



Thèse

présentée pour obtenir le grade de docteur

de l'Ecole Nationale Supérieure
des Télécommunications

Spécialité : Signal et Images

JAOUHAR AYADI

Estimation et égalisation aveugles de canal pour
les communications mobiles

Soutenue le 28 septembre 1999 devant le jury composé de

Claude Gueguen

Président

Madeleine Bonnet

Rapporteurs

Pierre Comon

Jean-Jacques Fuchs

Examineurs

Meriem Jaidane

Dirk Slock

Directeur de Thèse

Ecole Nationale Supérieure des Télécommunications

A ma famille.

Remerciements

Je voudrais remercier le professeur Claude Gueguen, directeur du Groupe des Ecoles des Télécommunications, pour la confiance qu'il a bien voulu m'accorder en m'accueillant au sein de l'Institut Eurécom pour préparer ce travail de thèse, et qui m'a fait l'honneur d'accepter la présidence de ce jury de thèse.

Je tiens à remercier Madame Madeleine Bonnet, professeur à l'Université René Descartes, et Monsieur Pierre Comon, directeur de recherche au Centre National de la Recherche Scientifique (CNRS) pour l'intérêt qu'ils ont porté à ma thèse en acceptant d'en être les rapporteurs et pour les discussions fructueuses qui ont enrichi mon travail.

Je remercie aussi Madame Meriem Jaidane, professeur à l'École Nationale d'Ingénieurs de Tunis, et Monsieur Jean-Jacques Fuchs, professeur à l'Université de Rennes 1, pour avoir accepté d'examiner mon travail de thèse.

J'exprime ma profonde gratitude à mon directeur de thèse, le professeur Dirk Slock pour m'avoir encadré tout au long de ma thèse, pour m'avoir fait bénéficier de ses grandes connaissances scientifiques et pour m'avoir fait partager son enthousiasme infini pour la recherche. Je le remercie pour ses précieux conseils et sa rigueur scientifique. Qu'il trouve ici l'expression de ma reconnaissance pour la disponibilité, l'amitié et la confiance qu'il m'a accordées.

Je remercie vivement mes parents, mes frères et mes soeurs qui grâce à leurs encouragements, amour et affection ont participé à l'accomplissement de ce travail de thèse.

Enfin, je voudrais remercier tous mes collègues de travail pour leur gentillesse et leur disponibilité et tous ceux qui ont contribué, de près ou de loin, à l'avancement de mes travaux.

Résumé

Nous traitons le problème de l'estimation et de l'égalisation aveugles de canal dans le contexte des communications mobiles. Les méthodes d'identification que nous développons utilisent un modèle multi-canal pour décrire le canal de transmission entre l'émetteur et le récepteur. Nous commençons par la description des sous-espaces signal et bruit correspondant à l'estimation des paramètres du canal, et nous discutons des différentes paramétrisations du sous-espace bruit ainsi que leur relation avec l'engendrement de ce dernier. Nous présentons, ensuite, des méthodes d'estimation déterministes qui incluent les approches d'ajustement de sous-espaces pour lesquelles nous proposons des techniques alternatives qui évitent la décomposition propre de la matrice de covariance. Une autre classe intéressante de méthodes déterministes est celle qui est basée sur le critère de maximum de vraisemblance. Nous proposons deux algorithmes pour optimiser ce critère. Le premier algorithme permet de débruiter le critère considéré et par conséquent d'obtenir des estimées sans biais du canal. Pour ce qui est du deuxième algorithme, il se base sur une reformulation de la fonction de coût considérée comme un critère quadratique pondéré de façon à ce que son gradient corresponde exactement au gradient du critère de maximum de vraisemblance. Asymptotiquement, ces deux algorithmes sont globalement convergents et leur convergence nécessite quelques itérations. La performance du deuxième algorithme reste étroitement proche de la borne de Cramer-Rao.

Nous exploitons, ensuite, la connaissance a priori des filtres de transmission/réception. Nous étendons les méthodes aveugles pour incorporer cette connaissance et nous calculons les bornes de Cramer-Rao correspondant à ce cas. Nous constatons que les méthodes étendues aboutissent à des performances meilleures. En effet, les bornes de performance sont beaucoup plus basses que celles correspondant au cas purement aveugle, les critères à optimiser sont mieux conditionnés et des canaux qui ne sont pas identifiables par des méthodes aveugles peuvent être identifiés grâce à l'exploitation de cette connaissance.

Nous proposons deux procédures alternatives pour préserver la robustesse des méthodes d'estimation aveugle en absence de connaissance précise sur la longueur du canal. En fait, la première approche peut être vue comme un mécanisme de détection implicite de l'ordre du canal, par contre la seconde approche détecte cet ordre de façon explicite.

Nous développons une méthode d'ajustement de sous-espace pour identifier le canal lorsque le bruit additif n'est plus supposé blanc mais uniquement stationnaire de couleur inconnue. La méthode proposée exploite une structure de déplacement de la matrice de covariance du signal reçu dans laquelle la contribution du bruit disparaît. Nous étendons, ensuite, la méthode proposée pour exploiter la connaissance a priori des filtres de transmission/réception afin d'améliorer sa performance.

Nous proposons aussi une méthode d'estimation de canal basée sur le critère d'égalisation aveugle de forçage à zéro minimisant l'erreur quadratique moyenne. Nous avons prouvé, dans ce contexte, l'équivalence entre plusieurs critères d'égalisation dans le cas sans bruit et sous la contrainte d'absence du biais. Par conséquent, nous avons déduit qu'une étape de débruitage de la matrice de covariance du signal reçu est nécessaire

avant son utilisation dans la méthode d'estimation aveugle. L'avantage principal de cette méthode est qu'elle évite la décomposition propre de la matrice de covariance tout en aboutissant à des performances optimales par rapport à la borne de Cramer-Rao.

Table des matières

1	Introduction	1
1.1	Identification de Canal	1
1.1.1	Systèmes Mobiles	2
1.1.2	Identification de Canal par Séquence d'Apprentissage	2
1.1.3	Identification Aveugle de Canal	3
1.1.4	Algorithmes Aveugles Mono-entrée/Mono-sortie	4
1.2	Modèle Multi-canal Linéaire	5
1.2.1	Suréchantillonnage Temporel	6
1.2.2	Réseaux d'Antennes	7
1.2.3	Constellations Particulières	8
1.3	Formulation du problème	8
1.3.1	Représentation du Signal Reçu	9
1.3.2	Hypothèses	10
1.3.3	Identifiabilité	11
1.3.4	Modèles des Symboles	11
1.4	Contributions de la Thèse	12
2	Sous-espaces et Paramétrisations	19
2.1	Introduction	19
2.2	Sinusoïdes dans le Bruit	20
2.2.1	Prédiction Linéaire	20
2.2.2	Sous-espace Signal(SS)/Sous-espace Bruit(SB)	21
2.2.3	Ajustement du SS	22
2.2.4	Ajustement du SB	23
2.3	Identification Aveugle du Canal	23
2.3.1	Sous-espace Signal/Sous-espace Bruit	24
2.3.2	Caractérisation de la Matrice de Covariance	24
2.4	Paramétrisations du Sous-espace Bruit	25
2.4.1	Paramétrisation Linéaire Minimale du SB	26
2.4.2	Paramétrisations Linéaires du SB Basées sur le Canal	27
2.4.3	Paramétrisations et Génération du SB	29
2.4.4	Discussion	31

2.5	Conclusions	32
3	Méthodes Basées sur les Sous-espaces	35
3.1	Introduction	35
3.2	Statistiques de Second Ordre et Identification du Canal	36
3.2.1	Estimation Précise du SS	36
3.2.2	Simulations	37
3.3	Bornes de Cramer-Rao	38
3.3.1	Borne de Cramer-Rao Déterministe	40
3.3.2	Borne de Cramer-Rao Gaussienne	41
3.4	Subchannel Response Matching (SRM)	41
3.4.1	Cas Asymptotique	42
3.4.2	Formulation dans le cas “équilibré”	42
3.5	Méthodes d’Ajustement de Sous-espaces	44
3.5.1	Ajustement du SS	44
3.5.2	Ajustement du SB	44
3.6	Méthodes sans Décomposition Propre de la Matrice de Covariance	45
3.6.1	Méthode Alternative pour l’Ajustement du SS	46
3.6.2	Analyse des Performances	47
3.6.3	Simulations : Cas de l’Ajustement du SS	49
3.6.4	Méthode Alternative pour l’Ajustement du SB	49
3.6.5	Simulations : Cas de l’Ajustement du SB	50
3.7	Méthode Dérivée de la Prédiction Linéaire	51
3.7.1	Prédiction Linéaire	52
3.7.2	Approche de Base	52
3.8	Conclusions	52
4	Méthodes de Maximum de Vraisemblance	55
4.1	Introduction	55
4.2	Deterministic ML (DML) et Iterative Quadratic ML (IQML)	56
4.2.1	Deterministic ML (DML)	57
4.2.2	Iterative Quadratic ML (IQML)	58
4.3	Denoised IQML (DIQML) et Pseudo-Quadratic ML (PQML)	59
4.3.1	Denoised IQML (DIQML)	59
4.3.2	Pseudo-Quadratic ML (PQML)	61
4.4	Extensions au Cas Semi-aveugle	62
4.4.1	PQML Least-Squares (PQML-LS)	62
4.4.2	PQML Weighted-Least-Squares (PQML-WLS)	63
4.5	Simulations	65
4.5.1	Algorithmes Aveugles	65
4.5.2	Algorithmes Semi-aveugles	66
4.6	Discussion	68

4.6.1	Méthodes pour les Symboles Déterministes	68
4.6.2	Méthodes pour les Symboles Gaussiens	69
4.7	Exploitation de l'Alphabet Fini	70
4.7.1	Critère d'Egalisation	70
4.7.2	PQML-DD-WLS	71
4.7.3	Decision-Directed Least-Squares (DD-LS)	72
4.7.4	Simulations	72
4.8	Conclusions	73
5	Exploitation de la Connaissance des Filtrés de Transmission	75
5.1	Introduction	76
5.2	Connaissance des Filtrés de Transmission/Réception	77
5.2.1	Cas du Suréchantillonnage	77
5.2.2	Cas d'un Réseau d'Antennes	78
5.3	Méthodes Aveugles avec Connaissance des Filtrés de Transmission	79
5.3.1	Méthodes Aveugles	79
5.3.2	Méthodes Etendues	79
5.4	Aspects d'Identifiabilité	81
5.5	Bornes de Cramer-Rao avec Connaissance des Filtrés de TX/RX	82
5.6	Simulations	83
5.6.1	Techniques Basées sur les Sous-espaces	83
5.6.2	Techniques de Maximum de Vraisemblance	84
5.7	Conclusions	86
6	Robustesse à la Surdétermination de l'Ordre du Canal	89
6.1	Introduction	89
6.2	Ancrage de la Réponse Impulsionnelle	90
6.2.1	Prédiction Linéaire	90
6.2.2	Approche de Base	91
6.3	Robustesse	92
6.3.1	Première Approche: Détermination Implicite de l'Ordre	92
6.3.2	Deuxième Approche: Détermination Explicite de l'Ordre	93
6.4	Simulations	94
6.4.1	Cas de la Première Approche	94
6.4.2	Cas de la Deuxième Approche	96
6.5	Conclusions	98
7	Estimation Aveugle en Présence d'un Bruit Stationnaire de Couleur Inconnue	101
7.1	Introduction	101
7.2	Méthode d'Identification	102
7.2.1	Structure de Déplacement	103

7.2.2	Approche d'Ajustement de Sous-espace	104
7.3	Bornes de Cramer-Rao	104
7.3.1	Borne de Cramer-Rao Gaussienne	105
7.3.2	Borne de Cramer-Rao Déterministe	106
7.4	Connaissance du Filtre de Transmission (TX)	107
7.4.1	Méthode avec Connaissance du Filtre de TX (WTFXK)	107
7.4.2	Bornes de Cramer-Rao WTFXK	108
7.5	Simulations	108
7.5.1	Méthode d'Ajustement de Sous-espace	109
7.5.2	Méthode d'Ajustement de Sous-espace WTFXK	109
7.6	Conclusions	110
8	Estimation Multi-canal Aveugle par Egalisation MMSE ZF	113
8.1	Introduction	113
8.2	Méthode de Capon	114
8.2.1	Estimation des Directions d'Arrivées	115
8.2.2	Estimation Multi-canal	115
8.3	Relations entre Critères & Egaliseurs	116
8.3.1	Relations entre Critères	116
8.3.2	Egaliseurs	117
8.4	Egalisation Aveugle MMSE ZF	118
8.4.1	Interprétation en Termes de Prédiction Linéaire Bidirectionnelle	119
8.4.2	Analyse dans le Cas sans Bruit	119
8.4.3	Relation avec d'Autres Approches	120
8.4.4	Détection de l'Ordre du Canal	120
8.4.5	Débruitage de la Matrice de Covariance	121
8.5	Simulations	122
8.5.1	Performance de la Méthode	122
8.5.2	Détection de l'Ordre du Canal	122
8.6	Conclusions	123
	Conclusion Générale	125
A	Démonstrations pour les Egaliseurs de Longueur Infinie dans le Do-	
	maine Fréquentiel	129
A.1	Contrainte d'Absence du Biais	129
A.2	Contrainte ZF	130
	Bibliographie	133

Table des figures

1.1	Réseau cellulaire.	3
1.2	Système mono-entrée/mono-sortie à identifier.	5
1.3	Egalisation adaptative d'un système mono-entrée/mono-sortie.	5
1.4	Suréchantillonnage temporel avec $m = 2$	7
1.5	Réseau de $m = 2$ antennes.	7
1.6	Modèle multi-canal linéaire	9
3.1	Estimation précise du SS avec un nombre fini de données.	38
3.2	Estimation du premier moment de second ordre avec un nombre fini de données.	39
3.3	Performance de l'ajustement du SS sans et avec décomposition propre de R_{YY} (matrice de covariance estimée).	49
3.4	Performance de l'ajustement du SS sans décomposition propre de R_{YY} (matrice de covariance parfaitement estimée).	50
3.5	Performance de l'ajustement du SB sans et avec décomposition propre de R_{YY}	51
4.1	Performance des algorithmes DIQML et PQML : NMSE.	66
4.2	Performance des algorithmes DIQML et PQML : fonction de coût DML.	66
4.3	DIQML : Evolution de λ en fonction du nombre d'itérations effectuées.	67
4.4	PQML : Evolution de λ en fonction du nombre d'itérations effectuées.	67
4.5	Performance des algorithmes PQML, PQML-LS et PQML-WLS : NMSE.	68
4.6	Performance des algorithmes PQML-LS et PQML-WLS : NMSE.	69
4.7	Performance des algorithmes PQML-DD-WLS, WLS, DD-LS (symboles détectés) et DD-LS (symboles connus) : NMSE.	73
4.8	Importance de l'initialisation pour la méthode DD-LS (symboles détectés).	74
5.1	Système typique de communications numériques.	76
5.2	Canal global.	84
5.3	Performance des méthodes : SRM et SRM WTXFK.	85
5.4	Performance des méthodes : SSF et SSF WTXFK.	85
5.5	Comparaison des méthodes SRM WTXFK avec $\ \mathbf{c}\ = 1$ et $\ \mathbf{h}\ = 1$	86
5.6	Comparaison des méthodes : SRM WTXFK et IQML WTXFK.	87

6.1	SRM - première approche: comparaison des erreurs (NMSEs) obtenues avec les contraintes $\ \mathbf{h}\ _2 = 1$ et $\ \mathbf{h}(0)\ _2 = 1$	95
6.2	SRM - première approche: robustesse à la surestimation de l'ordre du canal avec les contraintes $\ \mathbf{h}\ _2 = 1$ et $\ \mathbf{h}(0)\ _2 = 1$	96
6.3	SSF - première approche: robustesse à la surestimation de l'ordre du canal avec les contraintes $\ \mathbf{h}\ _2 = 1$ et $\ \mathbf{h}(0)\ _2 = 1$ en utilisant une solution intermédiaire.	97
6.4	SSF - première approche: robustesse à la surestimation de l'ordre du canal avec les contraintes $\ \mathbf{h}\ _2 = 1$ et $\ \mathbf{h}(0)\ _2 = 1$ en utilisant la procédure proposée.	97
6.5	SRM - deuxième approche: robustesse à la surestimation de l'ordre du canal avec les contraintes $\ \mathbf{h}\ _2 = 1$ et $\ \mathbf{h}(0)\ _2 = 1$	98
6.6	SSF - deuxième approche: robustesse à la surestimation de l'ordre du canal avec les contraintes $\ \mathbf{h}\ _2 = 1$ et $\ \mathbf{h}(0)\ _2 = 1$	99
7.1	Les matrices R_{YY} , \underline{R}_{YY} et \overline{R}_{YY}	103
7.2	Paramètres de R_{VV} à estimer.	106
7.3	Performance de la méthode d'ajustement de sous-espace.	110
7.4	Filtre GMSK linéarisé.	111
7.5	Performance de la méthode d'ajustement de sous-espace WTXFK.	111
8.1	Prédiction linéaire bidirectionnelle.	119
8.2	Performance des différents algorithmes.	123
8.3	Détection de l'ordre du canal.	124

Notations

$(.)^*$:	complexe conjugué
$(.)^T$:	transposé
$(.)^H$:	transposé conjugué
$(.)^+$:	pseudo-inverse de Moore-Penrose
$(.)^\#$:	inverse généralisé
$H^\dagger(z)$:	$H^H(1/z^*)$
$\text{Re}\{.\}$:	partie réelle d'une variable complexe
$\text{Im}\{.\}$:	partie imaginaire d'une variable complexe
$\ \cdot\ _F$:	norme de Frobenius
$E\{.\}$:	moyenne statistique
$\ln(.)$:	logarithme népérien
δ_{ij}	:	symbole de Kronecker ($\delta_{ij} = 1$ si $i = j$ et $\delta_{ij} = 0$ si $i \neq j$)
$\lceil x \rceil$:	le plus petit entier supérieur ou égal à x
$\arg \min_x f(x)$:	l'argument qui minimise la fonction $f(x)$
$m \times n$:	dimension d'une matrice ayant m lignes et n colonnes
I	:	matrice identité
I_n	:	matrice identité de dimension $n \times n$
$0_{p \times q}$:	matrice de dimension $p \times q$ dont tous les éléments sont égaux à zéro
$\text{rang}(A)$:	dimension du sous-espace engendré par A
$\text{Image}\{A\}$:	sous-espace engendré par A
$\text{ker}\{A\}$:	$(\text{Image}\{A\})^\perp$, complément orthogonal de $\text{Image}\{A\}$
P_A	:	projection sur l'espace engendré par A , $P_A = A(A^H A)^+ A^H$
P_A^\perp	:	projection sur le complément orthogonal de l'espace engendré par A , $P_A^\perp = I - P_A$
$V_{\min}(A)$:	vecteur propre associé à la valeur propre minimale de A
$V_{\max}(A)$:	vecteur propre associé à la valeur propre maximale de A
$\text{tr}\{A\}$:	trace de la matrice carrée A
$\text{Sym}(A)$:	$\frac{1}{2}(A + A^H)$
$\det(A)$:	produit des valeurs propres de la matrice carrée A
\otimes	:	produit de Kronecker de deux matrices
$A \otimes B = \begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \\ \vdots & & \ddots \end{bmatrix} \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots \\ a_{21}B & a_{22}B & \\ \vdots & & \ddots \end{bmatrix}$		
$\text{diag}\{a_1, \dots, a_n\}$:	matrice ayant comme diagonale principale a_1, \dots, a_n et zéro ailleurs
$\text{blockdiag}\{.\}$:	matrice bloc diagonale

$$\text{blockdiag}\{A_1, A_2, \dots, A_n\} = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & A_n \end{bmatrix}$$

Abréviations

AIC	: Akaike's Information theoretic Criterion
AR	: AutoRégressif
BPSK	: Binary Phase Shift Keying
CDMA	: Code Division Multiple Access
CMA	: Constant Modulus Algorithm
CRB	: Cramer-Rao Bound
DD	: Decision-Directed
DIQML	: Denoised Iterative Quadratic Maximum Likelihood
DML	: Deterministic Maximum Likelihood
DOA	: Directions Of Arrival
FIR	: Finite Impulse Response
GMSK	: Gaussian Minimum Shift Keying
GML	: Gaussian Maximum Likelihood
GSM	: Global System for Mobile Communications
HOS	: Higher Order Statistics
IES	: Interférences Entre Symboles
i.i.d	: independant identically distributed
LCMV	: Linearly Constrained Minimum Variance
LD	: Lower Dimensional
LDU	: Lower Diagonal Upper
LS	: Least Squares
MA	: Moving Average
MAP	: Maximum A posteriori Probability
MDL	: Minimum Description Length
MDPL	: Méthode Dérivée de la Prédiction Linéaire
ML	: Maximum Likelihood
MMSE	: Minimum Mean Square Error
MOE	: Minimum Output Energy
MSE	: Mean Square Error
MSINR	: Maximum Signal to Interference plus Noise Ratio
MSK	: Minimum Shift Keying
NMSE	: Normalized Mean Square Error
NRMSE	: Normalized Root Mean Square Error
NSF	: Noise Subspace Fitting
PAM	: Pulse Amplitude Modulation
PQML	: Pseudo-Quadratic Maximum Likelihood
QAM	: Quadrature Amplitude Modulation
QPSK	: Quadrature Phase Shift Keying
RX	: Réception

SB	:	Sous-espace Bruit
SINR	:	Signal to Interference plus Noise Ratio
SNR	:	Signal to Noise Ratio
SRM	:	Subchannel Response Matching
SS	:	Sous-espace Signal
SSF	:	Signal Subspace Fitting
SVD	:	Singular Value Decomposition
TDMA	:	Time Division Multiple Access
TX	:	Transmission
UMMSE	:	Unbiased Minimum Mean Square Error
UMOE	:	Unbiased Minimum Output Energy
WLS	:	Weighted Least Squares
WTXFK	:	With Transmission Filter Knowledge
ZF	:	Zero Forcing

Chapitre 1

Introduction

Dans le cadre de cette thèse, nous nous intéressons aux problèmes d'estimation (ou d'identification) et d'égalisation de canal en utilisant les statistiques de second ordre. Ces problèmes sont rencontrés dans un contexte de communications numériques et plus particulièrement celui des communications mobiles. Nous commençons, dans une première partie de ce chapitre, par une description brève des communications mobiles. Puis nous présentons le modèle de canal qui sera adopté dans le développement des algorithmes d'identification/égalisation proposés dans cette thèse et nous expliquons les différentes situations dans lesquelles un tel modèle peut être envisagé. Ensuite, nous formulons le problème d'estimation de paramètres à résoudre ainsi que les hypothèses qui seront adoptées et nous décrivons les deux modèles possibles pour les symboles transmis selon que ces derniers sont considérés comme des quantités déterministes ou stochastiques. Dans la dernière partie de ce chapitre, nous présentons les différentes contributions de ce travail de thèse.

1.1 Identification de Canal

Les procédures d'identification/égalisation de canal qui nous intéressent, dans le cadre de cette thèse, sont qualifiées d'*aveugles*. Par ce terme nous entendons que ces algorithmes identifient le canal (entre l'émetteur et le récepteur) en disposant uniquement du signal reçu. Ces méthodes utilisent un modèle *multi-canal* qui peut être fréquemment rencontré dans les systèmes de communications mobiles. Les méthodes d'identification aveugles sont opposées aux méthodes basées sur les séquences d'apprentissage qui utilisent la connaissance de l'entrée et de la sortie du système afin de l'identifier. Nous commençons donc par une description brève des systèmes mobiles, puis nous expliquons le principe des méthodes d'identification de canal basées sur les séquences d'apprentissage, ensuite nous décrivons les algorithmes aveugles ainsi que les motivations qui rendent cette classe de méthodes intéressante.

1.1.1 Systèmes Mobiles

L'introduction des systèmes de communications numériques mobiles était à l'origine d'un grand effort de recherche dans le domaine du traitement du signal et des communications numériques. En communications mobiles, le canal physique est partagé par tous les utilisateurs et le spectre fréquentiel disponible pour les services mobiles est une ressource limitée. Par conséquent, les techniques qui permettent d'augmenter l'efficacité spectrale sont d'un intérêt considérable. Une approche utilisée pour augmenter l'efficacité spectrale dans les systèmes de communications mobiles émane du *concept cellulaire*, illustré dans la figure 1.1, tel que le système GSM (Global System for Mobile Communications). Dans un système cellulaire, une région géographique est divisée en petites régions appelées cellules. Chaque cellule est équipée d'une station de base qui permet de desservir les utilisateurs évoluant au sein de cette même cellule. Chaque cellule utilise un ensemble de fréquences qui sont différentes de celles utilisées dans les cellules voisines et seules les cellules qui sont suffisamment espacées peuvent réutiliser la même fréquence (dans la figure 1.1, la fréquence f_1 est utilisée dans deux cellules différentes). Cette réutilisation de fréquences permet d'augmenter la capacité en termes d'utilisateurs potentiels pouvant s'ajouter dans le réseau. Pour que cette capacité soit la plus élevée, la distance de réutilisation de fréquence doit être maintenue aussi petite que possible.

Une autre approche qui vient compléter le concept cellulaire est l'utilisation de réseaux d'antennes au niveau de la station de base [SBEM90],[AMVW91]. Les réseaux d'antennes peuvent amplifier les signaux dans une certaine direction et supprimer les signaux dans d'autres directions. Par conséquent, les systèmes utilisant les réseaux d'antennes peuvent tolérer plus d'interférences. Une propriété inhérente aux réseaux d'antennes découle du fait que le signal émis est reçu par plusieurs capteurs, ce qui facilite le traitement multi-dimensionnel de ce signal.

