mathbf{K}_P^d$ la “matrice de duplication” $P^2 \times \frac{P(P+1)}{2}$ telle que

$$\mathbf{r}_T = \mathbf{K}_P^d \mathbf{s}_T,$$

si

$$\sqrt{T} (\mathbf{r}_T - \mathbf{r}(\Theta)) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_r),$$

alors

$$\sqrt{T} (\mathbf{s}_T - \mathbf{s}(\Theta)) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_s) \quad \text{avec} \quad \mathbf{C}_s = \mathbf{K}_P^{d \#} \mathbf{C}_r \mathbf{K}_P^{d \# T}$$

où $\mathbf{K}_P^{d \#} \stackrel{\text{def}}{=} (\mathbf{K}_P^{d T} \mathbf{K}_P^d)^{-1} \mathbf{K}_P^{d T}$ représente la pseudo-inverse de Moore-Penrose de $\mathbf{K}_P^d$.

Nous allons maintenant montrer que si la matrice $\mathbf{C}_s$ est inversible, la matrice de covariance asymptotique de tout estimateur de Θ , continûment différentiable, est bornée inférieurement.

Soit $g(\mathbf{s})$ un estimateur de Θ considéré comme une application de $\mathbb{R}^M$ dans $\mathbb{R}^L$ supposée continûment différentiable

$$g(\mathbf{s}(\Theta) + \delta \mathbf{s}) = g(\mathbf{s}(\Theta)) + \mathbf{D}_s \delta \mathbf{s} + o(\delta \mathbf{s})$$

Puisque $g(\mathbf{s}(\Theta)) = \Theta$ pour tout Θ , nous pouvons écrire :

$$\begin{aligned}\Theta + \delta\Theta &= g(\mathbf{s}(\Theta + \delta\Theta)) = g\left(\mathbf{s}(\Theta) + \frac{\partial \mathbf{s}}{\partial \Theta} \delta\Theta + o(\delta\Theta)\right) \\ &= g(\mathbf{s}(\Theta)) + \mathbf{D}_s \frac{\partial \mathbf{s}}{\partial \Theta} \delta\Theta + o(\delta\Theta) \\ &= \Theta + \mathbf{D}_s \frac{\partial \mathbf{s}}{\partial \Theta} \delta\Theta + o(\delta\Theta)\end{aligned}\tag{3.2}$$

avec $\mathbf{D}_s \stackrel{\text{def}}{=} \frac{\partial g}{\partial \mathbf{s}}$.

Ceci montre d'une part que $\mathbf{S} \stackrel{\text{def}}{=} \frac{\partial \mathbf{s}}{\partial \Theta}$ est un inverse à droite de $\mathbf{D}_s$:

$$\mathbf{D}_s \frac{\partial \mathbf{s}}{\partial \Theta} = \mathbf{I}_L\tag{3.3}$$

et d'autre part que

$$\sqrt{T} (g(\mathbf{s}_T) - g(\mathbf{s}(\Theta))) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_\Theta) \quad \text{avec} \quad \mathbf{C}_\Theta = \mathbf{D}_s \mathbf{C}_s \mathbf{D}_s^T$$

grâce au théorème de Continuité (e.g. [5, théorème p. 122]) .

D'après le lemme 6.1, si la matrice $\mathbf{C}_s$ est inversible, alors la matrice de covariance asymptotique de $g(\mathbf{s}_T)$ est bornée inférieurement par $(\mathbf{S}^T \mathbf{C}_s^{-1} \mathbf{S})^{-1}$:

$$\mathbf{C}_\Theta = \mathbf{D}_s \mathbf{C}_s \mathbf{D}_s^T \geq (\mathbf{S}^T \mathbf{C}_s^{-1} \mathbf{S})^{-1}.$$

Une question primordiale est alors de savoir s'il existe des estimateurs dont la matrice de covariance asymptotique soit égale à cette borne, estimateurs qui par suite seraient asymptotiquement de variance minimale.

Nous démontrons en Annexe 2 le résultat suivant :

Résultat 6.2 *L'estimateur*

$$\hat{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\Theta))^T \mathbf{C}_s^{-1}(\Theta) (\mathbf{s}_T - \mathbf{s}(\Theta))\tag{3.4}$$

est asymptotiquement de variance minimale.

La minimisation (3.4) est lourde en charge de calcul car l'algorithme de minimisation doit calculer $\mathbf{C}_s^{-1}(\Theta)$ à chacune de ces itérations. Il est donc intéressant d'examiner si le remplacement de $\mathbf{C}_s^{-1}(\Theta)$ par $\mathbf{C}_{s,T}^{-1}$, un estimateur consistant de $\mathbf{C}_s^{-1}(\Theta)$, modifie la matrice de covariance asymptotique de l'estimé $\hat{\Theta}_T$ ainsi obtenu. Par une approche différente de Porat et Friedlander [6] nous démontrons en Annexe 2 le résultat suivant :

Résultat 6.3 *L'estimateur*

$$\tilde{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\Theta))^T \mathbf{C}_{s,T}^{-1} (\mathbf{s}_T - \mathbf{s}(\Theta))\tag{3.5}$$

où $\mathbf{C}_{s,T}$ est un estimateur consistant de $\mathbf{C}_s(\Theta)$, est asymptotiquement de variance minimale.

²Nous notons ponctuellement $\mathbf{C}_s(\Theta)$ la matrice de covariance asymptotiquement de $\mathbf{s}_T$ pour marquer le fait que cette matrice dépend de Θ .

Pour faire le lien avec l'estimateur COMET, supposons maintenant les observations *gaussiennes indépendantes*. Dans ce cas, la matrice de covariance asymptotique de $\mathbf{r}_T = \text{Vec}(\mathbf{R}_T)$ a pour expression

$$\mathbf{C}_r = (\mathbf{I}_{P^2} + \mathbf{K}_P^c) (\mathbf{R}(\Theta) \otimes_r \mathbf{R}(\Theta)) \quad ^3$$

Dans ce cas, puisque $\mathbf{s}_T = \mathbf{K}_P^d \mathbf{r}_T$ (cf. propriété (0.11)),

$$\mathbf{C}_s = \mathbf{K}_P^d \mathbf{K}_P^{\sharp} (\mathbf{I}_{P^2} + \mathbf{K}_P^c) (\mathbf{R}(\Theta) \otimes_r \mathbf{R}(\Theta)) \mathbf{K}_P^{\sharp T}$$

et, grâce à la propriété (0.12) de la matrice de duplication,

$$\mathbf{C}_s = 2 \mathbf{K}_P^d \mathbf{K}_P^{\sharp} (\mathbf{R}(\Theta) \otimes_r \mathbf{R}(\Theta)) \mathbf{K}_P^{\sharp T}$$

et

$$\mathbf{C}_s^{-1} = \frac{1}{2} \mathbf{K}_P^d \mathbf{K}_P^{\sharp T} (\mathbf{R}^{-1}(\Theta) \otimes_r \mathbf{R}^{-1}(\Theta)) \mathbf{K}_P^d.$$

Par suite

$$\begin{aligned} (\mathbf{s}_T - \mathbf{s}(\Theta))^T \mathbf{C}_s^{-1} (\mathbf{s}_T - \mathbf{s}(\Theta)) &= \frac{1}{2} (\mathbf{s}_T - \mathbf{s}(\Theta))^T \mathbf{K}_P^d \mathbf{K}_P^{\sharp T} (\mathbf{R}^{-1}(\Theta) \otimes_r \mathbf{R}^{-1}(\Theta)) \mathbf{K}_P^d (\mathbf{s}_T - \mathbf{s}(\Theta)) \\ &= \frac{1}{2} (\mathbf{r}_T - \mathbf{r}(\Theta))^T (\mathbf{R}^{-1}(\Theta) \otimes_r \mathbf{R}^{-1}(\Theta)) (\mathbf{r}_T - \mathbf{r}(\Theta)). \end{aligned} \quad (3.6)$$

En reportant (3.6) dans (3.4) l'expression de l'estimateur précédemment obtenu devient

$$\hat{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{r}_T - \mathbf{r}(\Theta))^T (\mathbf{R}^{-1}(\Theta) \otimes_r \mathbf{R}^{-1}(\Theta)) (\mathbf{r}_T - \mathbf{r}(\Theta)) \quad (3.7)$$

qui donne également, en remplaçant $\mathbf{R}(\Theta)$ par une estimée consistante $\mathbf{R}_T$,

$$\tilde{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{r}_T - \mathbf{r}(\Theta))^T (\mathbf{R}_T^{-1} \otimes_r \mathbf{R}_T^{-1}) (\mathbf{r}_T - \mathbf{r}(\Theta)) \quad (3.8)$$

qui n'est autre que l'expression (2.13).

En conclusion, on constate que sous l'hypothèse d'observations réelles gaussiennes indépendantes et centrées et à la condition que $\Theta \stackrel{\text{def}}{=} (\Theta_1^T, \Theta_2^T)^T$ où Θ_2 est un paramètre "linéaire", l'estimateur AMV n'est autre que l'estimateur COMET de matrice de pondération

$$\mathbf{W}_T = \mathbf{R}_T^{-1} \otimes_r \mathbf{R}_T^{-1}$$

qui donc, dans ce cas, est asymptotiquement de variance minimale.

6.3.2 Cas d'observations complexes gaussiennes circulaires indépendantes

Comme précédemment nous supposons tout d'abord les observations simplement *complexes circulaires, centrées et indépendantes* et nous ne ferons intervenir l'hypothèse gaussienne qu'en dernier lieu.

La matrice de covariance asymptotique de $\mathbf{R}_T$ étant cette fois régulière, nous utiliserons la statistique $\mathbf{s}_T = \text{Vec}(\mathbf{R}_T)$.

Corollaire 6.1 *Dans le cas où $\mathbf{s}_T$ utilise tous les termes de $\mathbf{R}_T$, la matrice de covariance asymptotique $\mathbf{C}_\Theta$ de tout estimateur Θ_T , fonction différentiable par rapport à $\Re(\mathbf{s}_T)$ et $\Im(\mathbf{s}_T)$, s'écrit :*

$$\mathbf{C}_\Theta = 4 \mathbf{D}_s \mathbf{C}_s \mathbf{D}_s^H$$

où $\mathbf{D}_s$ est définie comme dans le résultat 6.1.

³On retrouve bien le fait que $\mathbf{C}_r$ est singulière puisque $(\mathbf{I}_{P^2} + \mathbf{K}_P^c)(\mathbf{I}_{P^2} - \mathbf{K}_P^c) = \mathbf{O}$.

Preuve : Dans ce cas $\mathbf{s}_T = \text{Vec}(\mathbf{R}_T)$ et, compte tenu de la structure hermitienne de $\mathbf{R}_T$

$$\mathbf{s}_T^* = \text{Vec}^*(\mathbf{R}_T) = \text{Vec}(\mathbf{R}_T^*) = \text{Vec}(\mathbf{R}_T^T) = \mathbf{K}_P^c \mathbf{s}_T. \quad (3.9)$$

où $\mathbf{K}_P^c$ est la matrice de commutation d'ordre P (cf. formulaire).

De manière analogue nous avons

$$\mathbf{s}^*(\Theta) = \mathbf{K}_P^c \mathbf{s}(\Theta). \quad (3.10)$$

Grâce à ces propriétés nous pouvons écrire d'une part

$$\begin{aligned} \mathbf{C}'_s &\stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} T \mathbb{E} [(\mathbf{s}_T - \mathbf{s}(\Theta))(\mathbf{s}_T - \mathbf{s}(\Theta))^T] \\ &= \lim_{T \rightarrow \infty} T \mathbb{E} [(\mathbf{s}_T - \mathbf{s}(\Theta))(\mathbf{s}_T^* - \mathbf{s}^*(\Theta))^H] \\ &= \lim_{T \rightarrow \infty} T \mathbb{E} [(\mathbf{s}_T - \mathbf{s}(\Theta)) (\mathbf{K}_P^c (\mathbf{s}_T - \mathbf{s}(\Theta)))^H] \\ &= \mathbf{C}_s \mathbf{K}_P^c \end{aligned} \quad (3.11)$$

et d'autre part

$$\begin{aligned} \mathbf{C}_s^* &\stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} T \mathbb{E} [(\mathbf{s}_T - \mathbf{s}(\Theta))^*(\mathbf{s}_T - \mathbf{s}(\Theta))^T] \\ &= \lim_{T \rightarrow \infty} T \mathbb{E} [\mathbf{K}_P^c (\mathbf{s}_T - \mathbf{s}(\Theta)) (\mathbf{K}_P^c (\mathbf{s}_T - \mathbf{s}(\Theta)))^H] \\ &= \mathbf{K}_P^c \mathbf{C}_s \mathbf{K}_P^c \end{aligned} \quad (3.12)$$

Calculons maintenant $\mathbf{C}_\Theta$.

D'après le Résultat 6.1

$$\mathbf{C}_\Theta = \mathbf{D}_s \mathbf{C}_s \mathbf{D}_s^H + \mathbf{D}_s^* \mathbf{C}'_s \mathbf{D}_s^H + \mathbf{D}_s \mathbf{C}'_s \mathbf{D}_s^T + \mathbf{D}_s^* \mathbf{C}_s^* \mathbf{D}_s^T. \quad (3.13)$$

De part la propriété (3.10), pour tout $l = 1, \dots, L$, $\frac{\partial g_l}{\partial \mathbf{s}} = \frac{\partial g_l}{\partial \mathbf{s}^*} \mathbf{K}_P^c$ et par suite $\mathbf{D}_s^* = \mathbf{D}_s \mathbf{K}_P^c$ de même que $\mathbf{D}_s^T = \mathbf{K}_P^c \mathbf{D}_s^H$.

Grâce aux propriétés (3.11) et (3.12) et à la propriété (0.9) de la matrice de commutation on montre alors aisément que les différents termes du second membre de (3.13) sont égaux. ■

Comme précédemment nous allons maintenant montrer que si la matrice $\mathbf{C}_s$ est inversible, la matrice de covariance asymptotique de tout estimateur continûment différentiable au sens de (3.1) est bornée inférieurement.

Soit $g(\mathbf{s}_T)$ un estimateur de Θ considéré comme une application de $\mathbb{C}^M$ dans $\mathbb{R}^L$ supposée continûment différentiable au sens de (3.1). Nous pouvons écrire

$$g(\mathbf{s}(\Theta) + \delta \mathbf{s}) = g(\mathbf{s}(\Theta)) + [\mathbf{D}_s, \mathbf{D}_s^*] \begin{bmatrix} \delta \mathbf{s} \\ \delta \mathbf{s}^* \end{bmatrix} + o(\delta \mathbf{s})$$

avec $\mathbf{D}_s \stackrel{\text{def}}{=} \frac{\partial g}{\partial \mathbf{s}}$ au sens de (3.1). Or, nous avons vu plus haut que $\mathbf{D}_s^* = \mathbf{D}_s \mathbf{K}_P^c$. D'autre part, en posant $\mathbf{R}_T = \mathbf{R}(\Theta) + \delta \mathbf{R}_T$ où $\delta \mathbf{R}_T$ est une matrice hermitienne, nous avons

$$\delta \mathbf{s}_T^* = \text{Vec}^*(\delta \mathbf{R}_T) = \text{Vec}(\delta \mathbf{R}_T^*) = \text{Vec}(\delta \mathbf{R}_T^T) = \mathbf{K}_P^c \text{Vec}(\delta \mathbf{R}_T) = \mathbf{K}_P^c \delta \mathbf{s}_T.$$

Donc

$$g(\mathbf{s}(\Theta) + \delta \mathbf{s}) = g(\mathbf{s}(\Theta)) + 2\mathbf{D}_s \delta \mathbf{s} + o(\delta \mathbf{s})$$

Par suite, avec le même raisonnement que pour obtenir (3.2), on obtient

$$2\mathbf{D}_s\mathbf{S} = \mathbf{I}_L,$$

la matrice $2\mathbf{S}$ est un inverse à droite de $\mathbf{D}_s$. De part le lemme 6.1

$$\mathbf{D}_s\mathbf{C}_s\mathbf{D}_s^H \geq \frac{1}{4}(\mathbf{S}^H\mathbf{C}_s^{-1}\mathbf{S})^{-1}$$

et de part le corollaire 6.1 nous voyons que la matrice de covariance asymptotique de $g(\mathbf{s}_T)$ est bornée inférieurement par la matrice réelle symétrique $(\mathbf{S}^H\mathbf{C}_s^{-1}\mathbf{S})^{-1}$,

$$\mathbf{C}_\Theta = 4\mathbf{D}_s\mathbf{C}_s\mathbf{D}_s^H \geq (\mathbf{S}^H\mathbf{C}_s^{-1}\mathbf{S})^{-1}.$$

Pour se convaincre que la matrice $(\mathbf{S}^H\mathbf{C}_s^{-1}\mathbf{S})^{-1}$ est réelle symétrique il suffit de constater d'une part que $\mathbf{S}^H\mathbf{C}_s^{-1}\mathbf{S}$ est hermitienne et d'autre part que d'après (3.9) $\mathbf{S}^* = \mathbf{K}_P^c\mathbf{S}$ et $\mathbf{C}_s^T = \mathbf{K}_P^c\mathbf{C}_s\mathbf{K}_P^c$. Par suite $(\mathbf{S}^H\mathbf{C}_s^{-1}\mathbf{S})^T = \mathbf{S}^T\mathbf{C}_s^{-T}\mathbf{S}^* = \mathbf{S}^H\mathbf{K}_P^c\mathbf{C}_s^{-T}\mathbf{K}_P^c\mathbf{S} = \mathbf{S}^H\mathbf{C}_s^{-1}\mathbf{S}$.

Existe-t-il, comme dans le cas réel, des estimateurs asymptotiquement de variance minimale ?

Nous montrons en Annexe 3 les résultats suivants :

Résultat 6.4 *L'estimateur*

$$\hat{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\Theta))^H \mathbf{C}_s^{-1}(\Theta) (\mathbf{s}_T - \mathbf{s}(\Theta)) \quad (3.14)$$

est asymptotiquement de variance minimale.

Résultat 6.5 *L'estimateur*

$$\tilde{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\Theta))^H \mathbf{C}_{s,T}^{-1} (\mathbf{s}_T - \mathbf{s}(\Theta)) \quad (3.15)$$

où $\mathbf{C}_{s,T}$ est un estimateur consistant de $\mathbf{C}_s(\Theta)$, est asymptotiquement de variance minimale.

Pour faire le lien avec l'estimateur COMET, considérons le cas particulier d'observations *complexes gaussiennes circulaires indépendantes centrées*. Dans ce cas, la matrice de covariance asymptotique de $\mathbf{s}_T \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_T)$ est $\mathbf{C}_s(\Theta) = \mathbf{R}(\Theta) \otimes_c \mathbf{R}(\Theta)$ et $\mathbf{C}_{s,T} = \mathbf{R}_T \otimes_c \mathbf{R}_T$ en est une estimation consistante, par suite $\mathbf{C}_{s,T}^{-1} = \mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1}$. On constate alors que les expressions (2.8) et (3.15) sont identiques et que par suite l'estimateur COMET (2.8) est asymptotiquement de variance minimale.

Nous venons de voir que, dans le cas d'observations gaussiennes, réelles ou complexes circulaires, *indépendantes* où le paramètre inconnu Θ comporte une partie "linéaire", l'approche AMV conduit aux mêmes résultats que l'approche COMET. Ceci est possible grâce notamment à l'hypothèse d'indépendance des observations qui permet la régularité de la matrice de covariance asymptotique de la statistique dont nous avons eu besoin pour AMV : $\text{Vec}(\mathbf{R}_T)$ dans le cas complexe ou $\text{Vet}(\mathbf{R}_T)$ dans le cas réel. On peut maintenant se demander à quoi conduit l'approche AMV lorsque les observations ne sont pas indépendantes. C'est ce qui constitue l'objet du paragraphe suivant.

6.4 Algorithme AMV dans le cas d'observations non indépendantes

Rappelons que la mise en œuvre de l'approche AMV nécessite que les 2 conditions suivantes soient remplies

1. l'algorithme doit être continûment différentiable
2. la matrice de covariance asymptotique $\mathbf{C}_s$ de la statistique doit être régulière.

Dans le cas d'observations réelles, la 1ère condition ne posera pas de problème particulier et le choix de la statistique $\mathbf{s}_T$ sera uniquement influencé par la 2ème condition.

Par contre, dans le cas d'observations complexes, le problème de la différentiabilité de l'algorithme se pose et le choix de $\mathbf{s}_T$ sera influencé par les deux conditions simultanément.

Ceci nous conduit à nouveau à distinguer le cas réel du cas complexe.

6.4.1 Cas d'observations réelles

Le cas d'observations *réelles indépendantes et centrées* a été traité en paragraphe 6.3.1 et conduit, comme on l'a vu, à choisir comme statistique $\mathbf{s}_T = \text{Vet}(\mathbf{R}_T)$ et aux estimateurs AMV (3.4) et (3.5).

Considérons maintenant un exemple où les observations ne sont plus indépendantes. Si l'observation $\mathbf{y}_t = [y_t, y_{t-1}, \dots, y_{t-P+1}]^T$ résulte de l'échantillonnage régulier d'un processus réel y_t , stationnaire au sens strict, la matrice de covariance $\mathbf{R}(\Theta)$ est réelle symétrique, de structure Toeplitz. Cette matrice peut être estimée par la matrice de covariance empirique classique $\mathbf{R}_T = \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^T$ ou par la matrice symétrique Toeplitz $\mathbf{R}_T^{to}$ déduite de sa première colonne $\mathbf{r}_T = \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t y_t$. Il est démontré dans [16] que $\mathbf{R}_T$ et $\mathbf{R}_T^{to}$ ont la même matrice de covariance asymptotique $\mathbf{C}_R(\Theta)$. De ce fait le rang de $\mathbf{C}_R(\Theta)$ est égal au rang de l'ensemble des termes de $\mathbf{R}_T^{to}$ considérés comme des variables aléatoires, c'est à dire égal à P . La matrice $\mathbf{C}_R(\Theta)$ est donc singulière, ce qui exclut le choix de $\text{Vec}(\mathbf{R}_T)$ comme statistique pouvant conduire à un estimateur AMV. Pour répondre à la deuxième condition précédente, nous choisirons donc comme statistique $\mathbf{r}_T$, la 1ère colonne de $\mathbf{R}_T$. En appliquant le même raisonnement que celui tenu au paragraphe 6.3.1 on montre alors que si

$$\sqrt{T} (\mathbf{r}_T - \mathbf{r}(\Theta)) \xrightarrow{L} \mathcal{N}(\mathbf{0}, \mathbf{C}_r)$$

où $\mathbf{C}_r$ est une matrice régulière et si $g(\mathbf{r}_T)$ un estimateur de Θ continûment différentiable alors la matrice de covariance asymptotique de $g(\mathbf{r}_T)$ est bornée inférieurement par $(\mathbf{S}^T \mathbf{C}_r^{-1} \mathbf{S})^{-1}$ avec $\mathbf{S} \stackrel{\text{def}}{=} \frac{\partial \mathbf{r}}{\partial \Theta}$. D'autre part, en appliquant le même raisonnement que celui tenu en Annexe 2, on obtient les résultats suivants :

Résultat 6.6 *L'estimateur*

$$\hat{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{r}_T - \mathbf{r}(\Theta))^T \mathbf{C}_r^{-1}(\Theta) (\mathbf{r}_T - \mathbf{r}(\Theta)) \quad (4.1)$$

est asymptotiquement de variance minimale.

et

Résultat 6.7 *L'estimateur*

$$\tilde{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{r}_T - \mathbf{r}(\Theta))^T \mathbf{C}_{r,T}^{-1} (\mathbf{r}_T - \mathbf{r}(\Theta)) \quad (4.2)$$

où $\mathbf{C}_{r,T}$ est un estimateur consistant de $\mathbf{C}_r(\Theta)$, est asymptotiquement de variance minimale.

A nouveau, si le paramètre réel inconnu $\Theta = (\Theta_1^T, \Theta_2^T)^T$ est composé d'un paramètre d'intérêt Θ_1 et d'un paramètre de nuisance Θ_2 duquel la statistique $\mathbf{r}$ dépend linéairement sous la forme

$$\mathbf{r}(\Theta) = \Psi(\Theta_1) \Theta_2,$$

l'estimateur de Θ_1 déduit de (4.2) peut être obtenu d'une façon analogue à celle décrite dans le paragraphe (6.2.2) ce qui conduit à :

$$\tilde{\Theta}_{1,T} = \arg \max_{\Theta_1} \mathbf{g}^T(\Theta_1, \mathbf{R}_T) \mathbf{G}^{-1}(\Theta_1, \mathbf{R}_T) \mathbf{g}(\Theta_1, \mathbf{R}_T) \quad (4.3)$$

avec

$$\mathbf{g}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^T(\Theta_1) \mathbf{W}_T \mathbf{r}_T \quad , \quad \mathbf{G}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^T(\Theta_1) \mathbf{W}_T \Psi(\Theta_1) \quad \text{et} \quad \mathbf{W}_T = \mathbf{C}_{r,T}^{-1}.$$

6.4.2 Cas d'observation complexes

Le problème est ici un peu plus délicat. Si les observations ne sont pas indépendantes, comme dans le cas d'observations réelles du paragraphe précédent, la matrice de covariance asymptotique de $\text{Vec}(\mathbf{R}_T)$ peut ne pas être régulière. Par exemple, pour $\mathbf{y}_t = [y_t, y_{t-1}, \dots, y_{t-P+1}]^T$ où y_t est un processus complexe circulaire centré stationnaire, il a été démontré dans [16], dans le cas de processus constitués de sinusôides complexes et d'un bruit additif MA, que les matrices de covariance empiriques $\mathbf{R}_T$ et $\mathbf{R}_T^{to}$ (matrice hermitienne Toeplitz construite à partir de la première colonne $\frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^*$ de $\mathbf{R}_T$) ont la même matrice de covariance asymptotique $\mathbf{C}_R(\Theta)$. Par suite, le rang de $\mathbf{C}_R(\Theta)$ qui est le rang de l'ensemble des variables aléatoires qui constituent les termes de $\mathbf{R}_T^{to}$ est de $2P - 1$ et en conséquence $\mathbf{C}_R(\Theta)$ est singulière. Ceci exclut dans ce cas le choix de $\text{Vec}(\mathbf{R}_T)$ comme statistique pour obtenir un estimateur AMV.

Si maintenant $\mathbf{s}_T$ est une statistique constituée de termes de $\mathbf{R}_T$, de matrice de covariance asymptotique $\mathbf{C}_s$ régulière et si $g(\mathbf{s}_T)$ est un estimateur supposé simplement différentiable dans $\mathbb{R}^2$ au sens où

$$g(\mathbf{s} + \delta\mathbf{s}) = g(\mathbf{s}) + [\mathbf{D}_s, \mathbf{D}_s^*] \begin{bmatrix} \delta\mathbf{s} \\ \delta\mathbf{s}^* \end{bmatrix} + o(\delta\mathbf{s}), \quad (4.4)$$

d'après le résultat 6.1

$$\mathbf{C}_\Theta = [\mathbf{D}_s, \mathbf{D}_s^*] \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_s' \\ \mathbf{C}_s'^* & \mathbf{C}_s^* \end{bmatrix} [\mathbf{D}_s, \mathbf{D}_s^*]^H.$$

Avec le même raisonnement que pour obtenir (3.2), on obtient

$$[\mathbf{D}_s, \mathbf{D}_s^*] \begin{bmatrix} \delta\mathbf{S} \\ \delta\mathbf{S}^* \end{bmatrix} = \mathbf{I}_L$$

et, sous réserve que la matrice $\begin{bmatrix} \mathbf{C}_s & \mathbf{C}_s' \\ \mathbf{C}_s'^* & \mathbf{C}_s^* \end{bmatrix}$ soit régulière, la minoration de $\mathbf{C}_\Theta$ fournie par le lemme 6.1 est

$$\mathbf{C}_\Theta \geq \left([\mathbf{S}^H, \mathbf{S}^T] \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_s' \\ \mathbf{C}_s'^* & \mathbf{C}_s^* \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{S} \\ \mathbf{S}^* \end{bmatrix} \right)^{-1}.$$

Cependant, il n'existe pas nécessairement d'algorithme permettant d'atteindre asymptotiquement cette borne.

Par contre, si la statistique $\mathbf{s}_T$ est choisie de telle sorte qu'il existe une bijection entre $\mathbf{s}_T$ et $\mathbf{R}_T$ et s'il existe une matrice de permutation $\mathbf{K}_p'$ telle que

$$\mathbf{s}_T^* = \mathbf{K}_p' \mathbf{s}_T,$$

alors

$$\delta\mathbf{s}^* = \mathbf{K}_p' \delta\mathbf{s}$$

et (4.4) devient

$$g(\mathbf{s} + \delta\mathbf{s}) = g(\mathbf{s}) + \mathbf{D}_g \delta\mathbf{s} + o(\delta\mathbf{s}).$$

L'estimateur est donc différentiable dans $\mathbb{C}$ avec $\mathbf{D}_g \stackrel{\text{def}}{=} \mathbf{D}_s + \mathbf{D}_s^* \mathbf{K}_p'$. Par suite, avec le même raisonnement que pour obtenir (3.2),

$$\mathbf{D}_g \mathbf{S} = \mathbf{I}_L$$

et

$$\sqrt{T} (g(\mathbf{s}_T) - g(\mathbf{s}(\Theta))) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_\Theta)$$

avec

$$\mathbf{C}_\Theta = \mathbf{D}_g \mathbf{C}_s \mathbf{D}_g^H.$$

D'après le lemme 6.1, la matrice de covariance asymptotique de $g(\mathbf{s}_T)$ est bornée inférieurement par la matrice réelle symétrique $(\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})^{-1}$,

$$\mathbf{C}_\Theta = \mathbf{D}_g \mathbf{C}_s \mathbf{D}_g^H \geq (\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})^{-1}. \quad (4.5)$$

La justification du caractère symétrique réel de la matrice $(\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})^{-1}$ est en tout point similaire à celle donnée au paragraphe 6.3.2, la matrice $\mathbf{K}_p'$ jouant ici le rôle de la matrice $\mathbf{K}_p^c$.

Par ailleurs, un raisonnement similaire à celui développé en Annexe 3 permettrait de montrer les résultats suivants :

Résultat 6.8 *L'estimateur*

$$\hat{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\Theta))^H \mathbf{C}_s^{-1}(\Theta) (\mathbf{s}_T - \mathbf{s}(\Theta)) \quad (4.6)$$

est asymptotiquement de variance minimale.

et

Résultat 6.9 *L'estimateur*

$$\tilde{\Theta}_T = \arg \min_{\Theta \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\Theta))^H \mathbf{C}_{s,T}^{-1} (\mathbf{s}_T - \mathbf{s}(\Theta)) \quad (4.7)$$

où $\mathbf{C}_{s,T}$ est un estimateur consistant de $\mathbf{C}_s$, est asymptotiquement de variance minimale.

Si maintenant le paramètre réel inconnu $\Theta = (\Theta_1^T, \Theta_2^T)^T \in \mathbb{R}^L$ est composé d'un paramètre d'intérêt $\Theta_1 \in \mathbb{R}^K$ et d'un paramètre de nuisance $\Theta_2 \in \mathbb{R}^{L-K}$ duquel la statistique $\mathbf{s}$ dépend linéairement sous la forme $\mathbf{s}(\Theta) = \Psi(\Theta_1)\Theta_2$, l'estimateur de Θ_1 déduit de (4.7) peut être obtenu d'une façon analogue à celle décrite dans le paragraphe 6.2.2 :

- tout d'abord on détermine l'estimé du paramètre linéaire Θ_2 sous la forme

$$\tilde{\Theta}_{2,T} = [\Psi^H(\Theta_1) \mathbf{W}_T \Psi(\Theta_1)]^{-1} \Psi^H(\Theta_1) \mathbf{W}_T \mathbf{s}_T \quad (4.8)$$

avec $\mathbf{W}_T = \mathbf{C}_{s,T}^{-1}$, où $\mathbf{C}_{s,T}$ est une estimation consistante de $\mathbf{C}_s(\Theta)$, en justifiant comme précédemment que $\tilde{\Theta}_{2,T}$ est bien réel si $\mathbf{C}_{s,T}$ est de même structure, du point de vue partie réelle/partie imaginaire, que $\mathbf{C}_s(\Theta)$. La démonstration de ce point est donnée en Annexe 4,

- on montre ensuite que l'estimé de Θ_1 est donné par

$$\Theta_{1,T} = \arg \max_{\Theta_1} \mathbf{g}^H(\Theta_1, \mathbf{R}_T) \mathbf{G}^{-1}(\Theta_1, \mathbf{R}_T) \mathbf{g}(\Theta_1, \mathbf{R}_T) \quad (4.9)$$

avec

$$\mathbf{g}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^H(\Theta_1) \mathbf{W}_T \mathbf{s}_T \quad \text{et} \quad \mathbf{G}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^H(\Theta_1) \mathbf{W}_T \Psi(\Theta_1).$$

En conclusion, nous avons donc étendu la notion d'estimateur AMV du second ordre, classique dans le cas d'observations réelles, au cas d'observations complexes où certaines précautions doivent être prises quant à la définition de la statistique du second ordre à prendre en compte.

Remarque : Puisque $\mathbf{D}_g \mathbf{S} = \mathbf{I}_L$ implique que

$$\mathbf{D}_{g_1} \mathbf{S}_1 = \mathbf{I}_K$$

avec $\mathbf{S}_1 \stackrel{\text{def}}{=} \frac{\partial \mathbf{s}}{\partial \Theta_1}$ et où $\mathbf{D}_{g_1}$ désigne la différentielle d'un algorithme $\mathbf{s}_T \rightarrow \Theta_{1,T} = g_1(\mathbf{s}_T)$. Par application du lemme 6.1

$$\mathbf{D}_{g_1} \mathbf{C}_s^{-1} \mathbf{D}_{g_1}^H \geq (\mathbf{S}_1^H \mathbf{C}_s^{-1} \mathbf{S}_1)^{-1}.$$

Par suite,

$$\mathbf{C}_{\Theta_1} \geq (\mathbf{S}_1^H \mathbf{C}_s^{-1} \mathbf{S}_1)^{-1}. \quad (4.10)$$

Ainsi, si l'on ne s'intéresse qu'au paramètre Θ_1 d'intérêt, puisque

$$\mathbf{S}_1^H \mathbf{C}_s^{-1} \mathbf{S}_1 = [\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S}]_{(1:K, 1:K)}$$

et que, d'après [17, théorème 7.7.8 p.474], pour toute matrice $\mathbf{A}$ symétrique réelle définie positive

$$[\mathbf{A}^{-1}]_{(1:K, 1:K)} \geq [\mathbf{A}_{(1:K, 1:K)}]^{-1},$$

nous avons l'inégalité suivante :

$$(\mathbf{S}_1^H \mathbf{C}_s^{-1} \mathbf{S}_1)^{-1} \leq (\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})_{(1:K, 1:K)}^{-1},$$

ce qui implique que ce nouveau minorant $(\mathbf{S}_1^H \mathbf{C}_s^{-1} \mathbf{S}_1)^{-1}$ est inférieur au minorant classique déduit de $(\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})^{-1}$. Ce minorant $(\mathbf{S}_1^H \mathbf{C}_s^{-1} \mathbf{S}_1)^{-1}$ constitue donc une borne inférieure pour la variance asymptotique du paramètre Θ_1 seul. Malheureusement, nous ne connaissons pas d'estimateur $g_1(\mathbf{s}_T)$ qui atteigne cette borne. Nous verrons numériquement sur un exemple en estimation de fréquence de sinusôides que ces 2 bornes sont, dans ce cas, en fait extrêmement voisines.

6.5 Performances des algorithmes AMV

Nous nous situons dans ce paragraphe dans les hypothèses d'application de l'algorithme COMET et supposons en particulier que le paramètre inconnu Θ est constitué d'une partie "linéaire" Θ_2 correspondant à des termes de nuisance et d'une partie "non linéaire" Θ_1 correspondant au paramètre d'intérêt. Les algorithmes AMV "optimaux" (au sens où la matrice de pondération utilisée est l'inverse $\mathbf{C}_s^{-1}$ de la matrice de covariance asymptotique de la statistique ou une estimée consistante $\mathbf{C}_{s,T}^{-1}$ de cette matrice) sont, par définition, asymptotiquement de variance minimum. Nous nous proposons de donner une expression analytique de la borne inférieure $(\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})_{(1:K, 1:K)}^{-1}$ de la covariance asymptotique du paramètre Θ_1 . Cela nous permettra en outre d'obtenir, dans le cas particulier d'observations gaussiennes complexes circulaires indépendantes, la borne de Cramer Rao du paramètre Θ_1 d'une façon différente de celle utilisée au paragraphe 6.2.4.

6.5.1 Cas optimum

Avec $\mathbf{S} = [\mathbf{S}_1, \Psi(\Theta_1)]$ où $\mathbf{S}_1 \stackrel{\text{def}}{=} \frac{\partial \mathbf{s}}{\partial \Theta_1}$ et $\mathbf{W} = \mathbf{C}_s^{-1}(\Theta)$

$$\mathbf{C}_\Theta = \begin{bmatrix} \mathbf{S}_1^H \mathbf{W} \mathbf{S}_1 & \mathbf{S}_1^H \mathbf{W} \Psi(\Theta_1) \\ \Psi^H(\Theta_1) \mathbf{W} \mathbf{S}_1 & \Psi^H(\Theta_1) \mathbf{W} \Psi(\Theta_1) \end{bmatrix}^{-1}.$$

Grâce au lemme d'inversion matricielle partitionnée on obtient comme au paragraphe 6.2.4

$$\mathbf{C}_{\Theta_1} = \left(\mathbf{S}_1^H \mathbf{W}^{1/2} \mathbf{\Pi}_{\mathbf{W}^{1/2} \Psi(\Theta_1)}^\perp \mathbf{W}^{1/2} \mathbf{S}_1 \right)^{-1} \quad (5.1)$$

avec $\mathbf{\Pi}_{\mathbf{W}^{1/2} \Psi(\Theta_1)}^\perp \stackrel{\text{def}}{=} \mathbf{I} - \mathbf{\Pi}_{\mathbf{W}^{1/2} \Psi(\Theta_1)}$ où $\mathbf{\Pi}_{\mathbf{W}^{1/2} \Psi(\Theta_1)}$ est la matrice de projection orthogonale sur les colonnes de $\mathbf{W}^{1/2} \Psi(\Theta_1)$.

On en déduit donc

$$\frac{1}{T} \mathbf{C}_{\Theta_1} = \frac{1}{T} \left(\mathbf{S}_1^H \mathbf{W}^{1/2} \mathbf{\Pi}_{\mathbf{W}^{1/2} \Psi(\Theta_1)}^\perp \mathbf{W}^{1/2} \mathbf{S}_1 \right)^{-1} \quad (5.2)$$

qui n'est autre que l'expression (2.18) de la borne de Cramer Rao obtenue au paragraphe 6.2.4, dans le cas d'observations complexes gaussiennes circulaires indépendantes, de façon différente.

Notons que cette expression analytique est compliquée car elle fait appel au calcul d'une matrice de projection. Nous allons maintenant donner une autre expression de $\mathbf{C}_{\Theta_1}$ qui nous permettra, au paragraphe suivant, de donner une expression de la covariance asymptotique d'algorithmes du second ordre d'estimation de Θ_1 sous optimaux.

Dans le cas d'observations complexes non gaussiennes et/ou non indépendantes (cf. paragraphe 6.4.2) l'approche AMV "optimale", couplée à l'hypothèse $\mathbf{s} = \Psi(\Theta_1)\Theta_2$, conduit d'après (4.9) à l'estimateur

$$\Theta_{1,T} = \arg \max_{\alpha} \mathbf{s}_T^H \mathbf{G}(\alpha, T) \mathbf{s}_T \quad (5.3)$$

avec ici

$$\mathbf{G}(\alpha, T) \stackrel{\text{def}}{=} \mathbf{W}_T^H \Psi(\alpha) (\Psi^H(\alpha) \mathbf{W}_T \Psi(\alpha))^{-1} \Psi^H(\alpha) \mathbf{W}_T$$

où $\mathbf{s}_T$ vérifie $\mathbf{s}_T^* = \mathbf{K}_p^l \mathbf{s}_T$ et avec comme matrice de pondération $\mathbf{W}_T = \mathbf{C}_{s,T}^{-1}$, une estimée consistante de $\mathbf{C}_s(\Theta)$.

Pour faciliter le calcul de la matrice de covariance asymptotique de l'estimateur donné par (5.3) nous allons utiliser la propriété suivante démontrée en annexe 2 :

Les estimateurs

$$\Theta_T = \arg \min_{\alpha \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\alpha))^H \mathbf{C}_{s,T}^{-1} (\mathbf{s}_T - \mathbf{s}(\alpha)) \quad (5.4)$$

ont même covariance asymptotique pour tout estimateur consistant $\mathbf{C}_{s,T}$ de $\mathbf{C}_s(\Theta)$. Le choix particulier $\mathbf{C}_{s,T} = \mathbf{C}_s(\Theta)$ fournit une suite de variables aléatoires $\hat{\Theta}_T$ ayant toujours cette propriété. Bien sûr, dans ce cas $\hat{\Theta}_T$ n'est plus un estimateur (car $\hat{\Theta}_T$ est une fonction de Θ), mais cet artifice nous permet d'affirmer que la suite de variables aléatoires :

$$\hat{\Theta}_T = \arg \min_{\alpha \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\alpha))^H \mathbf{C}_s^{-1}(\Theta) (\mathbf{s}_T - \mathbf{s}(\alpha)) \quad (5.5)$$

a comme covariance asymptotique, la covariance optimale $(\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})^{-1}$. Par suite, en appliquant la technique COMET à (5.5) (cela est possible car la matrice de pondération $\mathbf{C}_s^{-1}(\Theta)$ ne dépend pas de α), la partie $\hat{\Theta}_{1,T}$ de $\hat{\Theta}_T$ est donnée par :

$$\hat{\Theta}_{1,T} = \arg \max_{\alpha} \mathbf{s}_T^H \mathbf{G}(\alpha) \mathbf{s}_T \quad (5.6)$$

avec ici $\mathbf{G}(\alpha) \stackrel{\text{def}}{=} \mathbf{C}_s^{-1}(\Theta) \Psi(\alpha) (\Psi^H(\alpha) \mathbf{C}_s^{-1}(\Theta) \Psi(\alpha))^{-1} \Psi^H(\alpha) \mathbf{C}_s^{-1}(\Theta)$ qui ne dépend plus ici de l'observation $(\mathbf{y}_t)_{t=1, \dots, T}$. Une approche fonctionnelle de cet algorithme nous permet d'écrire

$$\mathbf{C}_{\Theta_1} = \mathbf{D}_{\Theta_1, \mathbf{s}}^{alg} \mathbf{C}_s(\Theta) \left(\mathbf{D}_{\Theta_1, \mathbf{s}}^{alg} \right)^H$$

où $\mathbf{D}_{\Theta_1, \mathbf{s}}^{alg}$ désigne la matrice associée à la différentielle de l'algorithme par rapport à Θ_1 au point $\mathbf{s}(\Theta)$. Un calcul de perturbation développé en Annexe 5 montre que si $\Theta_1 = (\theta_1, \dots, \theta_K)^T$

$$\mathbf{D}_{\Theta_1, \mathbf{s}}^{alg} = -\mathcal{F}_1^{-1} \mathcal{F}_2$$

avec

$$\mathcal{F}_1(k, l) \stackrel{\text{def}}{=} \mathbf{s}^H \frac{\partial^2 \mathbf{G}(\Theta)}{\partial \theta_k \partial \theta_l} \mathbf{s}$$

et

$$\mathcal{F}_2(k, :) \stackrel{\text{def}}{=} \mathbf{s}^H \frac{\partial \mathbf{G}(\Theta)}{\partial \theta_k} + \mathbf{s}^T \frac{\partial \mathbf{G}^T(\Theta)}{\partial \theta_k} \mathbf{K}'_p \quad k, l = 1, \dots, K.$$

Dans le cas d'observations réelles non gaussiennes et/ou non indépendantes (cf. section 6.4.1) l'approche AMV "optimale", couplée à l'hypothèse $\mathbf{s} = \Psi(\Theta_1) \Theta_2$, conduit de la même façon à l'estimateur

$$\Theta_{1,T} = \arg \max_{\alpha} \mathbf{r}_T^T \mathbf{G}(\alpha, T) \mathbf{r}_T \quad (5.7)$$

avec $\mathbf{G}(\alpha, T) \stackrel{\text{def}}{=} \mathbf{W}_T^T \Psi(\alpha) (\Psi^T(\alpha) \mathbf{W}_T \Psi(\alpha))^{-1} \Psi^T(\alpha) \mathbf{W}_T$, où $\mathbf{r}_T$ est la 1ère colonne de $\mathbf{R}_T$ et avec comme matrice de pondération $\mathbf{W}_T = \mathbf{C}_{r,T}^{-1}$. Avec le même raisonnement que précédemment

$$\mathbf{C}_{\Theta_1} = \mathbf{D}_{\Theta_1, \mathbf{r}}^{alg} \mathbf{C}_r(\Theta) \left(\mathbf{D}_{\Theta_1, \mathbf{r}}^{alg} \right)^T$$

où $\mathbf{D}_{\Theta_1, \mathbf{r}}^{alg}$ désigne la matrice associée à la différentielle de l'algorithme par rapport à Θ_1 au point $\mathbf{r}(\Theta)$. Un calcul de perturbation développé en Annexe 5 montre que si $\Theta_1 = (\theta_1, \dots, \theta_K)^T$

$$\mathbf{D}_{\Theta_1, \mathbf{r}}^{alg} = -\mathcal{F}_1^{-1} \mathcal{F}_2$$

avec

$$\mathcal{F}_1(k, l) \stackrel{\text{def}}{=} \mathbf{r}^T \frac{\partial^2 \mathbf{G}(\Theta)}{\partial \theta_k \partial \theta_l} \mathbf{r}$$

et

$$\mathcal{F}_2(k, :) \stackrel{\text{def}}{=} 2 \mathbf{r}^T \frac{\partial \mathbf{G}(\Theta)}{\partial \theta_k} \quad k, l = 1, \dots, K.$$

6.5.2 Cas sous optimum

La mise en œuvre pratique de l'algorithme AMV optimal et de l'algorithme de type COMET qui s'en déduit nécessite la connaissance d'un estimateur consistant $\mathbf{C}_{s,T}$ de $\mathbf{C}_s(\Theta)$. Or cet estimateur n'est pas toujours disponible, en particulier dans le cas d'observations non gaussiennes et/ou non indépendantes. On peut alors envisager d'utiliser une matrice de pondération $\mathbf{W}_T$ plus simple, par exemple celle correspondant au modèle gaussien indépendant (par exemple, pour $\mathbf{R}_T$, $\mathbf{C}_{r,T}^{-1} = \mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1}$). L'algorithme

moindre carré non linéaire associé n'est alors plus AMV.

Ainsi, dans cet exemple, l'algorithme

$$\Theta_T = \arg \min_{\alpha \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\alpha))^H \mathbf{W}_T (\mathbf{s}_T - \mathbf{s}(\alpha))$$

où $\mathbf{W}_T$ est ici une estimée consistante de l'inverse $\mathcal{C}_s^{-1}(\Theta)$ de la matrice de covariance asymptotique associée au *modèle erroné gaussien complexe circulaire indépendant* conduit à un estimateur $\Theta_{1,T}$ dont la matrice de covariance asymptotique n'est plus donnée par

$$(\mathbf{S}^H \mathbf{C}_s^{-1}(\Theta) \mathbf{S})_{(1:K, 1:K)}^{-1}$$

ni par

$$(\mathbf{S}^H \mathcal{C}_s^{-1}(\Theta) \mathbf{S})_{(1:K, 1:K)}^{-1}.$$

Par contre le raisonnement du paragraphe précédent continue à s'appliquer avec ici

$$\mathbf{G}(\alpha) \stackrel{\text{def}}{=} \mathcal{C}_s^{-1}(\Theta) \Psi(\alpha) (\Psi^H(\alpha) \mathcal{C}_s^{-1}(\Theta) \Psi(\alpha))^{-1} \Psi^H(\alpha) \mathcal{C}_s^{-1}(\Theta).$$

6.6 Application de COMET à l'estimation de DOA

Nous nous proposons maintenant d'étudier les performances de l'algorithme COMET appliqué à l'estimation de DOA dans 2 contextes différents : lorsque les sources sont strictement bande étroite ou au contraire large bande et ceci sous différentes hypothèses de distribution et de corrélation temporelle de signaux sources.

6.6.1 Modèle et notations

Nous considérerons K signaux source spatialement non corrélés, de support spectral situé autour d'une fréquence f_0 , de largeur de bande B ($\frac{B}{2} < f_0$), d'enveloppe complexe s_t^k par rapport à cette fréquence f_0 , de puissance $\sigma_{s_k}^2$, caractérisés par leur DOA : $(\theta_k)_{k=1,\dots,K}$. L'enveloppe complexe $\mathbf{y}_t$ des signaux observés derrière un réseau de M capteurs, en présence de bruit gaussien spatialement non corrélé, de puissance non nécessairement identique sur les différents capteurs, a pour matrice de covariance spatiale

$$\mathbf{R}_y = \mathbb{E}(\mathbf{y}_t \mathbf{y}_t^H) = \sum_{k=1}^K \mathbf{a}(\theta_k) \mathbf{a}^H(\theta_k) \odot \mathbf{R}_{s_k} + \mathbf{R}_n$$

avec

$$\mathbf{R}_n \stackrel{\text{def}}{=} \text{Diag}(\sigma_1^2, \dots, \sigma_M^2),$$

$$\mathbf{a}(\theta_k) \stackrel{\text{def}}{=} [e^{-i2\pi f_0 \tau_{k,1}}, \dots, e^{-i2\pi f_0 \tau_{k,M}}]^T,$$

où $\tau_{k,m}$ est le retard sur le capteur m pour la source k par rapport à un point de référence, où $\mathbf{R}_{s_k}$ est la matrice $M \times M$ dont le terme (m, n) est la corrélation

$$[\mathbf{R}_{s_k}]_{m,n} = \mathbb{E}(s_{t-\tau_{k,m}}^k s_{t-\tau_{k,n}}^{k*})$$

de la source k , et où $\odot$ désigne le produit de matrices terme à terme.

Dans le cas de sources strictement bande étroite $\mathbf{R}_{s_k} = \sigma_{s_k}^2 \mathbf{1}\mathbf{1}^T$ et

$$\mathbf{R}_y = \sum_{k=1}^K \sigma_{s_k}^2 \mathbf{a}(\theta_k) \mathbf{a}^H(\theta_k) + \mathbf{R}_n.$$

Algorithme COMET bande étroite

Pour le modèle de K sources strictement bande étroite non corrélées spatialement et d'un réseau d'antenne connu quelconque, $\mathbf{R}_y$ est paramétrée par les $M + 2K$ paramètres réels (Θ_1, Θ_2) avec

$$\Theta_1 \stackrel{\text{def}}{=} (\theta_1, \dots, \theta_K)^T \quad \text{et} \quad \Theta_2 \stackrel{\text{def}}{=} (\sigma_{s_1}^2, \dots, \sigma_{s_K}^2, \sigma_1^2, \dots, \sigma_M^2)^T.$$

Nous supposons que (Θ_1, Θ_2) est strictement identifiable à partir de la matrice $\mathbf{R}_y$. Remarquons que cette condition d'identification dépend de la géométrie d'antenne. Il est en effet nécessaire que le nombre d'inconnues, ici $2K + M$, soit au plus égal au nombre d'équations réelles distinctes fournies par le réseau. Par exemple, dans le cas d'un réseau linéaire uniforme (LU) sans ambiguïté, la matrice $\mathbf{R}_y$ a une structure "presque" Toeplitz (les termes de la diagonale principale peuvent être différents). Ce réseau fournit donc $3M - 2$ équations réelles différentes et la condition d'identification se ramène à $K < M$. Dans le cas d'un réseau linéaire lacunaire sans redondance ni ambiguïté, le nombre d'équations réelles différentes est égal

à M^2 et la condition d'identification devient alors $K \leq \frac{M(M-1)}{2}$.