1.1.2 Identification de Canal par Séquence d'Apprentissage

Dans un contexte de communications mobiles, les canaux sont spécifiques du fait qu'ils peuvent varier rapidement dans le temps. A cause des limitations en largeur de bande et les multi-trajets de propagation, le canal de transmission introduit des distorsions sur le signal transmis pour engendrer des Interférences Entre Symboles (IES). Pour restituer la séquence de données émise, le récepteur doit identifier la distorsion introduite par le canal et l'*égaliser*. Les techniques d'identification classiques nécessitent l'utilisation de l'entrée et de la sortie du système à identifier, ce qui entraîne la transmission d'une séquence d'apprentissage, c'est à dire une séquence fixe (qui ne transporte pas d'information utile) et qui est connue par l'émetteur et le récepteur. En connaissant le signal d'entrée et en mesurant le signal à la sortie du canal, le récepteur peut déterminer le canal. Pour illustrer cette procédure, nous considérons le cas du système cellulaire GSM. Dans ce système, l'information est transmise sous forme de paquets de 148 symboles. Chaque paquet contient, au milieu, une séquence d'apprentissage de 26 symboles et l'in-

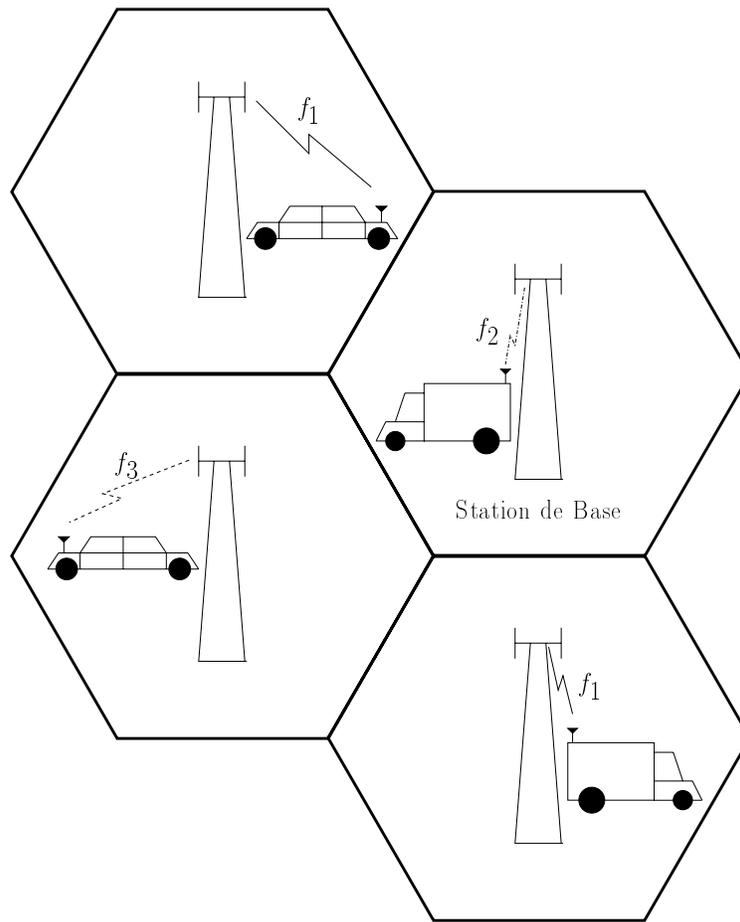


FIG. 1.1 – Réseau cellulaire.

formation utile est située avant et après cette séquence d'apprentissage. A la réception, le récepteur localise d'abord la séquence d'apprentissage et identifie ensuite le canal de transmission.

L'utilisation de séquences d'apprentissage réduit le débit de transmission, surtout lorsque la séquence d'apprentissage doit être retransmise souvent à cause des variations rapides qui peuvent avoir lieu en communications mobiles. Contrairement à ces techniques, les égaliseurs aveugles n'utilisent pas de séquences d'apprentissage.

1.1.3 Identification Aveugle de Canal

Plusieurs travaux (tels que [Sat75],[God80],[TA83]) ont pu démontrer que le canal peut être identifié en disposant uniquement de ses sorties (sans avoir à connaître les entrées correspondantes). En d'autres termes, même si le récepteur ne connaît pas les symboles transmis dans le paquet de données, il peut utiliser la structure du signal reçu pour identifier le canal. Cette approche est appelée *identification aveugle*.

Plusieurs arguments plaident en faveur de l'identification aveugle. En effet, il est clair que l'utilisation de la séquence d'apprentissage consomme une partie de la capacité du

canal. Par conséquent, l'élimination (ou la réduction) de cette séquence d'apprentissage permettra d'utiliser la partie de la capacité du canal qui lui est destinée pour transporter l'information utile, ce qui augmente l'efficacité spectrale. Une autre motivation pour l'identification aveugle provient du fait que d'une part, dans le cas des canaux qui varient rapidement dans le temps, l'apprentissage n'est pas efficace, d'autre part, lorsque les évanouissements dus aux multi-trajets sont sévères pendant la durée de la séquence d'apprentissage, les estimées du canal peuvent être de mauvaise qualité. Dans certains scénarios de communications numériques, on ne peut pas avoir de séquence d'apprentissage. C'est le cas de l'annulation des interférences, où une station de base reçoit le signal d'une station mobile évoluant dans la cellule desservie par cette même station de base en même temps que le signal d'un utilisateur évoluant dans une cellule voisine (interfère). Pour supprimer cette interférence, il est intéressant de connaître le canal entre l'interfère et la station de base en question. Comme, généralement, la station de base ne connaît pas la séquence d'apprentissage de l'interfère, l'unique façon de déterminer ce canal est d'utiliser des méthodes d'estimation aveugle de canal. L'identification aveugle trouve aussi sa motivation dans les scénarios multi-utilisateur, où l'apprentissage nécessite la synchronisation, ce qui n'est pas faisable dans ce contexte. Dans le cas des réseaux distribués, l'envoi de séquences d'apprentissage à chaque fois qu'une liaison est à établir peut ne pas être un procédé réaliste.

Ces motivations pour l'identification aveugle sont à l'origine de plusieurs travaux de recherche qui ont permis de proposer des algorithmes d'identification aveugle de canal. En se basant sur la description entrée/sortie du système à identifier, ces algorithmes peuvent être divisés en deux classes : une classe qui regroupe les méthodes traditionnelles qui identifient des systèmes mono-entrée/mono-sortie, et une autre classe qui inclue les techniques exploitant une structure mono-entrée/multi-sortie du système. Les algorithmes d'identification de canal proposés dans cette thèse appartiennent à la deuxième classe.

1.1.4 Algorithmes Aveugles Mono-entrée/Mono-sortie

Dans la figure 1.2, le système mono-entrée/mono-sortie à identifier est illustré, $a(k)$ étant la séquence d'entrée discrète, $y(k)$ représente le signal reçu discret, $v(k)$ est le bruit additif et $h(k)$ désigne la réponse impulsionnelle discrète du canal à identifier. Le problème de l'identification aveugle consiste donc en la formulation suivante : en disposant uniquement des sorties mesurées du système, $y(k)$, déterminer $h(k)$ et la séquence d'entrée $a(k)$. Généralement ce problème est résolu de façon adaptative en plaçant un filtre qui va filtrer le signal reçu. Cette situation est illustrée dans la figure 1.3, où le filtre $f(k)$ est utilisé pour compenser (égaliser) les distorsions introduites par $h(k)$ sur le signal $a(k)$ et donc sa sortie $\hat{a}(k)$, doit être la plus proche de $a(k)$ au sens d'un certain critère. L'élément de décision $dec(\cdot)$ génère un signal d'erreur $e(k)$ qui servira à adapter le filtre $f(k)$. Les méthodes formulées dans ce contexte exploitent, en définissant l'élément de décision $dec(\cdot)$, certaines caractéristiques des signaux de communications utilisés. Parmi ces méthodes, nous distinguons l'algorithme CMA (Constant Modulus

Algorithm) [JSE⁺98], qui utilise le fait que certains signaux de communications sont à module constant. Dans [Hay94], d'autres exemples d'applications qui utilisent l'identification/déconvolution aveugle sont présentés.

Le canal illustré dans la figure 1.2, lorsqu'il est à phase non minimale, ne peut pas être identifié à partir de ses statistiques de second ordre, et son identification nécessite donc l'utilisation des statistiques d'ordre supérieur à deux : Higher Order Statistics (HOS), dans le cas des signaux stationnaires. L'idée introduite dans des travaux récents tels que [TXK91],[Slo94a],[LXTK96],[TP98], consiste à utiliser l'information supplémentaire qui existe dans une représentation *multi-canal* du système à identifier. Une telle représentation est, en effet, fréquemment rencontrée dans les systèmes de communications mobiles qui utilisent des réseaux d'antennes à la station de base. Dans ce qui suit, nous décrivons le modèle multi-canal ainsi que les différents cas où il peut être envisagé.

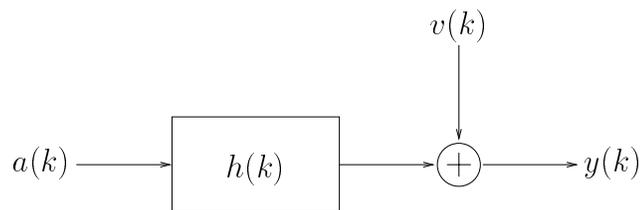


FIG. 1.2 – *Système mono-entrée/mono-sortie à identifier.*

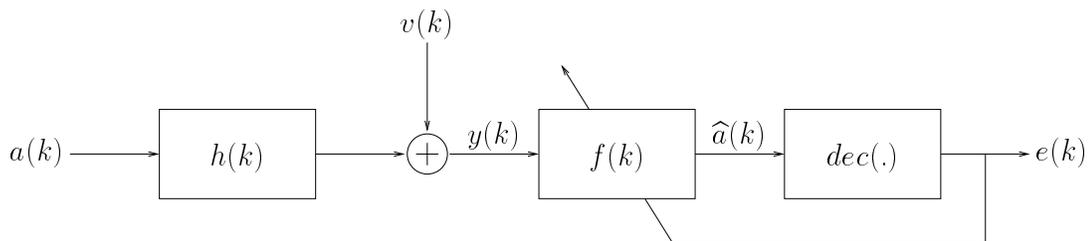


FIG. 1.3 – *Egalisation adaptative d'un système mono-entrée/mono-sortie.*

1.2 Modèle Multi-canal Linéaire

Nous considérons une modulation numérique linéaire à travers un canal à bruit additif gaussien. Le signal reçu peut être écrit comme

$$y(t) = \sum_k a(k)h(t - kT) + v(t) \quad (1.1)$$

où les $a(k)$ sont les symboles transmis, T est la période des symboles et $h(t)$ est la réponse impulsionnelle (souvent appelée réponse impulsionnelle du canal pour des raisons de simplicité) de la cascade : canal de propagation, filtre de mise en forme (filtre d'émission) et filtre de réception. Supposons que $\{a(k)\}$ et $\{v(t)\}$ sont stationnaires (au sens

large), le processus $\{y(t)\}$ est cyclostationnaire avec la période T . Gardner [Gar91a] a démontré que le canal peut être identifié à partir des statistiques de second ordre du signal cyclostationnaire $\{y(t)\}$. Si le canal est connu, on peut passer le signal reçu à travers un filtre adapté et échantillonner la sortie à la période des symboles. Ces échantillons fourniront des statistiques suffisantes pour la détection des symboles transmis. Si $\{y(t)\}$ est échantillonné avec la période T , le processus échantillonné obtenu est stationnaire (au sens large) et ses statistiques de second ordre ne contiennent pas d'information sur la phase du canal. Tong, Xu et Kailath [TXK91] ont proposé de *suréchantillonner* le signal reçu avec la période $\Delta = T/m$, $m > 1$.

Dans ce qui suit, nous supposons que le canal $h(t)$ est de durée finie. Tong *et al.* ont démontré que le canal peut être identifié à partir des statistiques de second ordre du signal reçu suréchantillonné. Les canaux discrets parallèles que nous allons considérer proviennent de la représentation d'un canal ayant une certaine diversité au rythme des symboles. La diversité considérée peut être *spectrale*, dans laquelle le suréchantillonnage par rapport à la période des symboles exploite l'excès en largeur de bande du canal; ou *spatiale* obtenue à partir de réseaux d'antennes réceptrices (et/ou des différentes polarisations) dans le contexte des communications mobiles. Une troisième possibilité concerne le cas où la constellation des symboles est unidimensionnelle (par exemple PAM ou BPSK) et les symboles transmis sont modulés. Dans ce cas, la réponse impulsionnelle possède une composante réelle et une composante imaginaire alors que l'entrée est purement réelle. Par conséquent, en considérant uniquement des signaux réels, nous obtenons un système à une entrée et deux sorties.

1.2.1 Suréchantillonnage Temporel

Nous considérons le cas où la représentation multi-canal est obtenue par l'intermédiaire d'un suréchantillonnage du signal reçu $\{y(t)\}$. Considérons, par exemple le cas d'un facteur de suréchantillonnage $m = 2$. La fréquence d'échantillonnage est, dans ce cas, égale à $\frac{2}{T}$ et les instants d'échantillonnage du signal exprimé dans (1.1) sont $T(k + \frac{j}{2})$ avec $j \in \{0, 1\}$. Pour $j = 0$, introduisons la notation $h_1(k) = h(kT)$ pour la réponse impulsionnelle discrète du canal, $y_1(k) = y(kT)$ pour le signal reçu discret et $v_1(k) = v(kT)$ pour la version discrète du bruit, nous avons donc

$$y_1(k) = \sum_i a(k-i)h_1(i) + v_1(k). \quad (1.2)$$

De façon similaire, en considérant $j = 1$ et en définissant $h_2(k) = h(kT + \frac{T}{2})$, $y_2(k) = y(kT + \frac{T}{2})$ et $v_2(k) = v(kT + \frac{T}{2})$, nous obtenons

$$y_2(k) = \sum_i a(k-i)h_2(i) + v_2(k). \quad (1.3)$$

Notons que (1.2) et (1.3) sont deux représentations discrètes du même système à temps continu décrit par (1.1). Cette conversion d'un système à temps continu en deux systèmes

à temps discret ayant la même séquence d'entrée est décrite par la figure 1.4. Ainsi, le processus suréchantillonné $\{y(k)\}$ peut être décrit comme un processus scalaire cyclostationnaire ou comme un processus vectoriel stationnaire (à une entrée et deux sorties).

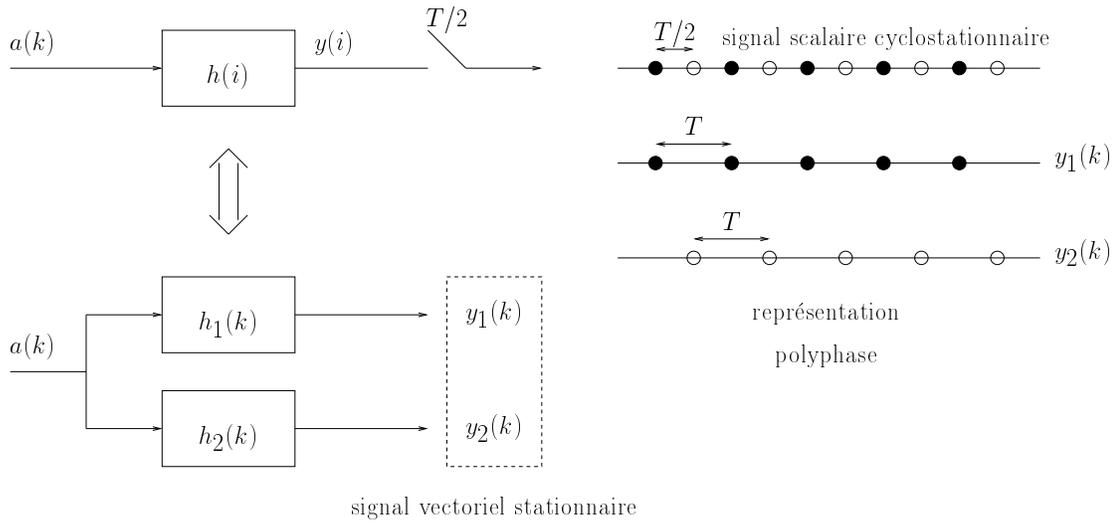


FIG. 1.4 – Suréchantillonnage temporel avec $m = 2$.

1.2.2 Réseaux d'Antennes

L'utilisation de réseaux d'antennes en communications mobiles permet d'amplifier les signaux faibles, de lutter contre les évanouissements que le signal peut subir et d'annuler les interférences [BS92],[Ott96]. Avec l'utilisation de réseaux d'antennes au niveau de la station de base, il est évident que la transmission du mobile vers la station de base peut être modélisée avec m canaux parallèles (où m désigne dans ce contexte le nombre d'antennes). Chaque canal est décrit par l'expression (1.1) et il est différent d'une antenne à l'autre (la différence dépend de l'environnement de propagation, l'espacement entre les antennes...). Le cas de deux antennes est illustré dans la figure 1.5.

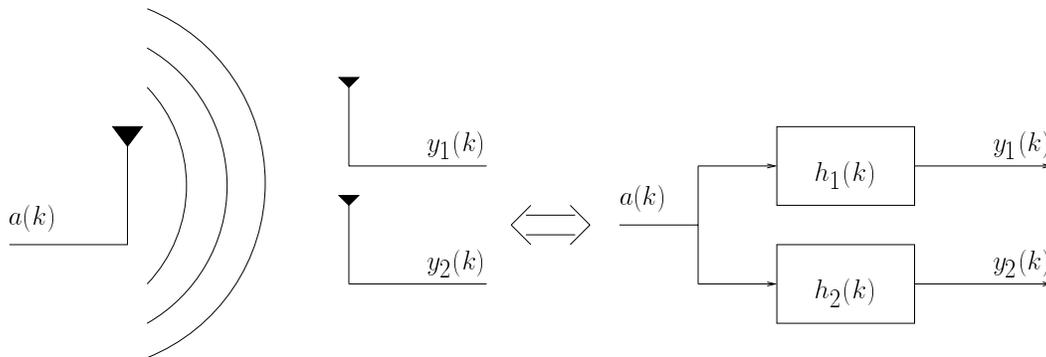


FIG. 1.5 – Réseau de $m = 2$ antennes.

1.2.3 Constellations Particulières

Dans certains cas, il est possible d'exploiter la structure de la constellation à laquelle appartient la séquence $\{a(k)\}$ pour obtenir une représentation multi-canal du signal reçu [KOS96]. Dans le cas où la constellation est unidimensionnelle (par exemple PAM ou BPSK), $a(k)$ est réelle, si nous considérons l'expression (1.1), nous pouvons séparer sa partie réelle et imaginaire comme suit

$$\begin{aligned} y_1(t) &= \operatorname{Re}\{y(t)\} = \sum_k a(k) \operatorname{Re}\{h(t - kT)\} + \operatorname{Re}\{v(t)\}, \\ y_2(t) &= \operatorname{Im}\{y(t)\} = \sum_k a(k) \operatorname{Im}\{h(t - kT)\} + \operatorname{Im}\{v(t)\}. \end{aligned} \quad (1.4)$$

Donc, grâce à la caractéristique de la constellation (réelle) de $\{a(k)\}$, nous pouvons obtenir deux canaux parallèles ayant la même séquence d'entrée. Dans [TS98a], cette séparation en parties réelle et imaginaire du signal est faite dans le cas d'une modulation MSK ou GMSK (voir aussi [GC98] pour une exploitation plus fine de la modulation MSK).

Nous avons élaboré dans cette section les différents scénarios qui nous permettent d'obtenir une représentation multi-canal. Dans la section qui suit nous présentons la formulation mathématique du problème d'estimation de canal utilisant cette représentation.

1.3 Formulation du problème

Nous supposons, dans tous les cas, que le canal est à réponse impulsionnelle finie (FIR) avec une durée approximative égale à NT . Considérons le cas du suréchantillonnage¹ du signal reçu avec une fréquence de suréchantillonnage égale à $\frac{m}{T}$. Les instants d'échantillonnage pour le signal donné par (1.1) sont $t_0 + T(k + \frac{j}{m})$, k étant un entier naturel et $j = 0, 1, \dots, m-1$. Nous introduisons la description polyphase du signal reçu : $y_j(k) = y(t_0 + T(k + \frac{j}{m}))$ pour $j = 0, 1, \dots, m-1$ (les $y_j(k)$ sont les m phases du signal reçu). De façon similaire, nous définissons les composantes polyphases de la réponse impulsionnelle du canal et du bruit additif. Dans le cas de la figure 1.4 (où $m = 2$), les deux composantes polyphases sont les deux séquences d'échantillons paire et impaire. En principe, il suffit d'introduire $t_0 \in [0, T[$ pour traiter le cas général. Cependant, nous devons considérer $t_0 = t'_0 + dT$ où $t'_0 \in [0, T[$ et d est choisi comme étant l'entier naturel le plus petit tel que $\mathbf{h}(0) = [h(t'_0 + dT) \cdots h(t'_0 + (d + \frac{m-1}{m})T)]^T \neq 0$ (T désigne transposée). Le canal étant causal, d est donc positif et représente un retard inhérent. La transformée en z du canal à la période des symboles $\frac{m}{T}$ est

$$\mathbf{H}(z) = \sum_{j=1}^m z^{-(j-1)} \mathbf{H}_j(z^m) \quad (1.5)$$

1. Nous supposons que le facteur de suréchantillonnage est un entier naturel supérieur ou égal à 2, le cas d'un suréchantillonnage avec un nombre rationnel est traité dans [Slo93].

où $H_j(z)$ est la fonction du transfert correspondant au $j^{\text{ème}}$ sous-canal.

1.3.1 Représentation du Signal Reçu

Bien que les trois procédures décrites dans la section §1.2 permettent d'obtenir une représentation multi-canal du système selon différentes façons, le système à m canaux obtenu peut être décrit dans tous les cas par le même modèle vectoriel illustré dans la figure 1.6.

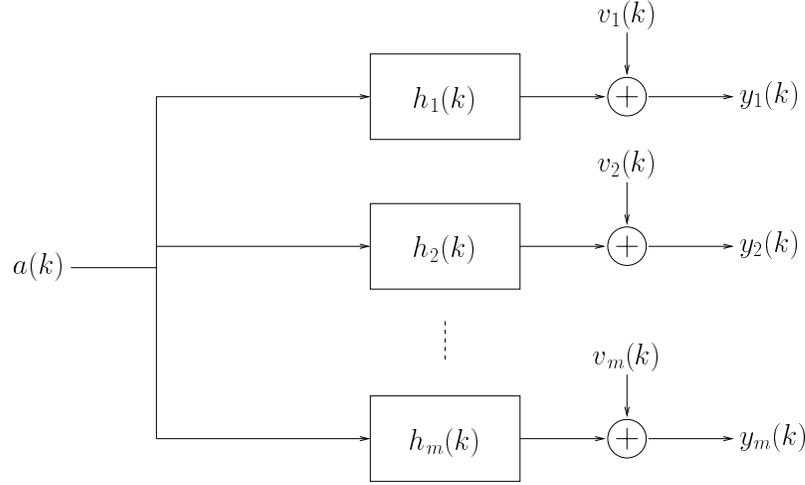


FIG. 1.6 – *Modèle multi-canal linéaire*

Dans ce qui suit, nous supposons que les m canaux sont obtenus en utilisant une de ces trois procédures ou une combinaison de deux d'entre elles ou les trois procédures en même temps. Le signal reçu peut être représenté sous forme vectorielle comme suit

$$\mathbf{y}(k) = \sum_{i=0}^{N-1} \mathbf{h}(i)a(k-i) + \mathbf{v}(k) = \mathbf{H}A_N(k) + \mathbf{v}(k),$$

$$\mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_m(k) \end{bmatrix}, \quad \mathbf{h}(i) = \begin{bmatrix} h_1(i) \\ \vdots \\ h_m(i) \end{bmatrix}, \quad \mathbf{v}(k) = \begin{bmatrix} v_1(k) \\ \vdots \\ v_m(k) \end{bmatrix}, \quad (1.6)$$

$$\mathbf{H} = [\mathbf{h}(N-1) \cdots \mathbf{h}(0)] = \begin{bmatrix} h_1(N-1) & h_1(N-2) & \cdots & h_1(0) \\ h_2(N-1) & h_2(N-2) & \cdots & h_2(0) \\ \vdots & \vdots & \vdots & \vdots \\ h_m(N-1) & h_m(N-2) & \cdots & h_m(0) \end{bmatrix},$$

$$A_N(k) = [a(k-N+1) \cdots a(k)]^T.$$

Nous formulons l'hypothèse de la durée finie du canal, NT , comme suit :

$$\text{(HFIR)} : \mathbf{h}(0) \neq 0, \mathbf{h}(N-1) \neq 0, \text{ et } \mathbf{h}(i) = 0 \text{ pour } i < 0 \text{ ou } i \geq N. \quad (1.7)$$

Pour le système décrit par (1.6), sa fonction de transfert 1-entrée/m-sortie est donnée par

$$\mathbf{H}(z) = \sum_{i=0}^{N-1} \mathbf{h}(i)z^{-i} = \begin{bmatrix} \mathbf{H}_1(z) \\ \vdots \\ \mathbf{H}_m(z) \end{bmatrix}. \quad (1.8)$$

Si les m canaux ne possèdent pas de zéros en commun, nous avons $\mathbf{H}(z) \neq 0, \forall z$. Cependant, si tous les m canaux possèdent un zéro commun, z_0 , alors $\mathbf{H}(z) = 0$ lorsque $z = z_0$. Les fonctions vectorielles polynômiales (appelées aussi fonctions rationnelles polynômiales) engendrent un *sous-espace vectoriel rationnel* de dimension un (les propriétés de ces sous-espaces peuvent être trouvées dans [For75],[Kai80]).