La matrice $\mathbf{R}_y$ étant linéaire par rapport à Θ_2 , sa vectorialisation s'écrit sous la forme

$$\mathbf{r}_y \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_y) = \Psi_1(\Theta)\Theta_2.$$

avec

$$\Psi_1(\Theta) = [\mathbf{a}(\theta_1) \otimes_c \mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K) \otimes_c \mathbf{a}(\theta_K), \mathbf{e}_1 \otimes_c \mathbf{e}_1, \dots, \mathbf{e}_M \otimes_c \mathbf{e}_M],$$

où $\mathbf{e}_k$ désigne le k -ième vecteur unité de $\mathbb{R}^M$.

Algorithme COMET large bande

Pour le modèle de sources large bande, le paramétrage de $\mathbf{R}_y$ nécessite un plus grand nombre de termes et il n'est identifiable que dans le cas d'une seule source de spectre symétrique par rapport à f_0 et pour un réseau d'antenne quelconque connu. En effet, par exemple dans le cas d'une seule source de spectre non symétrique et d'un réseau LU, la matrice complexe Hermitienne Toeplitz $\mathbf{R}_{s_1}$ nécessite $2(M-1)$ termes réels pour le paramétrage des ses éléments non diagonaux. La matrice diagonale $\mathbf{R}_n$ et les éléments diagonaux de $\mathbf{R}_{s_1}$ en nécessitant M , il faut au total $3M-1$ paramètres réels pour paramétrer $\mathbf{R}_y$. Or le nombre d'équations réelles distinctes est de $3M-2$ et le paramétrage n'est alors pas identifiable. Pour qu'il le soit, il est nécessaire que la matrice $\mathbf{R}_{s_1}$ soit réelle et symétrique et donc que le spectre de la source soit symétrique par rapport à f_0 . Dans ce cas, les termes de la diagonale principale sont égaux à $\sigma_{s_1}^2$ et la partie hors diagonale est paramétrée par ses termes $(r_{m,n})_{m,n \in \mathcal{M}}$ associés à l'ensemble des valeurs $\tau_{1,n} - \tau_{1,m}$ distinctes. L'ensemble $\mathcal{M}$ dépend de la géométrie de l'antenne.

Par exemple, pour un réseau LU, $\mathcal{M} = \{(2,1), \dots, (M,1)\}$ alors que pour un réseau circulaire uniforme, $\mathbf{R}_{s_1}$ n'ayant pas de structure supplémentaire $\mathcal{M} = \{(m,n); 1 \leq m < n \leq M\}$.

Par suite $\mathbf{R}_y$ est paramétrée par les paramètres réels (Θ_1, Θ_2) avec

$$\Theta_1 \stackrel{\text{def}}{=} \theta_1 \quad \text{et} \quad \Theta_2 \stackrel{\text{def}}{=} ((r_{m,n})_{m,n \in \mathcal{M}}, \sigma_{s_1}^2 + \sigma_1^2, \dots, \sigma_{s_1}^2 + \sigma_M^2)^T$$

et (Θ_1, Θ_2) est strictement identifiable à partir de la matrice $\mathbf{R}_y$. De plus $\mathbf{R}_y$ étant linéaire par rapport à Θ_2 , sa vectorialisation s'écrit à nouveau sous la forme

$$\mathbf{r}_y \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_y) = \Psi_2(\Theta_1)\Theta_2.$$

Par exemple, dans le cas d'un réseau LU

$$\Psi_2(\Theta_1) = \left[\begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_M \end{bmatrix}, \mathbf{e}_1 \otimes_c \mathbf{e}_1, \dots, \mathbf{e}_M \otimes_c \mathbf{e}_M \right]$$

où, pour tout $m = 1, \dots, M$, $\mathbf{A}_m$ est une matrice $M \times (M-1)$ telle que

$$[\mathbf{A}_m]_{(m,:)} = \mathbf{0}^T \quad \text{et} \quad [\mathbf{A}_m]_{(k,l)} = \begin{cases} e^{i2\pi f_0 l \tau} & \text{si } k+l = m \\ e^{-i2\pi f_0 l \tau} & \text{si } k-l = m \end{cases}$$

où τ représente le retard entre deux capteurs adjacents.

Dans ces 2 types de situations nous sommes dans les conditions d'application de l'algorithme COMET original et, à partir de la matrice de covariance empirique $\mathbf{R}_T \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^H$ de l'observation, l'estimateur de Θ_1 sera donné par (2.15)

$$\Theta_{1,T} = \arg \max_{\Theta_1} \mathbf{g}^H(\Theta_1, \mathbf{R}_T) \mathbf{G}^{-1}(\Theta_1, \mathbf{R}_T) \mathbf{g}(\Theta_1, \mathbf{R}_T)$$

avec

$$\mathbf{g}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^H(\Theta_1) \mathbf{W}_T \mathbf{s}_T, \quad \mathbf{G}(\Theta_1, \mathbf{R}_T) \stackrel{\text{def}}{=} \Psi^H(\Theta_1) \mathbf{W}_T \Psi(\Theta_1) \quad \text{et} \quad \mathbf{s}_T \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_T)$$

Dans le cas d'observations $(\mathbf{y}_t)_{t=1,\dots,T}$ indépendantes, un estimateur consistant $\mathbf{W}_T$ de $\mathbf{C}_R^{-1}(\Theta)$ est disponible ce qui permet d'avoir un algorithme COMET qui soit AMV.

Dans le cas gaussien complexe circulaire

$$\mathbf{W}_T^{-1} = \mathbf{R}_T \otimes_c \mathbf{R}_T$$

et dans le cas général d'observations $\mathbf{y}_T$ complexes indépendantes non nécessairement gaussiennes

$$\mathbf{W}_T^{-1} = \frac{1}{T} \sum_{t=1}^T \left((\mathbf{y}_t \otimes_c \mathbf{y}_t - \text{Vec}(\mathbf{R}_T)) (\mathbf{y}_t \otimes_c \mathbf{y}_t - \text{Vec}(\mathbf{R}_T))^H \right).$$

car

$$\text{Vec}(\mathbf{R}_T) = \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \otimes_c \mathbf{y}_t$$

6.6.2 Performances asymptotiques de COMET en DOA

Borne de Cramer-Rao

Dans le cas particulier d'observations complexes circulaires gaussiennes indépendantes, l'algorithme COMET est asymptotiquement gaussien et efficace et (2.19) nous donne indirectement l'expression $CRB_{\Theta_1} = \frac{1}{T} \mathbf{C}_{\Theta_1}$ de la borne de Cramer-Rao du paramètre Θ_1 seul. Nous remarquons que dans le cas particulier d'une seule source, bande étroite ou large bande à spectre symétrique et d'une géométrie d'antenne quelconque, une expression de CRB_{θ_1} , commune au cas bande étroite et large bande, peut être directement déduite de l'expression classique (voir, e.g., [7, rel. (15.52)]) de la matrice de Fisher associée au paramètre $\Theta \stackrel{\text{def}}{=} (\Theta_1, \Theta_2)$:

$$[\mathbf{I}(\Theta)]_{k,l} = T \text{Tr} \left(\mathbf{R}_y^{-1} \frac{\partial \mathbf{R}_y}{\partial \theta_k} \mathbf{R}_y^{-1} \frac{\partial \mathbf{R}_y}{\partial \theta_l} \right). \quad (6.1)$$

En effet dans ce cas $\mathbf{R}_y$ se met sous la forme :

$$\mathbf{R}_y = \Delta_{\theta_1} \mathbf{R}_{\Theta_2} \Delta_{\theta_1}^H,$$

avec $\Delta_{\theta_1} \stackrel{\text{def}}{=} \text{Diag}(e^{i2\pi f_0 \tau_{1,1}}, \dots, e^{i2\pi f_0 \tau_{1,M}})$ et $\mathbf{R}_{\Theta_2} = \mathbf{R}_{s_1} + \text{Diag}(\sigma_1^2, \dots, \sigma_M^2)$. En utilisant une démonstration tout à fait similaire à celle présentée dans le paragraphe 5.5.1, on démontre que la matrice de Fisher (6.1) associée à notre modélisation se découple sous la forme :

$$\mathbf{I}(\Theta) = \begin{bmatrix} I(\theta_1) & \mathbf{0}^T \\ \mathbf{0} & \mathbf{I}(\Theta_2) \end{bmatrix}$$

avec

$$[\mathbf{I}(\Theta_2)]_{k,l} = T \text{Tr} \left(\mathbf{R}_{\Theta_2}^{-1} \frac{\partial \mathbf{R}_{\Theta_2}}{\partial \theta_k} \mathbf{R}_{\Theta_2}^{-1} \frac{\partial \mathbf{R}_{\Theta_2}}{\partial \theta_l} \right) \quad k, l \geq 2$$

et

$$I(\theta_1) = 2T \text{Tr}(\Delta_{\theta_1}' \mathbf{R}_{\Theta_2}^{-1} \Delta_{\theta_1}' \mathbf{R}_{\Theta_2} - \Delta_{\theta_1}'^2)$$

où

$$\Delta'_{\theta_1} \stackrel{\text{def}}{=} 2\pi f_0 \text{Diag}(\tau'_{1,1}, \dots, \tau'_{1,M})$$

et

$$\tau'_{1,m} \stackrel{\text{def}}{=} \frac{d\tau_{1,m}}{d\theta_1}, \quad m = 1, \dots, M.$$

D'où :

$$CRB_{\theta_1} = \frac{1}{2T} \left(\text{Tr}(\Delta'_{\theta_1} \mathbf{R}_{\Theta_2}^{-1} \Delta'_{\theta_1} \mathbf{R}_{\Theta_2} - \Delta'_{\theta_1}{}^2) \right)^{-1}. \quad (6.2)$$

En bande étroite, cette expression devient une forme analytique interprétable, car dans ce cas le lemme d'inversion matricielle donne :

$$\mathbf{R}_{\Theta_2}^{-1} = \mathbf{\Sigma} - (\sigma_{s_1}^{-2} + \sum_{m=1}^M \sigma_m^{-2})^{-1} \tilde{\mathbf{\Sigma}}$$

avec $\mathbf{\Sigma} \stackrel{\text{def}}{=} \text{Diag}(\sigma_1^{-2}, \dots, \sigma_M^{-2})$ et $[\tilde{\mathbf{\Sigma}}]_{m,n} \stackrel{\text{def}}{=} \sigma_m^{-2} \sigma_n^{-2}$, $m, n = 1, \dots, M$. Ce qui permet d'obtenir l'expression :

$$CRB_{\theta_1} = \frac{1}{2T} ([\sigma_{s_1}^2 - (\sum_{m=1}^M \sigma_m^{-2})^{-1}] \mathbf{a}'_{\theta_1} \mathbf{\Sigma} \mathbf{a}'_{\theta_1} - \sigma_{s_1}^2 (\sum_{m=1}^M \sigma_m^{-2})^{-1} |\mathbf{a}_{\theta_1}^H \mathbf{\Sigma} \mathbf{a}'_{\theta_1}|^2)^{-1} \quad (6.3)$$

avec $\mathbf{a}'_{\theta_1} \stackrel{\text{def}}{=} \frac{d\mathbf{a}_{\theta_1}}{d\theta_1}$. Nous signalons qu'a été porté récemment à notre connaissance une expression générale [9] de la borne de Cramer-Rao du paramètre Θ_1 seul dans le cas du modèle stochastique bande étroite à matrice de covariance du bruit paramétrée de façon quelconque, obtenue directement à partir de l'expression (6.1).

Nous allons maintenant étudier les performances de l'algorithme COMET en DOA dans certaines situations particulières : lorsque les signaux sources ne sont pas gaussiens, lorsqu'ils sont temporellement corrélés ou lorsque le rapport S/B sur certains capteurs se dégrade.

Covariance asymptotique

Dans le cas d'observations *gaussiennes complexes circulaires indépendantes*, l'algorithme COMET avec $\mathbf{W}_T = \mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1}$ est asymptotiquement efficace et

$$\sqrt{T} (\Theta_{1,T} - \Theta_1) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_{\Theta_1})$$

avec $\mathbf{C}_{\Theta_1}$ donné de façon générale par (2.19) et dans le cas particulier d'une source par $T CRB_{\theta_1}$ de (6.2).

Par contre, dans le cas d'observations complexes *non gaussiennes et/ou non indépendantes*, l'algorithme COMET utilisant la matrice de pondération $\mathbf{W}_T = \mathbf{R}_T^{-1} \otimes_c \mathbf{R}_T^{-1}$ ne sera plus ni efficace, ni AMV mais nous avons toujours

$$\sqrt{T} (\Theta_{1,T} - \Theta_1) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_{\Theta_1})$$

avec, d'après le corollaire 6.1,

$$\mathbf{C}_{\Theta_1} = \left[4 \mathbf{D}_R^{COMET} \mathbf{C}_R(\Theta) (\mathbf{D}_R^{COMET})^H \right]_{1:K, 1:K} \quad (6.4)$$

où $\mathbf{D}_R^{COMET}$ est donnée en annexe 3 par (8.5)

$$\mathbf{D}_R^{COMET} = \frac{1}{2} (\mathbf{S}^H \mathbf{W}(\Theta) \mathbf{S})^{-1} \mathbf{S}^H \mathbf{W}(\Theta)$$

avec $\mathbf{W}(\Theta) = \mathbf{R}_y^{-1}(\Theta) \otimes_c \mathbf{R}_y^{-1}(\Theta)$.

Dans le cas d'un bruit temporellement non corrélé et indépendant des signaux source, l'expression de $\mathbf{C}_R(\Theta)$ qui apparaît dans (6.4) est donnée en bande étroite par l'expression (3.3) du chapitre 4, mais pour une source large bande, l'analyse développée au chapitre 4 doit être modifiée comme suit.

En posant

$$\mathbf{y}_t \stackrel{\text{def}}{=} \mathbf{s}_t + \mathbf{n}_t = \mathbf{s}_{\tau,t} \odot \mathbf{a}(f_0) + \mathbf{n}_t \quad (6.5)$$

avec

$$\mathbf{s}_{\tau,t} \stackrel{\text{def}}{=} (s_{t-\tau_1}, \dots, s_{t-\tau_M})^T$$

où s_t représente l'enveloppe complexe de la source par rapport à la fréquence f_0 de la porteuse et $\{\tau_1, \dots, \tau_M\}$ désignent les retards du signal source sur les différents capteurs et

$$\mathbf{a}(f_0) \stackrel{\text{def}}{=} (a_1(f_0), \dots, a_M(f_0))^T \stackrel{\text{def}}{=} (e^{-i2\pi f_0 \tau_1}, \dots, e^{-i2\pi f_0 \tau_M})^T,$$

la matrice $\mathbf{C}_R(\Theta)$ s'écrit, d'après l'expression (3.3) du chapitre 4,

$$\mathbf{C}_R(\Theta) = \mathbf{C}_{R_s} + \mathbf{C}_{R_n} + \mathbf{R}_s \otimes_c \mathbf{R}_n + \mathbf{R}_n \otimes_c \mathbf{R}_s \quad (6.6)$$

avec

$$\mathbf{R}_s = \mathbf{a}(f_0) \mathbf{a}^H(f_0) \odot \mathbf{R}_{s_1} \quad \text{où} \quad \mathbf{R}_{s_1} \stackrel{\text{def}}{=} \mathbb{E}[\mathbf{s}_{\tau,t} \mathbf{s}_{\tau,t}^H],$$

$$\mathbf{C}_{R_n} = \mathbf{R}_n \otimes_c \mathbf{R}_n \quad \text{où} \quad \mathbf{R}_n = \text{Diag}(\sigma_1^2, \dots, \sigma_M^2)$$

(le bruit est supposé gaussien complexe circulaire, temporellement et spatialement non corrélé), et

$$\mathbf{C}_{R_s} \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \frac{1}{T} \text{Cov}(\text{Vec}(\mathbf{R}_s(T))).$$

Or $\mathbf{R}_s(T) = \frac{1}{T} \sum_{t=1}^T \mathbf{s}_t \mathbf{s}_t^H$ donc

$$\text{Vec}(\mathbf{R}_s(T)) = \frac{1}{T} \sum_{t=1}^T (\mathbf{s}_{\tau,t} \odot \mathbf{a}(f_0)) \otimes_c (\mathbf{s}_{\tau,t} \odot \mathbf{a}(f_0))$$

et $\mathbf{C}_{R_s}$ s'obtient, comme au chapitre 4 relation (3.4), à l'aide de la densité spectrale de puissance à temps discret $S_{s_1}(f)$ ($f_e = 1$) de l'enveloppe complexe de la source large bande

$$\mathbf{C}_{R_s} = \int_{-1/2}^{1/2} \mathbf{S}_s(f) \otimes_c \mathbf{S}_s(f) df + \mathbf{Q}_s$$

avec ici

$$[\mathbf{S}_s(f)]_{k,l} = a_k(f_0) a_l^*(f_0) S_{s_1}(f) e^{-i2\pi f(\tau_k - \tau_l)} \quad k, l = 1, \dots, M$$

soit

$$\mathbf{S}_s(f) = S_{s_1}(f) (\mathbf{a}(f + f_0) \mathbf{a}^H(f + f_0)).$$

D'autre part, grâce à (6.5) et à (3.4) du chapitre 4,

$$[\mathbf{Q}_s]_{M(j-1)+i, M(l-1)+k} = a_i(f_0) a_j^*(f_0) a_l(f_0) a_k^*(f_0) \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t_1=1}^T \sum_{t_2=1}^T \text{Cum}(s_{t_1-\tau_i}, s_{t_1-\tau_j}^*, s_{t_2-\tau_l}, s_{t_2-\tau_k}^*),$$

d'où

$$\mathbf{Q}_s \stackrel{\text{def}}{=} ((\mathbf{a}(f_0) \mathbf{a}^H(f_0)) \otimes_c (\mathbf{a}(f_0) \mathbf{a}^H(f_0))) \odot \mathbf{Q}_{s_1} \quad (6.7)$$

avec

$$[\mathbf{Q}_{s_1}]_{M(j-1)+i, M(l-1)+k} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t_1=1}^T \sum_{t_2=1}^T \text{Cum}(s_{t_1-\tau_i}, s_{t_1-\tau_j}^*, s_{t_2-\tau_l}, s_{t_2-\tau_k}^*) \quad (6.8)$$

et, grâce à [8, lemme 4.1 p.101],

$$[\mathbf{Q}_{s_1}]_{M(j-1)+i, M(l-1)+k} = \sum_{t=-\infty}^{\infty} \text{Cum}(s_{t-\tau_i}, s_{t-\tau_j}^*, s_{t-\tau_l}, s_{t-\tau_k}^*). \quad (6.9)$$

Une démonstration en tout point similaire à celle utilisée en annexe 1 du chapitre 4 conduisant à l'expression (9.7) de ce chapitre, nous donne alors

$$[\mathbf{Q}_{s_1}]_{M(j-1)+i, M(l-1)+k} = \int_{-1/2}^{+1/2} \int_{-1/2}^{+1/2} \rho_{i,j,k,l}(f, -f', f') df df'$$

où

$$\rho_{i,j,k,l}(f, f', f'') \stackrel{\text{def}}{=} \sum_{t, t', t''} \text{Cum}(s_{t-\tau_i}, s_{t-\tau_j}^*, s_{t'-\tau_l}, s_{t''-\tau_k}^*) e^{i2\pi(ft + f't' + f''t')}$$

désigne “l'intertrispectre” de $(s_{t-\tau_i}, s_{t-\tau_j}^*, s_{t-\tau_l}, s_{t-\tau_k}^*)$.

Remarque 1 : Dans le cas particulier d'une source large bande gaussienne et d'observations issues d'un échantillonnage à la fréquence de Nyquist et par suite non indépendantes, $\mathbf{Q}_s = \mathbf{O}$ et $\mathbf{C}_R(\Theta)$ est donnée par (6.6) avec

$$\mathbf{C}_{R_s} = \int_{-1/2}^{1/2} |S_{s_1}(f)|^2 (\mathbf{a}(f + f_0) \mathbf{a}^H(f + f_0)) \otimes_c (\mathbf{a}(f + f_0) \mathbf{a}^H(f + f_0)) df.$$

En particulier, si la DSP de s_t est uniforme dans une bande de fréquence de largeur B centrée sur une fréquence f_c ,

$$\mathbf{C}_{R_s} = \frac{\sigma_s^4}{B^2} \int_{f_c-B/2}^{f_c+B/2} (\mathbf{a}(f + f_0) \mathbf{a}^H(f + f_0)) \otimes_c (\mathbf{a}(f + f_0) \mathbf{a}^H(f + f_0)) df.$$

où σ_s^2 représente la puissance de s_t . On en déduit alors aisément le terme général de $\mathbf{C}_{R_s}$

$$[\mathbf{C}_{R_s}]_{M(j-1)+i, M(l-1)+k} = \frac{\sigma_s^4}{B} \text{Sinc}(\pi B(\tau_i - \tau_j + \tau_l - \tau_k)) e^{-i2\pi(f_0+f_c)(\tau_i-\tau_j+\tau_l-\tau_k)}.$$

Remarque 2 : Dans le cas particulier d'une source large bande *non gaussienne* et d'observations indépendantes ⁴, $\mathbf{S}_s(f) = \mathbf{R}_s$ et

$$\mathbf{C}_{R_s} = \mathbf{R}_s \otimes_c \mathbf{R}_s + \mathbf{Q}_s$$

avec $\mathbf{Q}_s$ donnée par (6.7). Dans la situation que nous considérons ici, si $t_1 \neq t_2$, les signaux sources $s_{t_1-\tau}$ et $s_{t_2-\tau}$ sont indépendants. De ce fait, compte tenu de la stationnarité de s_t , (6.8) devient

$$[\mathbf{Q}_{s_1}]_{M(j-1)+i, M(l-1)+k} = \text{Cum}(s_{-\tau_i}, s_{-\tau_j}^*, s_{-\tau_l}, s_{-\tau_k}^*).$$

Cette dernière expression, fonction des statistiques du quatrième ordre du signal source s_t , dépend de la façon dont cette source est générée. Dans nos simulations, pour pouvoir faire varier facilement la largeur de bande B du signal, nous avons opté pour une approximation du signal sous la forme

$$s_t = \sum_{m=0}^{L-1} a_m e^{i2\pi f_m t} \quad (6.10)$$

avec $f_m = \frac{-L+2m+1}{2(L-1)}B$ et $L \gg 1$ ($L = 100$ dans nos simulations), où $(a_m)_{m=0, \dots, L-1}$ sont des variables aléatoires complexes indépendantes centrées, de parties réelles et imaginaires indépendantes, de même loi, telle que $E[|a_m|^2] = \frac{B}{L} S_{s_1}(f_m)$. Le cumulante du quatrième ordre normalisé de a_m , $m = 0, \dots, L-1$ est noté κ_s .

Nous notons que cette approximation de s_t n'est seulement qu'une approximation au 2^{ème} ordre de la représentation spectrale $s_t = \int_{-\frac{B}{2}}^{\frac{B}{2}} e^{i2\pi f t} d\mu(f)$. En conséquence, les propriétés au 4^{ème} ordre seraient différentes si le signal source était généré de façon différente.

La modélisation (6.10) conduit à

$$\text{Cum}(s_{-\tau_i}, s_{-\tau_j}^*, s_{-\tau_l}, s_{-\tau_k}^*) = \sum_{m=0}^{L-1} e^{-i2\pi f_m(\tau_i - \tau_j + \tau_l - \tau_k)} \text{Cum}(a_m, a_m^*, a_m, a_m^*)$$

avec

$$\text{Cum}(a_m, a_m^*, a_m, a_m^*) = \frac{B^2}{L^2} S_{s_1}^2(f_m) \kappa_s$$

d'où

$$\text{Cum}(s_{-\tau_i}, s_{-\tau_j}^*, s_{-\tau_l}, s_{-\tau_k}^*) = \frac{B^2}{L^2} \kappa_s \sum_{m=0}^{L-1} e^{-i2\pi f_m(\tau_i - \tau_j + \tau_l - \tau_k)} S_{s_1}^2(f_m).$$

En particulier, dans le cas d'une source de puissance σ_s^2 , de spectre uniforme sur $[-\frac{B}{2}, \frac{B}{2}]$, en posant $\tau_{i,j,k,l}(\theta) \stackrel{\text{def}}{=} \tau_i - \tau_j + \tau_l - \tau_k$, cette expression devient

$$\begin{aligned} \text{Cum}(s_{-\tau_i}, s_{-\tau_j}^*, s_{-\tau_l}, s_{-\tau_k}^*) &= \frac{\sigma_s^4}{L^2} \kappa_s \sum_{m=0}^{L-1} e^{-i2\pi \frac{-L+2m+1}{2(L-1)} B \tau_{i,j,k,l}(\theta)} \\ &= \frac{\sigma_s^4}{L^2} \kappa_s \frac{\sin \pi B \tau_{i,j,k,l}(\theta)}{\sin \pi \frac{B}{L} \tau_{i,j,k,l}(\theta)} \\ &\approx \frac{\sigma_s^4 \kappa_s \sin \pi B \tau_{i,j,k,l}(\theta)}{L \pi B \tau_{i,j,k,l}(\theta)} \quad \text{pour } L \gg 1. \end{aligned} \quad (6.11)$$

⁴c'est à dire par exemple issues d'un échantillonnage effectué à une fréquence très inférieure à la fréquence de Nyquist.

Remarquons que la matrice $\mathbf{Q}_s$ dépend de la modélisation choisie. En particulier, dans le cas d'un signal source constituée de L cisoïdes indépendantes, $\mathbf{Q}_s$ est une fonction décroissante de L . Ceci est en accord avec le théorème central limite qui implique la convergence en loi de s_t vers une loi gaussienne, de cumulant d'ordre quatre nul, lorsque M tend vers l'infini.

Pour illustrer ces résultats, la FIG.6.1 représente, en fonction du rapport $\frac{B}{f_0}$ de la largeur de bande B du signal à la fréquence f_0 de la porteuse pour une source gaussienne de spectre plat entre $[f_0 - \frac{B}{2}, f_0 + \frac{B}{2}]$, pour un réseau linéaire uniforme (LU) de 5 capteurs, une taille de snapshot $T = 160$ et un $S/B = 20$ dB,

- l'EQM estimée par méthode de Monte Carlo (100 réalisations) de l'algorithme COMET large bande appliqué à des échantillons gaussiens indépendants (*),
- l'EQM estimée par méthode de Monte Carlo (100 réalisations) de l'algorithme COMET bande étroite appliqué à des échantillons gaussiens indépendants (o),
- l'EQM estimée par méthode de Monte Carlo (100 réalisations) de l'algorithme COMET large bande appliqué à des échantillons gaussiens issus d'un échantillonnage à la fréquence de Nyquist (x),
- l'EQM théorique asymptotique (en T) de l'algorithme COMET large bande appliqué à des échantillons gaussiens non indépendants au sens précédent (— —),
- la borne de Cramer-Rao (-)
- l'EQM estimée (100 réalisations) de l'algorithme de focalisation de Friedlander [12] (+) (les snapshots sont constitués de 10 groupes indépendants de 16 échantillons prélevés à la fréquence de Nyquist).

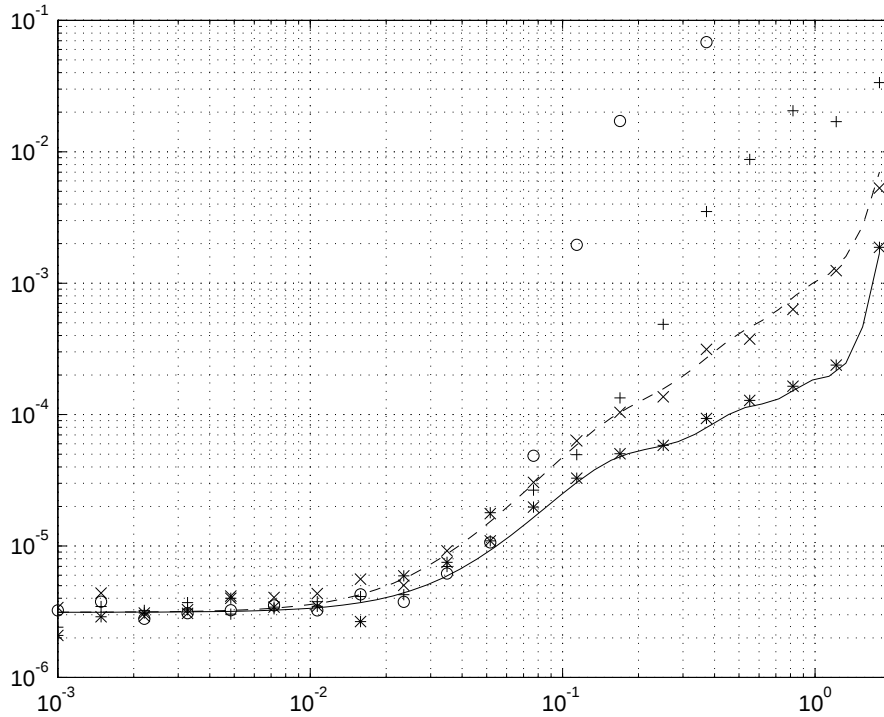


FIG. 6.1: Borne de Cramer-Rao (— —), EQM de COMET large bande appliqué à des échantillons gaussiens indépendants (*), ou issus d'un échantillonnage à la fréquence de Nyquist (x), de l'algorithme COMET bande étroite (o) et de l'algorithme de focalisation de Friedlander (+) en fonction de $\frac{B}{f_0}$ pour une source gaussienne de spectre uniforme, un réseau LU de 5 capteurs, $T = 160$ et $S/B = 20$ dB.

Cette figure, qui est à rapprocher de la figure FIG.5.12 du chapitre 5, permet de constater

- une bonne adéquation entre calcul théorique et estimation par méthode de Monte Carlo pour toutes les valeurs de $\frac{B}{f_0}$ jusqu'à $\frac{B}{f_0} = 2$ pour seulement $T = 160$,
- l'efficacité de l'algorithme COMET large bande dans le cas d'observations indépendantes,
- la non efficacité de l'algorithme COMET large bande dans le cas où les observations sont issues d'un échantillonnage à la fréquence de Nyquist,
- une bonne robustesse de l'algorithme COMET bande étroite jusqu'à $\frac{B}{f_0} \approx 10^{-1}$
- la non efficacité de l'algorithme de focalisation de Friedlander pour $\frac{B}{f_0} > 10^{-1}$.

6.6.3 Sensibilité de COMET à la distribution et à la corrélation temporelle des signaux sources

Nous avons montré précédemment (cf. chapitre 4 paragraphe 4.7.3) que, si le bruit est temporellement non corrélé, l'algorithme COMET bande étroite défini sous l'hypothèse a priori de non corrélation spatiale des sources (condition 2b du chapitre 4 paragraphe 4.4) n'est pas sensible à la distribution des sources à condition que celles ci soient indépendantes.

Cette propriété, par contre, n'est pas vérifiée par l'algorithme COMET large bande comme le montre la FIG.6.2. Cette figure présente les erreurs quadratiques moyennes (EQM) sur l'estimation, à partir de $T = 100$ observations indépendantes, de la direction d'arrivée (pulsation spatiale) d'une source gaussienne ($\circ$) ou distribuée selon une loi "normal inverse gaussian" (NIG) [13] de cumulants normalisés d'ordre 4 égal à 150 ($*$) en fonction du S/B pour l'algorithme COMET large bande. On constate une nette dégradation des performances de cet algorithme lorsque la source n'est pas gaussienne.

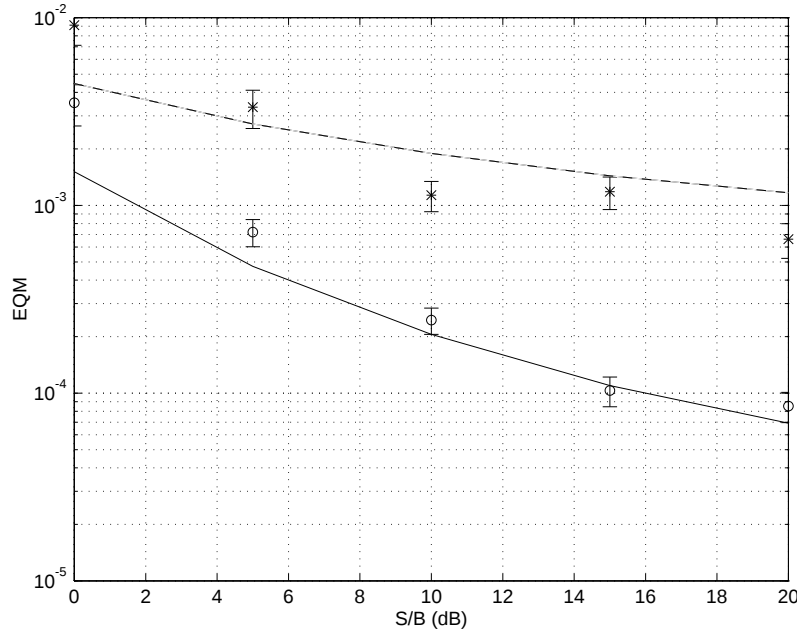


FIG. 6.2: EQM théorique et estimée (100 réalisations) de l'algorithme COMET large bande dans le cas d'une source gaussienne (- et $\circ$) ou distribuée selon une loi NIG [13] de cumulants normalisés d'ordre 4 égal à 150 (- - et $*$) en fonction du S/B pour un réseau LU de 10 capteurs et $\frac{B}{f_0} = 1$.

Toujours sous l'hypothèse d'un bruit temporellement non corrélé, nous avons également montré (cf. chapitre 4 résultat 4.8) que, dans le cas d'une source ou de plusieurs sources gaussiennes à spectres deux

à deux disjoints, l'algorithme COMET bande étroite n'est pas sensible à la corrélation temporelle des sources . Les FIG.6.3 et FIG.6.4 ci-dessous illustrent ces propriétés.

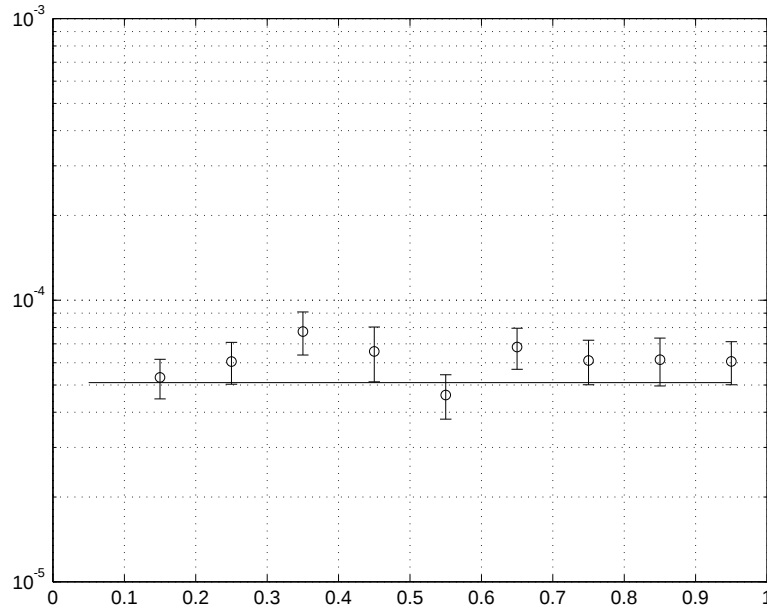


FIG. 6.3: EQM asymptotique (-) et estimée (o) de COMET bande étroite en fonction de la largeur de bande de la source.

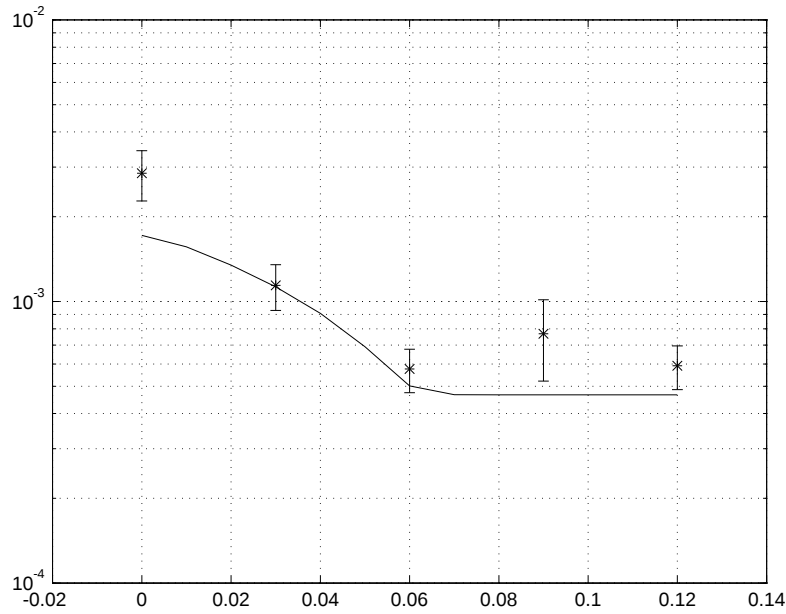


FIG. 6.4: EQM asymptotique (-) et estimée (*) de COMET bande étroite dans le cas de 2 sources de largeur de bande $B = 0.06$ en fonction de l'écart entre leur fréquence centrale.

La FIG.6.3 présente l'EQM asymptotique (-) et estimée (o) par méthode de Monte Carlo (100 réalisations et intervalles de confiance à 99 %) sur l'estimation, à partir de $T = 100$ observations non-indépendantes (prélevées à la fréquence de Nyquist), de la direction d'arrivée (pulsation spatiale) d'une source gaussienne de spectre centré, d'azimut 70° , par l'algorithme COMET bande étroite en fonction de la largeur de bande de cette source.

La FIG.6.4 présente l'EQM asymptotique (-) et estimée (*) par méthode de Monte Carlo (100 réalisations et intervalles de confiance à 99 %) sur l'estimation, par l'algorithme COMET bande étroite à partir de $T = 100$ observations non-indépendantes, de la direction d'arrivée de 2 sources gaussiennes de même largeur de bande $B = 0.06$, d'azimuts respectifs 50° et 65° , en fonction de l'écart entre les fréquences centrales de leur spectre. Pour chacune de ces deux simulations le réseau d'antenne est linéaire uniforme, composé de 5 capteurs espacés de $\frac{\lambda_0}{2}$ et le rapport S/B est égal à 10 dB. On constate sur la figure FIG.6.4 que les performances de l'algorithme COMET bande étroite se dégradent dès que les spectres des deux sources commencent à se chevaucher. Cette figure est à rapprocher de la figure FIG.4.7 du chapitre 4 illustrant l'effet de la Toeplitzation où un phénomène similaire est observé.

6.6.4 Robustesse par rapport au S/B de certains capteurs

Dans le cas d'une seule source bande étroite ou large bande, d'un réseau quelconque et d'observations gaussiennes circulaires indépendantes, nous avons la propriété suivante :

Résultat 6.10 *La loi asymptotique de l'estimée obtenue par l'algorithme COMET lorsque les puissances $(\sigma_m^2)_{m \in C_d}$, $C_d \subset \{1, \dots, M\}$ de certains capteurs tendent vers l'infini est celle obtenue à partir du réseau privé de ces capteurs déficients.*

Preuve : L'estimateur COMET étant consistant, il suffit de démontrer que la limite de la borne de Cramer-Rao obtenue lorsque $(\sigma_m^2)_{m \in C_d} \rightarrow \infty$ est égale à la borne de Cramer-Rao associée au réseau privé de ces capteurs déficients. Et pour ce faire, il suffit de démontrer cette propriété de proche en proche en examinant un seul capteur déficient à la fois. Si le capteur déficient est le premier capteur, en reportant les expressions du partitionnement de $\mathbf{R}_\Phi$

$$\mathbf{R}_{\Theta_2} = \begin{bmatrix} \sigma_{s_1}^2 + \sigma_1^2 & \tilde{\mathbf{r}}_{\Theta_2}^H \\ \tilde{\mathbf{r}}_{\Theta_2} & \tilde{\mathbf{R}}_{\Theta_2} \end{bmatrix}.$$

et de son inverse (e.g. [3, pp.413])

$$\mathbf{R}_{\Theta_2}^{-1} = \begin{bmatrix} \delta & (\sigma_{s_1}^2 + \sigma_1^2)^{-1} \tilde{\mathbf{r}}_{\Theta_2}^H \Delta^{-1} \\ (\sigma_{s_1}^2 + \sigma_1^2)^{-1} \Delta^{-1} \tilde{\mathbf{r}}_{\Theta_2} & \Delta^{-1} \end{bmatrix}.$$

avec

$$\delta \stackrel{\text{def}}{=} (\sigma_{s_1}^2 + \sigma_1^2)^{-1} + (\sigma_{s_1}^2 + \sigma_1^2)^{-2} \tilde{\mathbf{r}}_{\Theta_2}^H \Delta^{-1} \tilde{\mathbf{r}}_{\Theta_2}$$

et

$$\Delta \stackrel{\text{def}}{=} \tilde{\mathbf{R}}_{\Theta_2} - \tilde{\mathbf{r}}_{\Theta_2} (\sigma_{s_1}^2 + \sigma_1^2)^{-1} \tilde{\mathbf{r}}_{\Theta_2}^H$$

dans l'expression (6.2) de la borne de Cramer-Rao, on obtient sans difficulté :

$$\lim_{\sigma_1^2 \rightarrow \infty} \text{Tr}(\Delta'_{\theta_1} \mathbf{R}_{\Theta_2}^{-1} \Delta'_{\theta_1} \mathbf{R}_{\Theta_2} - \Delta'^2_{\theta_1}) = \text{Tr}(\tilde{\Delta}'_{\theta_1} \tilde{\mathbf{R}}_{\Theta_2}^{-1} \tilde{\Delta}'_{\theta_1} \tilde{\mathbf{R}}_{\Theta_2} - \tilde{\Delta}'^2_{\theta_1})$$

où $\tilde{\Delta}'_{\theta_1}$ est la matrice diagonale Δ'_{θ_1} privée de sa première ligne et première colonne. Dans le cas d'un capteur déficient quelconque m , il suffit d'appliquer le partitionnement précédent aux matrices $\mathbf{P} \mathbf{R}_{\Theta_2} \mathbf{P}$ et $\mathbf{P} \mathbf{R}_{\Theta_2}^{-1} \mathbf{P}$ où $\mathbf{P}$ est la matrice de permutation qui permute les colonnes 1 et m . ■

Les FIG.6.5 et FIG.6.6 ci-dessous illustrent cette propriété. La FIG.6.5 présente l'EQM sur l'estimation, à partir de $T = 320$ observations indépendantes, de la direction d'arrivée (pulsation spatiale) d'une source gaussienne par l'algorithme COMET large bande en fonction du rapport de la puissance du signal sur

la puissance du bruit sur le 1^{er} capteur, dans le cas où le signal utile sur ce capteur est présent (*), comparées à la borne de Cramer-Rao dans cette situation (-) et à la borne de Cramer-Rao correspondant au réseau privé du 1^{er} capteur (—) pour un réseau LU de 3 capteurs, $\frac{B}{f_0} = 1$ et un S/B = 20dB pour les autres capteurs. On a représenté également sur cette figure, l'EQM dans la situation précédente (puissance de bruit variable sur le 1^{er} capteur) mais avec un signal utile nul sur ce 1^{er} capteur (o).

La FIG.6.6 présente la borne de Cramer-Rao en fonction du rapport de la puissance du signal sur la puissance du bruit sur respectivement le 1^{er}, 2^{eme} ou 3^{eme} capteur d'un réseau LU de 5 capteurs recevant les signaux de 2 sources de même puissance avec un S/B = 20dB pour les autres capteurs. Les droites horizontales représentent la BCR du réseau privé du capteur déficient.

On peut noter la très bonne robustesse de l'algorithme COMET large bande en cas de défaillance complète de l'un des capteurs (puissance de bruit très importante et signal utile coupé) comme le montre la FIG.6.5. Cette propriété a été observée, également dans le cas de l'algorithme COMET bande étroite, à l'aide de simulations, dans différentes situations (nombres différents de capteurs et de sources).

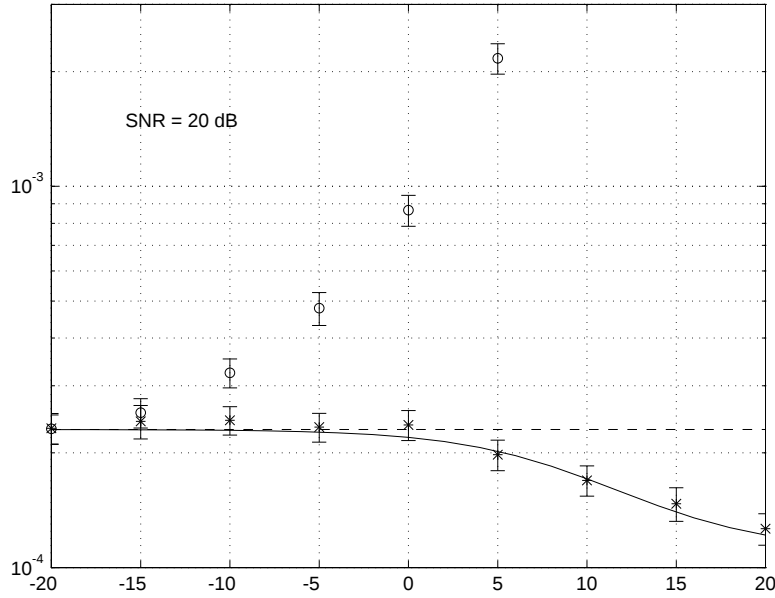


FIG. 6.5: EQM sur la DOA d'une source large bande ($\frac{B}{f_0} = 1$) en fonction du S/B du 1^{er} capteur d'un réseau LU de 3 capteurs dans le cas où le signal utile est présent (*) ou absent (o), comparées à la borne de Cramer-Rao sous la 1^{ere} hypothèse (-) et à la borne de Cramer-Rao correspondant au réseau privé du 1^{er} capteur (—).

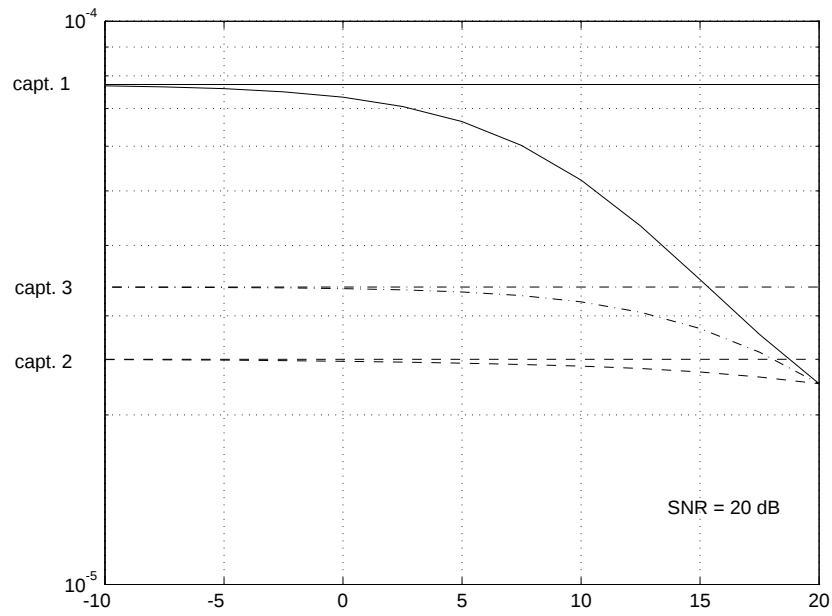


FIG. 6.6: Bornes de Cramer-Rao pour l'estimation de la DOA de 2 sources bande étroite en fonction du S/B de différents capteurs d'un réseau LU de 5 capteurs comparées à la BCR (constante) du réseau privé du capteur déficient.

6.7 Application de COMET à l'estimation de fréquence de sinusôides

Nous nous proposons maintenant d'étudier les performances d'un estimateur COMET modifié appliqué à l'estimation de fréquence de sinusôides en présence d'un bruit MA. Cet estimateur, *asymptotiquement de variance minimale*, sera noté AMV dans la suite. Nous comparerons en particulier ses performances à celles de l'algorithme MUSIC dans le cas d'un bruit blanc et à celles d'un algorithme de type "Pisarenko étendu", noté MPD, proposé par Sherman et Frazho dans [14] dans le cas d'un bruit MA. Nous étudierons également la robustesse de cet algorithme confronté à une erreur sur l'ordre du modèle.

6.7.1 Modèle et notations

Nous considérons dans cette partie un processus complexe constitué de la somme de K sinusôides et d'un bruit MA d'ordre Q :

$$y_t = x_t + n_t \quad \text{avec} \quad x_t = \sum_{k=1}^K a_k e^{i2\pi f_k t} e^{i\phi_k} \quad \text{et} \quad n_t = \sum_{q=0}^Q b_q u_{t-q}$$

Dans ce modèle, $(u_t)_{t \in \mathbb{N}}$ est une suite de variables aléatoires complexes, circulaires, non nécessairement gaussiennes, centrées, indépendantes et identiquement distribuées, telles que $E|u_t|^4 < \infty$; nous noterons $c_u \stackrel{\text{def}}{=} E|u_t|^2$, et $\kappa_u \stackrel{\text{def}}{=} \text{Cum}(u_t, u_t^*, u_t, u_t^*)$. Les constantes $(a_k)_{k=1, \dots, K}$ et $(b_q)_{q=1, \dots, Q}$, respectivement réelles et complexes, sont supposées inconnues. Les fréquences inconnues $(f_k)_{k=1, \dots, K}$ sont supposées réelles et distinctes dans l'intervalle $]-\frac{1}{2}, +\frac{1}{2}[$. Les $(\phi_k)_{k=1, \dots, K}$ sont des variables aléatoires réelles, équidistribuées sur $[0, 2\pi[$ et supposées mutuellement indépendantes et indépendantes des $(u_t)_{t \in \mathbb{N}}$. Le processus aléatoire y_t est alors stationnaire au sens large.