Considérons une séquence de M échantillons du signal reçu. L'équation qui décrit le système s'écrit comme :

$$\mathbf{Y}_M(k) = \mathcal{T}_M(\mathbf{h}) \mathbf{A}_{M+N-1}(k) + \mathbf{V}_M(k),$$

où $\mathbf{Y}_M(k) = \begin{bmatrix} \mathbf{y}(k-M+1) \\ \vdots \\ \mathbf{y}(k) \end{bmatrix}$, $\mathbf{V}_M(k) = \begin{bmatrix} \mathbf{v}(k-M+1) \\ \vdots \\ \mathbf{v}(k) \end{bmatrix}$, $\mathbf{h} = \begin{bmatrix} \mathbf{h}(N-1) \\ \vdots \\ \mathbf{h}(0) \end{bmatrix}$. (1.9)

La matrice $\mathcal{T}(\mathbf{h})$ est une matrice bloc Toeplitz, remplie avec les coefficients du canal \mathbf{h} , possédant M lignes bloc et $[\mathbf{H} \ 0_{m \times (M-1)}]$ comme première ligne bloc :

$$\mathcal{T}_M(\mathbf{h}) = \begin{bmatrix} \mathbf{h}(N-1) & \mathbf{h}(N-2) & \cdots & \mathbf{h}(0) & 0 & \cdots & 0 \\ 0 & \mathbf{h}(N-1) & \mathbf{h}(N-2) & \cdots & \mathbf{h}(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{h}(N-1) & \mathbf{h}(N-2) & \cdots & \mathbf{h}(0) \end{bmatrix}. \quad (1.10)$$

Avec $k = M-1$, l'expression (1.9) peut être simplifiée à

$$\mathbf{Y} = \mathcal{T}(\mathbf{h}) \mathbf{A} + \mathbf{V}. \quad (1.11)$$

1.3.2 Hypothèses

En plus de l'hypothèse (HFIR) exprimée dans (1.7), nous supposons que les symboles transmis $\{a(k)\}$ sont indépendants identiquement distribués (i.i.d.). Dans tous les chapitres, sauf le chapitre 7 (où le bruit additif sera supposé stationnaire de couleur inconnue), le bruit additif $\mathbf{v}(k)$ est supposé blanc, gaussien et circulaire dont les covariances sont données par :

$$r_{\mathbf{v}\mathbf{v}}(k-i) = E \mathbf{v}(k)\mathbf{v}(i)^H = \sigma_v^2 I_m \delta_{ki}. \quad (1.12)$$

Nous supposons que $mM > M+N-1$: dans ce cas, la matrice de convolution du canal \mathcal{T} possède plus de lignes que de colonnes. Le canal est dit irréductible si $\mathbf{H}_i(z)$, $i = 1, \dots, m$ n'ont pas de zéros communs, et réductible dans les autres cas. L'espace colonne de \mathcal{T} est appelé l'espace signal et son complément orthogonal l'espace bruit. L'espace signal est, linéairement, paramétrisé par le canal \mathbf{h} .

1.3.3 Identifiabilité

L'identifiabilité du canal (et des symboles) dans la représentation multi-canal peut être déduite en considérant l'équation (1.9) dans le cas sans bruit ($\mathbf{v}(k) \equiv 0$) [Slo97]. Dans ce cas, l'expression (1.9) représente un ensemble de mM équations ayant mN inconnues (les paramètres de \mathbf{H}) auxquels s'ajoutent $M + N - 1$ inconnues (les symboles $A_{M+N-1}(k)$). L'ensemble de ces équations peut, en principe, être résolu si $mM \geq mN + M + N - 1$, et donc $M \geq \frac{(m+1)N-1}{m-1}$. Pour n'importe quelle longueur de canal N , une longueur de données M qui satisfait cette contrainte peut être trouvée, pourvu que le nombre de canaux m soit supérieur ou égal à 2.

Notons aussi que l'identifiabilité dans le contexte aveugle (déterministe) est identique à l'identifiabilité à un facteur scalaire multiplicatif. Ceci peut être constaté en multipliant l'entrée du système à identifier par un facteur scalaire c et en divisant la réponse impulsionnelle par ce même facteur scalaire. Le système ayant pour entrée $ca(k)$ et pour réponse impulsionnelle $\mathbf{h}(k)/c$ possède la même sortie que le système ayant pour entrée $a(k)$ et pour réponse impulsionnelle $\mathbf{h}(k)$. Ceci est dû au fait que lorsqu'on observe les statistiques de la sortie du système, on ne peut pas déterminer le module et la phase des signaux de son entrée.

1.3.4 Modèles des Symboles

Selon les hypothèses faites sur les symboles transmis (quantités déterministes ou stochastiques), nous distinguons deux types de modèles des symboles.

Dans le *modèle déterministe*, les symboles émis A et les coefficients du canal \mathbf{h} sont supposés être des quantités déterministes inconnues. L'estimation conjointe de \mathbf{h} et A est faite à partir de l'équation 1.11. D'un point de vue identifiabilité, le module et la phase du canal ne peuvent pas être déterminés. En choisissant la contrainte de non trivialité sur les coefficients du canal, $\|\mathbf{h}\| = 1$, l'ambiguïté au niveau du canal ne concerne plus que sa phase. Les méthodes d'estimation multi-canal ainsi que les bornes de performance (bornes de Cramer-Rao), qui sont dérivées dans le cadre de ce modèle, sont qualifiées de déterministes. Dans les chapitres suivants, nous décrivons la classe des méthodes déterministes, nous calculons la borne de Cramer-Rao déterministe et nous comparons les performances de ces méthodes entre elles et par rapport à la borne de Cramer-Rao.

Dans le cas du *modèle gaussien*, les symboles transmis sont considérés comme des variables aléatoires i.i.d. gaussiennes : $a(k) \sim \mathcal{N}(0, \sigma_a^2)$. Les paramètres à estimer conjointement sont les coefficients du canal et la puissance du bruit σ_v^2 . Contrairement au modèle déterministe où le canal peut être déterminé uniquement à un facteur scalaire près, dans le cas de symboles gaussiens le module du canal peut être estimé. Le canal \mathbf{h} est identifiable à un facteur de phase continue près. L'hypothèse de symboles gaussiens permet donc de renforcer le problème d'identification par rapport au modèle déterministe en réduisant le degré de l'ambiguïté d'estimation sur le canal. Dans [dCS97a], les méthodes de maximum de vraisemblance basées sur le modèle gaussien aboutissent à des performances

supérieures à celles qui se basent sur le modèle déterministe. Une approche de base dans le modèle gaussien, est GML (Gaussian Maximum Likelihood). Une autre approche est la méthode “covariance matching” [GH97]. Dans le chapitre 3, nous calculons la borne de Cramer-Rao gaussienne et une discussion sur les méthodes d’estimation multi-canal gaussiennes est présentée dans le chapitre 4.

1.4 Contributions de la Thèse

Dans cette section, nous présentons un aperçu sur chaque chapitre de la thèse ainsi que les contributions qu’il contient.

Chapitre 2

Ce chapitre trace les liens qui existent entre les sous-espaces et les paramétrisations. Nous partons de l’exemple des sinusoïdes dans le bruit pour clarifier les sous-espaces signal et bruit correspondants. Nous déduisons, ensuite, les similitudes qui existent entre ce cas et l’identification aveugle. Nous caractérisons ainsi le sous-espace signal et le sous-espace-bruit dans le cas de l’identification aveugle multi-canal. Nous décrivons les différentes paramétrisations du sous-espace bruit et leur relation avec la génération (partielle ou totale) de ce sous-espace.

Nous introduisons une nouvelle paramétrisation du sous-espace bruit que nous appelons $\mathbf{H}_{bal,min}^{\dagger}(z)$ et qui contient exactement m lignes (m étant le nombre de canaux considérés). Nous démontrons que l’utilisation de cette paramétrisation dans une méthode d’estimation aveugle multi-canal permet d’engendrer le sous-espace bruit de façon complète. Nous comparons, à la fin de ce chapitre, notre paramétrisation avec celle d’Abed-Meraim et Hua ([AMH97a],[AMH97b]) qui possède $(2m - 3)$ lignes et permet aussi d’engendrer la totalité du sous-espace bruit pourvu que les deux premiers sous-canaux $H_1(z)$ et $H_2(z)$ n’aient pas de zéros communs.

Chapitre 3

Ce chapitre utilise les résultats du chapitre 2. Dans un contexte de symboles déterministes, nous décrivons des méthodes d’estimation multi-canal aveugles qui utilisent explicitement la notion de sous-espaces. Nous illustrons le fait que si les statistiques de second ordre aboutissent à de bonnes performances en termes d’estimées du canal, ce n’est autre qu’une conséquence de la bonne estimation du sous-espace signal avec un nombre fini de données. Nous calculons, ensuite, les bornes de Cramer-Rao (déterministe et gaussienne) qui serviront de mesure de performance pour les méthodes décrites dans cette thèse. Nous formulons la méthode SRM (Subchannel Response Matching) dans le cas de la paramétrisation $\mathbf{H}_{bal,min}^{\dagger}(z)$ et nous décrivons les deux méthodes : ajustement du sous-espace signal et ajustement du sous-espace bruit.

Les méthodes d'ajustement de sous-espaces présentent l'avantage d'être des procédures d'estimation de canal non itératives. Un inconvénient inhérent à ces méthodes est leur coût de calcul élevé dû à la décomposition propre de la matrice de covariance du signal reçu. Motivés par la réduction de ce coût de calcul, nous proposons des méthodes alternatives pour l'ajustement du sous-espace signal et l'ajustement du sous-espace bruit, qui ne nécessitent pas la décomposition propre de la matrice de covariance. L'analyse de performance des méthodes proposées montre que les performances asymptotiques des méthodes alternatives restent étroitement proches de celles des méthodes classiques.

Certains résultats décrits dans ce chapitre ont été présentés à la conférence *the 13th international conference on Digital Signal Processing (DSP 97)*:

L. Deneire, J. Ayadi and D.T.M. Slock. "Subspace Fitting Without Eigendecomposition". In *Proc. of the 13th international conference on Digital Signal Processing: DSP 97*, Santorini, Greece, July 1997.

Nous avons aussi dérivé les versions optimalement pondérées des méthodes alternatives pour l'ajustement du sous-espace signal et l'ajustement du sous-espace bruit sans décomposition propre de la matrice de covariance.

Ces résultats ont été présentés à la conférence *the IEEE 5th Symposium on Communications and Vehicular Technology in the Benelux*:

L. Deneire, J. Ayadi and D.T.M. Slock. "Weighted and Unweighted Subspace Fitting without Eigendecomposition". In *Proc. of the IEEE 5th Symposium on Communications and Vehicular Technology in the Benelux*, Twente, Netherlands, October 1997.

Chapitre 4

Dans ce chapitre, nous nous intéressons à l'estimation aveugle basée sur le critère du maximum de vraisemblance. Notre motivation trouve son origine dans le fait que l'estimateur de maximum de vraisemblance est asymptotiquement efficace (asymptotiquement, cet estimateur atteint sa borne de Cramer-Rao). Par conséquent, asymptotiquement, la performance d'une méthode d'estimation basée sur ce critère est supposée être supérieure à celles des autres méthodes déterministes.

Le critère DML (Deterministic Maximum Likelihood) dans le cas de l'estimation aveugle multi-canal est présenté. Une méthode pour résoudre ce critère est la stratégie IQML (Iterative Quadratic ML). L'inconvénient principal de cette méthode est qu'en présence du bruit, elle aboutit à des estimées biaisées du canal. Pour enlever ce biais, nous proposons de "débruiter" le critère ML et de le résoudre selon une stratégie que nous appelons DIQML (Denoised IQML). Nous avons aussi introduit une formulation différente du critère ML de façon à ce que son gradient corresponde exactement au gradient du critère ML. Nous appelons cette approche PQML (Pseudo-Quadratic ML). La relation qui existe entre les approches DIQML et PQML a été établie. Les performances

excellentes des deux algorithmes DIQML et PQML sont dues à un facteur d'ajustement particulier pour lequel nous avons introduit un choix judicieux.

Motivés par les performances des algorithmes DIQML et PQML, nous avons essayé de bénéficier de cette robustesse aveugle pour dériver des algorithmes semi-aveugles en combinant la fonction de coût aveugle avec une fonction de coût basée sur une séquence d'apprentissage. La première approche semi-aveugle que nous avons dérivée est PQML-LS (PQML Least Squares) qui se base sur un point de vue purement déterministe. La seconde approche est PQML-WLS (PQML Weighted LS) qui combine un point de vue déterministe et un point de vue stochastique. Les résultats obtenus montrent que PQML-WLS aboutit à des performances meilleures que celles de PQML-LS et que ces performances sont étroitement proches de la borne de Cramer-Rao Correspondante.

Ces résultats ont été présentés à la conférence *the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 98)* :

J. Ayadi, E. de Carvalho and D.T.M. Slock. “Blind and Semi-blind Maximum Likelihood Methods for FIR Multichannel Identification”. In *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing: ICASSP 98*, Seattle, Washington, USA, May 1998.

Dans la seconde partie de ce chapitre, nous exploitons l'alphabet fini des symboles transmis. Nous incorporons la connaissance a priori de l'alphabet fini en combinant un critère d'estimation aveugle avec un critère d'égalisation MMSE-DD (Minimum Mean Square Error Decision Directed). Les résultats obtenus nous permettent de constater qu'une bonne initialisation de l'algorithme considéré est nécessaire et qu'une simple stratégie DD-LS, quand elle est bien initialisée, aboutit à la meilleure performance.

Les résultats établis dans cette partie du chapitre ont été présentés à la conférence *the European Signal Processing Conference (EUSIPCO 98)* :

J. Ayadi and D.T.M. Slock. “Blind Multichannel Estimation Exploiting the Finite Symbol Alphabet”. In *Proc. of the European Signal Processing Conference: EUSIPCO 98*, Island of Rhodes, Greece , September 1998.

Chapitre 5

Dans ce chapitre, nous considérons l'identification aveugle multi-canal exploitant la connaissance a priori des filtres de transmission. En effet, le canal total (ou global) inclut le filtre d'émission (TX), le canal de propagation et le filtre de réception (RX). Comme les filtres de TX/RX sont des entités connues dans un système de communications numériques, les paramètres à estimer sont en réalité ceux qui correspondent au canal de propagation. Nous commençons par récapituler toutes les méthodes d'estimation multi-canal (celles qui sont basées sur les sous-espaces et celles de maximum de vraisemblance) qui sont formulées sous une forme unifiée dans laquelle chaque méthode devient paramétrisée linéairement par le canal. Comme la connaissance des filtres de

TX/RX est exprimée en termes de la réponse impulsionnelle du canal, nous décrivons le formalisme d'exploitation de cette connaissance et nous étendons les méthodes aveugles en incorporant cette information. Nous calculons, ensuite, les bornes de Cramer-Rao avec connaissance a priori sur les filtres de TX/RX (déterministe et gaussienne).

Les résultats obtenus montrent que l'exploitation de la connaissance des filtres de TX/RX aboutit à une amélioration significative des performances en termes de borne de Cramer-Rao, de performance de la méthode elle-même et d'identifiabilité. En effet, des canaux qui ne sont pas identifiables par des méthodes purement aveugles, peuvent être identifiés en exploitant la connaissance des filtres de TX/RX. Cette amélioration de performance est due à la réduction du nombre des paramètres à estimer dans le cas de l'exploitation de la connaissance des filtres de TX/RX. Cette réduction nécessitera moins de puissance du signal ou une longueur du paquet de données plus courte pour identifier le canal. D'un autre côté, pour les canaux mal-conditionnés (qui ne sont donc pas identifiables par les méthodes purement aveugles), lorsqu'on exploite la connaissance des filtres de TX/RX, la matrice du critère à optimiser se trouve bien conditionnée et par conséquent l'estimation du canal aboutit à de bonnes performances.

Les résultats établis dans ce chapitre ont été présentés à la conférence *IEEE First Signal Processing Advances in Wireless Communications Workshop (SPAWC 97)* et ont fait l'objet d'un papier de journal accepté pour être publié à *Signal Processing (EURASIP) - Special Issue on Signal Processing Technologies for Short-Burst Wireless Communications*:

J. Ayadi and D.T.M. Slock. "Cramer-Rao Bounds and Methods for Knowledge Based Estimation of Multiple FIR Channels". In *Proc. of the IEEE First Signal Processing Advances in Wireless Communications Workshop: SPAWC 97*, Paris, France, April 1997.

J. Ayadi and D.T.M. Slock. "Blind Channel Estimation Exploiting Transmission Filter Knowledge". To appear in *Signal Processing (EURASIP) - Special Issue on Signal Processing Technologies for Short-Burst Wireless Communications*.

Chapitre 6

Les méthodes d'estimation aveugle multi-canal déterministes présentées dans les chapitres 3 et 4 présentent, toutes, l'inconvénient du manque de robustesse à la surdétermination de l'ordre du canal: dès que l'ordre est surdéterminé, l'estimée du canal obtenue n'est pas consistante. Certaines solutions proposées pour pallier ce problème consistent à détecter l'ordre du canal en dehors de la procédure d'estimation puis adopter cet ordre dans l'algorithme d'identification. D'autres solutions suggèrent d'effectuer la procédure d'estimation à tous les ordres possibles du canal et d'en considérer celui qui aboutit à la meilleure estimation. Nous nous sommes inspirés de la prédiction linéaire (qui est robuste à la surdétermination de l'ordre du canal) pour reformuler le critère d'estimation de façon à ce qu'il devienne analogue à celui de la prédiction linéaire.

Une des propriétés qui aboutit à la robustesse à la surestimation de l'ordre du canal, est l'ancrage de la réponse impulsionnelle à son premier coefficient vectoriel $\mathbf{h}(0)$. Cette opération d'ancrage doit, ensuite, être suivie par un traitement approprié effectué sur la matrice du critère à optimiser. Ce traitement peut s'opérer selon deux approches. La première approche exploite la structure de la matrice du critère et peut se résumer en une détection implicite de l'ordre du canal. Cependant la seconde approche présente un mécanisme de détection explicite de l'ordre du canal en se basant sur l'algorithme de Levinson multi-canal adapté au cas singulier. Les résultats de simulations montrent que les deux approches permettent de préserver la robustesse des méthodes d'identification multi-canal à la surdétermination de l'ordre du canal.

Les résultats obtenus dans le cas de la première approche ont été présentés à la conférence *the 13th international conference on Digital Signal Processing (DSP 97)*:

J. Ayadi, L. Deneire and D.T.M. Slock. "Channel-based Blind Multichannel Identification Without Order Overestimation Problems". In *Proc. of the 13th international conference on Digital Signal Processing: DSP 97*, Santorini, Greece, July 1997.

Les résultats obtenus dans le cas de la seconde approche ont été présentés à la conférence *16^{ème} colloque GRETSI sur le traitement du signal et des images*:

J. Ayadi, L. Deneire and D.T.M. Slock. "Approches Robustes à la Surdétermination de l'Ordre du Canal pour l'Estimation Aveugle de Canaux Multiples RIF". In *Proc. of the 16^{ème} colloque GRETSI sur le traitement du signal et des images*, Grenoble, France, September 1997.

Chapitre 7

Dans ce chapitre, nous considérons le cas de l'estimation aveugle en présence d'un bruit stationnaire de couleur inconnue. Nous décrivons, dans ce contexte, une méthode d'ajustement de sous-espace pour identifier le canal. L'idée de base consiste à remarquer que dans le cas du suréchantillonnage, le signal d'intérêt est cyclostationnaire, alors que le bruit est stationnaire. Nous considérons le cas où l'aspect multi-canal provient d'un réseau de n antennes et du suréchantillonnage du signal reçu avec un facteur m . Dans ce cas, la matrice de covariance du bruit stationnaire est bloc Toeplitz dont les blocs sont de dimension $n \times n$. Cependant les matrices de covariance de la partie signal et du signal reçu total sont bloc Toeplitz dont les blocs sont de dimension $mn \times mn$.

Nous considérons une structure de déplacement de la matrice de covariance, dans laquelle la contribution du bruit disparaît, pour développer une méthode d'identification aveugle du canal basée sur l'ajustement de sous-espace. Nous calculons, ensuite, les bornes de Cramer-Rao (déterministe et gaussienne) pour ce cas spécifique du modèle où le bruit est supposé être uniquement stationnaire. Comme l'élimination du bruit stationnaire dans l'approche d'ajustement de sous-espace est basée sur le suréchantillonnage et

par conséquent sur l'excès en largeur de bande, nous étendons la méthode proposée pour exploiter la connaissance des filtres de TX/RX afin d'améliorer sa performance.

Les résultats obtenus dans ce chapitre ont été présentés à la conférence *the 9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing (SSAP 98)* :

J. Ayadi and D.T.M. Slock. "Blind Multichannel Identification in the Stationary Colored Noise Case". In *Proc. of the 9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing: SSAP 98*, Portland, Oregon, USA, September 1998.

Chapitre 8

Dans ce chapitre, nous décrivons une méthode d'estimation de canal basée sur l'égalisation aveugle MMSE ZF (MMSE Zero Forcing). Nous partons du principe de Capon [Cap69] pour la formation de voies LCMV (Linearly Constrained Minimum Variance) et nous indiquons comment l'appliquer de façon correcte à l'égalisation multi-canal. Nous montrons ensuite l'équivalence entre le critère MMSE et MOE (Minimum Output Energy) dans le cas sans biais et nous prouvons, dans le cas sans bruit, l'égalité entre les égaliseurs : UMOE (MOE non biaisé), MOE ZF, UMMSE (MMSE non biaisé) et MMSE ZF. Alors que le but est de faire MMSE ZF, il est plus facile mais équivalent d'abord le problème par UMMSE appliqué à la composante signal du signal reçu. Donc la matrice de covariance doit être "débruitée" avant son utilisation dans la méthode d'estimation aveugle. L'avantage essentiel de notre méthode est qu'elle évite la décomposition propre de la matrice de covariance, tout en aboutissant à des performances équivalentes à celles de l'algorithme PQML.

Les résultats établis dans cette partie de ce chapitre ont été présentés à la conférence *the IEEE Second Signal Processing Advances in Wireless Communications Workshop (SPAWC 99)* :

J. Ayadi and D.T.M. Slock. "Multichannel Estimation by Blind MMSE ZF Equalization". In *Proc. of the IEEE Second Signal Processing Advances in Wireless Communications Workshop: SPAWC 99*, Annapolis, Maryland, USA, May 1999.

Nous proposons ensuite une procédure simple pour incorporer la détection de l'ordre du canal dans notre approche. En effet, la méthode proposée permet de trouver le canal comme étant le vecteur propre correspondant à la valeur propre maximale d'une matrice de covariance d'erreurs de prédiction dont le rang est, théoriquement, égal à un. Par conséquent, la détection de l'ordre du canal revient à trouver l'ordre pour lequel le rang de cette matrice s'approche le plus d'une matrice de rang un.

Les résultats obtenus dans cette partie de ce chapitre ont été présentés aux conférences *17^{ème} colloque GRETSI sur le traitement du signal et des images* et *the IEEE Vehicular Technology Conference (VTC 99)* :

Jaouhar Ayadi and Dirk T.M. Slock. "Estimation Multi-canal Aveugle par Egalisation MMSE ZF avec Détection Conjointe de l'Étalement Temporel". In *Proc.*

of the 17^{ème} colloque GRETSI sur le traitement du signal et des images, Vannes, France, September 1999.

J. Ayadi and D.T.M. Slock. “Blind Multichannel Estimation and Joint Order Detection by MMSE ZF Equalization”. In *Proc. of the IEEE Vehicular Technology Conference: VTC 99*, Amsterdam, Netherlands, September 1999.

Chapitre 2

Sous-espaces et Paramétrisations

Dans ce chapitre, nous expliquons la notion de sous-espaces dans le cas de l'identification aveugle multi-canal. Ces sous-espaces sont décrits par des paramétrisations dont certaines sont linéaires en termes de réponse impulsionnelle du canal. Nous décrivons les différentes paramétrisations pouvant être utilisées ainsi que leur relation avec la génération du sous-espace en question. Nous introduisons, ensuite, une nouvelle paramétrisation linéaire du sous-espace bruit qui permet d'engendrer complètement ce sous-espace et nous la comparons à la paramétrisation introduite dans [AMH97a],[AMH97b] qui, à son tour, permet d'engendrer le sous-espace bruit.

2.1 Introduction

L'équation (1.11) qui décrit le système met en évidence une relation matricielle qui régit les signaux de sortie et ceux d'entrée. Cette relation permet de faire appel à des techniques de l'algèbre linéaire pour résoudre le problème de l'estimation des paramètres du système. La notion de *sous-espace* est fondamentale en algèbre linéaire. Un sous-espace se définit comme étant un sous-ensemble d'un espace vectoriel qui vérifie certaines propriétés. Un exemple qui illustre bien la notion de sous espace est celui des sinusoides dans le bruit [Slo97]. Dans ce qui suit, nous présentons un bref rappel sur ce cas ainsi que les différents sous-espaces et les méthodes d'estimation paramétriques qui sont utilisés pour estimer les fréquences des différents sinusoides. Nous déduisons, ensuite, les similitudes qui existent entre le cas des sinusoides dans le bruit et celui de l'identification aveugle des paramètres au niveau des sous-espaces introduits. Nous analysons les différentes paramétrisations introduites dans le cas de l'estimation aveugle multi-canal et nous discutons les différents liens qui existent entre elles, ainsi que la relation qui existe entre l'utilisation d'une paramétrisation donnée et la génération du sous-espace correspondant.

2.2 Sinusoïdes dans le Bruit

Considérons un signal x_k qui s'écrit comme la somme de sinusoïdes,

$$x_k = \sum_{i=1}^M A_i \cos(\omega_i k + \phi_i), \quad y_k = x_k + v_k. \quad (2.1)$$

Nous supposons que les amplitudes A_i et les fréquences $f_i = \omega_i/2\pi$ sont des inconnues déterministes et que les phases ϕ_i sont des inconnues aléatoires indépendantes entre elles et uniformément distribuées sur $[0, 2\pi[$. Le signal x_k est donc un processus stationnaire de moyenne nulle. Le signal y_k est une version bruitée du signal x_k , où nous supposons que le bruit v_k est blanc, indépendant et ayant une variance σ_v^2 . Pour la composante purement signal x_k , la densité spectrale de puissance

$$S_{xx}(f) = \sum_{i=1}^M \frac{A_i^2}{2} [\delta(f - f_i) + \delta(f + f_i)] \quad (2.2)$$

est de mesure nulle. Les fréquences et les phases peuvent être, facilement, déterminées à partir de $S_{xx}(f)$.

2.2.1 Prédiction Linéaire

Une somme de sinusoïdes satisfait l'équation suivante

$$P(q)x_k = 0, \quad P(z) = \prod_{i=1}^M (1 - 2 \cos \omega_i z^{-1} + z^{-2}), \quad (2.3)$$

q^{-1} étant l'opérateur retard défini tel que $q^{-1}x_k = x_{k-1}$. Par conséquent le signal x_k est parfaitement prédictible à partir de ses $2M$ échantillons précédents. Le polynôme $P(z)$, et donc les ω_i , peuvent être déterminés par prédiction linéaire (méthode de Prony). Les équations normales qui décrivent ce problème sont les suivantes :

$$PR_{XX} = \begin{bmatrix} 0 & \cdots & 0 & \sigma_v^2 \end{bmatrix}, \quad \sigma_v^2 = 0, \quad (2.4)$$

où

$$\begin{aligned} R_{XX} &= EXX^T, X = [x_0 \cdots x_{2M}]^T, \\ P &= \begin{bmatrix} P_{2M} & \cdots & P_1 & P_0 \end{bmatrix}, \\ P_0 &= 1, \quad P_i = P_{2M-i}, \quad i = 0, \dots, M-1. \end{aligned} \quad (2.5)$$

2.2.2 Sous-espace Signal(SS)/Sous-espace Bruit(SB)

Considérons le vecteur $X_k = [x_0 x_1 \cdots x_k]^T$. Ce vecteur peut s'écrire sous forme de produit matriciel comme suit

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 1 & 0 \\ \cos \omega_1 & \sin \omega_1 & \cdots & \cos \omega_M & \sin \omega_M \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos \omega_1 k & \sin \omega_1 k & \cdots & \cos \omega_M k & \sin \omega_M k \end{bmatrix} \begin{bmatrix} A_1 \cos \phi_1 \\ -A_1 \sin \phi_1 \\ \vdots \\ A_M \cos \phi_1 \\ -A_M \sin \phi_1 \end{bmatrix}. \quad (2.6)$$

Par conséquent, $X_k = \mathcal{V}S$ où $\mathcal{V} = \mathcal{V}(\Omega)$, avec $\Omega = [\omega_1 \cdots \omega_M]$, est une matrice bloc de Vandermonde (en commençant par la première colonne, deux colonnes consécutives de \mathcal{V} forment une colonne bloc). Notons que deux éléments bloc consécutifs d'une colonne bloc de \mathcal{V} sont proportionnels, en effet

$$\begin{bmatrix} \cos \omega_i k & \sin \omega_i k \end{bmatrix} = \begin{bmatrix} \cos \omega_i (k-1) & \sin \omega_i (k-1) \end{bmatrix} \begin{bmatrix} \cos \omega_i & \sin \omega_i \\ -\sin \omega_i & \cos \omega_i \end{bmatrix} \quad (2.7)$$

L'espace image de \mathcal{V} est appelé *sous-espace signal* et son complément orthogonal est appelé *sous-espace bruit*.

$$\begin{aligned} \text{Image} \{ \mathcal{V} \} &= \text{SS (Sous-espace Signal)}, \\ (\text{Image} \{ \mathcal{V} \})^\perp &= \text{SB (Sous-espace Bruit)}. \end{aligned} \quad (2.8)$$

Pour le signal bruité, nous obtenons

$$Y_k = X_k + V_k = \mathcal{V}S + V_k \Rightarrow R_{YY} = \mathcal{V}R_{SS}\mathcal{V}^T + \sigma_v^2 I, \quad (2.9)$$

où Y_k et V_k sont définis de façon similaire à X_k et $R_{SS} = \frac{1}{2} \text{diag} \{ A_1^2, A_1^2, \dots, A_M^2, A_M^2 \}$. Notons que la composante signal X_k du signal Y_k peut "vivre" uniquement dans le SS, alors que la composante bruit V_k vit dans le SB. Considérons la décomposition propre de la matrice R_{YY} , où les valeurs propres λ_i sont ordonnées dans un ordre décroissant ($\lambda_1 \geq \lambda_2 \geq \dots$). La matrice de covariance R_{YY} peut s'écrire comme suit :

$$R_{YY} = \sum_{i=1}^{2M} \lambda_i V_i V_i^H + \sum_{i=2M+1}^{k+1} \lambda_i V_i V_i^H = V_S \Lambda_S V_S^H + V_N \Lambda_N V_N^H, \quad (2.10)$$

où $V_S = [V_1 \cdots V_{2M}]$, $V_N = [V_{2M+1} \cdots V_{k+1}]$, $\Lambda_S = \text{diag} \{ \lambda_1 \cdots \lambda_{2M} \}$ et $\Lambda_N = \sigma^2 I_{k+1-2M}$. Supposons que \mathcal{V} et R_{SS} sont de rang plein (tous les $\omega_i \in]0, \pi[$ sont différents, tous les $A_i > 0$), les vecteurs propres V_S et V_N sont orthogonaux : $V_S^H V_N = 0$, et $\lambda_i > \sigma_v^2$, $i = 1 \cdots 2M$. D'où les descriptions équivalentes du SS et du SB suivantes

$$\begin{aligned} \text{SS} &= \text{Image} \{ V_S \} = \text{Image} \{ \mathcal{V}(\Omega) \}, \\ \text{SB} &= \text{Image} \{ V_N \} \text{ tel que } V_N^T \mathcal{V} = 0. \end{aligned} \quad (2.11)$$

En conclusion, nous disposons d'une description paramétrique du SS et nous pouvons déterminer le SS et le SB à partir de la décomposition propre de la matrice de covariance R_{YY} .

Nous pouvons aussi trouver une description du SB basée sur les quantités de la prédiction linéaire. Nous avons $P(q) \cos \omega_i k = 0$ et $P(q) \sin \omega_i k = 0$, d'où $\mathcal{G}(P)^T \mathcal{V} = 0$, où $\mathcal{G}(P)$ est défini comme suit

$$\mathcal{G}(p)^T = \mathcal{T}_{k+1-2M}(P) = \begin{bmatrix} P_{2M} & \cdots & P_1 & P_0 & 0 & \cdots & 0 \\ 0 & P_{2M} & \cdots & P_1 & P_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & P_{2M} & \cdots & P_1 & P_0 \end{bmatrix}. \quad (2.12)$$

$\mathcal{G}(p)^T$ est une matrice Toeplitz de rang plein égal à la dimension du SB. Par conséquent,

$$\text{SB} = \text{Image}(\mathcal{G}(p)). \quad (2.13)$$

Notons que le SS, $\mathcal{V}(\Omega)$, et le SB, $\mathcal{G}(p)$ sont paramétrisés par M paramètres indépendants.

Dans ce qui suit, nous utilisons le SS et le SB pour dériver des méthodes, dites d'*ajustement* de sous-espaces, pour résoudre le problème d'estimation des fréquences à partir des statistiques de second ordre.

2.2.3 Ajustement du SS

Nous disposons de deux descriptions théoriques équivalentes du SS: $\mathcal{V}(\Omega)$ et V_S , les deux matrices possèdent le même espace colonne, ainsi une matrice des deux peut être transformée en l'autre. La matrice V_S peut être déterminée à partir de la matrice de covariance R_{YY} (estimée). En ajustant $\mathcal{V}(\Omega)$ à V_S on peut déterminer Ω . En d'autres termes, on cherche à déterminer Ω comme étant l'ensemble des paramètres qui ajuste *au mieux* $\mathcal{V}(\Omega)$ à V_S au sens d'un certain critère. Nous considérons donc le problème de minimisation suivant

$$\min_{\Omega, T} \|\mathcal{V}(\Omega) - V_S T\|_F \quad (2.14)$$

où T est une matrice de transformation carrée et F désigne la norme de Frobenius définie comme $\|A\|_F^2 = \text{tr} A^H A$, et tr désigne trace. La valeur minimale du critère vaut zéro si V_S est exact. Le critère de minimisation exprimé dans (2.14) est séparable. En particulier, il est quadratique en T . La minimisation par rapport à T donne: $T = V_S^T \mathcal{V}$ et $\mathcal{V} - V_S T = P_{V_S}^\perp \mathcal{V}$ où $P_{V_S}^\perp = I - P_{V_S}$ et P_{V_S} est la matrice de projection sur le sous-espace signal (V_S). D'où

$$\begin{aligned} \min_T \|\mathcal{V}(\Omega) - V_S T\|_F^2 &= \|P_{V_S}^\perp \mathcal{V}\|_F^2 = \text{tr} \{ \mathcal{V}^T P_{V_S}^\perp \mathcal{V} \} = \text{tr} \{ \mathcal{V}^T P_{V_S} \mathcal{V} \} = \|V_N^T \mathcal{V}\|_F^2 \\ &= \sum_{i=2M+1}^k \|V_i^T \mathcal{V}\|^2 = \sum_{j=1}^M \sum_{i=2M+1}^k |V_i(\omega_j)|^2 \end{aligned} \quad (2.15)$$

où $V_i(\omega) = \begin{bmatrix} 1 & e^{j\omega} & \dots & e^{j\omega k} \end{bmatrix} V_i$ est la transformée de Fourier de la séquence d'éléments de V_i . L'expression dans (2.15) doit être minimisée par rapport à ω_i . Une solution approximative consiste à chercher les ω_i comme étant les M plus grand pics de

$$\frac{1}{\sum_{i=2M+1}^k |V_i(\omega)|^2}. \quad (2.16)$$

Cette méthode est appelée "Multiple Signal Classification" (MUSIC) [BK83].

2.2.4 Ajustement du SB

De façon similaire au SS, nous disposons de deux descriptions théoriques équivalentes du SB qui sont $\mathcal{G}(P)$ et V_N (déterminé à partir de R_{YY}). En ajustant $\mathcal{G}(P)$ à V_N , nous pouvons déterminer P . Nous introduisons le critère de minimisation suivant

$$\min_{P,T} \|\mathcal{G}(P) - V_N T\|_F. \quad (2.17)$$

En minimisant par rapport à T en premier lieu, nous obtenons : $T = V_N^T \mathcal{G}$ et $\mathcal{G} - V_N T = P_{V_N}^\perp \mathcal{G}$. Donc (2.17) devient

$$\min_{P,T} \|\mathcal{G}(P) - V_N T\|_F = \|P_{V_N}^\perp \mathcal{G}\|_F^2 = \text{tr} \{ \mathcal{G}^T P_{V_N}^\perp \mathcal{G} \} = \text{tr} \{ \mathcal{G}^T P_{V_S} \mathcal{G} \} = \|V_S^T \mathcal{G}\|_F^2 = \sum_{i=1}^{2M} \|\mathcal{G}^T V_i\|^2 \quad (2.18)$$

Nous pouvons écrire, d'après la commutativité de la convolution, $\mathcal{G}^T V_i = \mathcal{H}_i P^T$, où la matrice $\mathcal{H}_i = \mathcal{H}(V_i)$ est une matrice de Hankel. La symétrie de P peut être exprimée comme $P = P J$, où J est la matrice carrée possédant des uns sur son antidiagonale principale et des zéros ailleurs. Nous pouvons garantir la symétrie de la solution en exprimant explicitement la symétrie de P . Ainsi la minimisation du critère exprimé dans (2.18) par rapport à P devient

$$\min_P P \left[\left(\sum_{i=2}^{2M} \mathcal{H}_i^T \mathcal{H}_i \right) + J \left(\sum_{i=1}^{2M} \mathcal{H}_i^T \mathcal{H}_i \right) J \right] P^T \quad (2.19)$$

tel que $P_0 = 1$ ou $\|P\| = 1$. Cette approche a été proposée dans [KM95].

2.3 Identification Aveugle du Canal

Considérons le cas où le bruit additif est blanc et indépendant de moyenne nulle et supposons que $E \mathbf{v} \mathbf{v}(k)^H = \sigma_v^2 I_m$ où σ_v^2 est la variance inconnue (dans le cas d'un bruit complexe, les parties réelle et imaginaire seront supposées non corrélées). Un bruit coloré avec une structure de corrélation connue mais une variance inconnue peut aussi

être utilisé. Pour un vecteur de mesures, \mathbf{Y} , de longueur quelconque, L , l'équation (1.9) peut s'écrire comme

$$\mathbf{Y}_L(k) = \mathcal{T}_L(\mathbf{h})A_{L+N-1}(k) + \mathbf{V}_L. \quad (2.20)$$

Par conséquent, la structure de la matrice de covariance du signal reçu $\mathbf{y}(k)$ est

$$R_L^{\mathbf{y}} = E\mathbf{Y}_L(k)\mathbf{Y}_L(k)^H = \mathcal{T}_L(\mathbf{h})R_{L+N-1}^a\mathcal{T}_L(\mathbf{h})^H + \sigma_v^2 I_{mL}. \quad (2.21)$$

La variance du bruit σ_v^2 peut être identifiée comme étant la plus petite valeur propre de $R_L^{\mathbf{y}}$ (c'est la méthode de Pisarenko [Sch91]). Dans le cas où le bruit additif est corrélé dont la séquence de corrélations est connue à un facteur de variance inconnu près $R_{mL}^v = \sigma_v^2 \overline{R}_{mL}^c$, l'approche que nous venons de décrire peut être facilement généralisée. En effet, nous pouvons considérer $\overline{R}_{mL}^{-v/2} R_L^{\mathbf{y}} \overline{R}_{mL}^{-vH/2}$ qui possède une forme similaire à celle donnée par (2.21), ayant les mêmes valeurs propres et dont les vecteurs propres sont transformés par $\overline{R}_{mL}^{-vH/2}$

$$R_L^{\mathbf{y}} V_i = \lambda_i \overline{R}_{mL}^c V_i. \quad (2.22)$$

En particulier, la plus petite valeur propre généralisée dans (2.22) est de nouveau σ_v^2 . Dans ce qui suit, nous supposons que $\overline{R}_{mL}^c = I_{mL}$.

2.3.1 Sous-espace Signal/Sous-espace Bruit

Considérons la matrice bloc Toeplitz $\mathcal{T}_L(\mathbf{h})$ de dimension $mL \times (L + N - 1)$. Dans le cas sans bruit, $\mathbf{Y}_L(k) = \mathcal{T}_L(\mathbf{h})A_{L+N-1}(k)$, si les colonnes de $\mathcal{T}_L(\mathbf{h})$ sont linéairement indépendantes, le vecteur $\mathbf{Y}_L(k)$ se trouve *confiné* dans un espace de dimension plus petite que le nombre de composantes de ce même vecteur. La condition à satisfaire par les colonnes de la matrice $\mathcal{T}_L(\mathbf{h})$ pour que cette dernière soit de rang plein, est donnée par le lemme suivant (voir [BKAK78] pour la démonstration)

Lemme 1 *En supposant que le canal est de longueur finie et que les $H_i(z), i = 1, \dots, m$ sont premiers entre eux, si $L \geq \underline{L} = \left\lceil \frac{N-1}{m-1} \right\rceil$ (avec probabilité 1 pour un canal FIR complètement aléatoire) alors*

$$\text{rang}(\mathcal{T}_L(\mathbf{h})) = L + N - 1.$$

Donc, avec les hypothèses du lemme, $\mathcal{T}_L(\mathbf{h})$ est de rang colonne plein. Le complément orthogonal de l'espace engendré par les colonnes de $\mathcal{T}_L(\mathbf{h})$ est de dimension $mL - (L + N - 1)$. L'espace colonne de $\mathcal{T}_L(\mathbf{h})$ est appelé *sous-espace signal* et son complément orthogonal est appelé *sous-espace bruit*.

2.3.2 Caractérisation de la Matrice de Covariance

Considérons la décomposition propre de la matrice de covariance $R_L^{\mathbf{y}}$ dans laquelle les valeurs propres (réelles et positives) sont ordonnées dans un ordre décroissant

$$R_L^{\mathbf{y}} = \sum_{i=1}^{L+N-1} \lambda_i V_i V_i^H + \sum_{i=L+N}^{mL} \lambda_i V_i V_i^H = \mathcal{V}_S \Lambda_S \mathcal{V}_S^H + \mathcal{V}_N \Lambda_N \mathcal{V}_N^H \quad (2.23)$$

où $\Lambda_N = \sigma_v^2 I_{mL-(L+N-1)}$. L'ensemble des vecteurs propres V_S et V_N sont orthonormaux: $V_S^H V_N = 0$, et $\lambda_i > \sigma_v^2$, $i = 1, \dots, L + N - 1$. Nous avons donc les descriptions équivalentes suivantes des sous-espaces signal et bruit

$$\begin{aligned} \text{SS} &= \text{Image} \{ \mathcal{V}_S \} = \text{Image} \{ \mathcal{T}_L(\mathbf{h}) \}, \\ \text{SB} &= \text{Image} \{ \mathcal{V}_N \} = (\text{Image} \{ \mathcal{T}_L(\mathbf{h}) \})^\perp. \end{aligned} \quad (2.24)$$

La description de la paramétrisation du sous-espace bruit que nous considérons ici est basée sur la réponse impulsionnelle du canal. En particulier, nous avons

$$\mathcal{V}_N^H \mathcal{T}_L(\mathbf{h}) = 0, \quad (\text{Image} \{ \mathcal{T}_L(\mathbf{h}) \})^\perp \mathcal{V}_S = 0. \quad (2.25)$$

L'approche sous-espace introduite par Moulines *et al.* [MDCM95] utilise les équations d'orthogonalité données par $\mathcal{V}_N^H \mathcal{T}_L(\mathbf{h}) = 0$ pour résoudre le problème d'estimation de canal.

Le SS est linéairement paramétrisé par le canal \mathbf{h} . Une paramétrisation du SB se déduit de $(\text{Image} \{ \mathcal{T}_L(\mathbf{h}) \})^\perp$. Bien que la matrice du canal $\mathcal{T}_L(\mathbf{h})$ soit bien spécifique en termes de structure, elle ne donne pas d'information sur la structure de son complément orthogonal, le SB. Dans ce qui suit, nous présentons les différentes paramétrisations du SB, nous analysons les liens entre elles et nous expliquons leur relation avec la génération du SB.

2.4 Paramétrisations du Sous-espace Bruit

Nous disposons d'une paramétrisation linéaire (en termes de la réponse impulsionnelle du canal) du SS. Néanmoins, nous cherchons à obtenir des paramétrisations linéaires du SB. Ceci se justifie par le fait que certaines méthodes d'estimation aveugle multi-canal utilisent explicitement les paramétrisations du SB (voir les chapitre 4 et 5). Pour dériver ces paramétrisations, nous partons de la structure de la matrice de convolution du canal $\mathcal{T}_L(\mathbf{h})$. Chaque ligne bloc de cette matrice est identique à la précédente mais translatée vers la droite d'une colonne. Supposons qu'il existe un vecteur ligne g_1 orthogonal à $\mathcal{T}_L(\mathbf{h})$

$$g_1 \mathcal{T}_L(\mathbf{h}) = 0. \quad (2.26)$$

D'après la structure de $\mathcal{T}_L(\mathbf{h})$, si les m derniers éléments dans g_1 sont nuls, nous pouvons construire un autre vecteur g_2 orthogonal à $\mathcal{T}_L(\mathbf{h})$ en translatant vers la droite les éléments de g_1 . Si les derniers éléments nuls dans g_1 sont en nombre supérieur à m , ce processus peut être répété. Slock [Slo94c] exploite cette structure cyclique et construit un prédicteur linéaire de la composante purement signal du signal reçu pour dériver une paramétrisation linéaire minimale du SB. Dans la section qui suit, nous décrivons la dérivation de cette paramétrisation.

2.4.1 Paramétrisation Linéaire Minimale du SB

Pour un vecteur aléatoire de moyenne nulle $Y = \begin{bmatrix} y_1^H & y_2^H & \cdots & y_M^H \end{bmatrix}^H$, nous introduisons la notation $y_{1:M} = Y$. Considérons l'orthogonalisation de Gram-Schmidt des composantes de Y . L'estimée au sens des moindres carrés (LS) \hat{y}_i de y_i étant donné $y_{1:i-1}$ ainsi que l'erreur d'estimation qui lui est associée \tilde{y}_i sont données par

$$\hat{y}_i = \hat{y}_i|_{y_{1:i-1}} = R_{y_i y_{1:i-1}} R_{y_{1:i-1} y_{1:i-1}}^{-1} y_{1:i-1}, \quad \tilde{y}_i = \tilde{y}_i|_{y_{1:i-1}} = y_i - \hat{y}_i \quad (2.27)$$

où $R_{ab} = Eab^H$, a et b étant deux vecteurs colonnes aléatoires. Le processus d'orthogonalisation de Gram-Schmidt consiste à générer successivement les éléments de $\tilde{Y} = \begin{bmatrix} \tilde{y}_1^H & \tilde{y}_2^H & \cdots & \tilde{y}_M^H \end{bmatrix}^H$ en commençant par $\tilde{y}_1 = y_1$. Nous pouvons écrire

$$LY = \tilde{Y} \quad (2.28)$$

où L est une matrice triangulaire inférieure dont les éléments diagonaux sont égaux à un. A partir de (2.28), nous obtenons

$$E(LY)(LY)^H = E\tilde{Y}\tilde{Y}^H \Rightarrow LR_{\tilde{Y}\tilde{Y}}L^H = D = R_{\tilde{Y}\tilde{Y}}^{-1}. \quad (2.29)$$

Puisque les éléments \tilde{y}_i ne sont pas corrélés, D est une matrice diagonale. L'équation (2.29) peut être écrite comme la factorisation triangulaire UDL de $R_{\tilde{Y}\tilde{Y}}^{-1}$

$$R_{\tilde{Y}\tilde{Y}}^{-1} = L^H D^{-1} L. \quad (2.30)$$

Considérons le cas sans bruit et supposons que les symboles ne sont pas corrélés : $R_M^a = \sigma_a^2 I_M$ (la paramétrisation du SB que nous obtiendrons reste aussi valable lorsque les symboles sont corrélés). Considérons l'orthogonalisation de Gram-Schmidt des éléments consécutifs (scalaires) du vecteur $\mathbf{Y}_L(k)$. Nous commençons par effectuer la factorisation UDL de la matrice $\left(R_L^{\mathbf{y}}\right)^{-1}$ et nous obtenons les filtres des erreurs de prédiction et les variances consécutifs. Aucune singularité n'est rencontrée jusqu'à la ligne bloc \underline{L} dans laquelle nous traitons les éléments de $\mathbf{y}(k - L + 1)$. Puisque $\mathcal{T}_{\underline{L}}(\mathbf{h})$ est de rang colonne plein, nous obtenons $\underline{m} = m\underline{L} - (\underline{L} + N - 1) \in \{0, 1, \dots, m - 2\}$ singularités. Si $\underline{m} > 0$, après orthogonalisation les éléments scalaires de \mathbf{Y} deviennent nuls : $\tilde{y}_i(k - \underline{L} + 1) = 0, i = m + 1 - \underline{m}, \dots, m$. Ainsi l'élément correspondant dans la matrice diagonale D devient aussi égal à zéro. Les lignes correspondantes dans la matrice triangulaire inférieure L seront désignées par *filtres de prédiction singuliers*.

Pour $L = \underline{L} + 1$, $\mathcal{T}_{\underline{L}+1}(\mathbf{h})$ possède m lignes supplémentaires par rapport à $\mathcal{T}_{\underline{L}}(\mathbf{h})$ mais uniquement une seule colonne supplémentaire. Le rang colonne croit donc de un. Par conséquent, $\tilde{y}_1(k - \underline{L})$ n'est pas nul en général alors que $\tilde{y}_i(k - \underline{L}) = 0, i = 2, \dots, m$ (nous supposons que $h_1(N - 1) \neq 0$, ceci peut être toujours garanti en changeant l'ordre des canaux puisque $\mathbf{h}(N - 1) \neq 0$). Notons que, puisque $\mathcal{T}_{\underline{L}}(\mathbf{h})$ est de rang colonne plein, l'orthogonalisation de $y_1(k - \underline{L})$ par rapport à $\mathbf{Y}_{\underline{L}}(k)$ est équivalente à celle de $y_1(k - \underline{L})$ par rapport à $A_{\underline{L}+N-1}(k)$. Puisque les symboles a_k sont supposés être non corrélés, nous

ne gardons que la composante de $y_1(k - \underline{L})$ dans la direction de $a_{k-\underline{L}-N+1}$: $y_1(k - \underline{L}) = h_1(N - 1)a_{k-\underline{L}-N+1}$. Comme l'erreur de prédiction est blanche, l'augmentation de la longueur du prédicteur n'améliorera plus la qualité de la prédiction. Donc $y_1(k - \underline{L}) = h_1(N - 1)a_{k-\underline{L}-N+1}$, $L \geq \underline{L}$ et les filtres de prédiction dans les lignes correspondantes de la matrice L seront des versions translatées de façon appropriée du filtre de prédiction dans la ligne $m\underline{L} + 1$. De façon similaire, pour les erreurs de prédiction égales à zéro, l'augmentation de la longueur du prédicteur n'améliorera plus la qualité de la prédiction. Donc $y_i(k - L) = 0, i = 2, \dots, m, L \geq \underline{L}$. Les autres filtres de prédiction qui sont en bas de la matrice L sont des versions translatées de façon appropriées des $m - 1$ premiers filtres de prédiction singuliers. En plus, les entrées dans ces $m - 1$ premiers filtres de prédiction singuliers qui apparaissent sous les "1s" (éléments "diagonaux") sont égales à zéro (pour les mêmes raisons précédentes). Par conséquent, nous obtenons une erreur de prédiction blanche de rang un avec une prédiction d'ordre fini. Par suite, le processus $\mathbf{y}(k)$ est *autorégressif* (AR). Après un effet transitoire fini, la matrice L devient une matrice triangulaire inférieure bloc Toeplitz.

Considérons, maintenant, $L > \underline{L}$, et construisons la matrice \mathcal{G}_L de dimension égale à $((m - 1)(L - \underline{L}) + \underline{m}) \times (mL)$ en collectant les filtres de prédiction singuliers consécutifs de la matrice L . L'espace ligne de la matrice \mathcal{G}_L est (la transposition) du SB. En effet, chaque filtre de prédiction singulier appartient au SB puisque $\mathcal{G}_L \mathcal{T}(\mathbf{h}) = 0$, toutes les lignes dans \mathcal{G}_L sont linéairement indépendantes (car elles sont un sous-ensemble d'une matrice triangulaire dont les éléments diagonaux sont égaux à un), et le nombre de lignes dans \mathcal{G}_L est égal à la dimension du SB. La matrice \mathcal{G}_L est une matrice bloc Toeplitz dans laquelle $m - 1 - \underline{m}$ lignes ont été omises. En fait, la matrice \mathcal{G}_L est paramétrisée par les $m - 1$ premiers filtres de prédiction singuliers. Si nous mettons les entrées des $m - 1$ filtres de prédiction dans un vecteur colonne G_N , nous pouvons écrire $\mathcal{G}_L(G_N)$. La longueur de G_N est

$$\underline{m}((\underline{L} - 1)m + m - \underline{m}) + (m - 1 - \underline{m})((L - 1)m + m - \underline{m} + 1) = mN - 1 \quad (2.31)$$

qui est égale au nombre de degrés de liberté réel de \mathbf{H} (le canal peut être déterminé à un facteur scalaire près, d'où le "-1"). En conclusion, $\mathcal{G}_L(G_N)$ représente une paramétrisation linéaire minimale du SB.