En posant $\mathbf{y}_t \stackrel{\text{def}}{=} (y_t, \dots, y_{t-M+1})^T$, la matrice de covariance d'ordre M de ce processus, $\mathbf{R}(\Theta) \stackrel{\text{def}}{=} E(\mathbf{y}_t \mathbf{y}_t^H)$ s'écrit :

$$\mathbf{R}(\Theta) = \sum_{k=1}^K a_k^2 \mathbf{e}(f_k) \mathbf{e}^H(f_k) + \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_Q & 0 & \cdots & 0 \\ \gamma_1^* & \gamma_0 & \cdots & \gamma_{Q-1} & \gamma_Q & \ddots & \cdots \\ \vdots & \ddots & \ddots & \gamma_{Q-2} & \gamma_{Q-1} & \ddots & 0 \\ \gamma_Q^* & \gamma_{Q-1}^* & \ddots & \ddots & \ddots & \ddots & \gamma_Q \\ 0 & \gamma_Q^* & \gamma_{Q-1}^* & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \gamma_1^* & \gamma_0 & \gamma_1 \\ 0 & \cdots & 0 & \gamma_Q^* & \cdots & \gamma_1^* & \gamma_0 \end{pmatrix}.$$

avec

$$\mathbf{e}(f_k) \stackrel{\text{def}}{=} (1, e^{-i2\pi f_k}, \dots, e^{-i2\pi(M-1)f_k})^T$$

et

$$(\gamma_k)_{k=0, \dots, Q} \stackrel{\text{def}}{=} E(n_t n_{t-k}^*).$$

La matrice $\mathbf{R}(\Theta)$ dépend des $L = 2(Q + K) + 1$ paramètres réels $\Theta = (\Theta_1^T, \Theta_2^T)^T$ avec

$$\Theta_1 \stackrel{\text{def}}{=} (f_1, \dots, f_K)^T$$

qui seront les paramètres d'intérêt et

$$\Theta_2 \stackrel{\text{def}}{=} [a_1^2, \dots, a_K^2, \gamma_0, \Re(\gamma_1), \dots, \Re(\gamma_Q), \Im(\gamma_1), \dots, \Im(\gamma_Q)]^T$$

qui constitueront les paramètres de nuisance. Remarquons que la première colonne $\mathbf{r}(\Theta)$ de $\mathbf{R}(\Theta)$ est linéaire par rapport à Θ_2 : $\mathbf{r}(\Theta) = \Psi'(\Theta_1)\Theta_2$ avec

$$\Psi'(\Theta_1) = \begin{pmatrix} \mathbf{e}(f_1), \dots, \mathbf{e}(f_K), & \mathbf{I}_{Q+1}, & \mathbf{0}^T \\ & \mathbf{O}_{M-Q-1, 2Q+1} & -i\mathbf{I}_Q \end{pmatrix}.$$

La matrice $\Psi'(\Theta_1)$ est de rang plein colonne si et seulement si $M - Q - 1 \geq K$. Cette condition est équivalente au fait que le nombre de paramètres réels inconnus est inférieur ou égal au nombre d'équations disponibles, i.e. $2(Q + K) + 1 \leq 1 + 2(M - 1)$. Cette condition nécessaire n'est toutefois pas suffisante pour assurer l'identifiabilité du paramètre Θ à partir de $\mathbf{R}(\Theta)$. Nous supposons néanmoins dans la suite que si la condition précédente est remplie, le paramètre Θ est effectivement identifiable à partir de $\mathbf{R}(\Theta)$.

La matrice de covariance empirique de l'observation, $\mathbf{R}_T \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^H$ est calculée à partir de T snapshots eux mêmes issus de $N = T + M - 1$ échantillons de l'observation. L'échantillonnage étant régulier et l'observation stationnaire au sens large, $\mathbf{R}(\Theta)$ est de structure Toeplitz et nous désignerons par $\mathbf{R}_T^{to}$ la matrice obtenue par moyennage le long des diagonales de $\mathbf{R}_T$.

La distribution de probabilité asymptotique de ces matrices est caractérisée par le résultat suivant [16] :

Théorème 1 *Vec($\mathbf{R}_T$) et Vec($\mathbf{R}_T^{to}$) ont la même distribution asymptotique gaussienne caractérisée par la distribution asymptotique de $\mathbf{r}_T$, la première colonne de $\mathbf{R}_T$:*

$$\sqrt{T} (\mathbf{r}_T - \mathbf{r}(\Theta)) \xrightarrow{\mathcal{L}} \mathcal{N}_c(\mathbf{0}; \mathbf{C}_r, \mathbf{C}_r'). \quad (7.1)$$

avec

$$\begin{aligned} \mathbf{C}_r &= \int_{-1/2}^{+1/2} S_n^2(f) \mathbf{e}(f) \mathbf{e}^H(f) df + 2 \sum_{k=1}^K a_k^2 S_n(f_k) \mathbf{e}(f_k) \mathbf{e}^H(f_k) + \kappa_u \boldsymbol{\gamma} \boldsymbol{\gamma}^H, \\ \mathbf{C}_r' &= \int_{-1/2}^{+1/2} S_n^2(f) \mathbf{e}(f) \mathbf{e}^T(f) df + 2 \sum_{k=1}^K a_k^2 S_n(f_k) \mathbf{e}(f_k) \mathbf{e}^T(f_k) + \kappa_u \boldsymbol{\gamma} \boldsymbol{\gamma}^T, \end{aligned}$$

où $\boldsymbol{\gamma}$ est le $M \times 1$ vecteur $(\gamma_0, \dots, \gamma_Q, 0, \dots, 0)^H$ et $S_n(f)$, la densité spectrale de puissance de n_t .

6.7.2 L'algorithme AMV

Nous sommes ici dans le cadre d'observations complexes circulaires non gaussiennes ni indépendantes et, afin de satisfaire les conditions développées au paragraphe 6.4.2 pour obtenir un estimateur AMV de Θ , nous proposons de choisir la statistique $\mathbf{s}_T$, équivalente à $\mathbf{R}_T$, suivante :

$$\mathbf{s}_T = \begin{bmatrix} r_{0,T} \\ \mathbf{r}_T' \\ \mathbf{r}_T'^* \end{bmatrix}$$

où $\mathbf{r}_T = \begin{bmatrix} r_{0,T} \\ \mathbf{r}_T' \end{bmatrix}$ représente la 1ère colonne de $\mathbf{R}_T$ et $r_{0,T}$, le 1er terme de $\mathbf{r}_T$.

Notons qu'alors $\mathbf{s}_T^* = \mathbf{K}_p' \mathbf{s}_T$ où $\mathbf{K}_p'$ est la matrice de permutation $\begin{pmatrix} 1 & \mathbf{0}^T & \mathbf{0}^T \\ \mathbf{0} & \mathbf{O} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \mathbf{O} \end{pmatrix}$.

La distribution asymptotique de $\mathbf{s}_T$ est déduite de celle de $\mathbf{r}_T$ grâce au théorème de continuité [5, théorème p. 122]) :

$$\sqrt{T} (\mathbf{s}_T - \mathbf{s}(\Theta)) \xrightarrow{\mathcal{L}} \mathcal{N}_c(\mathbf{0}; \mathbf{C}_s(\Theta), \mathbf{C}_s'(\Theta))$$

avec

$$\mathbf{C}_s(\Theta) = \begin{pmatrix} c_r & \mathbf{c}_r^H & \mathbf{c}_r^T \\ \mathbf{c}_r & \mathbf{C}_{r'} & \mathbf{C}'_{r'} \\ \mathbf{c}_r^* & \mathbf{C}'_{r'}^* & \mathbf{C}_{r'}^* \end{pmatrix} \quad \text{et} \quad \mathbf{C}'_s(\Theta) = \mathbf{C}_s(\Theta) \mathbf{K}_p' = \begin{pmatrix} c_r & \mathbf{c}_r^T & \mathbf{c}_r^H \\ \mathbf{c}_r & \mathbf{C}'_{r'} & \mathbf{C}_{r'} \\ \mathbf{c}_r^* & \mathbf{C}_{r'}^* & \mathbf{C}'_{r'}^* \end{pmatrix} \quad (7.2)$$

où $\mathbf{C}_r$ and $\mathbf{C}'_r$ sont partitionnées de la manière suivante

$$\mathbf{C}_r = \begin{pmatrix} c_r & \mathbf{c}_r^H \\ \mathbf{c}_r & \mathbf{C}_{r'} \end{pmatrix} \quad \text{et} \quad \mathbf{C}'_r = \begin{pmatrix} c_r & \mathbf{c}_r^T \\ \mathbf{c}_r & \mathbf{C}'_{r'} \end{pmatrix}.$$

Notons enfin que $\mathbf{s}(\Theta)$ dépend aussi linéairement du paramètre de nuisance Θ_2 et peut s'écrire

$$\mathbf{s}(\Theta) = \Psi(\Theta_1) \Theta_2$$

ce qui nous permet de définir un estimateur de Θ_1 de type COMET selon l'expression (4.9) qui sera donc asymptotiquement de variance minimale à condition que $\mathbf{W}_T = \mathbf{C}_{s,T}^{-1}$ où $\mathbf{C}_{s,T}$ est une estimation consistante de $\mathbf{C}_s(\Theta)$.

6.7.3 Analyse des performances et de la robustesse

Nous nous proposons d'apprécier l'amélioration apportée par l'estimateur AMV par rapport aux performances des estimateurs MDP et MUSIC en comparant les variances asymptotiques respectives de ces estimateurs.

Covariance asymptotique de l'estimateur AMV

Le calcul des performances de l'algorithme AMV développé au paragraphe 6.5 s'appliquent ici. L'expression de la covariance asymptotique $\mathbf{C}_{\Theta_1}$ donnée par (5.1) s'applique avec ici $\mathbf{W} = \mathbf{C}_s^{-1}(\Theta)$ donné par (7.2).

Remarque 1 : Il est démontré en annexe 7 qu'un algorithme de type AMV, défini en utilisant comme matrice de pondération celle associée au modèle de bruit gaussien, est insensible à la nature de la distribution de probabilité du bruit.

Remarque 2 : La mise en œuvre pratique de cet estimateur nécessite d'une part de disposer d'un point de départ pour initialiser la recherche d'extremum, d'autre part, comme nous venons de le voir, de disposer d'une estimation consistante de $\mathbf{C}_s$. Ce double problème peut être résolu, comme nous l'avons fait, en utilisant les résultats d'une estimation préalable du paramètre Θ obtenus par une méthode de type décomposition de Pisarenko modifiée (MPD) telle que celle proposée par Sherman et Frazho dans [14] qui ne nécessite que l'estimation de la matrice $\mathbf{R}_T$. L'estimation préalable Θ_T^{MPD} de Θ ainsi obtenue étant consistante, il en va de même pour l'estimation de $\mathbf{C}_s(\Theta_T^{MPD})$ qui s'en déduit.

Le "domaine asymptotique" dans lequel la covariance de Θ_1 peut être approximée par $\frac{1}{T} \mathbf{C}_{\Theta_1}$ dépend d'un certain nombre de facteurs dont en particulier le rapport S/B. La FIG.6.7 illustre ce phénomène dans le cas d'une sinusoïde de fréquence $f_1 = 0.2$ dans un bruit MA d'ordre $Q = 4$ et une durée de snapshot $M = 8$. Les variances théoriques $\frac{1}{T} \mathbf{C}_{\Theta_1}$ et estimées (100 Monte Carlo et intervalles de confiance à 99%) y sont comparées en fonction du nombre de snapshots T pour différentes valeurs du rapport S/B. On constate que le nombre de snapshots correspondant au domaine asymptotique doit être d'autant plus important que le rapport S/B est élevé contrairement au cas classique $y_t = x_t + n_t$ où x_t est déterministe et $(n_t)_t$ est une suite de variables aléatoires indépendantes et identiquement distribuées où le "domaine asymptotique" est atteint pour T grand et/ou S/B élevé. Cette constatation a déjà été faite dans [15], dans le cas d'une observation constituée de sinusoïdes et d'un bruit blanc. Par ailleurs, en pointillé figure la borne $\frac{1}{T} (\mathbf{S}_1^H \mathbf{C}_s^{-1} \mathbf{S}_1)^{-1}$ mentionnée au paragraphe 6.4.2 ce qui permet de constater qu'elle est très voisine de $\frac{1}{T} \mathbf{C}_{\Theta_1}$.

Covariance asymptotique de l'estimateur MPD

La matrice de covariance asymptotique de l'estimateur MPD peut être obtenue en utilisant, comme nous l'avons déjà fait, une approche fonctionnelle consistant à considérer l'algorithme comme une relation fonctionnelle $\mathbf{R}_T \xrightarrow{MPD} \Theta_T$, prolongement de l'application $\mathbf{R}(\Theta) \rightarrow \Theta$. En utilisant un théorème classique de continuité on obtient :

$$\sqrt{T} (\Theta_T^{MPD} - \Theta) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{C}_\Theta^{MPD}) \quad (7.3)$$

avec

$$\mathbf{C}_\Theta^{MPD} = \mathbf{D}_{\Theta, \mathbf{R}}^{MPD} \mathbf{C}_R (\mathbf{D}_{\Theta, \mathbf{R}}^{MPD})^H$$

où $\mathbf{D}_{\Theta, \mathbf{R}}^{MPD}$ représente la dérivée de cet estimateur. Un rappel succinct du principe de l'algorithme MPD est présenté en Annexe 6 ainsi que le calcul de la différentielle $\mathbf{D}_{\Theta, \mathbf{R}}^{MPD}$ de cet algorithme.

La matrice de covariance asymptotique de l'estimateur MUSIC est quand à elle fournie dans [18].

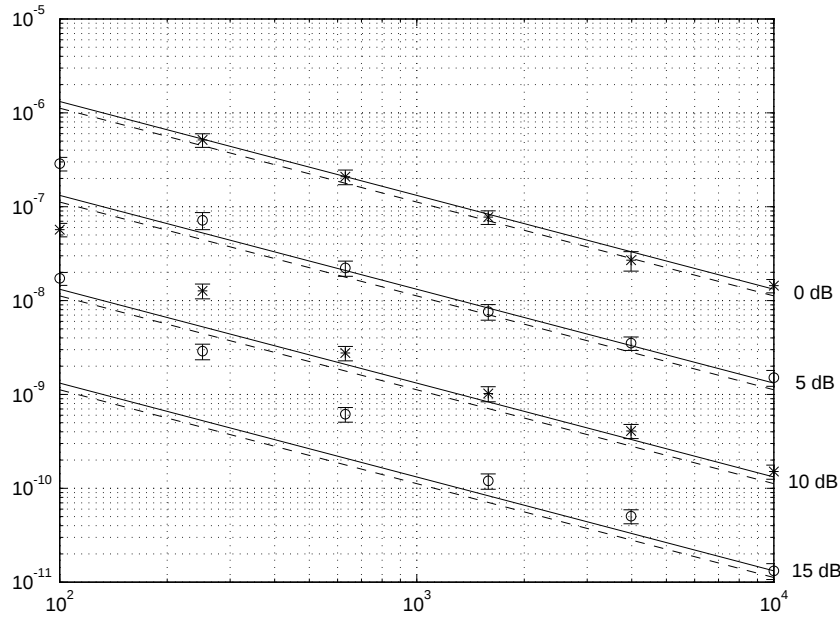


FIG. 6.7: Variance théorique et estimée pour l'estimateur AMV pour différentes valeurs du rapport S/B en fonction du nombre T de snapshots.

Comparaison des performances

La FIG.6.8 représente les variances asymptotiques et estimées (100 Monte Carlo et intervalles de confiance à 99%) des algorithmes AMV et MPD dans le cas de 2 sinusoïdes (de fréquence $f_1 = 0.1$ et $f_2 = 0.3$), d'un bruit MA d'ordre $Q = 1$, de coefficients MA, $b_0 = b_1 = 0.707$ et de $T = 2000$ snapshots en fonction du rapport S/B pour différentes valeurs de la durée de snapshot M . A noter que contrairement à l'approche AMV pour laquelle la contrainte sur M est $M \geq K + Q + 1$, l'estimateur MPD, utilisant une matrice de Hankel construite à partir des blocs triangulaires "nuls" de $\mathbf{R}_T$ impose un ordre fixé $M = 2K + Q + 2$. La comparaison entre les deux estimateurs doit donc être faite a priori pour la même valeur de M soit ici $M = 7$ et est bien sur en faveur de AMV. On constate toutefois que l'algorithme AMV autorise des ordres plus petits fournissant néanmoins de meilleures performances que l'algorithme MPD.

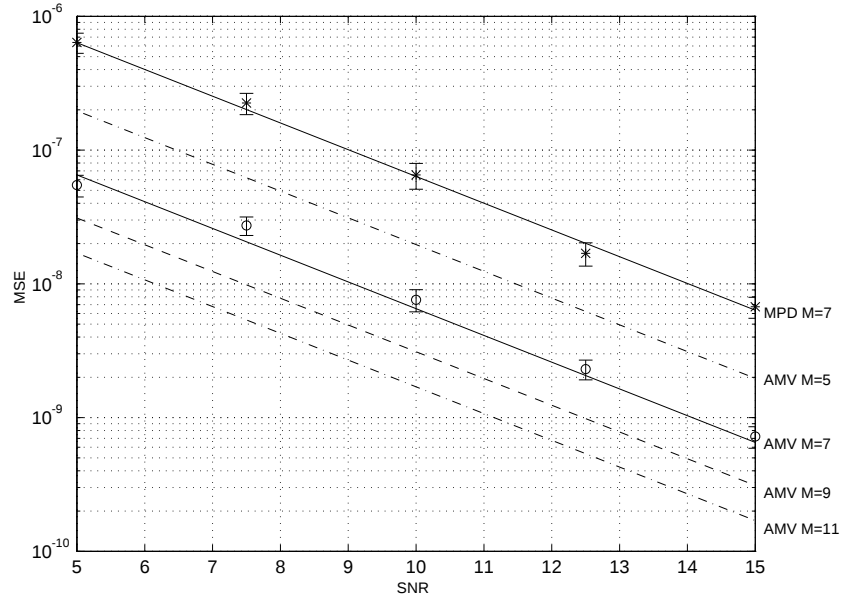


FIG. 6.8: Variance théorique et estimée pour les estimateurs MDP ($M = 7$) et AMV ($M = 5, 7, 9$ et 11) en fonction du rapport S/B.

La FIG.6.9 permet de comparer les performances asymptotiques des algorithmes AMV et MUSIC dans le cas où le bruit est blanc. Elle représente les variances théoriques $\frac{1}{T}\mathbf{C}_{\Theta_1}^{MUSIC}$ et $\frac{1}{T}\mathbf{C}_{\Theta_1}^{AMV}$ de ces estimateurs en fonction du rapport S/B lors de l'estimation de la fréquence d'une seule sinusoïde pour différentes valeurs de l'ordre M de la matrice de covariance empirique de l'observation. Nous pouvons constater que *l'estimateur MUSIC n'est pas de variance minimum* mais que les performances de ces deux algorithmes restent néanmoins assez voisines, l'écart étant pratiquement nul pour $M = 2$ et se creusant au fur et à mesure que l'ordre de la matrice de covariance de l'observation augmente.

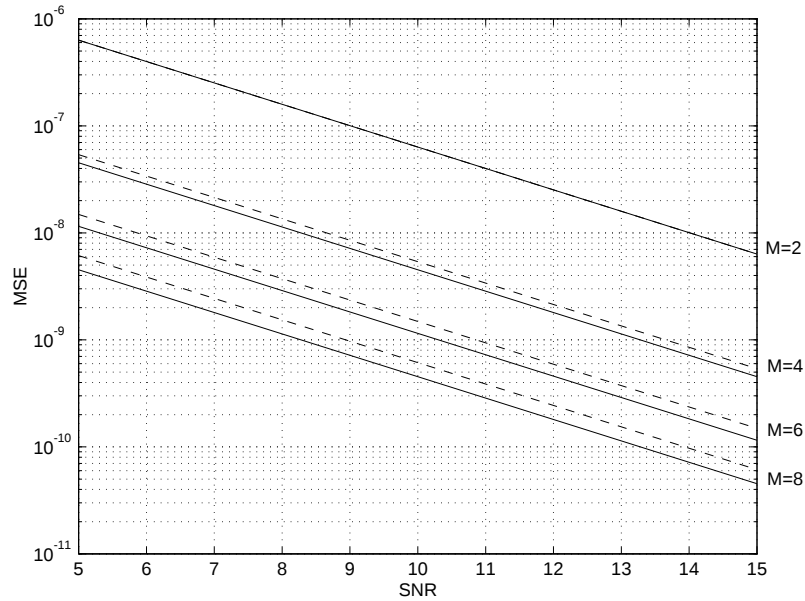


FIG. 6.9: Variance asymptotique des algorithmes AMV (-) et MUSIC (- -) en fonction du rapport S/B dans le cas d'une seule fréquence pour différentes valeurs de M .

L'étude suivante illustre l'influence de la valeur de la densité spectrale de puissance (DSP) à la fréquence de la sinusoïde sur l'estimation de cette dernière. Nous sommes dans la situation où $K = 1$, où le bruit est MA d'ordre 1 avec $b_0 = 0.707$ et $b_1 = -0.218 - 0.672i$ pour que cette DSP s'annule en $f = 0.2$ et où $M = 4$, $T = 2000$ et le rapport S/B = 5 dB. La FIG.6.10 représente en pointillé la forme de la DSP du bruit MA et en trait plein la DSP du bruit blanc de même puissance. La FIG.6.11 présente, en fonction de la fréquence f_1 de la sinusoïde, la variance asymptotique (1) et estimée (100 Monte Carlo et intervalles de confiance à 99%) de l'algorithme AMV en présence de bruit MA ; la variance asymptotique (3) de l'algorithme AMV ainsi que celle de l'algorithme MUSIC (2) en présence d'un bruit blanc de même puissance.

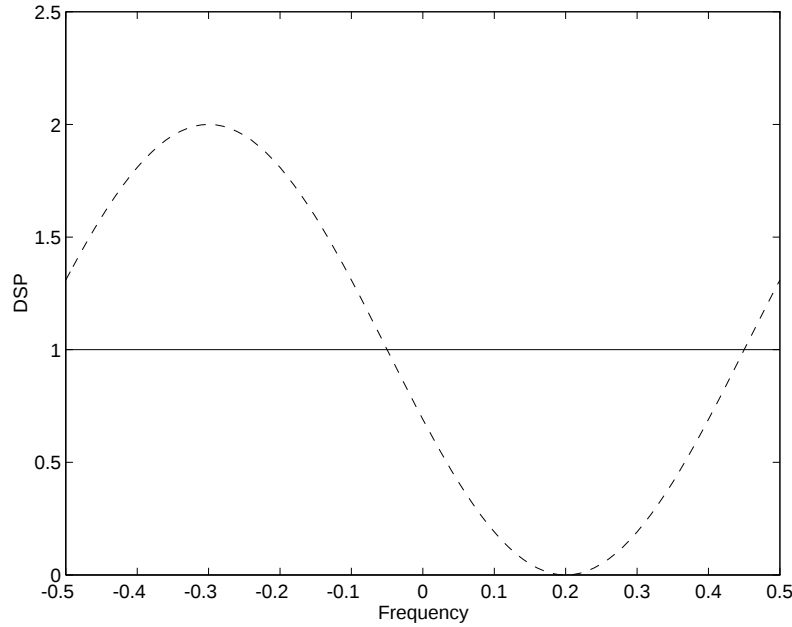


FIG. 6.10: Densités spectrales de puissance du bruit MA (- -) et du bruit blanc (-).

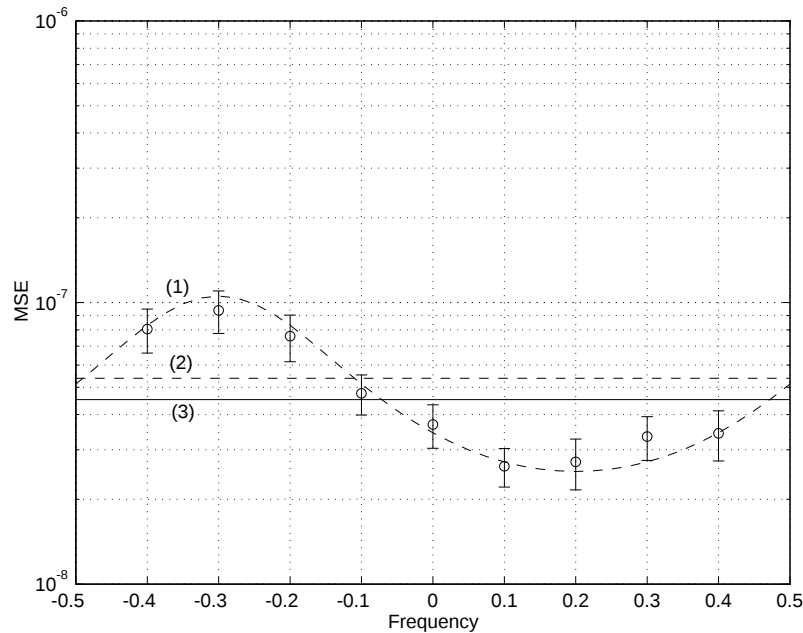


FIG. 6.11: Variance théorique et estimée pour l'estimateur AMV dans le cas d'un bruit MA (1) et pour les estimateurs AMV (3) et MUSIC (2) dans le cas d'un bruit blanc en fonction de la fréquence de la sinusoïde.

Nous constatons à nouveau que l'estimateur MUSIC n'est pas de variance minimum et que les performances de l'algorithme AMV se dégradent lorsque la fréquence f_1 est au voisinage du maximum de la DSP du bruit. Ce résultat est semblable aux performances de l'estimateur des moindres carrés non linéaire du modèle déterministe [19] pour lequel la variance asymptotique est proportionnelle à $S_n(f_1)/a_1^2$.

Sensibilité des algorithmes AMV et MPD à l'écart entre les fréquences

Afin d'apprécier la sensibilité des algorithmes AMV et MPD à l'écart entre les fréquences, nous considérons la situation suivante : le signal est constitué de 2 sinusoides de fréquence $f_1 = -0.1$ et $f_2 = f_1 + \Delta f$ avec Δf variant de 0.01 à 0.4 et d'un bruit MA d'ordre $Q = 1$ et de coefficients MA : $b_0 = b_1 = 0.707$. La matrice de covariance empirique est estimée à partir de $T = 2000$ snapshots de durée $M = 7$ et le rapport S/B = 10 dB. La FIG.6.12 représente les variances asymptotiques et estimées (100 réalisations et intervalles de confiance à 99%) de ces deux algorithmes. On peut constater, sur cette étude, que l'algorithme AMV est moins sensible à la proximité des fréquences que l'algorithme MPD. D'autre part, pour $\Delta f \leq 0.02$, une EQM estimée de l'ordre de $2 \cdot 10^{-3}$ traduit le fait que l'algorithme MPD ne fournit plus d'estimations correctes des fréquences. Or l'algorithme AMV utilise ces estimations d'une part pour évaluer la matrice de pondération $\mathbf{W}_T = \mathbf{C}_{s,T}^{-1}$ et d'autre part pour initialiser sa recherche. Ceci explique que pour $\Delta f \leq 0.02$, l'algorithme AMV soit également mis en échec.

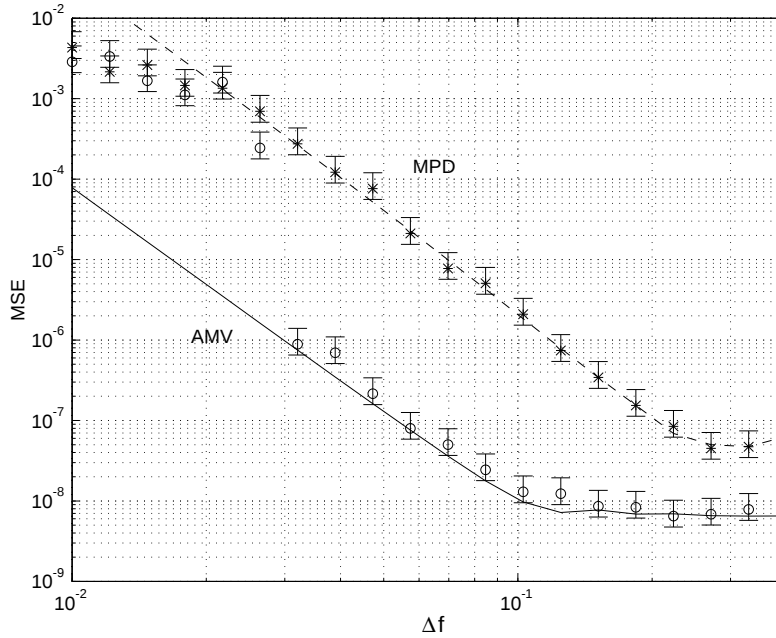


FIG. 6.12: Variance théorique et estimée pour les estimateurs AMV et MPD dans l'estimation de deux fréquences $f_1 = -0.1$ et $f_2 = f_1 + \Delta f$ avec un bruit MA d'ordre $Q = 1$ en fonction de l'écart Δf entre les deux fréquences.

Robustesse aux erreurs de modèle

L'algorithme AMV proposé ne s'applique théoriquement que dans le cas d'un bruit MA d'ordre connu. Il convient d'étudier sa robustesse aux erreurs de modèle. En fait, pour tout processus aléatoire purement non déterministe n_t , $\lim_{k \rightarrow \infty} \gamma_k = 0$. Il existe donc un entier Q tel que $\gamma_k \approx 0 \quad \forall k > Q$. La matrice de covariance de y_t devient alors $\mathbf{R}(\Theta) + \delta \mathbf{R}$ où $\delta \mathbf{R}$ est constituée des "petits" termes $\gamma_{Q+1}, \dots, \gamma_{M-1}$ résultants de l'erreur de modèle. L'estimateur $\Theta_{1,T}$ sera alors biaisé et le biais asymptotique sera donné par :

$$\mathbf{E}(\Theta_{1,T}) - \Theta_1 = \mathbf{D}_s^{\text{AMV}_1} \delta \mathbf{s} + O(\|\delta \mathbf{s}\|^2) + O\left(\frac{1}{T}\right), \quad (7.4)$$

où $\delta \mathbf{s} = (0, \dots, 0, \gamma_{Q+1}^*, \dots, \gamma_{M-1}^*, 0, \dots, 0, \gamma_{Q+1}, \dots, \gamma_{M-1})^T$ et où $\mathbf{D}_{\mathbf{s}}^{\text{AMV}_1}$ représente la dérivée de l'algorithme AMV. L'expression de cette dérivée est (cf. Annexe 3)

$$\mathbf{D}_{\mathbf{s}}^{\text{AMV}_1} = \frac{1}{2} [(\mathbf{S}^H \mathbf{C}_s^{-1}(\Theta) \mathbf{S})^{-1} \mathbf{S}^H \mathbf{C}_s^{-1}(\Theta)]_{(1:K,:)}.$$

La covariance asymptotique quant à elle est donnée par :

$$\text{Cov}(\Theta_{1,T}) = \frac{1}{T} \mathbf{D}_{\mathbf{s}+\delta \mathbf{s}}^{\text{AMV}_1} \mathbf{C}_{\mathbf{s}+\delta \mathbf{s}}(\Theta) \left(\mathbf{D}_{\mathbf{s}+\delta \mathbf{s}}^{\text{AMV}_1} \right)^H + O(\|\delta \mathbf{s}\|^4) + O\left(\frac{1}{T^2}\right) + O(\|\delta \mathbf{s}\|^2)O\left(\frac{1}{T}\right).$$

Pour illustrer ceci considérons la situation suivante : une sinusoïde de fréquence $f_1 = 0.2$ dans un bruit AR d'ordre 1 avec un rapport S/B = 5 dB, $T = 2000$ échantillons de taille $M = 5$. Le bruit est considéré à tort comme MA d'ordre 1 par les deux algorithmes.

La FIG.6.13 représente le biais asymptotique théorique

$$\mathbf{D}_{\mathbf{s}}^{\text{alg}} \delta \mathbf{s}$$

et le biais estimé par $I = 100$ réalisations indépendantes

$$\left(\frac{1}{I} \sum_{i=1}^I \Theta_{1,T}^{(i)} \right) - \Theta_1$$

pour les algorithmes AMV et MDP en fonction du coefficient du bruit AR. Nous pouvons constater que les valeurs fournies par les estimations sont en accord avec les valeurs asymptotiques théoriques sauf pour les faibles valeurs du coefficient AR. Pour expliquer ce phénomène, remarquons que le terme $O(\|\delta \mathbf{s}\|^2)$ du développement (7.4) diminue corrélativement avec le coefficient AR du bruit alors que le terme $O(\frac{1}{T})$ de ce développement, restant constant, devient prépondérant sur $O(\|\delta \mathbf{s}\|^2)$. La FIG.6.14 qui compare, dans la même situation, le biais asymptotique théorique $\mathbf{D}_{\mathbf{s}}^{\text{alg}} \delta \mathbf{s}$ au biais asymptotique au sens suivant

$$\lim_{T \rightarrow \infty} E(\Theta_{1,T}) - \Theta_1 = \text{alg}(\mathbf{R}(\Theta) + \delta \mathbf{R}),$$

corrobores cette explication.

Enfin la FIG.6.15 présente l'erreur quadratique moyenne asymptotique et estimée des estimateurs AMV et MDP dans la situation précédente en fonction du coefficient AR du bruit.

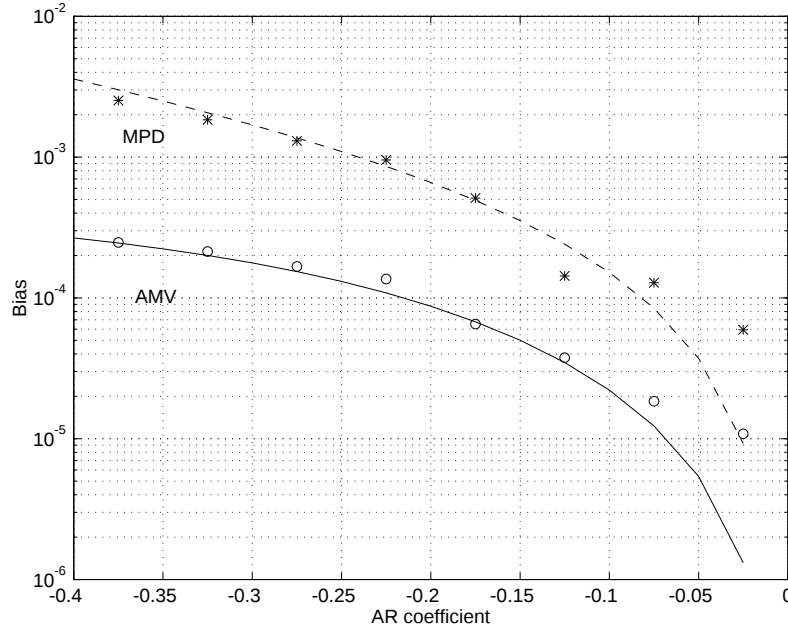


FIG. 6.13: Biais théorique et estimé (les algorithmes sont appliqués à la matrice de covariance empirique) pour les estimateurs AMV et MDP en fonction du coefficient AR du bruit.

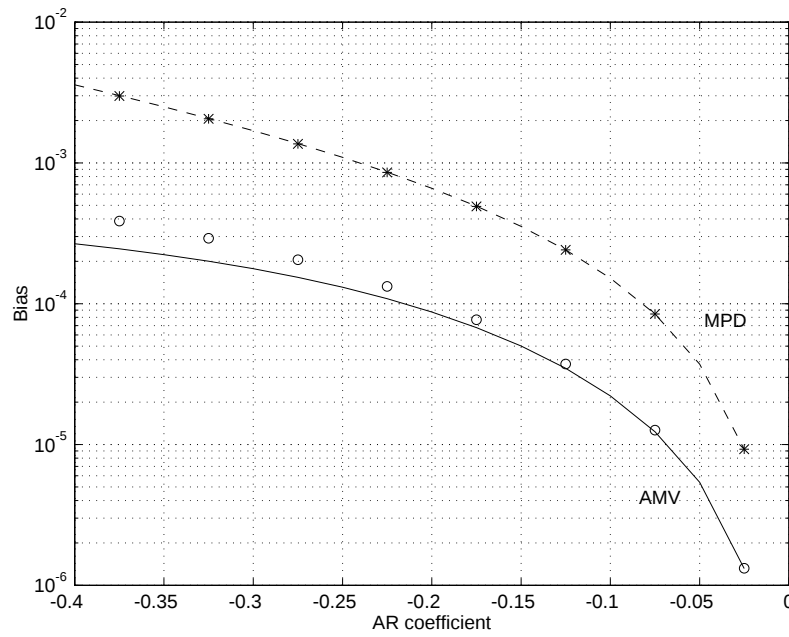


FIG. 6.14: Biais théorique et estimé (les algorithmes sont appliqués à la matrice de covariance exacte) pour les estimateurs AMV et MPD en fonction du coefficient AR du bruit.

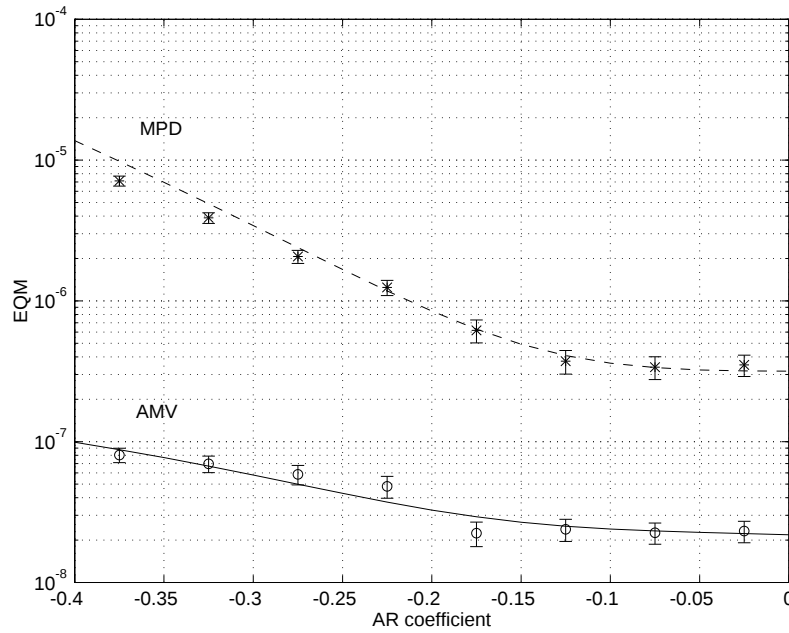


FIG. 6.15: Erreur quadratique moyenne théorique et estimée pour les estimateurs AMV et MPD en fonction du coefficient AR du bruit.

Ces figures montrent une très bonne robustesse de l'algorithme AMV à ce type d'erreur de modèle, comparativement meilleure que celle de l'algorithme MPD.

Performances d'algorithmes sous-optimaux

Dans le but de simplifier la mise en œuvre de l'algorithme AMV "optimal" et de l'algorithme de type COMET qui s'en déduit nous pouvons envisager trois types d'algorithmes sous-optimaux.

Le premier, que nous noterons AMV_r , est un algorithme de type moindres carrés non linéaires opérant

non plus sur $\mathbf{s}_T$ mais simplement sur $\mathbf{r}_T$ et utilisant comme matrice de pondération $\mathbf{W}_T = \mathbf{C}_{r,T}^{-1}$ où $\mathbf{C}_{r,T}$ est une estimation consistante de la matrice de covariance asymptotique de $\mathbf{r}_T$, c'est à dire l'algorithme

$$\Theta_{1,T} = \arg \min_{\alpha} (\mathbf{r}_T - \mathbf{r}(\alpha))^H \mathbf{C}_{r,T}^{-1} (\mathbf{r}_T - \mathbf{r}(\alpha)). \quad (7.5)$$

Cet algorithme semblerait être à première vue un algorithme AMV. En fait il est sous optimal car la démonstration donnée en 6.4.2 ne s'y applique pas.

Le second, que nous noterons AMV_s^g , opère comme l'algorithme AMV optimal sur $\mathbf{s}_T$ mais utilise comme matrice de pondération $\mathbf{W}_T$ celle correspondant au modèle d'observations complexes gaussiennes circulaires indépendantes qui peut être facilement déduite de $\mathbf{C}_{R,T} = \mathbf{R}_T \otimes_c \mathbf{R}_T$:

$$\mathbf{W}_T = \mathcal{C}_{s,T}^{-1} = \begin{pmatrix} r_{0,T}^2 & r_{0,T} \mathbf{r}'_T{}^H & r_{0,T} \mathbf{r}'_T{}^T \\ r_{0,T} \mathbf{r}'_T & r_{0,T} \mathbf{R}'_T & \mathbf{r}'_T \mathbf{r}'_T{}^T \\ r_{0,T} \mathbf{r}'_T{}^* & \mathbf{r}'_T{}^* \mathbf{r}'_T{}^H & r_{0,T} \mathbf{R}'_T{}^* \end{pmatrix}^{-1}$$

avec $\mathbf{R}_T$ partitionnée de la façon suivante :

$$\mathbf{R}_T = \begin{pmatrix} r_{0,T} & \mathbf{r}'_T{}^H \\ \mathbf{r}'_T & \mathbf{R}'_T \end{pmatrix}.$$

Le troisième enfin, noté AMV_r^g opère sur $\mathbf{r}_T$ et utilise comme précédemment la matrice de pondération correspondant au modèle d'observations complexes gaussiennes circulaires indépendantes

$$\mathbf{W}_T = \mathcal{C}_{r,T}^{-1} = (r_{0,T} \mathbf{R}_T)^{-1}$$

Les FIG.6.16 à FIG.6.19 permettent de comparer, en fonction du rapport S/B ou de l'ordre M de la matrice de covariance empirique, les performances de l'algorithme AMV optimal et celles de ces algorithmes sous-optimaux lorsque le signal comporte une ou deux sinusoïdes. Sur chacune de ces figures sont également présentées les performances des algorithmes de type moindres carrés non linéaires *sans* matrice de pondération, opérant sur $\mathbf{s}_T$ ou $\mathbf{r}_T$ et notés respectivement LS_s et LS_r ainsi que les performances de l'algorithme MPD.

Les FIG.6.16 et FIG.6.17 correspondent à la situation suivante : le signal est constitué d'une sinusoïde de fréquence $f_1 = 0.3$ et d'un bruit MA d'ordre $Q = 1$, de coefficients $b_0 = b_1 = 0.707$ et le nombre de snapshots est $T = 2000$. La variance asymptotique des algorithmes est représentée en FIG.6.16 en fonction du rapport S/B (en dB) pour $M = 5$ et en FIG.6.17 en fonction de M pour un rapport S/B = 5 dB.

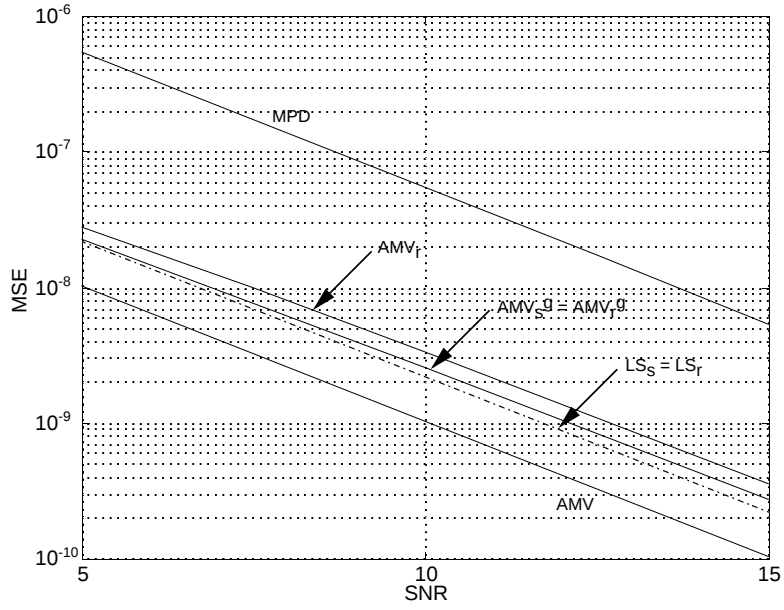


FIG. 6.16: Variance asymptotique de l'algorithme AMV optimal, des algorithmes AMV sous-optimaux AMV_r , AMV_r^g et AMV_s^g et des algorithmes LS_s , LS_r et MPD, en fonction du rapport S/B ($M = 5$).

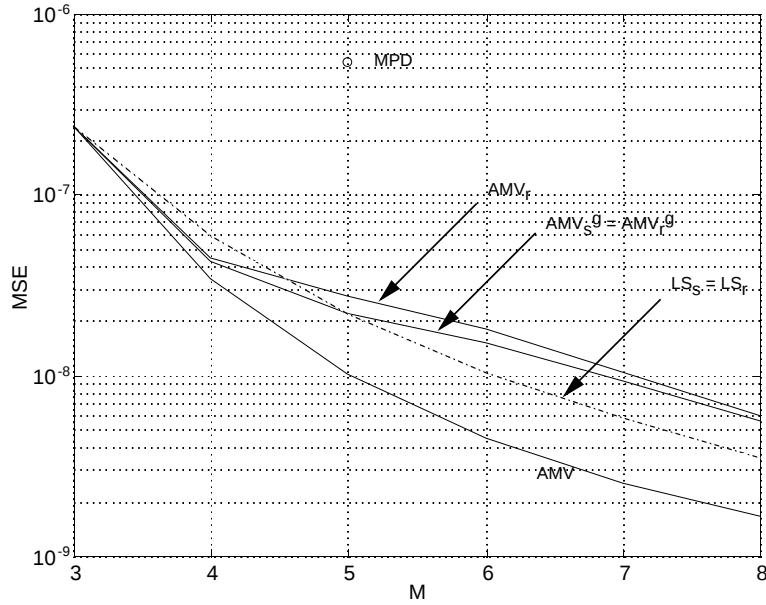


FIG. 6.17: Variance asymptotique de l'algorithme AMV optimal, des algorithmes AMV sous-optimaux AMV_r , AMV_r^g et AMV_s^g et des algorithmes LS_s , LS_r et MPD, en fonction de l'ordre M de la matrice de covariance empirique ($S/B = 5$ dB).

Les FIG.6.18 et FIG.6.19 correspondent à la situation suivante : le signal est constitué de deux sinusoides de fréquence $f_1 = 0.1$ et $f_2 = 0.25$ et d'un bruit MA d'ordre $Q = 1$, de coefficients $b_0 = b_1 = 0.707$ et le nombre de snapshots est $T = 2000$. La variance asymptotique des algorithmes est représentée en FIG.6.18 en fonction du rapport S/B (en dB) pour $M = 7$ et en FIG.6.19 en fonction de M pour un rapport S/B = 10 dB.

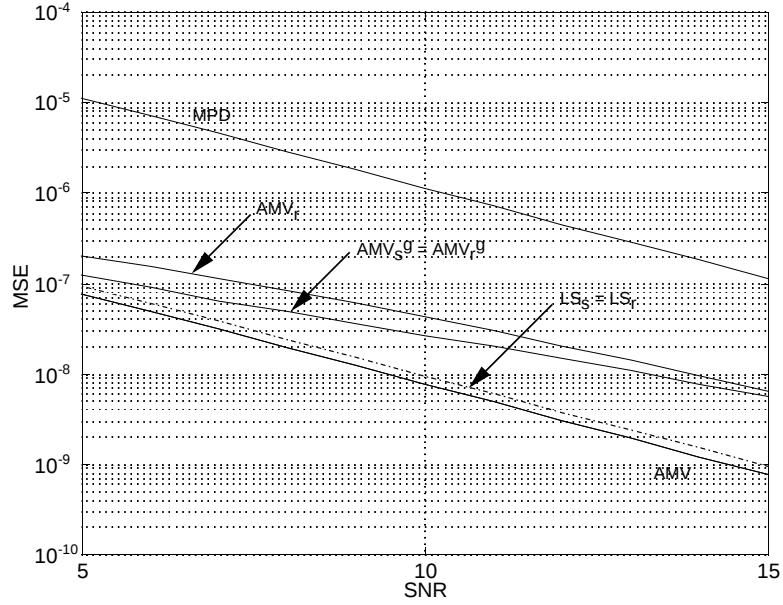


FIG. 6.18: Variance asymptotique de l'algorithme AMV optimal, des algorithmes AMV sous-optimaux AMV_r , AMV_r^g et AMV_s^g et des algorithmes LS_s , LS_r et MPD, en fonction du rapport S/B ($M = 7$).

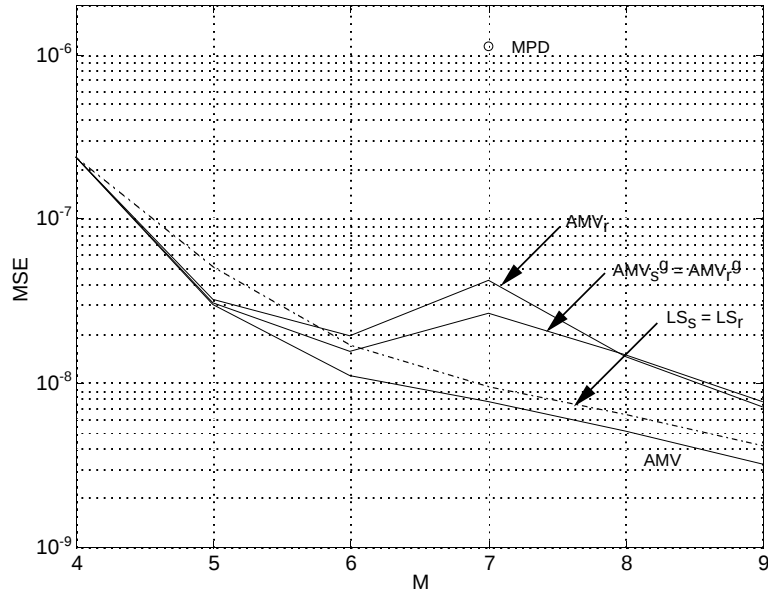


FIG. 6.19: Variance asymptotique de l'algorithme AMV optimal, des algorithmes AMV sous-optimaux AMV_r , AMV_r^g et AMV_s^g et des algorithmes LS_s , LS_r et MPD, en fonction de l'ordre M de la matrice de covariance empirique ($S/B = 10$ dB).

Les FIG.6.16 à FIG.6.19 permettent de constater :

- L'algorithme AMV_r (7.5) qui pouvait intuitivement être un algorithme AMV car il existe une bijection entre $\mathbf{r}_T$ et $\mathbf{R}_T$, n'est pas un algorithme AMV. En d'autres termes, pour passer du cas réel au cas complexe circulaire, il ne suffit pas d'adapter l'algorithme en remplaçant l'opération de transposition matricielle en l'opération de transposition hermitienne.

- Les algorithmes AMV_s^g et AMV_r^g , dont les performances apparaissent comme identiques, ne sont pas non plus, en général, des algorithmes AMV.
- Les performances des algorithmes LS_s et LS_r , apparaissant également comme identiques, sont assez proches de celles de l’algorithme AMV.
- Dans le cas particulier où $M = K + Q + 1$ (ordre minimum de $\mathbf{R}(\Theta)$ pour que Θ soit identifiable), tous ces algorithmes sont des algorithmes AMV. Ce résultat numérique, observé dans toutes les situations examinées, n’a cependant pas pu être démontré analytiquement.
En supposant cette propriété générale, cette valeur particulière de M permettrait donc d’atteindre les performances optimales au moyen d’algorithmes plus simples que l’algorithme AMV “optimal”, en particulier grâce à l’algorithme LS_r .

6.8 Conclusion

Dans la première partie de ce chapitre nous avons rappelé le principe de l’algorithme COMET original, présenté dans [1] dans le cadre d’observations complexes circulaires gaussiennes et indépendantes.

Après avoir étendu cet estimateur au cas d’observations réelles gaussiennes indépendantes, nous l’avons interprété comme un cas particulier, dans ces deux contextes, d’une classe plus générale d’estimateurs asymptotiquement de variance minimale.