La paramétrisation du SB que nous venons de décrire est basée sur les quantités de la prédiction linéaire. Dans la section qui suit, nous nous intéressons à des paramétrisations linéaires du SB basées sur la réponse impulsionnelle du canal.

2.4.2 Paramétrisations Linéaires du SB Basées sur le Canal

Notre intérêt pour la dérivation des paramétrisations linéaires du SB basées sur la réponse impulsionnelle du canal se justifie par le fait que les fonctions de coût à minimiser dans certaines méthodes d'estimation aveugles multi-canal deviennent quadratiques (en terme des paramètres du canal à estimer) lorsqu'on utilise ces paramétrisations. Par conséquent, le problème de minimisation de ces fonctions de coût peut avoir des solutions

analytiques (ce qui évite de faire appel à des procédures de résolution itératives). Dans ce qui suit, nous adoptons un point de vue purement fréquentiel. Considérons le cas sans bruit et supposons que les symboles ne sont pas corrélés et possèdent une variance égale à σ_a^2 . La matrice de densité spectrale de puissance du processus vectoriel stationnaire $\mathbf{y}(k)$ est

$$S_{\mathbf{y}\mathbf{y}}(z) = \sigma_a^2 \mathbf{H}(z) \mathbf{H}^\dagger(z) \quad (2.32)$$

où $\mathbf{H}^\dagger(z) = \mathbf{H}^H(1/z^*)$. Dans le but d'expliquer la dérivation des paramétrisations linéaires du SB, considérons tout d'abord le cas de deux canaux : $m = 2$. Nous pouvons remarquer qu'en absence de bruit, nous avons $\mathbf{H}_2(z)y_1(k) - \mathbf{H}_1(z)y_2(k) = 0$, ce qui peut s'écrire sous forme matricielle comme $[\mathbf{H}_2(z) \quad -\mathbf{H}_1(z)] \mathbf{y}(k) = \mathbf{H}^{\perp\dagger}(z) \mathbf{y}(k) = \mathbf{H}^{\perp\dagger}(z) \mathbf{H}(z) a(k) = 0$. Pour un nombre de canaux égal à m quelconque, $\mathbf{H}^{\perp\dagger}(z)$ vérifie

$$\mathbf{H}^{\perp\dagger}(z) \mathbf{H}(z) = 0. \quad (2.33)$$

Pour $m > 2$, la construction de $\mathbf{H}^{\perp\dagger}(z)$ se fait en considérant des paires de (sous-)canaux. Le choix de $\mathbf{H}^{\perp\dagger}(z)$ est loin d'être unique. En effet, le sous-ensemble de paires à considérer, qui représente le nombre de lignes dans $\mathbf{H}^{\perp\dagger}(z)$, peut être choisi selon plusieurs façons. Le nombre de paires minimal est égal à $m-1$ alors que le nombre de paires maximal vaut $\frac{m(m-1)}{2}$, avec des paramétrisations correspondantes $\mathbf{H}_{min}^\perp(z)$ et $\mathbf{H}_{max}^\perp(z)$. Notons que le choix de $\mathbf{H}_{min}^\perp(z)$ n'est pas unique.

Nous avons introduit dans [AS97] la notion de paramétrisation *équilibrée* du SB : $\mathbf{H}_{bal}^\perp(z)$. L'aspect "équilibré" provient du fait que nous attribuons le même poids aux sous-canaux dans la construction de $\mathbf{H}^\perp(z)$. Nous qualifions $\mathbf{H}^\perp(z)$ de paramétrisation équilibrée si elle vérifie

$$\text{tr}\{\mathbf{H}^{\perp\dagger}(z) \mathbf{H}^\perp(z)\} = \alpha \mathbf{H}^\dagger(z) \mathbf{H}(z) \quad (2.34)$$

pour un certain scalaire réel α . Le nombre minimal de lignes dans $\mathbf{H}_{bal}^{\perp\dagger}(z)$ est égal à m auquel cas $\alpha = 2$. Le choix d'une telle paramétrisation $\mathbf{H}_{bal,min}^\perp(z)$ n'est pas unique. Notons que $\mathbf{H}_{max}^\perp(z)$ est équilibrée avec un scalaire $\alpha = m-1$. Dans la littérature ([AMCG⁺97], [BR94], [Hua96], [Smi96]), les méthodes d'estimation aveugle utilisant une paramétrisation linéaire du SB se basent généralement sur $\mathbf{H}_{max}^\perp(z)$. Certains travaux tels que ceux décrits dans [XLTK95] utilisent $\mathbf{H}_{min}^\perp(z)$. A titre d'exemple, nous avons

$$\mathbf{H}_{min}^{\perp\dagger}(z) = \begin{bmatrix} -\mathbf{H}_2(z) & \mathbf{H}_1(z) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{H}_m(z) & 0 & \cdots & \mathbf{H}_1(z) \end{bmatrix}, \quad (2.35)$$

$$\mathbf{H}_{bal,min}^{\perp\dagger}(z) = \begin{bmatrix} -\mathbf{H}_2(z) & \mathbf{H}_1(z) & 0 & \cdots & 0 \\ 0 & -\mathbf{H}_3(z) & \mathbf{H}_2(z) & \cdots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \mathbf{H}_m(z) & 0 & \cdots & 0 & -\mathbf{H}_1(z) \end{bmatrix}. \quad (2.36)$$

Le choix d'une paramétrisation se répercute directement sur la génération du SB : un choix particulier peut engendrer la totalité du SB, alors qu'un autre choix peut donner lieu à un ensemble non complet de vecteurs engendrant le SB. Nous proposons de discuter et d'analyser ces différents choix dans la section qui suit.

2.4.3 Paramétrisations et Génération du SB

Considérons un vecteur d'observation \mathbf{Y} de longueur M . Par construction, la paramétrisation linéaire minimale introduite par Slock [Slo94c] produit le nombre exact de vecteurs linéairement indépendants nécessaires pour engendrer le SB. Ce nombre n'est autre que la dimension du SB qui est égale à $mM - (M + N - 1)$. Dans ce qui suit, nous nous concentrons sur les paramétrisations linéaires du SB basées sur le canal. Une paramétrisation $\mathbf{H}^{\perp\dagger}(z)$ peut être écrite comme la transformée en z de la séquence matricielle de dimension $p \times m$ ($m - 1 \leq p \leq m(m - 1)/2$) : $\{\mathbf{h}^{\perp}(0), \dots, \mathbf{h}^{\perp}(N - 1)\}$, comme suit

$$\mathbf{H}^{\perp\dagger}(z) = \sum_{i=0}^{N-1} \mathbf{h}^{\perp}(i) z^{-i}, \quad (2.37)$$

où les $\mathbf{h}^{\perp}(i)$, $i = 0, \dots, N - 1$ s'écrivent en fonction des éléments du vecteur $\mathbf{h}(i)$. Par exemple, dans le cas de la paramétrisation donnée par (2.36) : $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$, où $p = m$, nous avons

$$\mathbf{h}^{\perp}(i) = \begin{bmatrix} -h_2(i) & h_1(i) & 0 & \cdots & 0 \\ 0 & -h_3(i) & h_2(i) & \cdots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ h_m(i) & 0 & \cdots & 0 & -h_1(i) \end{bmatrix}, \quad i = 0, \dots, N - 1. \quad (2.38)$$

Soit \mathbf{h}^{\perp} la matrice de dimension $p \times mN$ définie comme $\mathbf{h}^{\perp} = \begin{bmatrix} \mathbf{h}^{\perp}(N - 1) & \cdots & \mathbf{h}^{\perp}(0) \end{bmatrix}$. Nous avons $\mathbf{h}^{\perp} \mathbf{h} = 0$, et l'équation donnée par (2.33) se traduit par une convolution temporelle qui s'écrit sous la forme

$$\mathcal{T}_{M-N+1}(\mathbf{h}^{\perp}) \mathcal{T}(\mathbf{h}) = 0. \quad (2.39)$$

Comme les colonnes de $\mathcal{T}(\mathbf{h})$ engendrent le SS, les lignes de $\mathcal{T}_{M-N+1}(\mathbf{h}^{\perp})$ appartiennent au SB. Par exemple, la paramétrisation $\mathbf{H}_{min}^{\perp\dagger}(z)$ donnée par (2.35) produit $(m - 1)(M - N + 1)$ vecteurs linéairement indépendants du SB. En fait, lorsqu'on utilise cette paramétrisation, on a encore besoin de $(m - 2)(N - 1)$ vecteurs linéairement indépendants (entre eux et entre les $(m - 1)(M - N + 1)$ vecteurs précédents) pour engendrer tout le SB. Concernant $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$, nous pouvons énoncer le théorème suivant

Théorème 1 *L'utilisation de $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ engendre le SB, et on a :*

$$si \mathbf{H}^{\perp\dagger}(z) = \mathbf{H}_{bal,min}^{\perp\dagger}(z) \Rightarrow Image \{ \mathcal{T}_{M-N+1}(\mathbf{h}^{\perp}) \} = (Image \{ \mathcal{T}_M(\mathbf{h}) \})^{\perp} = SB.$$

Démonstration : nous démontrons ce théorème dans le cas de $\mathbf{H}_{bal,min}^{\perp\perp}(z)$ donnée par (2.36) mais le résultat reste valable pour n'importe quel choix de $\mathbf{H}_{bal,min}^{\perp\perp}(z)$.

Pour $m = 2$, $\mathbf{H}_{bal,min}^{\perp\perp}(z)$ et $\mathbf{H}_{min}^{\perp\perp}(z)$ sont identiques, la matrice de convolution du canal correspondante $\mathcal{T}_{M-N+1}(\mathbf{h}^\perp)$ possède $M - N + 1$ lignes indépendantes qui engendrent le SB (de dimension $2M - (M + N - 1) = M - N + 1$).

Pour $m \geq 3$, l'idée de base est de compter le nombre de dépendances entre les lignes de $\mathcal{T}_{M-N+1}(\mathbf{h}^\perp)$. Considérons le cas de $m = 3$ canaux, nous avons l'égalité suivante :

$$\begin{bmatrix} \mathbf{H}_3(z) & \mathbf{H}_1(z) & \mathbf{H}_2(z) \end{bmatrix} \mathbf{H}_{bal,min}^{\perp\perp}(z) = \begin{bmatrix} \mathbf{H}_3(z) & \mathbf{H}_1(z) & \mathbf{H}_2(z) \end{bmatrix} \begin{bmatrix} -\mathbf{H}_2(z) & \mathbf{H}_1(z) & 0 \\ 0 & -\mathbf{H}_3(z) & \mathbf{H}_2(z) \\ \mathbf{H}_3(z) & 0 & -\mathbf{H}_1(z) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}. \quad (2.40)$$

Dans le cas de $m = 4$ canaux, nous avons

$$\mathbf{H}_{bal,min}^{\perp\perp}(z) = \begin{bmatrix} -\mathbf{H}_2(z) & \mathbf{H}_1(z) & 0 & 0 \\ 0 & -\mathbf{H}_3(z) & \mathbf{H}_2(z) & 0 \\ 0 & 0 & -\mathbf{H}_4(z) & \mathbf{H}_3(z) \\ \mathbf{H}_4(z) & 0 & 0 & -\mathbf{H}_1(z) \end{bmatrix} \quad (2.41)$$

et l'égalité exprimée dans (2.40) devient

$$\begin{bmatrix} \mathbf{H}_3(z)\mathbf{H}_4(z) & \mathbf{H}_4(z)\mathbf{H}_1(z) & \mathbf{H}_1(z)\mathbf{H}_2(z) & \mathbf{H}_2(z)\mathbf{H}_3(z) \end{bmatrix} \mathbf{H}_{bal,min}^{\perp\perp}(z) = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}. \quad (2.42)$$

Pour $m \geq 3$, désignons le vecteur ligne qui multiplie $\mathbf{H}_{bal,min}^{\perp\perp}(z)$ par le vecteur $\mathbf{g}^\perp(z)$. Ce vecteur satisfait

$$\mathbf{g}^\perp(z)\mathbf{H}_{bal,min}^{\perp\perp}(z) = \mathbf{0}_{1 \times m}. \quad (2.43)$$

Dans le domaine temporel, la relation (2.43) donne le nombre exact de dépendances entre les lignes de la matrice $\mathcal{T}(\mathbf{h}^\perp)$. Chaque élément du vecteur $\mathbf{g}^\perp(z)$ est un produit de $(m - 2)$ réponses impulsionnelles de sous-canaux $\mathbf{H}_i(z)$. L'ordre de ce produit est égal à $(m - 2)N - (m - 3) = (m - 2)(N - 1) + 1 = K$. Donc la relation (2.43) s'écrit dans le domaine temporel comme la convolution suivante :

$$\mathcal{T}_{M-N+1-K+1}(\mathbf{g}^\perp)\mathcal{T}_{M-N+1}(\mathbf{h}^\perp) = \mathbf{0}. \quad (2.44)$$

L'équation (2.44) donne le nombre de dépendances entre les lignes de la matrice $\mathcal{T}_{M-N+1}(\mathbf{h}^\perp)$. Ce nombre est égal à $M - N + 1 - K + 1 = M - (m - 1)(N - 1)$. Par suite, le nombre des vecteurs lignes linéairement indépendants dans la matrice $\mathcal{T}_{M-N+1}(\mathbf{h}^\perp)$ est

$$m(M - N + 1) - M + (m - 1)(N - 1) = mM - (M + N - 1) \quad (2.45)$$

qui est égal à la dimension du SB, et donc les lignes de $\mathcal{T}_{M-N+1}(\mathbf{h}^\perp)$ engendrent le SB, ce qui prouve que l'utilisation de $\mathbf{H}_{bal,min}^{\perp\perp}(z)$ engendrent le SB.

corollaire 1 *L'utilisation de $\mathbf{H}_{bal,max}^{\perp\perp}(z)$ engendrent le SB, et on a :*

$$si \mathbf{H}^{\perp\perp}(z) = \mathbf{H}_{bal,max}^{\perp\perp}(z) \Rightarrow Image \{ \mathcal{T}_{M-N+1}(\mathbf{h}^\perp) \} = (Image \{ \mathcal{T}_M(\mathbf{h}) \})^\perp = SB.$$

Démonstration : puisque les lignes de $\mathcal{T}(\mathbf{h}^\perp)$ correspondant à $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ sont un sous-ensemble des lignes de $\mathcal{T}(\mathbf{h}^\perp)$ correspondant à $\mathbf{H}_{bal,max}^{\perp\dagger}(z)$, et comme l'utilisation de $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ engendre le SB, alors l'utilisation de $\mathbf{H}_{bal,max}^{\perp\dagger}(z)$ engendre aussi le SB.

Ceci constitue une preuve plus simple que celle présentée dans [Hua96] pour démontrer que la considération du nombre de lignes maximum dans $\mathbf{H}^{\perp\dagger}(z)$ permet d'engendrer tout le SB. En fait le mécanisme de génération du SB, peut être décrit comme suit : une paramétrisation $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ inclut une paramétrisation $\mathbf{H}_{min}^{\perp\dagger}(z)$. C'est cette dernière paramétrisation qui va produire les $(m-1)(M-N+1)$ premiers vecteurs linéairement indépendants du SB. La ligne en plus dans $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ par rapport à $\mathbf{H}_{min}^{\perp\dagger}(z)$ va produire $M-N+1$ vecteurs. N'importe quel sous-ensemble contenant $(m-2)(N-1)$ vecteurs de ces derniers viendra compléter les $(m-1)(M-N+1)$ premiers vecteurs pour engendrer tout le SB, les $(M-N+1-(m-2)(N-1))$ vecteurs restants deviennent immédiatement des combinaisons linéaires des vecteurs qui ont engendré le SB.

Récemment, les auteurs de [AMH97a] ont proposé une nouvelle paramétrisation du SB et ont déduit un choix judicieux des lignes de la matrice correspondante, $\mathcal{T}(\mathbf{h}^\perp)$, pour produire le nombre exact des vecteurs linéairement indépendants pour engendrer le SB. Nous proposons, dans la section qui suit, de situer cette paramétrisation par rapport aux paramétrisations que nous avons introduites et de discuter leurs intérêts respectifs.

2.4.4 Discussion

D'un point de vue fréquentiel, la paramétrisation linéaire proposée par Abed-Meraim et Hua (voir [AMH97a] et [AMH97b]), possède $(2m-3)$ lignes (et donc possède plus de lignes que $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ quand $m > 3$). Cette paramétrisation s'écrit, quand $m > 2$, comme

$$\mathbf{H}_{AH}^{\perp\dagger}(z) = \left[\begin{array}{ccccc} -\mathbf{H}_2(z) & \mathbf{H}_1(z) & 0 & \cdots & 0 \\ -\mathbf{H}_3(z) & 0 & \mathbf{H}_1(z) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -\mathbf{H}_m(z) & 0 & \cdots & 0 & \mathbf{H}_1(z) \\ \hline 0 & -\mathbf{H}_3(z) & \mathbf{H}_2(z) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -\mathbf{H}_m(z) & 0 & \cdots & \mathbf{H}_2(z) \end{array} \right] = \left[\begin{array}{c} \mathbf{H}_{min}^{\perp\dagger}(z) \\ \hline \mathbf{H}_{second}^{\perp\dagger}(z) \end{array} \right]. \quad (2.46)$$

Notons que pour $m = 3$, la paramétrisation $\mathbf{H}_{AH}^{\perp\dagger}(z)$ est équilibrée. La matrice de convolution correspondant à $\mathbf{H}_{AH}^{\perp\dagger}(z)$ s'écrit comme

$$\mathcal{T}(\mathbf{h}^\perp) = \left[\begin{array}{c} \mathcal{T}(\mathbf{h}_{min}^\perp) \\ \hline \mathcal{T}(\mathbf{h}_{second}^\perp) \end{array} \right], \quad (2.47)$$

où la matrice $\mathcal{T}(\mathbf{h}_{min}^\perp)$ correspond à $\mathbf{H}_{min}^{\perp\dagger}(z)$ et la matrice $\mathcal{T}(\mathbf{h}_{second}^\perp)$ correspond à $\mathbf{H}_{second}^{\perp\dagger}(z)$. Abed-Meraim et Hua proposent de faire un choix judicieux des lignes de $\mathcal{T}(\mathbf{h}^\perp)$ pour engendrer tout le SB. Ce choix se fait de la manière suivante : considérer toutes les lignes de $\mathcal{T}(\mathbf{h}_{min}^\perp)$ qui sont en nombre de $(m-1)(M-N+1)$, puis à chaque ligne des $(m-2)$ lignes de $\mathbf{H}_{second}^{\perp\dagger}(z)$, considérer les $(N-1)$ premières lignes correspondantes dans $\mathcal{T}(\mathbf{h}_{second}^\perp)$. Le nombre total de lignes conservées est donc égal à

$$(m-1)(M-N+1) + (m-2)(N-1) = mM - (M+N-1) \quad (2.48)$$

qui n'est autre que la dimension du SB. Dans [AMH97b], les auteurs montrent qu'un tel choix fournit l'ensemble exact des vecteurs linéairement indépendants nécessaires pour engendrer le SB, pourvu que les deux premiers sous-canaux $H_1(z)$ et $H_2(z)$ ne partagent pas de zéros communs. Ce dernier point réduit l'intérêt de l'utilisation de cette paramétrisation dans une méthode d'estimation multi-canal. En effet, la paramétrisation $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ est minimale et son utilisation engendre le SB sans avoir à contraindre $H_1(z)$ et $H_2(z)$ à être premiers entre eux. En fait, cette dernière contrainte imposée à $H_1(z)$ et $H_2(z)$ pour engendrer tout le SB est due à l'absence de l'aspect "équilibré" dans la paramétrisation $\mathbf{H}_{AH}^{\perp\dagger}$. En effet, dans la construction de cette paramétrisation, on attache plus de poids à $H_1(z)$ qui joue un rôle de pivot dans $\mathbf{H}_{min}^{\perp\dagger}(z)$ et à $H_2(z)$ ayant ce même rôle de pivot dans $\mathbf{H}_{second}^{\perp\dagger}(z)$. En conclusion, c'est cet aspect "équilibré", sur lequel se base la construction de $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ où on attache le même poids à tous les sous-canaux, qui permet d'engendrer tout le SB sans avoir à contraindre les réponses impulsionnelles des sous-canaux.

2.5 Conclusions

Ce chapitre explique la notion de sous-espace dans le but de préparer le terrain aux méthodes d'identification aveugle multi-canal. Partant de l'exemple des sinusoïdes dans le bruit, nous avons expliqué les sous-espaces signal et bruit inhérents à ce cas. De tels sous-espaces ont évoqué des paramétrisations correspondantes et des méthodes d'estimation basées sur le principe d'ajustement. Un raisonnement analogue dans le cas de l'identification aveugle multi-canal permet de caractériser les sous-espaces signal et bruit associés à ce cas. Nous avons analysé les différentes paramétrisations linéaires du SB. Dans une première étape, nous avons expliqué la paramétrisation linéaire minimale du SB, puis nous avons présenté dans une deuxième partie les différentes paramétrisations du SB basées sur le canal. L'étude de la relation existant entre les différentes paramétrisations et la génération du SB, nous a conduit à mettre en évidence une paramétrisation qualifiée d'équilibrée et dont l'utilisation engendre systématiquement tout le SB. Les chapitres 3 et 4 utiliseront les résultats que nous venons de décrire. En effet, certaines méthodes d'estimation multi-canal aveugle se baseront sur ces paramétrisations dans la formulation du critère à minimiser. L'influence de la paramétrisation utilisée se situera

au niveau du conditionnement de la matrice du critère à optimiser et au niveau de la complexité numérique de la méthode d'estimation.

Chapitre 3

Méthodes Basées sur les Sous-espaces

Nous nous intéressons, dans ce chapitre, à décrire des méthodes d'estimation du canal (utilisant les statistiques de second ordre) dans les cas où les symboles sont supposés être des inconnues déterministes. Les méthodes qui seront discutées utilisent explicitement la notion de sous-espaces introduite dans le chapitre précédent. Nous commençons donc par justifier la suffisance des statistiques de second ordre pour l'estimation aveugle multi-canal. Puis nous calculons les bornes de Cramer-Rao aveugles (dans le cas du modèle déterministe et gaussien). Ensuite, nous présentons les différentes méthodes d'identification du canal qui incluent les techniques d'ajustement de sous-espaces. L'inconvénient principal de ces techniques d'ajustement de sous-espaces est leur coût élevé dû à la décomposition propre de la matrice de covariance du signal reçu. Pour y remédier, nous proposons des méthodes alternatives d'ajustement de sous-espaces qui évitent la décomposition propre de la matrice de covariance. L'analyse des performances de ces méthodes, à laquelle s'ajoutent les résultats de simulation, montrent qu'asymptotiquement les performances des méthodes alternatives sont identiques à celles des versions classiques.

3.1 Introduction

Récemment, un grand effort de recherche a été dédié à l'estimation et/ou l'égalisation aveugle basée sur les statistiques de second ordre du signal reçu. Cet effort est généralement justifié par la complexité réduite du calcul des moments de second ordre. Laquelle complexité rend l'utilisation des méthodes d'identification basées sur les statistiques de second ordre plus justifiée que celle des techniques d'identification basées sur les statistiques d'ordre supérieur à deux (HOS). Dans la première section de ce chapitre, nous essayons de mettre en lumière les raisons pour lesquelles l'utilisation des statistiques de deuxième ordre donnent des estimées précises du canal. Ensuite, nous présentons les différentes méthodes d'estimation aveugle de canal qui se basent sur la notion de sous-espaces.

3.2 Statistiques de Second Ordre et Identification du Canal

L'avantage de l'aspect multi-canal introduit en identification aveugle est que les statistiques de second ordre sont suffisantes pour estimer le canal. Ces techniques sont souvent opposées aux méthodes basées sur les statistiques d'ordre supérieur à deux (HOS) qui exploitent les cumulants d'ordre quatre du signal reçu. Nous pensons que le fait que les statistiques de second ordre donnent des bonnes performances en termes d'estimées du canal, n'est pas lié au moment de second ordre mais ceci n'est autre qu'une conséquence de l'estimation précise du SS avec un nombre fini de données. Notre constatation est justifiée par le raisonnement présenté dans la section qui suit.