Nous avons ensuite étendu l’étude des estimateurs AMV au cas d’observations non gaussiennes et/ou non indépendantes. En particulier, nous avons montré que certaines précautions devaient être prises quant au choix de la statistique à utiliser pour obtenir des estimateurs AMV dans le cas d’observations complexes circulaires non indépendantes.

Nous avons également donné, sous des hypothèses générales, une nouvelle expression de la matrice de covariance asymptotique des estimées fournies par ces algorithmes AMV qui permet notamment d’évaluer les performances d’algorithmes sous-optimaux déduits des algorithmes AMV.

Dans une seconde partie de ce chapitre, nous avons étudié la mise en œuvre et l’application de l’algorithme COMET classique au problème de l’estimation de DOA dans le cas de sources bande étroite ou large bande et d’un bruit gaussien, spatialement non corrélé et non uniforme sur les capteurs.

Nous avons montré que dans le cas large bande, COMET ne pouvait s’appliquer qu’à une source de spectre symétrique par rapport à la fréquence f_0 de démodulation.

Nous avons évalué les performances en DOA de l’algorithme COMET utilisant la matrice de pondération du modèle d’observations gaussiennes indépendantes.

- Dans le cas d’observations gaussiennes complexes circulaires indépendantes, là où l’algorithme COMET est efficace, nous avons donné, dans le cas d’une seule source large bande ou bande étroite, l’expression générale de la borne de Cramer-Rao de laquelle nous avons déduit, dans le cas bande étroite, une expression plus simple et interprétable.
- Dans le cas d’observations complexes circulaires non gaussiennes et/ou non indépendantes, et que l’algorithme COMET n’est ni efficace ni asymptotiquement de variance minimale, nous avons donné l’expression générale de la matrice de covariance asymptotique du paramètre directionnel seul en considérant notamment les cas particuliers d’une source large bande gaussienne et d’observations non indépendantes d’une part et d’une source large bande non gaussienne et d’observations indépendantes d’autre part.

Nous avons alors étudié la sensibilité de l’algorithme COMET à la distribution statistique et à la corrélation temporelle des signaux sources en rappelant et en illustrant numériquement certaines propriétés de l’algorithme COMET bande étroite établies au chapitre 4 et en montrant que, contrairement à l’algorithme COMET bande étroite, l’algorithme COMET large bande est sensible à la distribution de

la source.

En étudiant enfin l'effet de la défaillance d'un ou de plusieurs capteurs sur les performances de l'algorithme COMET, nous avons montré la très bonne robustesse de cet algorithme face à ce problème.

Dans la dernière partie de ce chapitre, nous avons étudié l'application d'un algorithme de type COMET modifié, asymptotiquement de variance minimale, déduit de notre étude théorique initiale, à l'estimation de fréquence de sinusoides dans un bruit additif MA.

Nous avons comparé les performances de cet algorithme AMV à celles d'un algorithme de type "Pisarenko étendu" (MPD) dans le cas d'un bruit additif MA et à celles de l'algorithme MUSIC dans le cas d'un bruit blanc additif en fonction du rapport S/B. Ces comparaisons sont bien sur en faveur de l'algorithme AMV dans les deux cas mais avec un très net avantage à l'algorithme AMV par rapport à l'algorithme MPD dans le premier cas.

Nous avons également montré que l'algorithme AMV est moins sensible à la proximité des fréquences à estimer et plus robuste aux erreurs de modèle (bruit AR supposé MA par les algorithmes) que l'algorithme MPD.

Nous avons enfin proposé un certain nombre de versions sous-optimales de l'algorithme AMV et comparé leurs performances à celles de l'algorithme original.

Annexe

Annexe 1 : Preuve de (2.9)

Pour montrer que $\tilde{\Theta}_{2,T}$ est réel remarquons que (2.9) peut également s'écrire

$$\tilde{\Theta}_{2,T} = [(\Psi^H(\Theta_1)\mathbf{J}^H)(\mathbf{J}^{-H}\mathbf{W}_T\mathbf{J}^{-1})(\mathbf{J}\Psi(\Theta_1))]^{-1}(\Psi^H(\Theta_1)\mathbf{J}^H)(\mathbf{J}^{-H}\mathbf{W}_T\mathbf{J}^{-1})(\mathbf{J}\mathbf{s}_T).$$

Or $\mathbf{J}\mathbf{s}_T = \mathbf{\Gamma}_T$ est réel de même que $\mathbf{J}\Psi(\Theta_1)$ car pour tout Θ_2 réel $\mathbf{J}\mathbf{s} = (\mathbf{J}\Psi(\Theta_1))\Theta_2$ est réel. Pour terminer, il reste à montrer que le terme $\mathbf{J}^{-H}\mathbf{W}_T\mathbf{J}^{-1}$, qui n'est autre que le 2ème membre de (2.7), est aussi réel. Cela se montrerait comme au paragraphe 6.2.2 avec cette fois $(z_t)_{t \in \mathbb{N}}$, une suite de variables aléatoires complexes centrées telle que $\mathbf{W}_T^{-1}$ soit la matrice de covariance asymptotique de la matrice de covariance empirique vectorisée de z_t , de loi de probabilité quelconque.

Annexe 2 : Preuve des résultats 6.2 et 6.3

Pour montrer, dans le cas réel, que l'estimateur

$$\hat{\Theta}_T = \arg \min_{\alpha \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\alpha))^T \mathbf{C}_s^{-1}(\alpha) (\mathbf{s}_T - \mathbf{s}(\alpha))$$

est asymptotiquement de variance minimale nous allons commencer par évaluer la matrice de covariance asymptotique d'un estimateur plus général,

$$\Theta_T = \arg \min_{\alpha \in D_\Theta} V_T(\alpha) \stackrel{\text{def}}{=} \arg \min_{\alpha \in D_\Theta} (\mathbf{s}_T - \mathbf{s}(\alpha))^T \mathbf{W}(\alpha) (\mathbf{s}_T - \mathbf{s}(\alpha)) \quad (8.1)$$

de matrice de pondération $\mathbf{W}(\alpha)$ supposée simplement symétrique réelle définie positive, puis appliquer le résultat général obtenu au cas particulier où $\mathbf{W}(\alpha) = \mathbf{C}_s^{-1}(\alpha)$.

D'après le résultat 6.1, la matrice de covariance asymptotique de Θ_T est $\mathbf{C}_\Theta = \mathbf{D}_s \mathbf{C}_s \mathbf{D}_s^T$ où $\mathbf{D}_s$ désigne la matrice associée à la différentielle de l'estimateur (8.1).

Pour calculer $\mathbf{D}_s$, remarquons que $V_T(\alpha)$ est minimum en $\Theta_T = \Theta + \delta\Theta$ donc $\frac{\partial V_T(\alpha)}{\partial \alpha}|_{\Theta+\delta\Theta} = \mathbf{0}$ et par suite, $\forall i = 1, \dots, L$

$$\begin{aligned} - \frac{\partial \mathbf{s}^T(\alpha)}{\partial \alpha_i}|_{\Theta+\delta\Theta} \mathbf{W}(\Theta + \delta\Theta) (\mathbf{s}_T - \mathbf{s}(\Theta + \delta\Theta)) + (\mathbf{s}_T - \mathbf{s}(\Theta + \delta\Theta))^T \frac{\partial \mathbf{W}(\alpha)}{\partial \alpha_i}|_{\Theta+\delta\Theta} (\mathbf{s}_T - \mathbf{s}(\Theta + \delta\Theta)) \\ - (\mathbf{s}_T - \mathbf{s}(\Theta + \delta\Theta))^T \mathbf{W}(\Theta + \delta\Theta) \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i}|_{\Theta+\delta\Theta} = 0. \end{aligned} \quad (8.2)$$

Or

$$\begin{aligned} \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i}|_{\Theta+\delta\Theta} &= \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i}|_{\Theta} + \sum_{j=1}^L \frac{\partial^2 \mathbf{s}(\alpha)}{\partial \alpha_j \partial \alpha_i}|_{\Theta} \delta\theta_j + o(\delta\Theta) \\ \frac{\partial \mathbf{W}(\alpha)}{\partial \alpha_i}|_{\Theta+\delta\Theta} &= \frac{\partial \mathbf{W}(\alpha)}{\partial \alpha_i}|_{\Theta} + \sum_{j=1}^L \frac{\partial^2 \mathbf{W}(\alpha)}{\partial \alpha_j \partial \alpha_i}|_{\Theta} \delta\theta_j + o(\delta\Theta) \end{aligned}$$

$$\mathbf{s}_T = \mathbf{s}(\Theta) + \delta \mathbf{s} \quad \text{et} \quad \mathbf{s}(\Theta + \delta \Theta) = \mathbf{s}(\Theta) + \sum_{j=1}^L \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_j} \Big|_{\Theta} \delta \theta_j + o(\delta \Theta).$$

Compte tenu de ces expressions, (8.2) s'écrit alors, $\forall i = 1, \dots, L$,

$$\begin{aligned} \frac{\partial \mathbf{s}^T(\alpha)}{\partial \alpha_i} \Big|_{\Theta} \mathbf{W}(\Theta) \sum_{j=1}^L \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i} \Big|_{\Theta} \delta \theta_j + \sum_{j=1}^L \frac{\partial \mathbf{s}^T(\alpha)}{\partial \alpha_i} \Big|_{\Theta} \delta \theta_j \mathbf{W}(\Theta) \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i} \Big|_{\Theta} \\ = \frac{\partial \mathbf{s}^T(\alpha)}{\partial \alpha_i} \Big|_{\Theta} \mathbf{W}(\Theta) \delta \mathbf{s} + \delta \mathbf{s}^T \mathbf{W}(\alpha) \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i} \Big|_{\Theta} + o(\delta s). \end{aligned} \quad (8.3)$$

En posant $\left[\frac{\partial \mathbf{s}}{\partial \Theta} \right]_{k,l} \stackrel{\text{def}}{=} \frac{\partial s_k}{\partial \theta_l}$, (8.3) s'écrit sous forme matricielle

$$\left(\frac{\partial \mathbf{s}^T}{\partial \Theta} \mathbf{W}(\Theta) \frac{\partial \mathbf{s}}{\partial \Theta} \right) \delta \Theta = \frac{\partial \mathbf{s}^T}{\partial \Theta} \mathbf{W}(\Theta) \delta \mathbf{s} + o(\delta s).$$

En posant enfin $\mathbf{S} \stackrel{\text{def}}{=} \frac{\partial \mathbf{s}}{\partial \Theta}$, nous obtenons la différentielle de l'algorithme :

$$\mathbf{D}_s = (\mathbf{S}^T \mathbf{W}(\Theta) \mathbf{S})^{-1} \mathbf{S}^T \mathbf{W}(\Theta).$$

Remarque : Si on remplace la matrice de pondération $\mathbf{W}(\alpha)$ par une estimation consistante $\mathbf{W}_T$ de la matrice $\mathbf{W}(\Theta)$ au sens où $\mathbf{W}_T = \mathbf{W}(\Theta) + o(\mathbf{s}_T - \mathbf{s}(\Theta))$, l'expression (8.2) devient

$$-\frac{\partial \mathbf{s}^T(\alpha)}{\partial \alpha_i} \Big|_{\Theta + \delta \Theta} \mathbf{W}_T (\mathbf{s}_T - \mathbf{s}(\Theta + \delta \Theta)) - (\mathbf{s}_T - \mathbf{s}(\Theta + \delta \Theta))^T \mathbf{W}_T \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i} \Big|_{\Theta + \delta \Theta} = 0$$

dont le développement conduit comme précédemment à l'expression (8.3). Le calcul de la différentielle de ce nouvel algorithme conduit donc au même résultat que précédemment et les performances asymptotiques de ces deux algorithmes seront les mêmes.

Si maintenant on applique le résultat général précédent au cas particulier $\mathbf{W}(\Theta) = \mathbf{C}_s^{-1}(\Theta)$, il vient

$$\begin{aligned} \mathbf{C}_{\Theta} &= \mathbf{D}_s \mathbf{C}_s \mathbf{D}_s^T \\ &= (\mathbf{S}^T \mathbf{C}_s^{-1} \mathbf{S})^{-1} \mathbf{S}^T \mathbf{C}_s^{-1} \mathbf{C}_s \mathbf{C}_s^{-1} \mathbf{S} (\mathbf{S}^T \mathbf{C}_s^{-1} \mathbf{S})^{-1} \\ &= (\mathbf{S}^T \mathbf{C}_s^{-1} \mathbf{S})^{-1} \end{aligned}$$

ce qui montre que l'estimateur $\hat{\Theta}_T$ est asymptotiquement de variance minimale. D'autre part, compte tenu de la remarque précédente, l'estimateur $\hat{\Theta}_T$ obtenu en remplaçant $\mathbf{C}_s$ par une estimation consistante $\mathbf{C}_{s,T}$ a même matrice de covariance asymptotique que $\hat{\Theta}_T$ et est donc également asymptotiquement de variance minimale.

Annexe 3 : Preuve des résultats 6.4 et 6.5

Pour montrer, dans le cas complexe circulaire, que l'estimateur

$$\hat{\Theta}_T = \arg \min_{\alpha \in D_{\Theta}} (\mathbf{s}_T - \mathbf{s}(\alpha))^H \mathbf{C}_s^{-1}(\alpha) (\mathbf{s}_T - \mathbf{s}(\alpha))$$

est asymptotiquement de variance minimale nous commencerons, comme dans l'Annexe précédente, par évaluer la matrice de covariance asymptotique d'un estimateur plus général,

$$\Theta_T = \arg \min_{\alpha \in D_{\Theta}} V_T(\alpha) = \arg \min_{\alpha \in D_{\Theta}} (\mathbf{s}_T - \mathbf{s}(\alpha))^H \mathbf{W}(\alpha) (\mathbf{s}_T - \mathbf{s}(\alpha))$$

de matrice de pondération $\mathbf{W}(\alpha)$ supposée simplement hermitienne définie positive, puis nous appliquerons le résultat général obtenu au cas particulier où $\mathbf{W}(\alpha) = \mathbf{C}_s^{-1}(\alpha)$.

Si $\mathbf{s}_T = \text{Vec}(\mathbf{R}_T)$, d'après le corollaire 6.1,

$$\mathbf{C}_\Theta = 4\mathbf{D}_s \mathbf{C}_s(\Theta) \mathbf{D}_s^H$$

L'expression de $\mathbf{D}_s$ peut être obtenue d'une façon analogue à celle utilisée dans l'Annexe précédente. L'expression (8.3) devient, $\forall i = 1, \dots, L$,

$$\begin{aligned} \frac{\partial \mathbf{s}^H(\alpha)}{\partial \alpha_i} \Big|_\Theta \mathbf{W}(\Theta) \sum_{j=1}^L \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i} \Big|_\Theta \delta \theta_j &+ \sum_{j=1}^L \frac{\partial \mathbf{s}^H(\alpha)}{\partial \alpha_i} \Big|_\Theta \delta \theta_j \mathbf{W}(\Theta) \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i} \Big|_\Theta \\ &= \frac{\partial \mathbf{s}^H(\alpha)}{\partial \alpha_i} \Big|_\Theta \mathbf{W}(\Theta) \delta \mathbf{s} + \delta \mathbf{s}^H \mathbf{W}(\Theta) \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha_i} \Big|_\Theta + o(\delta \mathbf{s}) \end{aligned}$$

soit

$$\begin{aligned} \left(\frac{\partial \mathbf{s}^H(\alpha)}{\partial \alpha_i} \Big|_\Theta \mathbf{W}(\Theta) \frac{\partial \mathbf{s}(\alpha)}{\partial \alpha} \Big|_\Theta + \frac{\partial \mathbf{s}^T(\alpha)}{\partial \alpha_i} \Big|_\Theta \mathbf{W}^T(\Theta) \frac{\partial \mathbf{s}^*(\alpha)}{\partial \alpha} \right) \delta \Theta \\ = \frac{\partial \mathbf{s}^H(\alpha)}{\partial \alpha_i} \Big|_\Theta \mathbf{W}(\Theta) \delta \mathbf{s} + \frac{\partial \mathbf{s}^T(\alpha)}{\partial \alpha_i} \Big|_\Theta \mathbf{W}^T(\Theta) \delta \mathbf{s}^* + o(\delta \mathbf{s}). \end{aligned}$$

Soit encore, sous forme matricielle, en posant $[\mathbf{S}]_{k,l} \stackrel{\text{def}}{=} \frac{\partial \mathbf{s}_k}{\partial \theta_l}$

$$(\mathbf{S}^H \mathbf{W}(\Theta) \mathbf{S} + \mathbf{S}^T \mathbf{W}^T(\Theta) \mathbf{S}^*) \delta \Theta = [\mathbf{S}^H \mathbf{W}(\Theta), \mathbf{S}^T \mathbf{W}^T(\Theta)] \begin{bmatrix} \delta \mathbf{s} \\ \delta \mathbf{s}^* \end{bmatrix} + o(\delta \mathbf{s}). \quad (8.1)$$

Puisque $\mathbf{s}(\Theta) = \text{Vec}(\mathbf{R}(\Theta))$ et que $\mathbf{R}(\Theta)$ est hermitienne nous avons

$$\mathbf{s}^*(\Theta) = \mathbf{K}_M^c \mathbf{s}(\Theta) \quad \text{et} \quad \mathbf{s}^T(\Theta) = \mathbf{s}^H(\Theta) \mathbf{K}_M^c.$$

De même, puisque $\mathbf{s}_T = \text{Vec}(\mathbf{R}_T)$ et que $\mathbf{R}_T$ est hermitienne nous avons

$$\mathbf{s}_T^*(\Theta) = \mathbf{K}_M^c \mathbf{s}_T \quad \text{et} \quad \mathbf{s}_T^T = \mathbf{s}_T^H \mathbf{K}_M^c.$$

Nous en déduisons, d'une part que

$$\mathbf{S}^* = \mathbf{K}_M^c \mathbf{S} \quad \text{et} \quad \mathbf{S}^T = \mathbf{S}^H \mathbf{K}_M^c \quad (8.2)$$

et d'autre part que

$$\mathbf{C}_s^*(\Theta) = \mathbf{K}_M^c \mathbf{C}_s(\Theta) \mathbf{K}_M^c \quad \text{d'où} \quad (\mathbf{C}_s^*(\Theta))^{-1} = \mathbf{K}_M^c \mathbf{C}_s^{-1}(\Theta) \mathbf{K}_M^c. \quad (8.3)$$

Plus généralement d'ailleurs, pour toute matrice $\mathbf{W}$, structurée du point de vue parties réelles et imaginaires comme $\mathbf{C}_s^{-1}(\Theta)$, nous avons

$$\mathbf{W}^* = \mathbf{K}_M^c \mathbf{W} \mathbf{K}_M^c. \quad (8.4)$$

En conséquence, si la matrice $\mathbf{W}(\Theta)$ est structurée du point de vue parties réelles et imaginaires comme $\mathbf{C}_s^{-1}(\Theta)$, les propriétés (8.2) et (8.3) appliquées à (8.1) montrent que la différentielle de l'estimateur peut s'écrire

$$\mathbf{D}_s = \frac{1}{2} (\mathbf{S}^H \mathbf{W}(\Theta) \mathbf{S})^{-1} \mathbf{S}^H \mathbf{W}(\Theta). \quad (8.5)$$

En particulier, si $\mathbf{W}(\Theta) = \mathbf{C}_s^{-1}(\Theta)$, il vient

$$\begin{aligned}\mathbf{C}_\Theta &= 4 \mathbf{D}_s \mathbf{C}_s(\Theta) \mathbf{D}_s^H \\ &= (\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})^{-1} \mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{C}_s \mathbf{C}_s^{-1} \mathbf{S} (\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})^{-1} \\ &= (\mathbf{S}^H \mathbf{C}_s^{-1} \mathbf{S})^{-1}\end{aligned}$$

ce qui montre que l'estimateur $\hat{\Theta}_T$ est asymptotiquement de variance minimale.

Remarque : Comme dans le cas réel traité en annexe 2, si on remplace la matrice de pondération $\mathbf{W}(\alpha)$ par une estimation consistante $\mathbf{W}_T$ de la matrice $\mathbf{W}(\Theta)$ au sens où $\mathbf{W}_T = \mathbf{W}(\Theta) + o(\mathbf{s}^T - \mathbf{s}(\Theta))$, le calcul de la différentielle de ce nouvel algorithme conduit au même résultat (8.5) que précédemment et les performances asymptotique de ces algorithmes seront les mêmes. En conséquence, l'estimateur $\tilde{\Theta}_T$ obtenu en remplaçant $\mathbf{C}_s$ par une estimation consistante $\hat{\mathbf{C}}_{s,T}$ a même matrice de covariance asymptotique que $\hat{\Theta}_T$ et est donc également de variance minimale.

Annexe 4 : Preuve de (4.8)

Pour montrer que $\tilde{\Theta}_{2,T}$ est réel on peut emprunter la même démarche que celle suivie dans l'annexe 1 précédente. Comme $\mathbf{s}_T^* = \mathbf{K}_p' \mathbf{s}_T$, l'application qui à $\mathbf{s}_T$ associe $\mathbf{\Gamma}_T$, le vecteur constitué des parties réelles et imaginaires de $\mathbf{s}_T$ est linéaire et inversible. Il existe donc une matrice inversible $\mathbf{J}$ telle que $\mathbf{\Gamma}_T = \mathbf{J} \mathbf{s}_T$. $\tilde{\Theta}_{2,T}$ peut aussi s'écrire

$$\tilde{\Theta}_{2,T} = [(\Psi^H(\Theta_1) \mathbf{J}^H)(\mathbf{J}^{-H} \mathbf{W}_T \mathbf{J}^{-1})(\mathbf{J} \Psi(\Theta_1))]^{-1} (\Psi^H(\Theta_1) \mathbf{J}^H)(\mathbf{J}^{-H} \mathbf{W}_T \mathbf{J}^{-1})(\mathbf{J} \mathbf{s}_T).$$

où $\mathbf{J} \mathbf{s}_T$ et $\mathbf{J} \Psi(\Theta_1)$ sont réels. Il reste à montrer que $\mathbf{J} \mathbf{W}_T^{-1} \mathbf{J}^H = \mathbf{J} \mathbf{C}_{s,T} \mathbf{J}^H$ est réel. Or $\mathbf{C}_{s,T}$ est supposée structurée comme $\mathbf{C}_s$ (du point de vue partie réelle, partie imaginaire) c'est à dire comme $\mathbf{s}_T \mathbf{s}_T^H$ et $\mathbf{J} \mathbf{s}_T \mathbf{s}_T^H \mathbf{J}^H = \mathbf{\Gamma}_T \mathbf{\Gamma}_T^H$ est réel. Donc $\mathbf{J} \mathbf{W}_T^{-1} \mathbf{J}^H = \mathbf{J} \mathbf{C}_{s,T} \mathbf{J}^H$ est bien réel.