3.2.1 Estimation Précise du SS

Rappelons que la structure de la matrice de covariance du signal reçu \mathbf{Y} de longueur M est donnée par

$$R_{YY} = E\mathbf{Y}\mathbf{Y}^H = \mathcal{T}(\mathbf{h})R_{AA}\mathcal{T}(\mathbf{h})^H + \sigma_v^2 I_{mM}, \quad (3.1)$$

et qu'une décomposition propre de cette matrice de covariance permet de faire apparaître les contributions des sous-espaces signal et bruit

$$R_{YY} = \sum_{i=1}^{M+N-1} \lambda_i V_i V_i^H + \sum_{i=M+N}^{mM} \lambda_i V_i V_i^H = \mathcal{V}_S \Lambda_S \mathcal{V}_S^H + \mathcal{V}_N \Lambda_N \mathcal{V}_N^H. \quad (3.2)$$

Pour étudier la précision de l'estimation du SS, nous utilisons la distribution asymptotique des vecteurs propres du SS de l'estimée empirique de la matrice de covariance. En effet, le SS et le SB peuvent être estimés de façon consistante à partir de l'estimée empirique de la matrice de covariance donnée par

$$\widehat{R}_{YY} = \frac{1}{T} \sum_{i=1}^T \mathbf{Y}_M(i) \mathbf{Y}_M^H(i) = \widehat{\mathcal{V}}_S \widehat{\Lambda}_S \widehat{\mathcal{V}}_S^H + \widehat{\mathcal{V}}_N \widehat{\Lambda}_N \widehat{\mathcal{V}}_N^H. \quad (3.3)$$

A partir de la théorie asymptotique des composantes principales [OVS93], nous avons le résultat suivant :

Théorème 2 *les $M + N - 1$ vecteurs propres correspondant aux $M + N - 1$ valeurs propres les plus grandes de \widehat{R}_{YY} sont asymptotiquement (pour T suffisamment grand) distribués selon une loi normale de moyenne et de covariances données par*

$$\begin{aligned} E\widehat{V}_k &= V_k + O(T^{-1}), \\ E(\widehat{V}_k - V_k)(\widehat{V}_l - V_l)^H &= \delta_{kl} \frac{\lambda_k}{T} \sum_{i=1, i \neq k}^{mM} \frac{\lambda_i}{(\lambda_i - \lambda_k)^2} V_i V_i^H + o(T^{-1}), \\ E(\widehat{V}_k - V_k)(\widehat{V}_l - V_l)^T &= (1 - \delta_{kl}) \frac{-\lambda_k \lambda_l}{T(\lambda_k - \lambda_l)^2} V_l V_k^T + o(T^{-1}). \end{aligned}$$

Une mesure de l'erreur d'estimation du SS peut être obtenue en considérant $E\|\widehat{\mathcal{V}}_S - \mathcal{V}_S\|_F^2 = \sum_k E(\widehat{V}_k - V_k)(\widehat{V}_k - V_k)^H$. La matrice de covariance de l'erreur $E\widetilde{V}_k\widetilde{V}_k^H = E(\widehat{V}_k - V_k)(\widehat{V}_k - V_k)^H$ peut être écrite comme la somme de deux contributions : la première contribution est celle due à l'erreur au niveau des vecteurs propres du SS, que nous désignons par $\epsilon_{\widehat{V}_k}^S$, et la seconde contribution est celle associée aux vecteurs propres du SB, que nous appelons $\epsilon_{\widehat{V}_k}^N$. Nous obtenons

$$\begin{aligned} E\widetilde{V}_k\widetilde{V}_k^H &= \frac{\lambda_k}{T} \sum_{i=1, i \neq k}^{mM} \frac{\lambda_i}{(\lambda_i - \lambda_k)^2} V_i V_i^H + o(T^{-1}) \\ &= \frac{\lambda_k}{T} \sum_{i=1, i \neq k}^{M+N-1} \frac{\lambda_i}{(\lambda_i - \lambda_k)^2} V_i V_i^H + \frac{\lambda_k}{T} \sum_{i=M+N}^{mM} \frac{\lambda_i}{(\lambda_i - \lambda_k)^2} V_i V_i^H + o(T^{-1}) \\ &= \epsilon_{\widehat{V}_k}^S + \epsilon_{\widehat{V}_k}^N + o(T^{-1}). \end{aligned} \quad (3.4)$$

Dans l'expression (3.4), la contribution de l'erreur $\epsilon_{\widehat{V}_k}^N$, associée au SB, tend vers zéro lorsque le SNR tend vers l'infini, alors que la contribution de l'erreur $\epsilon_{\widehat{V}_k}^S$, associée au SS, n'a pas d'influence sur l'estimée du SS puisque la composante $\widetilde{V}_k^S = (\widehat{V}_k - V_k)^S$ reste dans le SS. Ceci montre que le SS de la matrice de covariance empirique \widehat{R}_{YY} (qui est estimé avec un nombre fini de données) est "parfaitement" estimé lorsque le SNR $\rightarrow \infty$.

3.2.2 Simulations

A titre d'illustration du résultat que nous venons d'énoncer, nous étudions le comportement de l'erreur d'estimation du SS et des moments de second ordre (tous deux estimés à partir d'un nombre fini de données) en fonction du SNR. La matrice de covariance empirique \widehat{R}_{YY} est estimée à partir d'un paquet de données de longueur 200. Le nombre de canaux est $m = 3$ et la séquence des symboles est une séquence i.i.d. BPSK. Le canal est complexe, généré de façon aléatoire, ayant pour longueur $N = 3$:

$$\mathbf{H} = \begin{bmatrix} -0.9712 + 0.9603i & 0.6087 + 0.1392i & 0.0243 - 1.2639i \\ 1.0701 + 0.8570i & 0.0974 + 0.3818i & -0.9887 - 0.6548i \\ -0.1217 + 2.1088i & 1.0908 - 0.6647i & -0.6495 + 0.4929i \end{bmatrix}. \quad (3.5)$$

La matrice de covariance empirique considérée est de taille $M = 10$. L'erreur quadratique moyenne normalisée (NMSE) sur le SS est calculée en prenant la moyenne sur 300 réalisations Monte-Carlo :

$$\text{NMSE1} = \frac{1}{300} \sum_{i=1}^{300} \left(1 - \frac{\|\widehat{\mathcal{V}}_S^{(i)H} \mathcal{V}_S\|_F^2}{M + N - 1} \right) \quad (3.6)$$

où $\widehat{\mathcal{V}}_S^{(i)}$ représente l'estimée du SS donnée par les $M + N - 1$ vecteurs propres correspondant aux $M + N - 1$ valeurs propres les plus grandes de la matrice de covariance empirique \widehat{R}_{YY} dans la $i^{\text{ème}}$ réalisation.

Dans le but de comparer la qualité de l'estimation du SS à celle des moments de second ordre, nous considérons le premier élément bloc de la matrice de covariance empirique $\hat{r}_{YY}^{(i)}(0)$ de la $i^{\text{ème}}$ réalisation et nous définissons le NMSE comme :

$$\text{NMSE2} = \frac{1}{300} \sum_{i=1}^{300} \|\hat{r}_{YY}^{(i)}(0) - r_{YY}(0)\|_F^2 / \|r_{YY}(0)\|_F^2 \quad (3.7)$$

où $r_{YY}(0)$ est le premier élément bloc de la matrice de covariance calculée en utilisant les statistiques exactes $R_{YY} = \sigma_a^2 \mathcal{T}(\mathbf{h}) \mathcal{T}^H(\mathbf{h}) + \sigma_v^2 I_{mM}$. Dans la Figure 3.1, nous traçons l'erreur NMSE1 en fonction du SNR : il est clair qu'avec un nombre fini de données nous pouvons estimer de façon précise le SS, alors que dans la Figure 3.2, l'erreur NMSE2 se stabilise à partir d'un SNR ≈ 20 dB. Ceci implique qu'avec un nombre fini de données, l'estimation des moments de second ordre ne peut pas se faire de façon parfaite (une erreur d'estimation y est inhérente).

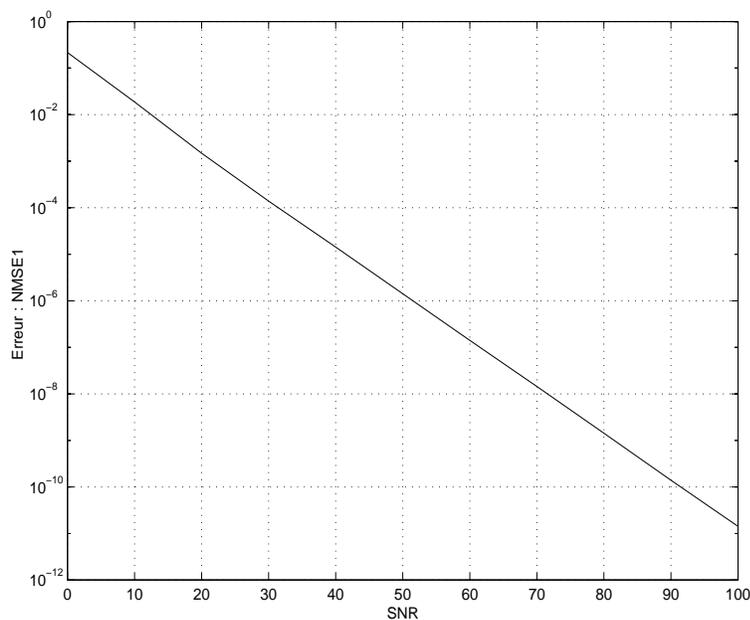


FIG. 3.1 – Estimation précise du SS avec un nombre fini de données.

3.3 Bornes de Cramer-Rao

Dans cette section, nous nous intéressons à dériver des bornes de performance pour les méthodes d'estimation multi-canal (qui seront décrites dans ce chapitre et les chapitres suivants). La mesure de performance que nous considérons est la borne de Cramer-Rao (CRB) [Kay93]. En effet, la borne de Cramer-Rao représente une borne inférieure de la matrice de covariance de l'erreur d'estimation. Cette borne nous permet d'une part, de fixer la limite de ce qu'on peut espérer atteindre comme performance et d'autre part, elle nous permet de juger de la qualité d'estimation par rapport à cette limite.

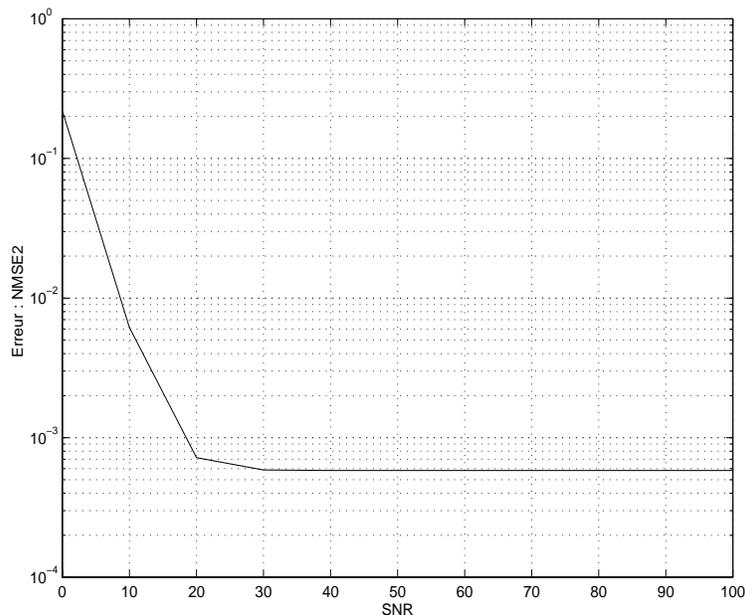


FIG. 3.2 – Estimation du premier moment de second ordre avec un nombre fini de données.

Soit θ le vecteur des paramètres complexes à estimer, et soit $\theta_R = \begin{bmatrix} \text{Re}(\theta)^H & \text{Im}(\theta)^H \end{bmatrix}^H$ le vecteur des paramètres réels qui lui sont associés. La matrice d'information de Fisher associée à θ_R est :

$$J_{\theta_R \theta_R} = E_{\mathbf{Y}|\theta_R} \left(\frac{\partial \ln f(\mathbf{Y}|\theta_R)}{\partial \theta_R} \right) \left(\frac{\partial \ln f(\mathbf{Y}|\theta_R)}{\partial \theta_R} \right)^T, \quad (3.8)$$

\mathbf{Y} étant le vecteur d'observations et $f(\mathbf{Y}|\theta_R)$ est la fonction de densité de probabilité qui lui est associée. Soient $\hat{\theta}_R$ une estimée non biaisée des paramètres, $\tilde{\theta}_R = \theta_R - \hat{\theta}_R$ l'erreur d'estimation et $C_{\tilde{\theta}_R \tilde{\theta}_R} = E \tilde{\theta}_R \tilde{\theta}_R^H$ la matrice de covariance de l'erreur d'estimation. $J_{\theta_R \theta_R}^{-1}$ est la borne de Cramer-Rao :

$$C_{\tilde{\theta}_R \tilde{\theta}_R} \geq CRB_{\tilde{\theta}_R} = J_{\theta_R \theta_R}^{-1}. \quad (3.9)$$

Comme nous manipulons des quantités complexes, il serait plus facile de considérer l'opération de dérivation complexe, par rapport à $\theta = \alpha + j\beta$, définie¹ comme

$$\frac{\partial}{\partial \theta} = \frac{1}{2} \left(\frac{\partial}{\partial \alpha} - j \frac{\partial}{\partial \beta} \right). \quad (3.10)$$

La matrice d'information de Fisher $J_{\varphi\psi}$ associée aux paramètres complexes φ, ψ (qui sont des parties de θ) est définie comme suit :

$$J_{\varphi\psi} = E_{Y|\theta} \left(\frac{\partial \ln f(Y|\theta)}{\partial \varphi^*} \right) \left(\frac{\partial \ln f(Y|\theta)}{\partial \psi^*} \right)^H. \quad (3.11)$$

1. Une autre définition de l'opération de dérivation complexe qui n'utilise pas le facteur $\frac{1}{2}$ peut être adoptée (voir [Com86] et les références incluses).

Introduisons le vecteur des paramètres complexes étendu $\theta_C = [\theta^T \ \theta^H]^T$. La matrice $J_{\theta_C \theta_C}$ contient la même information que la matrice $J_{\theta_R \theta_R}$.

Si $J_{\theta\theta^*} = 0$, la matrice $J_{\theta\theta}$ peut être considérée comme une matrice d'information de Fisher complexe, et la matrice de covariance de l'erreur d'estimation non biaisée, $\tilde{\theta} = \theta - \hat{\theta}$, est $C_{\tilde{\theta}\tilde{\theta}} \geq J_{\theta\theta}^{-1}$, qui est la borne de Cramer-Rao. Si $J_{\theta\theta^*} \neq 0$, $J_{\theta\theta}^{-1}$ est aussi une borne de $C_{\tilde{\theta}\tilde{\theta}}$, mais qui n'est pas aussi étroite que la borne de Cramer-Rao réelle $\text{CRB} = C_{\tilde{\theta}_C \tilde{\theta}_C} \geq J_{\theta_C \theta_C}^{-1}$. Dans le cas du problème de l'estimation aveugle du canal, considéré ici, un problème d'identifiabilité apparaît puisque le canal peut être identifié à une constante scalaire près. Ceci se traduit par une *singularité* de la matrice d'information de Fisher. Pour le calcul des bornes de Cramer-Rao, nous remplaçons les inverses des matrices d'information de Fisher $J_{\theta_C \theta_C}^{-1}$ par les pseudo-inverses de Moore-Penrose $J_{\theta_C \theta_C}^+$. Pour que ces inverses restent bien des bornes de Cramer-Rao (valables), le facteur du canal doit être ajusté selon la façon expliquée dans [dCS97b].

3.3.1 Borne de Cramer-Rao Déterministe

Les symboles transmis et le canal sont considérés comme des inconnues déterministes. Le vecteur des paramètres (complexes) est : $\theta = \begin{bmatrix} A^T & \mathbf{h}^T \end{bmatrix}^T$. La fonction de densité de probabilité qui lui est associée est

$$f(\mathbf{Y}/\theta) = \frac{1}{(\pi\sigma_v^2)^{mM}} e^{-\frac{1}{\sigma_v^2} \|\mathbf{Y} - \mathcal{T}A\|^2}. \quad (3.12)$$

Dans [SP94], la matrice d'information de Fisher pour l'estimation de \mathbf{h} à partir de (3.12) a été dérivée. Son expression est

$$J(\mathbf{h}) = \frac{1}{\sigma_v^2} \mathcal{A}_{M,N}^H P_{\mathcal{T}}^\perp \mathcal{A}_{M,N} \quad (3.13)$$

où $\mathcal{A}_{M,N} = A_{M,N} \otimes I_m$ tel que $\mathcal{A}_{M,N} \mathbf{h} = \mathcal{T}(\mathbf{h})A$ et \otimes désigne le produit de Kronecker. La matrice $A_{M,N}$ est une matrice de Hankel définie comme

$$A_{M,N} = \begin{bmatrix} a(-N+1) & a(-N+2) & \cdots & a(0) \\ a(-N+2) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a(M-N) & \cdots & \cdots & a(M-1) \end{bmatrix}. \quad (3.14)$$

La borne de Cramer-Rao aveugle déterministe est

$$\text{CRB} = J^+(\mathbf{h}) = \sigma_v^2 \left[\mathcal{A}_{M,N}^H P_{\mathcal{T}}^\perp \mathcal{A}_{M,N} \right]^+. \quad (3.15)$$

où $^+$ désigne la pseudo-inverse de Moore-Penrose.

3.3.2 Borne de Cramer-Rao Gaussienne

Dans le cas où les symboles transmis sont supposés gaussiens, le vecteur de paramètres à estimer est $\theta = \mathbf{h}$. Nous avons $\mathbf{Y} \sim \mathcal{N}(\mathcal{T}(\mathbf{h})A, R_{YY})$, $R_{YY} = \sigma_a^2 \mathcal{T}(\mathbf{h})\mathcal{T}^H(\mathbf{h}) + \sigma_v^2 I$. La fonction de densité de probabilité correspondante est

$$f(\mathbf{Y}/\mathbf{h}) = \frac{1}{\pi^{mM} \det R_{YY}} e^{-(\mathbf{Y} - \mathcal{T}(\mathbf{h})A)^H R_{YY}^{-1} (\mathbf{Y} - \mathcal{T}(\mathbf{h})A)}. \quad (3.16)$$

La matrice d'information de Fisher $J_{\varphi\psi}$ est définie comme suit

$$J_{\varphi\psi} = E_{\mathbf{Y}/\mathbf{h}} \left(\frac{\partial \ln f(\mathbf{Y}/\mathbf{h})}{\partial \varphi^*} \right) \left(\frac{\partial \ln f(\mathbf{Y}/\mathbf{h})}{\partial \psi^*} \right)^H \quad (3.17)$$

et nous considérons $J_{\mathbf{h}\mathbf{h}}$ et $J_{\mathbf{h}\mathbf{h}^*}$. Dans le cas déterministe, $J_{\mathbf{h}\mathbf{h}^*} = 0$. Dans ce cas, la matrice $J_{\mathbf{h}\mathbf{h}}$ peut être considérée comme une matrice d'information de Fisher et $C_{\mathbf{h}}^- \geq J_{\mathbf{h}\mathbf{h}}^+$ la CRB complexe. Si $J_{\mathbf{h}\mathbf{h}^*} \neq 0$, ce qui est le cas lorsque les symboles sont supposés gaussiens, $J_{\mathbf{h}\mathbf{h}}^+$ est aussi une borne de $C_{\mathbf{h}}^-$, mais qui n'est pas aussi étroite que la CRB réelle obtenue en considérant $\mathbf{h}_R = \begin{bmatrix} \text{Re}(\mathbf{h})^T & \text{Im}(\mathbf{h})^T \end{bmatrix}^T$ (les paramètres réels associés à \mathbf{h}). Nous obtenons :

$$J_R(\mathbf{h}_R) = 2 \begin{bmatrix} \text{Re}(J_{\mathbf{h}\mathbf{h}}) & -\text{Im}(J_{\mathbf{h}\mathbf{h}}) \\ \text{Im}(J_{\mathbf{h}\mathbf{h}}) & \text{Re}(J_{\mathbf{h}\mathbf{h}}) \end{bmatrix} + 2 \begin{bmatrix} \text{Re}(J_{\mathbf{h}\mathbf{h}^*}) & -\text{Im}(J_{\mathbf{h}\mathbf{h}^*}) \\ \text{Im}(J_{\mathbf{h}\mathbf{h}^*}) & \text{Re}(J_{\mathbf{h}\mathbf{h}^*}) \end{bmatrix} \quad (3.18)$$

et

$$J_{\mathbf{h}_i \mathbf{h}_j} = \text{tr} \left\{ R_{YY}^{-1} \left(\frac{\partial R_{YY}}{\partial \mathbf{h}_i^*} \right) R_{YY}^{-1} \left(\frac{\partial R_{YY}}{\partial \mathbf{h}_j^*} \right)^H \right\}, \quad (3.19)$$

$$J_{\mathbf{h}_i \mathbf{h}_j^*} = \text{tr} \left\{ R_{YY}^{-1} \left(\frac{\partial R_{YY}}{\partial \mathbf{h}_i^*} \right) R_{YY}^{-1} \left(\frac{\partial R_{YY}}{\partial \mathbf{h}_j^*} \right) \right\}, \quad (3.20)$$

$$\frac{\partial R_{YY}}{\partial \mathbf{h}_i^*} = \sigma_a^2 \mathcal{T}(\mathbf{h}) \mathcal{T}_M^H \left(\frac{\partial \mathbf{H}}{\partial \mathbf{h}_i} \right), \quad (3.21)$$

où la matrice $\mathcal{T}_M \left(\frac{\partial \mathbf{H}}{\partial \mathbf{h}_i} \right)$ est une matrice bloc Toeplitz, remplie avec les coefficients du canal \mathbf{h} , possédant M lignes bloc et $\begin{bmatrix} \frac{\partial \mathbf{H}}{\partial \mathbf{h}_i} & 0_{m \times (M-1)} \end{bmatrix}$ comme première ligne bloc.

Ces bornes vont servir de mesure de performance pour les méthodes d'identification multi-canal. Dans ce qui suit, nous décrivons des méthodes d'estimation de canal aveugles dont la borne de Cramer-Rao est exprimée dans (3.15).

3.4 Subchannel Response Matching (SRM)

L'approche SRM a été introduite dans [BR94] et correspond aussi à l'approche déterministe de Liu et Xu [LX94a],[LX94b],[XLTK95]. L'approche SRM a été aussi présentée dans [GN95] et sert d'initialisation dans l'algorithme de Hua [Hua96]. La technique SRM

utilise les paramétrisations linéaires du SB basées sur le canal que nous venons d'expliquer dans la section §2.4.3. Considérons, tout d'abord, le cas de deux canaux : $m = 2$. Dans le cas sans bruit, nous avons

$$\mathbf{H}_2(z)y_1(k) - \mathbf{H}_1(z)y_2(k) = [\mathbf{H}_2(z) \ -\mathbf{H}_1(z)] \mathbf{y}(k) = \mathbf{H}^{\perp\dagger}(z)\mathbf{H}(z)a(k) = 0. \quad (3.22)$$

En empilant les signaux reçus $\{\mathbf{y}(k)\}_{k=0\dots M-1}$, nous obtenons une expression de la forme $\mathcal{Y}\mathbf{h} = 0$, où \mathcal{Y} est une matrice structurée remplie avec les observations $\{\mathbf{y}(k)\}_{k=0\dots M-1}$. Sous la contrainte $\|\mathbf{h}\|_2 = 1$, le canal est trouvé comme $\mathbf{h} = V_{\min}(\mathcal{Y}^H\mathcal{Y})$ où $V_{\min}(A)$ désigne le vecteur propre associé à la plus petite valeur propre de la matrice A .

Pour $m > 2$, la paramétrisation $\mathbf{H}^{\perp\dagger}(z)$ peut être choisie comme l'une de celles décrites dans la section §2.4.3. Un tel choix se répercute sur les performances de la méthode SRM. La convolution $\mathbf{H}^{\perp}(z)\mathbf{y}(k)$ utilisant la séquence $\{\mathbf{y}(k)\}_{k=0\dots M-1}$ peut être écrite, sous forme matricielle, comme $\mathcal{T}(\mathbf{h}^{\perp})\mathbf{Y}$. Puisque dans le cas sans bruit nous obtenons $\mathcal{T}(\mathbf{h}^{\perp})\mathbf{Y} = 0$, le critère SRM minimise $\|\mathcal{T}(\mathbf{h}^{\perp})\mathbf{Y}\|_2^2$. Notons que la formulation du critère SRM discutée dans [Smi96] utilise la paramétrisation $\mathbf{H}_{\max}^{\perp\dagger}(z)$.

3.4.1 Cas Asymptotique

Par la loi des grands nombres, asymptotiquement le critère SRM qui correspond à $\|\mathcal{T}(\mathbf{h}^{\perp})\mathbf{Y}\|_2^2$ peut être remplacé par sa moyenne statistique, qui peut être écrite dans le domaine fréquentiel comme

$$\begin{aligned} \mathcal{J} &= \frac{1}{2\pi j} \oint \text{tr}\{\hat{\mathbf{H}}^{\perp\dagger}(z)\mathbf{S}\mathbf{y}\mathbf{y}(z)\hat{\mathbf{H}}^{\perp}(z)\} \frac{dz}{z} \\ &= \frac{\sigma_a^2}{2\pi j} \oint \mathbf{H}^{\dagger}(z)\hat{\mathbf{H}}^{\perp}(z)\hat{\mathbf{H}}^{\perp\dagger}(z)\mathbf{H}(z) \frac{dz}{z} + \frac{\sigma_v^2}{2\pi j} \oint \text{tr}\{\hat{\mathbf{H}}^{\perp\dagger}(z)\hat{\mathbf{H}}^{\perp}(z)\} \frac{dz}{z} \end{aligned} \quad (3.23)$$

Considérons le cas de $\mathbf{H}_{\text{bal}}^{\perp}(z)$. Dans ce cas, nous avons $\text{tr}\{\mathbf{H}^{\perp\dagger}(z)\mathbf{H}^{\perp}(z)\} = \alpha\mathbf{H}^{\dagger}(z)\mathbf{H}(z)$ pour un certain scalaire réel α . Par conséquent, la minimisation de (3.23) donne

$$\min_{\|\hat{\mathbf{h}}\|=1} \mathcal{J} = \alpha\sigma_v^2\|\hat{\mathbf{h}}\|^2 + \frac{\sigma_a^2}{2\pi j} \min_{\|\hat{\mathbf{h}}\|=1} \oint \mathbf{H}^{\dagger}(z)\hat{\mathbf{H}}^{\perp}(z)\hat{\mathbf{H}}^{\perp\dagger}(z)\mathbf{H}(z) \frac{dz}{z} \quad (3.24)$$

ce qui aboutit à la valeur correcte du canal $\hat{\mathbf{h}} = \mathbf{h}$ (et donc une estimée non biaisée du canal!) à un facteur scalaire près (et en supposant que l'ordre du canal est correctement choisi). Notons que la première étape de l'algorithme présenté par Hua [Hua96] n'est autre que la méthode SRM.

3.4.2 Formulation dans le cas "équilibré"

Nous cherchons dans cette section à formuler le critère SRM sous forme analytique dans le cas de la paramétrisation linéaire du SB : $\mathbf{H}_{\text{bal},\min}^{\perp\dagger}(z)$. Cet intérêt se justifie par le

fait que l'utilisation de cette paramétrisation dans la formulation du critère SRM aboutit à des estimées non biaisées du canal. La $i^{\text{ème}}$ ligne de $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ peut être écrite comme

$$\mathbf{H}_{bal,i}^{\perp\dagger}(z) = \mathbf{H}^T(z)\mathcal{P}_i,$$

$$\text{où } \mathcal{P}_{i+1} = \mathcal{C}\mathcal{P}_i\mathcal{C}^H, \mathcal{P}_1 = \begin{bmatrix} 0 & 1 & 0 & \cdots \\ -1 & 0 & \cdots \\ 0 & \vdots & \ddots \\ \vdots \end{bmatrix}, \mathcal{C} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 \\ 0 & \ddots & \vdots \\ \vdots & 0 & 1 & 0 \end{bmatrix}. \quad (3.25)$$

Pour cette paramétrisation $\mathbf{H}_{bal}^{\perp\dagger}(z)$, le critère SRM $\|\mathcal{T}(\mathbf{h}^\perp)\mathbf{Y}\|_2^2$ peut être écrit comme la minimisation par rapport à \mathbf{h} de

$$\begin{aligned} \text{tr}\{\mathcal{T}(\mathbf{h}^\perp)\mathbf{Y}\mathbf{Y}^H\mathcal{T}^H(\mathbf{h}^\perp)\} &= \text{tr}\{\mathbf{h}^\perp \left(\sum_{k=N-1}^{M-1} \mathbf{Y}_N(k)\mathbf{Y}_N^H(k) \right) \mathbf{h}^{\perp H}\} \\ &= (M-N+1)\text{tr}\{\mathbf{h}^\perp \widehat{R}_{YY} \mathbf{h}^{\perp H}\} \end{aligned} \quad (3.26)$$

où la $i^{\text{ème}}$ ligne de \mathbf{h}^\perp est $\mathbf{h}_i^\perp = \mathbf{h}^T \mathcal{S}_i$, $\mathcal{S}_i = I_N \otimes \mathcal{P}_i$. Par conséquent, le critère SRM exprimé dans (3.26) devient

$$\min_{\mathbf{h}} \mathbf{h}^H A \mathbf{h}, \text{ où } A = \sum_{i=1}^m \mathcal{S}_i \widehat{R}_{YY}^* \mathcal{S}_i^H. \quad (3.27)$$

Si la matrice exacte R_{YY} est utilisée, alors la contribution du bruit dans (3.27) devient

$$\mathbf{h}^H \left(\sum_{i=1}^m \mathcal{S}_i \sigma_v^2 I \mathcal{S}_i^H \right) \mathbf{h} = \sigma_v^2 \mathbf{h}^H \left(\sum_{i=1}^m \mathcal{S}_i \mathcal{S}_i^H \right) \mathbf{h} = \sigma_v^2 \mathbf{h}^H (2I) \mathbf{h} = 2\sigma_v^2 \|\mathbf{h}\|^2. \quad (3.28)$$

Donc la minimisation du critère (3.27) sous la contrainte $\|\mathbf{h}\| = 1$ aboutit à l'estimée consistante du canal $\mathbf{h} = V_{min}(A)$ (en supposant que l'ordre du canal est correctement choisi). Puisque $\sigma_v^2 = \lambda_{min}(R_{YY})$, la valeur propre minimale de R_{YY} , la contribution du bruit peut être éliminée en remplaçant \widehat{R}_{YY} par $\widehat{R}_{YY} - \lambda_{min}(\widehat{R}_{YY})I$ ou, encore mieux, en remplaçant la matrice A par la matrice $B = A - \lambda_{min}(A)I$ (le choix précédent ne rend pas la matrice A singulière avec un nombre fini de données). Avec cette modification, le critère (3.27) devient (asymptotiquement) "insensible" à la contribution du bruit et n'importe quelle normalisation de \mathbf{h} aboutira à une estimée consistante du canal. Notons, aussi, que l'utilisation de la paramétrisation linéaire $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ permet de retrouver une estimée de σ_v^2 comme $\widehat{\sigma}_v^2 = \lambda_{min}(A)/2$ pour $m > 2$ (pour le cas $m = 2$, $\widehat{\sigma}_v^2 = \lambda_{min}(A)$). Nous disposons ainsi d'une méthode pour estimer la puissance du bruit autre que la procédure empirique qui consiste à prendre la moyenne arithmétique des valeurs propres correspondant au SB.

3.5 Méthodes d'Ajustement de Sous-espaces

Dans la section §2.2, nous avons présenté l'exemple des sinusoïdes dans le bruit pour illustrer la notion des sous-espaces bruit et signal. L'importance d'une telle notion réside dans le fait que disposant d'une paramétrisation linéaire du SS ou du SB, nous sommes capables d'établir des procédures d'estimation des paramètres du canal. Ces méthodes d'estimation ont été présentées dans les sections §2.2.3 et §2.2.4 dans le cas des sinusoïdes dans le bruit, où des algorithmes basés sur le principe d'ajustement de sous-espaces ont été décrits. Dans ce qui suit, nous décrivons des méthodes d'estimation du canal analogues, dans le cas de l'estimation aveugle multi-canal.

3.5.1 Ajustement du SS

Nous supposons que la dimension correcte du SS, $M + N - 1$ (et donc l'ordre correct du canal, N) a été déterminée. Plusieurs techniques traitant ce problème ont été présentées dans la littérature. Nous supposons aussi que les valeurs propres de la matrice de covariance sont ordonnées dans un ordre décroissant comme dans (2.23). Le SS est défini comme le sous-espace engendré par les vecteurs propres correspondant aux $M + N - 1$ valeurs propres les plus grandes de la matrice de covariance. Considérons, maintenant, le problème d'ajustement de sous-espace suivant

$$\min_{\mathbf{h}, T} \|\mathcal{T}(\mathbf{h}) - \mathcal{V}_S T\|_F. \quad (3.29)$$

La matrice de transformation optimale est de nouveau donnée par

$$T = V_S^H \mathcal{T}(\mathbf{h}). \quad (3.30)$$

En utilisant (3.30) et la commutativité de la convolution, nous pouvons démontrer que le problème de minimisation donné par (3.29) est équivalent à

$$\min_{\mathbf{h}} \text{tr} \{ \mathcal{T}^H(\mathbf{h}) P_{\mathcal{V}_S}^\perp \mathcal{T}(\mathbf{h}) \} = \min_{\mathbf{h}} \mathbf{h}^H A \mathbf{h}, \quad (3.31)$$

où $P_{\mathcal{V}_S}^\perp = I - \mathcal{V}_S (\mathcal{V}_S^H \mathcal{V}_S)^{-1} \mathcal{V}_S^H$. La matrice A peut être déterminée à partir de $P_{\mathcal{V}_S}^\perp = P_{\mathcal{V}_N}$. Ce problème de minimisation doit être augmenté d'une contrainte de non trivialité sur le canal \mathbf{h} . Dans le cas où cette contrainte est la contrainte quadratique $\|\mathbf{h}\|_2 = 1$, le canal \mathbf{h} est trouvé comme $V_{\min}(A)$.

3.5.2 Ajustement du SB

De façon similaire à l'ajustement du SS, nous considérons dans le cas de l'ajustement du SB le critère de minimisation suivant:

$$\min_{\mathbf{h}, T} \|\mathcal{T}^H(\mathbf{h}^\perp) - \mathcal{V}_N T\|_F. \quad (3.32)$$

La génération du SB dépend de la paramétrisation linéaire utilisée et donc de $\mathcal{T}^H(\mathbf{h}^\perp)$. Dans le cas de $\mathbf{H}_{min}^{\perp\dagger}(z)$ par exemple, $\mathcal{T}^H(\mathbf{h}^\perp)$ n'engendre pas tout le SB et le critère (3.32) traduit un ajustement entre un sous-espace du SB de dimension réduite (par rapport à celle du SB) et le SB estimé.

De façon similaire au raisonnement de la section précédente, après optimisation par rapport à la matrice de transformation T , nous obtenons

$$\min_{\mathbf{h}} \text{tr} \{ \mathcal{T}(\mathbf{h}^\perp) P_{\mathcal{V}_N}^\perp \mathcal{T}^H(\mathbf{h}^\perp) \} = \text{tr} \{ \mathbf{h}^\perp B \mathbf{h}^{\perp H} \} = \mathbf{h}^H A \mathbf{h}. \quad (3.33)$$

Dans le cas de $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ donné par (2.36), la matrice B peut être déterminée à partir de $P_{\mathcal{V}_N}^\perp = P_{\mathcal{V}_S} = I - \mathcal{V}_N (\mathcal{V}_N^H \mathcal{V}_N)^{-1} \mathcal{V}_N^H$ et la matrice $A = \sum_{i=1}^m \mathcal{S}_i B^* \mathcal{S}_i^H$.

Notons que nous pouvons aussi considérer la paramétrisation linéaire minimale du SB basée sur la prédiction linéaire (qui correspond à la matrice $\mathcal{G}(G)$) dont la dérivation est expliquée dans la section §2.4.1. Dans ce cas le critère d'ajustement du sous-espace donné par (3.32) devient

$$\min_{G,T} \| \mathcal{G}^H(G) - \mathcal{V}_N T \|_F. \quad (3.34)$$

Bien que les méthodes d'ajustement du SS et du SB soient "élégantes" puisqu'elles permettent de trouver des solutions analytiques pour l'estimation du canal, leur inconvénient principal reste la nécessité de disposer d'une décomposition propre de la matrice de covariance (pour avoir $\hat{\mathcal{V}}_S$ dans le cas de l'ajustement du SS ou $\hat{\mathcal{V}}_N$ lorsqu'on veut ajuster le SB), d'où un coût de calcul élevé. Dans la section qui suit, nous nous intéressons à dériver des algorithmes d'ajustement de sous-espaces sans avoir à utiliser la décomposition propre de la matrice de covariance du signal reçu.

3.6 Méthodes sans Décomposition Propre de la Matrice de Covariance

Certains articles tels que [AMCG⁺97] et [ZT97] présentent une analyse de performance des algorithmes du type sous-espace basés sur la propriété d'orthogonalité entre la matrice de convolution du canal et le SB de la matrice de covariance du signal reçu. Une grande partie des algorithmes proposés dans la littérature pour l'ajustement de sous-espace se base sur la décomposition SVD (Singular Value Decomposition) [Sch91]. Cette décomposition rend leur utilisation dans une implémentation temps réel non appropriée. Certains travaux tel que [GvdV96] ont proposé des algorithmes sous-espaces rapides basés sur les décompositions de type QR et HQR (voir [GvdV96] et les références incluses). Nos travaux qui portent sur les méthodes sans décomposition propre de la matrice de covariance qui seront décrites dans cette section sont différents de ceux présentés dans [GvdV96]. Récemment, Kristensson *et al.* ont considéré dans [KOS96] le cas d'une constellation réelle dans un système de communication numérique. En exploitant la propriété de circularité du bruit, les auteurs proposent une approche alternative

d'ajustement de sous-espace qui n'utilise pas la décomposition propre de la matrice de covariance et dont la performance reste similaire à celle de l'ajustement de sous-espace classique. Dans [DAS97a], nous avons proposé deux approches d'ajustement du SS et du SB sans décomposition propre de la matrice de covariance. Bien que ces approches soient présentées dans le contexte de l'identification aveugle du canal, l'idée clé exploitée peut être utilisée dans n'importe quel problème d'ajustement de sous-espace. Nous décrivons dans ce qui suit la dérivation des algorithmes que nous avons proposés dans [DAS97a].

3.6.1 Méthode Alternative pour l'Ajustement du SS

En l'absence du bruit, la matrice de covariance du signal reçu peut être écrite comme

$$R_{YY} = \mathcal{T}(\mathbf{h})R_{AA}\mathcal{T}^H(\mathbf{h}) = \mathcal{V}_S\Lambda'_S\mathcal{V}_S^H + \mathcal{V}_N\Lambda'_N\mathcal{V}_N^H \quad (3.35)$$

où $\Lambda'_S = \Lambda_S - \sigma_v^2 I$ et $\Lambda'_N = 0$. A partir de cette expression, nous constatons que l'espace colonne de $\mathcal{T}(\mathbf{h})$ et l'espace colonne de R_{YY} sont identiques, ce qui permet d'introduire le critère de minimisation suivant

$$\min_{\mathbf{h}, Q} \|\mathcal{T}(\mathbf{h}) - \widehat{R}BQ\|_F^2 \quad (3.36)$$

où \widehat{R} est une estimée consistante de la matrice R_{YY} . La matrice B a la même dimension que la matrice $\mathcal{T}(\mathbf{h})$ et elle est fixée (nous verrons plus loin comment le choix de cette matrice peut se répercuter sur la méthode d'estimation). Notons que l'espace image de la matrice $F = \widehat{R}B$ donne une estimée du SS. Nous pouvons considérer \widehat{R} comme $\widehat{R} = \widehat{R}_{YY} - \widehat{\sigma}_v^2 I = \widehat{R} - \lambda_{\min}(\widehat{R}_{YY})I$, ce qui aboutit à une estimée meilleure pour R . Le problème de minimisation donnée par (3.36) est séparable en $\mathcal{T}(\mathbf{h})$ et Q . La minimisation par rapport à Q , en premier lieu, donne

$$Q = (F^H F)^{-1} F^H \mathcal{T}(\mathbf{h}). \quad (3.37)$$

La substitution de (3.37) dans (3.36) donne

$$\min_{\mathbf{h}} \|P_F^\perp \mathcal{T}(\mathbf{h})\|_F^2 = \min_{\mathbf{h}} \text{tr} \{ \mathcal{T}^H(\mathbf{h}) P_F^\perp \mathcal{T}(\mathbf{h}) \}. \quad (3.38)$$

Sous la contrainte $\|\mathbf{h}\| = 1$, le canal \mathbf{h} est trouvé comme

$$\widehat{\mathbf{h}} = \arg \max_{\|\mathbf{h}\|=1} \text{tr} \{ \mathcal{T}^H(\mathbf{h}) P_F \mathcal{T}(\mathbf{h}) \} = \arg \max_{\|\mathbf{h}\|=1} \mathbf{h}^H \mathcal{F} \mathbf{h} \quad (3.39)$$

où la matrice \mathcal{F} peut être déterminée à partir de P_F . La solution est donc $\widehat{\mathbf{h}} = V_{\max}(\mathcal{F})$. Etant donné que $R_{YY} = \mathcal{T}(\mathbf{h})R_{AA}\mathcal{T}^H(\mathbf{h})$, le choix de la matrice B doit satisfaire certaines conditions. Si les colonnes de B sont dans le SB, par exemple, alors $F = 0$ pour $\widehat{R} = R_{YY}$. Intuitivement, le meilleur choix de la matrice B consiste à la choisir comme $B = \mathcal{T}(\mathbf{h})$, ce qui correspond à un filtrage adapté de $\mathcal{T}^H(\mathbf{h})$ avec $\mathcal{T}(\mathbf{h})$ (la post-multiplication de B avec une matrice carrée non singulière ne change rien dans le critère

puisque cette matrice peut être absorbée dans la matrice Q). Ces considérations, nous amènent à proposer la *procédure en deux étapes* suivante :

- étape 1 : la matrice B est choisie comme une matrice de sélection arbitraire. Cette étape aboutit à une estimée consistante du canal (si la matrice $\mathcal{T}(\mathbf{h})B$ est non singulière).
- étape 2 : l'estimée consistante du canal est utilisée pour former $\mathcal{T}(\hat{\mathbf{h}})$ et résoudre de nouveau (3.36) en considérant $B = \mathcal{T}(\hat{\mathbf{h}})$.

En ce qui concerne la première étape de la procédure, un choix de $B = \begin{bmatrix} I & 0 \end{bmatrix}$ est étroitement lié à la méthode ‘‘Pisarenko rectangulaire’’ de Fuchs [Fuc96]. Nous avons, cependant, remarqué que considérer la matrice B avec la même structure bloc Toeplitz que $\mathcal{T}(\mathbf{h})$ mais remplie avec un canal généré de façon aléatoire aboutit à de bons résultats pour la procédure décrite précédemment.

3.6.2 Analyse des Performances

Nous proposons d'analyser les performances de la méthode et de la comparer avec l'approche d'ajustement du SS classique. Asymptotiquement, nous avons $\hat{R} = R$. Par conséquent, $F = RB = \mathcal{T}(\mathbf{h})R_{AA}\mathcal{T}^H(\mathbf{h})B$. En supposant que $R_{AA} > 0$, alors si $\mathcal{T}^H(\mathbf{h})B$ est non singulière

$$P_F = P_{\mathcal{T}(\mathbf{h})} = P_{\mathcal{V}_S}. \quad (3.40)$$

Si en plus nous avons une estimée consistante du canal, alors asymptotiquement nous pouvons considérer $B = \mathcal{T}(\mathbf{h})$. Dans ce cas, l'utilisation de $R = R_{YY}$ et par conséquent $F = R_{YY}\mathcal{T}(\mathbf{h})$ conduit aussi à (3.40). Nous appliquons dans ce qui suit une analyse de perturbation pour étudier le comportement asymptotique de la procédure en deux étapes.

Pour la première étape de l'algorithme, nous obtenons une estimée consistante du canal si $\mathcal{T}^H(\mathbf{h})B$ est non singulière. Considérons l'analyse de perturbation au premier ordre au niveau de la matrice de projection F . Nous pouvons écrire $\hat{F} = F + \Delta F$ et donc $P_{\hat{F}} = P_F + \Delta P_F$, avec

$$\Delta P_F = 2Sym \left(P_F^\perp \Delta F (F^H F)^{-1} F^H \right) \quad (3.41)$$

où $2Sym(X) = X + X^H$.

- Optimalité du choix $B = \mathcal{T}(\mathbf{h})$:

Considérons l'analyse de perturbation au premier ordre au niveau de la matrice R . Nous pouvons écrire $\hat{R} = R + \Delta R$. En utilisant la décomposition propre de R , nous obtenons au premier ordre :

$$\begin{aligned} \Delta R &= \Delta \mathcal{V}_S \Lambda'_S \mathcal{V}_S^H + \mathcal{V}_S \Delta \Lambda'_S \mathcal{V}_S^H + \mathcal{V}_S \Lambda'_S \Delta \mathcal{V}_S^H \\ &\quad + \Delta \mathcal{V}_N \Lambda'_N \mathcal{V}_N^H + \mathcal{V}_N \Delta \Lambda'_N \mathcal{V}_N^H + \mathcal{V}_N \Lambda'_N \Delta \mathcal{V}_N^H. \end{aligned} \quad (3.42)$$

Soit $B = \mathcal{V}_S B_S + \mathcal{V}_N B_N$ où nous supposons que la matrice B_S est non singulière. En utilisant $\hat{F} = \hat{R}B$, nous avons

$$\begin{aligned} \Delta P_F = & \\ 2Sym \left(P_{\mathcal{V}_N} \left[\Delta \mathcal{V}_S \Lambda'_S + \mathcal{V}_N \Delta \Lambda'_N B_N B_S^{-1} \right] (\mathcal{T}(\mathbf{h}) \mathcal{V}_S)^{-1} R_{AA}^{-1} (\mathcal{T}^H(\mathbf{h}) \mathcal{T}(\mathbf{h}))^{-1} \mathcal{T}^H(\mathbf{h}) \right) & \end{aligned} \quad (3.43)$$

ce qui montre que le choix de $B_N = 0$ est optimal.

- Equivalence asymptotique entre les deux méthodes d'ajustement du SS :
En utilisant $\hat{F} = \left(\hat{R}_{YY} - \lambda_{\min}(\hat{R}_{YY}) I \right) \mathcal{T}(\hat{\mathbf{h}})$ avec \hat{R}_{YY} et $\mathcal{T}(\hat{\mathbf{h}})$ comme des estimées consistantes, nous pouvons démontrer que ΔP_F est la même qu'avec $\hat{F} = \hat{\mathcal{V}}_S$. Par conséquent, nous obtenons au premier ordre

$$P_{(\hat{R}_{YY} - \lambda_{\min}(\hat{R}_{YY}) I) \mathcal{T}(\hat{\mathbf{h}})} = P_{(\hat{R}_{YY} - \sigma_v^2 I) \mathcal{T}(\mathbf{h})} = P_{\hat{\mathcal{V}}_S}. \quad (3.44)$$

Ceci montre que l'approche alternative d'ajustement du SS aboutit asymptotiquement à la même performance de la méthode d'ajustement du SS classique. Notons que lorsque les estimées de σ_v^2 et $\mathcal{T}(\hat{\mathbf{h}})$ sont de plus en plus consistantes, les erreurs d'estimation correspondantes n'ont pas d'influence au premier ordre.

- Méthode simplifiée :
Soit $\mathcal{T}(\hat{\mathbf{h}}) = \mathcal{T}(\mathbf{h}) + \mathcal{T}(\Delta \mathbf{h})$. Lorsque nous utilisons simplement $\hat{F} = \hat{R}_{YY} \mathcal{T}(\hat{\mathbf{h}})$, nous obtenons

$$\Delta P_F = P_{\mathcal{V}_N} \left(\Delta \mathcal{V}_S + \mathcal{T}(\Delta \mathbf{h}) (\mathcal{V}_S^H \mathcal{T}(\mathbf{h}))^{-1} \Lambda_N \Lambda_S^{-1} \right) \mathcal{V}_S^H. \quad (3.45)$$

L'utilisation de \hat{R}_{YY} au lieu de \hat{R} est à l'origine de l'apparition du second terme dans (3.45). L'importance de ce terme est proportionnelle à $\Lambda_N \Lambda_S^{-1}$. Par conséquent, ce terme est négligeable pour de hauts SNRs.

Le coût de calcul² de la méthode d'ajustement du SS sans décomposition propre de la matrice de covariance du signal reçu dépend fortement de la manière avec laquelle la matrice de projection, $P_F^\perp = I - F (F^H F)^{-1} F$, peut être calculée. Nous avons proposé dans [DAS97b] des algorithmes rapides, pour la méthode en question. Dans cette même contribution [DAS97b], nous avons proposé une version de la méthode d'ajustement du SS sans décomposition propre de la matrice de covariance optimalement pondérée et étroitement liée à l'approche sous-espace signal optimalement pondérée proposée par Gorokhov[Gor97b].

2. Une évaluation de la complexité de la méthode d'ajustement du SS sans décomposition propre de la matrice de covariance du signal reçu est présentée dans [DAS97b].

3.6.3 Simulations : Cas de l'Ajustement du SS

Pour illustrer la performance de la méthode d'ajustement du SS sans décomposition propre de la matrice de covariance du signal reçu, nous considérons un canal réel, généré de façon aléatoire, de longueur 6. La séquence de symboles d'entrée est QAM-4 et le nombre de canaux est $m = 4$. La matrice de covariance est calculée à partir d'un paquet de données de longueur 100. La mesure de performance est faite par l'intermédiaire de la racine de l'erreur quadratique moyenne normalisée (NRMSE) sur le canal, en prenant la moyenne sur 100 réalisations Monte-Carlo. Nous avons tracé dans la Figure 3.3 l'erreur NRMSE pour la première étape de l'algorithme, puis la seconde étape ainsi que l'erreur correspondant à la méthode d'ajustement du SS avec décomposition propre de la matrice de covariance. Les courbes obtenues montrent que l'algorithme proposé aboutit à la même performance que l'approche d'ajustement du SS avec décomposition propre de la matrice de covariance. Nous avons utilisé dans les simulations \hat{R}_{YY} et $\hat{R}_{YY} - \lambda_{\min}(\hat{R}_{YY})I$, dans les deux cas nous avons obtenu la même performance pour la méthode en question. Dans la figure 3.4, nous avons tracé l'erreur NRMSE correspondant à l'estimée du canal lorsqu'on utilise une matrice de covariance parfaitement estimée pour les deux étapes de l'algorithme. La comparaison de deux courbes illustre la prépondérance de l'erreur d'estimation de covariance sur l'erreur d'estimation du canal.

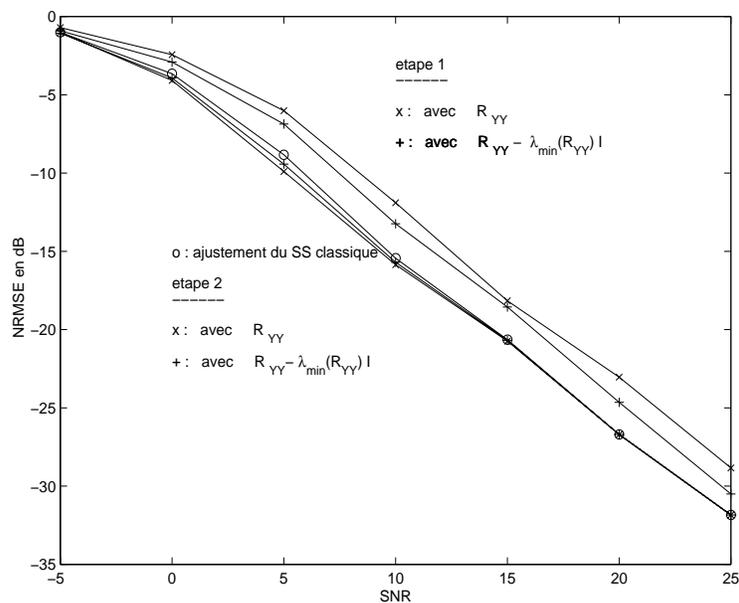


FIG. 3.3 – Performance de l'ajustement du SS sans et avec décomposition propre de R_{YY} (matrice de covariance estimée).