Annexe 5 : Calcul des différentielles des estimateurs (5.3) et (5.7)

Dans le cas d'observations complexes non gaussiennes et/ou non indépendantes, l'estimateur (5.3), de type COMET, de $\Theta_1 = (\theta_1, \dots, \theta_K)^T$ obtenu par une approche AMV est

$$\Theta_{1,T} = \arg \max_{\alpha} \mathbf{s}_T^H \mathbf{G}(\alpha) \mathbf{s}_T \quad (8.1)$$

avec $\mathbf{G}(\alpha) \stackrel{\text{def}}{=} \mathbf{C}_s^{-H}(\Theta) \Psi(\alpha) (\Psi^H(\alpha) \mathbf{C}_s^{-1}(\Theta) \Psi(\alpha))^{-1} \Psi^H(\alpha) \mathbf{C}_s^{-1}(\Theta)$.
D'après (8.1)

$$\frac{\partial}{\partial \alpha} (\mathbf{s}_T^H \mathbf{G}(\alpha) \mathbf{s}_T) |_{\Theta_{1,T}} = 0.$$

Donc, en notant $\Theta_{1,T} = \Theta_1 + \delta \Theta_1$ l'estimée associée à $\mathbf{s}_T = \mathbf{s} + \delta \mathbf{s}$

$$(\mathbf{s} + \delta \mathbf{s})^H \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} |_{\Theta_1 + \delta \Theta_1} (\mathbf{s} + \delta \mathbf{s}) = 0 \quad l = 1, \dots, K$$

D'autre part

$$\frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} |_{\Theta_1 + \delta \Theta_1} = \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} |_{\Theta_1} + \sum_{k=1}^L \frac{\partial^2 \mathbf{G}(\alpha)}{\partial \alpha_k \partial \alpha_l} |_{\Theta_1} \delta \theta_k + o(\delta \Theta_1) \quad l = 1, \dots, K.$$

Donc, compte tenu du fait que $\Theta_1 = \arg \max_{\alpha} \mathbf{s}^H \mathbf{G}(\alpha) \mathbf{s}$,

$$\mathbf{s}^H \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} \Big|_{\Theta_1} \mathbf{s} = 0 \quad l = 1, \dots, K,$$

nous avons

$$\begin{aligned} \left(\mathbf{s}^H \sum_{k=1}^L \frac{\partial^2 \mathbf{G}(\alpha)}{\partial \alpha_k \partial \alpha_l} \Big|_{\Theta_1} \mathbf{s} \right) \delta \theta_k &= -\mathbf{s}^H \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} \Big|_{\Theta_1} \delta \mathbf{s} - \delta \mathbf{s}^H \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} \Big|_{\Theta_1} \mathbf{s} \\ &= -\mathbf{s}^H \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} \Big|_{\Theta_1} \delta \mathbf{s} - \mathbf{s}^T \frac{\partial \mathbf{G}^T(\alpha)}{\partial \alpha_l} \Big|_{\Theta_1} \mathbf{K}'_p \delta \mathbf{s} \quad l = 1, \dots, K, \end{aligned} \quad (8.2)$$

soit, sous forme matricielle,

$$\mathcal{F}_1 \delta \Theta_1 = -\mathcal{F}_2 \delta \mathbf{s}$$

avec

$$\mathcal{F}_1(k, l) \stackrel{\text{def}}{=} \mathbf{s}^H \frac{\partial^2 \mathbf{G}(\Theta_1)}{\partial \theta_k \partial \theta_l} \mathbf{s}$$

et

$$\mathcal{F}_2(k, :) \stackrel{\text{def}}{=} \mathbf{s}^H \frac{\partial \mathbf{G}(\Theta_1)}{\partial \theta_k} + \mathbf{s}^T \frac{\partial \mathbf{G}^T(\Theta_1)}{\partial \theta_k} \mathbf{K}'_p \quad k, l = 1, \dots, K.$$

Dans le cas d'observations réelles, l'adaptation du raisonnement et du calcul précédent est immédiate. La relation (8.2) devient

$$\begin{aligned} \left(\mathbf{r}^T \sum_{k=1}^L \frac{\partial^2 \mathbf{G}(\alpha)}{\partial \alpha_k \partial \alpha_l} \Big|_{\Theta_1} \mathbf{r} \right) \delta \theta_k &= -\mathbf{r}^T \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} \Big|_{\Theta_1} \delta \mathbf{r} - \delta \mathbf{r}^T \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} \Big|_{\Theta_1} \mathbf{r} \\ &= -2\mathbf{r}^T \frac{\partial \mathbf{G}(\alpha)}{\partial \alpha_l} \Big|_{\Theta_1} \delta \mathbf{r}. \end{aligned}$$

Soit, sous forme matricielle

$$\mathcal{F}_1 \delta \Theta_1 = -\mathcal{F}_2 \delta \mathbf{r}$$

avec

$$\mathcal{F}_1(k, l) \stackrel{\text{def}}{=} \mathbf{r}^T \frac{\partial^2 \mathbf{G}(\Theta_1)}{\partial \theta_k \partial \theta_l} \mathbf{r}$$

et

$$\mathcal{F}_2(k, :) \stackrel{\text{def}}{=} 2 \mathbf{r}^T \frac{\partial \mathbf{G}(\Theta_1)}{\partial \theta_k} \quad k, l = 1, \dots, K.$$

Annexe 6 : L'algorithme de Pisarenko modifié (MPD)

Principe

Avec les notations du paragraphe 6.7.1, la matrice de covariance de l'observation $\mathbf{y}_t \stackrel{\text{def}}{=} (y_t, \dots, y_{t-M+1})^T$ s'écrit $\mathbf{R}_y = \mathbf{R}_x + \mathbf{R}_n$. Cette matrice, de structure Hermitienne, Toeplitz, est engendrée par sa première colonne : $(r_0 + \gamma_0, r_1 + \gamma_1, \dots, r_Q + \gamma_Q, r_{Q+1}, \dots, r_{M-1})^H$ où,

$$\forall l = 0, \dots, M-1 \quad r_l \stackrel{\text{def}}{=} \text{E}[x_t x_{t-l}^*] = \sum_{k=1}^K a_k^2 e^{i2\pi f_k l},$$

terme que l'on peut également écrire

$$r_l = \mathbf{a}^T \Delta^l \mathbf{a} \quad \text{avec} \quad \mathbf{a} \stackrel{\text{def}}{=} (a_1, \dots, a_K)^T \quad \text{et} \quad \Delta = \text{Diag}(e^{i2\pi f_1}, \dots, e^{i2\pi f_K})$$

et $\gamma_q \stackrel{\text{def}}{=} E[n_t n_{t-q}^*]$, $q = 0, \dots, Q$.

Les $M - Q - 1$ derniers termes de cette première colonne ne dépendent donc que des $\{a_k\}_{k=1, \dots, K}$ et des $\{f_k\}_{k=1, \dots, K}$. Pour extraire ces paramètres et en particulier les $\{f_k\}$ de ces termes, choisissons $M = Q + 2K + 2$ et appelons $\mathbf{H}$ la matrice de Hankel construite à partir de ces termes de la façon suivante :

$$\mathbf{H} = \begin{pmatrix} r_{Q+1} & r_{Q+2} & \cdots & r_{Q+K+1} \\ r_{Q+2} & r_{Q+3} & \cdots & r_{Q+K+2} \\ \vdots & \ddots & \ddots & \vdots \\ r_{Q+K+1} & r_{Q+K+2} & \cdots & r_{Q+2K+1} \end{pmatrix}.$$

L'estimation des $\{f_k\}$ va reposer sur le théorème suivant :

Théorème 2 *Les fréquences $\{f_k\}_{k=1, \dots, K}$ étant supposées deux à deux distinctes, $\{e^{i2\pi f_1}, \dots, e^{i2\pi f_K}\}$ sont les racines du polynôme $\mathcal{C}_K(z) = \sum_{l=0}^K c_l z^l$ si et seulement si $\mathbf{H} \mathbf{c} = \mathbf{0}$ où $\mathbf{c} \stackrel{\text{def}}{=} [c_0, \dots, c_K]^T$.*

Preuve : Supposons que $\sum_{l=0}^K c_l e^{i2\pi l f_k} = 0$ pour $k = 1, \dots, K$, alors $\sum_{l=0}^K c_l \Delta^l = \mathbf{0}$. En conséquence,

$$\mathbf{a}^T \left(\sum_{l=0}^K c_l \Delta^l \right) \Delta^{Q+k+1} \mathbf{a} = 0, \quad \forall k = 0, \dots, K,$$

donc

$$\sum_{l=0}^K \left(\mathbf{a}^T \Delta^{Q+k+1+l} \mathbf{a} \right) c_l = \sum_{l=0}^K r_{Q+1+k+l} c_l = 0, \quad \forall k = 0, \dots, K.$$

Donc $\mathbf{H} \mathbf{c} = \mathbf{0}$.

Réciproquement, supposons maintenant que $\mathbf{H} \mathbf{c} = \mathbf{0}$. Alors en particulier

$$\left(\mathbf{a}^T \left(\sum_{l=0}^K c_l \Delta^l \right) \right) \Delta^{Q+k+1} \mathbf{a} = 0 \quad \forall k = 0, \dots, K-1. \quad (8.1)$$

Or, comme $a_i \neq 0$, $i = 1, \dots, K$, $\{\Delta^{Q+k+1} \mathbf{a}\}_{k=0, \dots, K-1}$ constitue une base de $\mathbb{C}^K$. En effet, la matrice

$$[\Delta^{Q+1} \mathbf{a}, \dots, \Delta^{Q+K} \mathbf{a}] = \Delta^{Q+1} [\mathbf{Ia}, \Delta \mathbf{a}, \dots, \Delta^{K-1} \mathbf{a}] = \Delta^{Q+1} \text{Diag}(a_1, \dots, a_K) \mathbf{V}$$

où $\mathbf{V}$ est une matrice de Vandermonde $K \times K$, de terme général $[\mathbf{V}]_{m,n} = e^{i2\pi(n-1)f_m}$, $m, n = 1, \dots, K$, est régulière.

Donc (8.1) implique que $\mathbf{a}^T \left(\sum_{l=0}^K c_l \Delta^l \right) = \mathbf{0}^T$, soit $\sum_{l=0}^K c_l \Delta^l = \mathbf{0}$. ■

Mise en œuvre

Dans la pratique, la mise en œuvre de l'algorithme MPD peut être faite de la façon suivante :

1. A partir de la matrice de covariance empirique d'ordre $M = Q + 2K + 2$ de l'observation $\mathbf{R}_T \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^H$, on forme la matrice de Hankel $\mathbf{H}_T$ par lissage diagonal ⁵

$$r_{k,T} = \frac{1}{Q + 2K + 2 - k} \sum_{i=1}^{Q+2K+2-k} [\mathbf{R}_T]_{i,i+k}, \quad k = Q + 1, \dots, Q + 2K + 1.$$

2. La matrice $\mathbf{H}_T$ ainsi formée n'ayant a priori aucune raison d'être de rang déficient, on cherche par exemple le vecteur $\mathbf{c}_T$ estimé de $\mathbf{c}$ comme solution de

$$\min_{\mathbf{c}} \frac{\|\mathbf{H}_T \mathbf{c}\|}{\|\mathbf{c}\|}$$

soit le vecteur singulier à droite associé à la plus petite valeur singulière de $\mathbf{H}_T$.

3. On résout l'équation $\mathcal{C}_{K,T}(z) = 0$ où $\mathcal{C}_{K,T}(z)$ est le polynôme qui a pour coefficients, les composantes de $\mathbf{c}_T$.
4. Des racines $\mathbf{z}_T \stackrel{\text{def}}{=} (z_{1,T}, \dots, z_{K,T})^T$ ainsi obtenues, on tire les estimées des fréquences $\Theta_{1,T} = (f_{1,T}, \dots, f_{K,T})^T$ par $f_{k,T} = \frac{1}{2\pi} \arg(z_{k,T}) \quad k = 1, \dots, K$.

Dérivée de l'algorithme MPD

Pour le calcul de la dérivée de l'algorithme MDP qui, à $\text{Vec}(\mathbf{R}_T)$, associe $\Theta_1(T) \stackrel{\text{def}}{=} (f_{1,T}, \dots, f_{K,T})^T$, on peut considérer cet algorithme comme composé d'une suite d'applications

$$\text{Vec}(\mathbf{R}_T) \xrightarrow{\mathcal{F}_1} \text{Vec}(\mathbf{R}_T^{to}) \xrightarrow{\mathcal{F}_2} \text{Vec}(\mathbf{H}_T) \xrightarrow{\mathcal{F}_3} \mathbf{c}_T \xrightarrow{\mathcal{F}_4} \mathbf{z}_T \xrightarrow{\mathcal{F}_5} \Theta_{1,T} = \frac{1}{2\pi} \arg(\mathbf{z}_T)$$

- L'application $\mathcal{F}_1$ qui à $\text{Vec}(\mathbf{R}_T) \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_y) + \text{Vec}(\delta \mathbf{R}_y)$ associe $\text{Vec}(\mathbf{R}_T^{to}) \stackrel{\text{def}}{=} \text{Vec}(\mathbf{R}_y^{to}) + \text{Vec}(\delta \mathbf{R}_y^{to})$ est linéaire et il lui est associée une matrice de "Toeplization" $\mathbf{K}^{to}$. Donc

$$\text{Vec}(\delta \mathbf{R}_y^{to}) = \mathbf{K}^{to} \text{Vec}(\delta \mathbf{R}_y).$$

- L'application $\mathcal{F}_2$ qui, à $\text{Vec}(\mathbf{R}_T^{to})$, associe $\text{Vec}(\mathbf{H}_T) \stackrel{\text{def}}{=} \text{Vec}(\mathbf{H}) + \text{Vec}(\delta \mathbf{H})$, le "vectorisé" de la matrice de Hankel construite à partir de $\mathbf{R}_T^{to}$ selon le principe présenté précédemment, est également linéaire et il lui est associé une matrice $\mathbf{K}_H$. Donc

$$\text{Vec}(\delta \mathbf{H}) = \mathbf{K}_H \text{Vec}(\delta \mathbf{R}_y^{to}) \quad \text{et} \quad \text{Vec}(\delta \mathbf{H}^*) = \mathbf{K}_H \mathbf{K}_M^c \text{Vec}(\delta \mathbf{R}_y^{to})$$

où $\mathbf{K}_M^c$ est une matrice de commutation (cf. Formulaire, relation (0.8)).

- Etant donné $\mathbf{c}$ le vecteur singulier à droite, de norme unité, associé à l'unique valeur singulière 0 de $\mathbf{H}$, l'application $\mathcal{F}_3$ associe à $\text{Vec}(\mathbf{H}_T)$ le vecteur singulier à droite $\mathbf{c}_T \stackrel{\text{def}}{=} \mathbf{c} + \delta \mathbf{c}$ associé à la plus petite valeur singulière de $\mathbf{H}_T$ telle que $\mathbf{c}^H \mathbf{c}_T = 1$ ⁶.

Pour calculer la dérivée de l'application $\mathcal{F}_3$ nous utiliserons un théorème démontré par Magnus [4, Theorem 8, p.161] dans le cadre des éléments propres d'une matrice. Selon ce théorème, si λ_0

⁵On aurait pu également commencer par faire un "lissage Toeplitz" de la matrice $\mathbf{R}_T$ puis construire la matrice $\mathbf{H}_T$ à partir des derniers termes de la première colonne de la matrice lissée $\mathbf{R}_T^{to}$ ainsi obtenue.

⁶Le choix de cette normalisation du vecteur $\mathbf{c}_T$, qui n'est pas la seule envisageable, permet, comme le montre Magnus [4, Theorem 8, p.161] une expression plus simple de la dérivée de l'application.

est une valeur propre simple d'une matrice quelconque $\mathbf{Z}_0 \in \mathbb{C}^{n \times n}$ et si $\mathbf{u}_0$ est un vecteur propre associé, alors les applications λ et $\mathbf{u}$ définies pour tout $\mathbf{Z}$ dans un voisinage $N(\mathbf{Z}_0)$ de $\mathbf{Z}_0$, telles que

$$\lambda(\mathbf{Z}_0) = \lambda_0, \quad \mathbf{u}(\mathbf{Z}_0) = \mathbf{u}_0, \quad \mathbf{Z}\mathbf{u} = \lambda\mathbf{u} \quad \text{avec} \quad \mathbf{u}_0^H \mathbf{u} = 1$$

sont indéfiniment dérivables sur $N(\mathbf{Z}_0)$, les dérivées en $\mathbf{Z}_0$ étant

$$d\lambda = \frac{\mathbf{v}_0^H d\mathbf{Z} \mathbf{u}_0}{\mathbf{v}_0^H \mathbf{u}_0} \quad \text{et} \quad d\mathbf{u} = (\lambda_0 \mathbf{I} - \mathbf{Z}_0)^\# \left(\mathbf{I} - \frac{\mathbf{u}_0 \mathbf{v}_0^H}{\mathbf{v}_0^H \mathbf{u}_0} \right) d\mathbf{Z} \mathbf{u}_0,$$

où $\mathbf{v}_0$ est le vecteur propre associée à la valeur propre λ_0^* de $\mathbf{Z}_0^H$.

Notre problème concernant le vecteur singulier à droite associé à la valeur singulière 0 de $\mathbf{H}$, nous appliquons le théorème précédent à $\mathbf{Z}_0 = \mathbf{H}^H \mathbf{H}$, $\lambda_0 = 0$ et $\mathbf{u}_0 = \mathbf{v}_0 = \mathbf{c}$. Il vient alors :

$$\delta \mathbf{c} = -(\mathbf{H}^H \mathbf{H})^\# (\mathbf{I}_{K+1} - \mathbf{c} \mathbf{c}^H) (\delta \mathbf{H}^H \mathbf{H} + \mathbf{H}^H \delta \mathbf{H}) \mathbf{c} + o(\delta \mathbf{H}).$$

Or, grâce à (0.2) et (0.8),

$$\begin{aligned} \delta \mathbf{H}^H \mathbf{H} \mathbf{c} = \text{Vec}(\mathbf{I}_{K+1} \delta \mathbf{H}^H \mathbf{c}) &= (\mathbf{I}_{K+1} \otimes_c \mathbf{c}^H \mathbf{H}^H) \text{Vec}(\delta \mathbf{H}^H) \\ &= (\mathbf{I}_{K+1} \otimes_c \mathbf{c}^H \mathbf{H}^H) \mathbf{K}_{K+1}^c \text{Vec}(\delta \mathbf{H}^*) \end{aligned}$$

et

$$\mathbf{H}^H \delta \mathbf{H} \mathbf{c} = (\mathbf{H}^H \otimes_c \mathbf{c}^H) \text{Vec}(\delta \mathbf{H}).$$

Donc

$$\delta \mathbf{c} = \mathbf{A}_1 \text{Vec}(\delta \mathbf{H}) + \mathbf{A}_2 \text{Vec}(\delta \mathbf{H}^*) + o(\delta \mathbf{H})$$

avec

$$\begin{aligned} \mathbf{A}_1 &= -(\mathbf{H}^H \mathbf{H})^\# (\mathbf{I}_{K+1} - \mathbf{c} \mathbf{c}^H) (\mathbf{H}^H \otimes_c \mathbf{c}_0^H) \\ \mathbf{A}_2 &= -(\mathbf{H}^H \mathbf{H})^\# (\mathbf{I}_{K+1} - \mathbf{c} \mathbf{c}^H) (\mathbf{I}_{K+1} \otimes_c \mathbf{c}^H \mathbf{H}^H) \mathbf{K}_{K+1}^c. \end{aligned} \quad (8.2)$$

- A $\mathbf{c}_T$, l'application $\mathcal{F}_4$ fait correspondre $\mathbf{z}_T \stackrel{\text{def}}{=} \mathbf{z} + \delta \mathbf{z}$ où $\mathbf{z} = (z_1, \dots, z_K)^T$ est le vecteur dont les composantes sont les racines du polynôme $\mathcal{C}_K(z)$ dont les coefficients $c_0, \dots, c_K$ sont les composantes de $\mathbf{c}$.

Pour tout $k = 1, \dots, K$, $\sum_{n=0}^K (c_n + \delta c_n)(z_k + \delta z_k)^n = 0$. En développant cette relation et en ne retenant que les termes du premier ordre il vient

$$\sum_{n=0}^K \delta c_n z_k^n = -\mathcal{C}'_K(z_k) \delta z_k + o(\delta z_k)$$

où $\mathcal{C}'_K(z) \stackrel{\text{def}}{=} \frac{\partial \mathcal{C}_K(z)}{\partial z}$. Ceci, exprimé sous forme matricielle, nous donne

$$[\mathbf{1}, \mathbf{z}, \dots, \mathbf{z}^K] \delta \mathbf{c} = - \left(\text{Diag}(\mathcal{C}'_K(z_1), \dots, \mathcal{C}'_K(z_K)) \right) \delta \mathbf{z} + o(\delta \mathbf{z}).$$

Or $\mathcal{F}_4$ est une application holomorphe de $\mathbb{C}^{K+1}$ dans $\mathbb{C}^K$. Elle est donc différentiable dans $\mathbb{C}^{K+1}$ et par suite

$$\delta \mathbf{z} = \mathbf{D}_z \delta \mathbf{c} + o(\delta \mathbf{c})$$

avec

$$\mathbf{D}_z = - \left(\text{Diag}(\mathcal{C}'_K(z_1), \dots, \mathcal{C}'_K(z_K)) \right)^{-1} [\mathbf{1}, \mathbf{z}, \dots, \mathbf{z}^K].$$

- En posant $\Theta_{1,T} = \Theta_1 + \delta\Theta_1$, un calcul de perturbation effectué sur l'application $\mathcal{F}_5$ montre que :

$$\delta\Theta_1 = \frac{1}{4i\pi} (\mathbf{D}_A^* \delta\mathbf{z} - \mathbf{D}_A \delta\mathbf{z}^*) + o(\delta\mathbf{z})$$

avec

$$\mathbf{D}_A \stackrel{\text{def}}{=} \text{Diag}\left(\frac{z_1}{|z_1|^2}, \dots, \frac{z_K}{|z_K|^2}\right) = \text{Diag}(z_1, \dots, z_K).$$

Grâce à ces différents résultats intermédiaires, en utilisant la règle de composition des applications qui se traduit par le produit des matrices associées aux différentielles correspondantes et en utilisant les notations précédentes il vient :

$$\begin{aligned} \delta\Theta_1 &= \frac{1}{4i\pi} [(\mathbf{D}_A^* \mathbf{D}_z \mathbf{A}_1 - \mathbf{D}_A \mathbf{D}_z^* \mathbf{A}_2^*) \mathbf{K}_H + (\mathbf{D}_A^* \mathbf{D}_z \mathbf{A}_2 - \mathbf{D}_A \mathbf{D}_z^* \mathbf{A}_1^*) \mathbf{K}_H \mathbf{K}_{K+1}^c] \mathbf{K}^{to} \text{Vec}(\delta\mathbf{R}_y) \\ &+ o(\delta\mathbf{R}_y). \end{aligned} \quad (8.3)$$

Annexe 7 :

Il est démontré dans [16] que les performances de tout algorithme du second ordre d'estimation des fréquences Θ_1 de sinusoides dans un bruit MA, régulier au sens des conditions 1 et 2b du paragraphe 4.4.1 du chapitre 4, ne dépendent pas de la nature de la distribution de probabilité du bruit.

Supposons maintenant le bruit non nécessairement gaussien dans notre problème. En adoptant une démarche similaire à celle qui nous a conduit à l'expression (5.6) mais en utilisant comme matrice de pondération, la matrice $\mathbf{W}_g \stackrel{\text{def}}{=} \mathbf{C}_s^g(\Theta)^{-1}$ du modèle erroné de bruit gaussien, ceci définit un algorithme noté AMV' (la matrice $\mathbf{C}_s^g(\Theta)$ est construite comme $\mathbf{C}_s(\Theta)$ dans (7.2) mais sans les 3ème termes de $\mathbf{C}_r$ et de $\mathbf{C}'_r$ qui font intervenir le cumulants du 4ème ordre de u_t).

La matrice de covariance asymptotique de $\Theta_{1,T} = \text{AMV}'(\mathbf{R}_T)$ est alors donnée par

$$\mathbf{C}_{\Theta_1}^{AMV'} = \mathbf{D}_{\Theta_1, \mathbf{s}}^{AMV'} \mathbf{C}_s(\Theta) \left(\mathbf{D}_{\Theta_1, \mathbf{s}}^{AMV'} \right)^H$$

où $\mathbf{D}_{\Theta_1, \mathbf{s}}^{AMV'}$ désigne la matrice associée à la différentielle de l'algorithme AMV' par rapport à Θ_1 au point $\mathbf{s}(\Theta)$.

Comme les performances de cet algorithme ne dépendent pas de la nature de la distribution de probabilité du bruit, nous avons aussi

$$\mathbf{C}_{\Theta_1}^{AMV'} = \mathbf{D}_{\Theta_1, \mathbf{s}}^{AMV'} \mathbf{C}_s^g(\Theta) \left(\mathbf{D}_{\Theta_1, \mathbf{s}}^{AMV'} \right)^H.$$

Or le deuxième membre de cette expression n'est autre que la matrice de covariance asymptotique de $\Theta_{1,T}$ dans le cas où le bruit est gaussien.

Bibliographie

- [1] B. Ottersten, P. Stoica and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing-A Review Journal*, vol. 8, no. 3, pp. 185-210, July 1998.
- [2] P. Stoica, T. Söderström, "On reparameterization of loss functions used in estimation and the invariance principle," *Signal Process.*, vol. 17, pp.383-387, Aug. 1989.
- [3] B. Porat, *Digital processing of random signals, Theory and Methods*, Prentice Hall, 1993.
- [4] J.R. Magnus, H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley and Sons, 1999.
- [5] R.J. Serfling, *Approximation theorems of mathematical statistics*, John Wiley and Sons, 1980.
- [6] B. Friedlander and B. Porat, "Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from high-order moments", *IEEE Trans. Automat. Control*, vol. 35, no. 1, pp. 27-35, January 1990.
- [7] S.M. Kay, *Fundamentals of statistical signal processing*, New York: Prentice Hall, 1993.
- [8] O. Besson, P. Stoica, "Decoupled estimation of DOA and angular spread for spatially distributed source," *IEEE Trans. on Signal Processing*, vol. 48, no. 7, pp. 1872-1882, June 2000.
- [9] A. Gershman, M. Pesavento, P. Stoica and E. Larsson, "The stochastic CRB for array processing in unknown noise fields," *Proc. ICASSP Salt Lake City*, May 2001.
- [10] J.F. Cardoso, E. Moulines, "A robustness property of DOA estimators based on covariance," *IEEE Trans. on Signal Processing*, vol. 42, no. 11, pp. 3285-3287, Nov. 1994.
- [11] J.P. Delmas, Y. Meurisse, "Asymptotic performance analysis of DOA algorithms with temporally correlated narrow-band signals," *IEEE Trans. on Signal Processing*, vol. 48, no. 9, pp. 2669-2674, Sept. 2000.
- [12] B. Friedlander, A.J. Weiss, "Direction finding for wide-band signals using an interpolated array," *IEEE Trans. on Signal Processing*, vol. 41, no. 4, pp. 1618-1634, April 1993.
- [13] A. Hanssen, T.A. Oigard, "The normal inverse Gaussian distribution : A versatile model for heavy-tailed stochastic processes," *Proc. ICASSP Salt Lake City*, May 2001.
- [14] P.J. Sherman, A.E. Frazho, "High resolution spectral estimation of sinusoids in colored noise using a modified Pisarenko decomposition," *Proc. ICASSP Tokyo*, pp. 181-184, April 1986.
- [15] A. Eriksson, P. Stoica, T. Söderström, "Second-order properties of MUSIC and ESPRIT estimates of sinusoidal frequencies in high SNR scenarios," *Radar and Signal Processing. IEE PROCEEDINGS-F*, vol. 140, no. 4, pp. 266-272, August 1993.

- [16] J.P. Delmas, "Asymptotic normality of sample covariance matrix for mixed spectra time series : Application to sinusoidal frequencies estimation," *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1681-1687, May 2001.
- [17] R. A. Horn and C. R. Johnson *Matrix Analysis*, Cambridge University Press.
- [18] P. Stoica, T. Söderström, "Statistical Analysis of MUSIC and Subspace Rotation Estimates of Sinusoidal Frequencies," *IEEE Trans. on Signal Processing*, vol. 39, no. 8, pp. 1836-1847, Aug. 1991.
- [19] P. Stoica, A. Jakobsson and J. Li, "Cissoid parameter estimation in the colored noise case: asymptotic Cramer-Rao bound, maximum likelihood, and nonlinear least-squares," *IEEE, Trans. Signal Processing*, vol. 45, no. 8, pp. 2048-2059, August 1997.