3.6.4 Méthode Alternative pour l'Ajustement du SB

Une approche alternative pour l'ajustement du SB qui n'utilise pas la décomposition propre de la matrice de covariance du signal reçu peut être formulée en considérant le

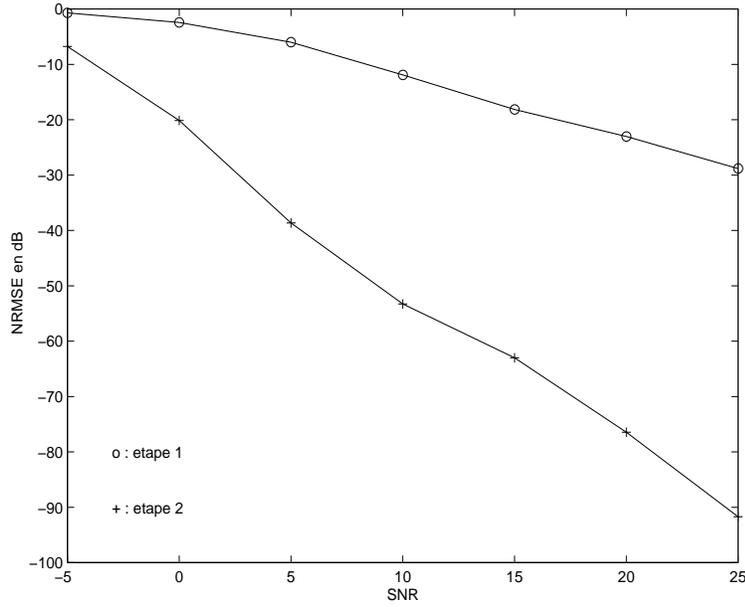


FIG. 3.4 – Performance de l’ajustement du SS sans décomposition propre de R_{YY} (matrice de covariance parfaitement estimée).

critère de minimisation suivant

$$\min_{\mathbf{h}} \|\mathcal{T}(\mathbf{h}^\perp) R_{YY}^{\frac{1}{2}}\|_F^2 \quad (3.46)$$

où la matrice racine carrée est de la forme $R_{YY}^{\frac{1}{2}} = \mathcal{V}_S \Lambda_S^{\frac{1}{2}} Q$ pour une certaine matrice unitaire Q (rappelons que $R_{YY} = \mathcal{T}(\mathbf{h}) R_{AA} \mathcal{T}^H(\mathbf{h}) = \mathcal{V}_S \Lambda_S' \mathcal{V}_S^H + \mathcal{V}_N \Lambda_N' \mathcal{V}_N^H$ avec $\Lambda_S' = \Lambda_S - \sigma_v^2 I$ et $\Lambda_N' = 0$). L’équivalence entre l’ajustement du SB classique et la méthode sans décomposition propre, que nous proposons, peut être démontrée en suivant un raisonnement similaire que celui de l’ajustement du SS. Le critère de minimisation donné par (3.46) est équivalent au critère SRM à une matrice de pondération près : $\Lambda_N^{\frac{1}{2}}$. En effet, nous avons

$$\begin{aligned} \min_{\mathbf{h}} \|\mathcal{T}(\mathbf{h}^\perp) R_{YY}^{\frac{1}{2}}\|_F^2 &= \min_{\mathbf{h}} \text{tr} \left\{ \mathcal{T}(\mathbf{h}^\perp) R_{YY}^{\frac{1}{2}} R_{YY}^{\frac{H}{2}} \mathcal{T}^H(\mathbf{h}^\perp) \right\} \\ &= \min_{\mathbf{h}} \text{tr} \left\{ \mathbf{h}^\perp \left(\sum_{k=N-1}^{M-1} R_{YY}(k) \right) \mathbf{h}^{\perp H} \right\} \\ &= \min_{\mathbf{h}} \mathbf{h}^H A \mathbf{h}, \end{aligned} \quad (3.47)$$

où la matrice A peut être déterminée à partir de la matrice $\sum_{k=N-1}^{M-1} R_{YY}(k)$.

3.6.5 Simulations : Cas de l’Ajustement du SB

Nous considérons un canal complexe, généré de façon aléatoire de longueur $N = 3$. Le nombre de sous-canaux est $m = 3$, la séquence des symboles est i.i.d. BPSK et la

longueur du paquet de données est égale à 210. La mesure de performance est faite par l'intermédiaire de la racine de l'erreur quadratique moyenne normalisée (NRMSE) sur le canal, en prenant la moyenne sur 100 réalisations Monte-Carlo. Dans la figure 3.5, nous avons comparé la performance de la méthode SRM (puisque la méthode SRM et l'approche d'ajustement du SB sans décomposition propre de la matrice de covariance sont équivalentes) à celle de l'ajustement du SB classique (courbe correspondant à NSF). Les deux courbes obtenues se superposent quasiment, ce qui montre que l'ajustement du SB sans décomposition propre de la matrice de covariance aboutit à des performances identiques à celles de la méthode avec décomposition propre de la matrice de covariance.

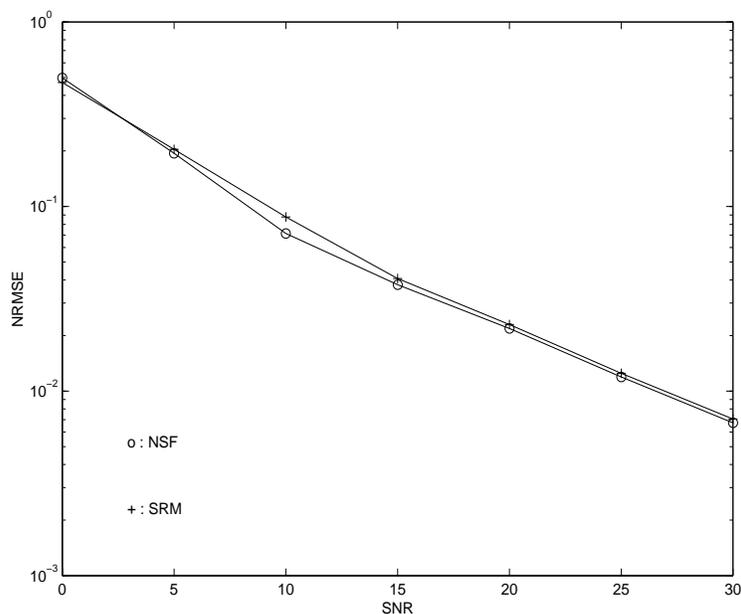


FIG. 3.5 – Performance de l'ajustement du SB sans et avec décomposition propre de R_{YY} .

3.7 Méthode Dérivée de la Prédiction Linéaire

Cette technique a été proposée dans [SP94] et utilise une paramétrisation du SB qui n'est pas linéaire en termes des réponses impulsionnelles des sous-canaux. La notion de sous-espace s'est manifestée dans la méthode SRM à travers la matrice $\mathcal{T}(\mathbf{h}^\perp)$ dont les lignes engendrent le SB de façon totale ou partielle, et elle est apparue de façon explicite dans les méthodes d'ajustement de sous-espaces à travers le SS ou le SB. Dans le cas de la méthode que nous évoquons dans cette section, la notion de sous-espace est implicite et se base sur la prédiction linéaire. Pour expliquer ce mécanisme, commençons tout d'abord par donner un aperçu sur l'approche de la prédiction linéaire dans le contexte de l'identification aveugle multi-canal.

3.7.1 Prédiction Linéaire

Introduite dans [Slo94a], l'approche de prédiction linéaire présente l'avantage de la robustesse à la surestimation de l'ordre du canal. Soit $\mathbf{P}(z) = \sum_{i=0}^L \mathbf{p}(i) z^{-i}$ avec $\mathbf{p}(0) = I_m$ le filtre de l'erreur de prédiction MMSE d'ordre L , dans le cas d'un signal reçu sans bruit $\mathbf{y}(k)$. Si $L \geq \underline{L} = \lceil \frac{N-1}{m-1} \rceil$, d'après [SP94], nous avons (avec probabilité 1, pour un canal FIR complètement aléatoire)

$$\mathbf{P}(z)\mathbf{H}(z) = \mathbf{h}(0). \quad (3.48)$$

L'équation (3.48) montre que $\mathbf{H}(z)$ et $\mathbf{P}(z)$, $\mathbf{h}(0)$ sont deux paramétrisations équivalentes.

3.7.2 Approche de Base

Considérons la matrice de rang plein $\mathbf{h}^\perp(0)$ de dimension $m \times (m-1)$ définie de la façon suivante: $\mathbf{h}^{\perp H}(0)\mathbf{h}(0) = 0$, alors (3.48) implique que $\overline{\mathbf{P}}(z) = \mathbf{h}^{\perp H}(0)\mathbf{P}(z)$ est un polynôme de dimension $(m-1) \times m$ qui satisfait

$$\overline{\mathbf{P}}(z)\mathbf{H}(z) = 0. \quad (3.49)$$

$\overline{\mathbf{P}}(z)$ paramétrise bien le SB. Ce polynôme (ou de façon équivalente $\mathbf{P}(z)$ et $\mathbf{h}(0)$) peut être estimé en utilisant la prédiction linéaire. Si $\overline{\mathbf{P}}(z)$ est estimé de façon robuste à la surestimation de l'ordre du canal, donc l'ordre de $\mathbf{H}(z)$ est connu et $\mathbf{H}(z)$ peut être estimé évidemment à partir de $\overline{\mathbf{P}}(z)$. Ceci peut être établi en considérant le critère de minimisation suivant :

$$\min_{\mathbf{h}} \frac{1}{2\pi j} \oint \mathbf{H}^\dagger(z) \overline{\mathbf{P}}^\dagger(z) \overline{\mathbf{P}}(z) \mathbf{H}(z) \frac{dz}{z}. \quad (3.50)$$

Puisque $\mathbf{H}(z) = \mathbf{Q}(z)\mathbf{h} = \begin{bmatrix} z^{-(N-1)}I_m & z^{-(N-2)}I_m & \dots & I_m \end{bmatrix} \mathbf{h}$, le problème de minimisation exprimé dans (3.50) peut être écrit comme

$$\min_{\mathbf{h}} \frac{1}{2\pi j} \oint \mathbf{h}^H \mathbf{Q}^\dagger(z) \overline{\mathbf{P}}^\dagger(z) \overline{\mathbf{P}}(z) \mathbf{Q}(z) \mathbf{h} \frac{dz}{z} = \min_{\mathbf{h}} \mathbf{h}^H \left(\frac{1}{2\pi j} \oint \mathbf{Q}^\dagger(z) \overline{\mathbf{P}}^\dagger(z) \overline{\mathbf{P}}(z) \mathbf{Q}(z) \frac{dz}{z} \right) \mathbf{h} \quad (3.51)$$

qui est de nouveau de la forme $\min_{\mathbf{h}} \mathbf{h}^H \mathbf{A} \mathbf{h}$.

3.8 Conclusions

En se basant sur les paramétrisations introduites dans le chapitre précédent, nous avons formulé et discuté plusieurs méthodes d'estimation du canal. La méthode la plus simple (et ayant un coût de calcul faible) parmi les techniques d'estimation utilisant la décomposition SVD est la méthode SRM. L'idée clé de cette méthode consiste à minimiser, au sens d'un certain critère, la contribution de la composante du vecteur d'observation contenue dans le SB. Un choix adéquat de la paramétrisation du SB permet

à la méthode SRM d'aboutir à des estimées non biaisées du canal. La seconde classe de méthodes d'estimation basées sur les sous-espaces est celle de l'ajustement du SS et du SB. Bien que les solutions obtenues par l'intermédiaire de ces méthodes soient élégantes en termes d'estimées du canal, l'inconvénient principal de ces approches reste leur coût de calcul élevé dû à la décomposition propre de la matrice de covariance du signal reçu. Pour remédier à ce problème, nous avons proposé des méthodes alternatives d'ajustement de sous-espaces qui ne nécessitent pas la décomposition propre de la matrice de covariance du signal reçu. L'analyse de performance théorique, à laquelle viennent s'ajouter les résultats de simulations, prouvent que les performances des méthodes sans décomposition propre de la matrice de covariance sont asymptotiquement identiques à leurs versions classiques.

Chapitre 4

Méthodes de Maximum de Vraisemblance

Dans ce chapitre, nous nous intéressons au développement d'algorithmes d'estimation du canal basés sur le critère du maximum de vraisemblance (ML). L'intérêt d'une telle estimation réside dans son efficacité asymptotique : l'estimateur ML est asymptotiquement efficace (asymptotiquement l'estimateur ML atteint sa borne de Cramer-Rao). Selon les hypothèses faites sur les symboles, nous pouvons distinguer deux contextes d'estimation ML. Dans le cas où les symboles sont modélisés comme des inconnues déterministes (la partie stochastique provient uniquement du bruit blanc additif gaussien), l'approche ML est dite déterministe (DML) et la borne de Cramer-Rao correspondante est déterministe. Si les symboles sont considérés comme des variables aléatoires qui suivent une distribution gaussienne, la méthode ML est qualifiée de gaussienne (GML) ou stochastique et lui correspond une borne de Cramer-Rao gaussienne. Notons que cette hypothèse permet de renforcer le problème d'estimation par rapport au modèle déterministe en termes de conditions d'identifiabilité (qui deviennent moins restrictives) et en termes de performance de la méthode ainsi que de la borne de Cramer-Rao correspondante.

Dans la dernière partie de ce chapitre, nous décrirons l'exploitation de l'alphabet fini dans le contexte de l'estimation aveugle multi-canal. L'incorporation de cette information sur les symboles transmis permet d'améliorer, de manière significative, la performance de l'estimation aveugle. En effet, cette performance devient équivalente à celle obtenue par une méthode d'identification de canal basée sur une séquence d'apprentissage.

4.1 Introduction

Dans [AdCS98], nous avons analysé l'estimation multi-canal basée sur les techniques ML. La fonction de coût ML est non linéaire (en termes des paramètres à estimer). Une façon simple pour optimiser cette fonction de coût est la stratégie "Iterative Quadratic ML" (IQML) [BM86],[KSS86], dans laquelle la fonction de coût est considérée comme une fonction quadratique pondérée. L'initialisation de cette procédure peut être faite par

l'une des méthodes d'estimation discutées dans le chapitre précédent (SRM, SSF, NSF). La solution devient, par conséquent, simple à calculer mais d'autres itérations sont nécessaires pour améliorer la pondération. Cependant, asymptotiquement, deux itérations sont suffisantes pour assurer la convergence de la méthode. Le problème réside dans le fait que l'estimée IQML converge vers l'estimée ML asymptotiquement en termes du SNR et non asymptotiquement en termes de la quantité de données. L'inconvénient principal de la méthode IQML est donc qu'en présence du bruit, elle fournit des estimées biaisées du canal [SLS97] qui, par conséquent, ne sont pas consistantes. Pour remédier à ce problème, nous avons introduit, dans une première étape, la méthode IQML débruitée (DIQML) qui fournit des estimées non biaisées du canal. Nous avons aussi introduit une formulation différente de la fonction de coût ML comme un critère quadratique pondéré de façon à ce que son gradient corresponde exactement au gradient du critère ML. Nous appelons cette approche "Pseudo-Quadratic ML" (PQML). Les deux techniques DIQML et PQML introduisent un second terme dans la fonction de coût par rapport à la fonction de coût IQML. Il s'avère, en plus, que la moyenne statistique du second terme dans la technique PQML est le second terme dans la méthode DIQML. Les deux techniques DIQML et PQML aboutissent à de très bonnes performances grâce à un facteur d'ajustement particulier pour lequel nous avons introduit un choix judicieux. Ce facteur d'ajustement correspond dans le cas de la technique DIQML à l'estimée de la variance du bruit. Notons qu'une méthode analogue à l'approche DIQML, basée sur une modification du critère IQML et qui aboutit à des estimées consistantes dans le cas de l'estimation des directions d'arrivée (DOA) a été introduite par Kristensson *et al.* dans [KJO98] puis elle a été analysée dans [Kri98],[KJO].

Dans ce qui suit, nous décrivons le problème DML dans le cas de l'estimation aveugle multi-canal et nous présentons l'estimateur DML et IQML. Ensuite, nous expliquons les modifications que nous avons introduites sur le critère IQML pour obtenir l'algorithme DIQML. L'analyse du nouvel algorithme fait apparaître des liens entre ce dernier et l'algorithme PQML dont nous décrivons la dérivation. Puis nous étendons cette dernière stratégie au cas semi-aveugle pour dériver des algorithmes d'estimation du canal semi-aveugle basés sur la combinaison de la technique PQML et un critère basé sur une séquence d'apprentissage. Nous présentons, ensuite, une analyse de performance et une comparaison entre les différents algorithmes discutés ainsi que les résultats des simulations obtenus.

4.2 Deterministic ML (DML) et Iterative Quadratic ML (IQML)

Le critère DML a été formulé pour l'estimation des directions d'arrivée dans le cas d'une antenne linéaire dans [Böh84]. Dans le contexte de l'estimation aveugle multi-canal, la méthode DML a été introduite dans [Slo94a] pour le cas de $m = 2$ canaux,

puis étendue pour un nombre m arbitraire de canaux dans [Slo94c]. Une extension au cas multi-utilisateurs a été présentée dans [Slo94b],[Slo95]. Une discussion de la méthode ML peut être trouvée dans [SP95], [dCS96]. Puisque le problème DML introduit une projection sur le SB, une solution attrayante (d'un point de vue coût de calcul) consiste à utiliser les paramétrisations linéaires du SB. Une paramétrisation intéressante du SB, pouvant être utilisée dans la formulation du critère DML, est celle que nous avons décrite dans la section §2.4.1 qui est minimale et est basée sur les quantités de prédiction de la composante signal du signal reçu.

4.2.1 Deterministic ML (DML)

Les symboles transmis $a(k)$ sont supposés être des quantités déterministes. Ceci implique que, non seulement, nous ne considérons pas la fonction de densité de probabilité réelle des symboles transmis mais aussi que nous les considérons comme des quantités déterministes. Les paramètres du canal sont à leur tour supposés être déterministes. Par conséquent, l'unique entité stochastique est supposée provenir du bruit additif. Ce dernier est statistiquement décrit comme suit : la séquence du bruit additif $v(k)$ est modélisée comme une variable aléatoire complexe i.i.d. avec une moyenne nulle et une variance σ_v^2 . Nous supposons aussi que les parties réelle et imaginaire du bruit sont indépendantes entre elles. La fonction de densité de probabilité complexe est donnée par

$$f(\mathbf{Y}|\mathbf{h}) = \frac{1}{(2\pi\sigma_v^2)^{\frac{mM}{2}}} e^{-\frac{1}{2\sigma_v^2}\|\mathbf{Y}-\mathcal{T}(\mathbf{h})A\|^2}. \quad (4.1)$$

$f(\mathbf{Y}|\mathbf{h})$ est la distribution conditionnelle gaussienne de \mathbf{Y} étant donné \mathbf{h} (en supposant que le bruit \mathbf{V} est blanc, gaussien et circulaire). La maximisation de la fonction de vraisemblance revient à la minimisation du problème des moindres carrés suivant :

$$\max_{A,\mathbf{h}} f(\mathbf{Y}|\mathbf{h}) \Leftrightarrow \min_{A,\mathbf{h}} \|\mathbf{Y} - \mathcal{T}(\mathbf{h})A\|^2. \quad (4.2)$$

Le problème donné par (4.2) est séparable. Pour \mathbf{h} fixé, la séquence des symboles transmis optimale est

$$A = (\mathcal{T}^H(\mathbf{h})\mathcal{T}(\mathbf{h}))^{-1} \mathcal{T}^H(\mathbf{h})\mathbf{Y}. \quad (4.3)$$

En remplaçant A donné par (4.3) dans (4.2), nous obtenons

$$\min_{\mathbf{h}} \|P_{\mathcal{T}(\mathbf{h})}^\perp \mathbf{Y}\|^2. \quad (4.4)$$

Le problème DML revient donc à une projection sur le complément orthogonal du SS : $(\text{Image}\{\mathcal{T}(\mathbf{h})\})^\perp$. D'où l'importance des paramétrisations linéaires du SB discutées dans la section §2.4, puisque selon le choix de la paramétrisation en question le SB peut être engendré de façon complète ou incomplète. Dans le cas de $\mathbf{H}_{bal,min}^{\perp\dagger}$ par exemple, nous avons $(\text{Image}\{\mathcal{T}(\mathbf{h})\})^\perp = \text{Image}\{\mathcal{T}^H(\mathbf{h}^\perp)\}$ et par conséquent $P_{\mathcal{T}(\mathbf{H})}^\perp = P_{\mathcal{T}^H(\mathbf{h}^\perp)}$ alors que dans le cas de $\mathbf{H}_{min}^{\perp\dagger}$ nous avons $\text{Image}\{\mathcal{T}^H(\mathbf{h}^\perp)\} \subset (\text{Image}\{\mathcal{T}(\mathbf{h})\})^\perp$ et donc

$P_{\mathcal{T}}^{\perp}(\mathbf{H}) \approx P_{\mathcal{T}^H}(\mathbf{h}^{\perp})$ mais l'erreur d'approximation disparaît asymptotiquement. Le choix de $\mathbf{H}_{max}^{\perp\dagger}$ comme paramétrisation linéaire du SB correspond à la méthode ML de Hua [Hua96]. Dans le cas où la paramétrisation $\mathbf{H}^{\perp\dagger}$ permet d'engendrer totalement le SB, le critère de minimisation donné par (4.4) devient

$$\begin{aligned}
 \min_{\mathbf{h}} \|P_{\mathcal{T}}^{\perp}(\mathbf{h})\mathbf{Y}\|^2 &= \min_{\mathbf{h}} \|P_{\mathcal{T}^H}(\mathbf{h}^{\perp})\mathbf{Y}\|^2 \\
 &= \min_{\mathbf{h}} \mathbf{Y}^H P_{\mathcal{T}^H}(\mathbf{h}^{\perp}) \mathbf{Y} \\
 &= \min_{\mathbf{h}} \mathbf{Y}^H \mathcal{T}^H(\mathbf{h}^{\perp}) [\mathcal{T}(\mathbf{h}^{\perp})\mathcal{T}^H(\mathbf{h}^{\perp})]^+ \mathcal{T}(\mathbf{h}^{\perp}) \mathbf{Y} \\
 &= \min_{\mathbf{h}} \mathbf{h}^H \left(\mathcal{Y}^H [\mathcal{T}(\mathbf{h}^{\perp})\mathcal{T}^H(\mathbf{h}^{\perp})]^+ \mathcal{Y} \right) \mathbf{h} \\
 &= \min_{\mathbf{h}} \mathbf{h}^H \mathbf{A} \mathbf{h}
 \end{aligned} \tag{4.5}$$

où la matrice \mathcal{Y} vérifie $\mathcal{T}(\mathbf{h}^{\perp})\mathbf{Y} = \mathcal{Y}\mathbf{h}$ et $+$ désigne la pseudo-inverse de Moore-Penrose. Notons que le problème de minimisation donné par (4.5) est non linéaire en termes des paramètres à estimer. Sous la contrainte quadratique $\|\mathbf{h}\|_2 = 1$, l'estimée du canal est donné par

$$\hat{\mathbf{h}}_{DML} = \arg \min_{\mathbf{h}} \text{tr} \left\{ P_{\mathcal{T}^H}(\mathbf{h}^{\perp}) \mathbf{Y} \mathbf{Y}^H \right\}. \tag{4.6}$$

Notons qu'en interprétant la méthode SRM, qui minimise $\|\mathcal{T}(\mathbf{h}^{\perp})\mathbf{Y}\|_2^2$, comme un problème des moindres carrés, le critère DML correspond à sa *version pondérée optimalement*. En effet, le bruit dans $\mathcal{T}(\mathbf{h}^{\perp})\mathbf{Y}$ est $\mathcal{T}(\mathbf{h}^{\perp})\mathbf{V}$ avec une matrice de covariance $\sigma_v^2 \mathcal{T}(\mathbf{h}^{\perp})\mathcal{T}^H(\mathbf{h}^{\perp})$. Asymptotiquement, n'importe quel choix de $\mathbf{H}^{\perp}(z)$ conduit à la même performance pour la méthode DML puisque n'importe quelle paramétrisation $\mathbf{H}^{\perp\dagger}(z)$ contient une paramétrisation $\mathbf{H}_{min}^{\perp\dagger}(z)$ et donc

$$P_{\mathbf{H}^{\perp}(z)} = P_{\mathbf{H}_{min}^{\perp}(z)} = P_{\mathbf{H}(z)}^{\perp}. \tag{4.7}$$

La résolution du critère donné par (4.5) peut être faite de manière itérative, où à chaque itération un problème quadratique apparaît.

4.2.2 Iterative Quadratic ML (IQML)

La stratégie IQML ([BM86],[KSS86]) considère le numérateur, $\mathcal{Y}^H \mathcal{Y}$, du critère donné par (4.5), et utilise l'estimée de \mathbf{h} obtenue à l'itération précédente dans le dénominateur, $\mathcal{R}(\mathbf{h}) = \mathcal{T}(\mathbf{h}^{\perp})\mathcal{T}^H(\mathbf{h}^{\perp})$, de ce même critère. L'initialisation de l'algorithme IQML peut être faite par la méthode SRM (d'où l'importance du choix de la paramétrisation linéaire du SB pour éviter le biais dans l'estimée du canal utilisée). Désignons, dans ce cas, l'estimée du canal \mathbf{h} obtenue à la $k^{\text{ème}}$ itération de la méthode IQML par $\hat{\mathbf{h}}_k$, la méthode IQML est décrite comme suit :

$$\begin{aligned}
 \hat{\mathbf{h}}_k &= \arg \min_{\mathbf{h}} \mathcal{C}_{IQML}^k(\mathbf{h}) \\
 \mathcal{C}_{IQML}^k(\mathbf{h}) &= \begin{cases} \mathbf{h}^H \mathcal{Y}^H \mathcal{Y} \mathbf{h} & \text{pour } k = 1, \\ \mathbf{h}^H \mathcal{Y}^H \mathcal{R}^+(\hat{\mathbf{h}}_{k-1}) \mathcal{Y} \mathbf{h} & \text{pour } k > 1. \end{cases}
 \end{aligned} \tag{4.8}$$

La minimisation de la fonction de coût $\mathcal{C}_{IQML}^k(\mathbf{h})$ est faite sous la contrainte quadratique $\|\mathbf{h}\| = 1$, cependant les estimées du canal obtenues par la procédure itérative donnée par (4.8) ne sont pas garanties d'être consistantes [SLS97]. En effet, pour des hautes valeurs du SNR l'estimée du canal obtenue par la méthode IQML est consistante. Cependant, pour des valeurs faibles du SNR la méthode IQML est biaisée parce que le vrai canal n'est pas un point stationnaire de l'algorithme IQML et par conséquent ce dernier aboutit à des performances médiocres même s'il est initialisé avec une estimée consistante du canal.

Nous proposons, dans ce qui suit, de modifier le critère IQML pour le *débruiter*. L'élimination de la contribution du bruit par l'intermédiaire de l'opération de débruitage va permettre au critère débruité, auquel nous appliquons la stratégie IQML, d'aboutir à des estimées consistantes du canal.

4.3 Denoised IQML (DIQML) et Pseudo-Quadratic ML (PQML)

Nous présentons deux stratégies d'estimation du canal basées sur le critère DML. Dans la première stratégie, nous proposons de modifier l'algorithme IQML dans le but de le débruiter et par conséquent obtenir des estimées consistantes du canal. La deuxième procédure, que nous appelons PQML, est naturellement asymptotiquement débruitée. Nous établissons les liens qui existent entre les deux approches et nous démontrons leur convergence globale. Nous supposons que les conditions d'identifiabilité aveugle sont vérifiées : le canal est irréductible (les sous-canaux ne partagent pas de zéros en commun), la séquence de symboles d'entrée a une excitation persistante et le paquet de données est suffisamment long.

4.3.1 Denoised IQML (DIQML)

L'algorithme IQML résout le problème de minimisation exprimé dans (4.5) itérativement : à chaque itération un problème quadratique apparaît. En effet, à chaque itération de l'algorithme IQML, le dénominateur \mathcal{R} , calculé grâce à l'itération précédente, est considéré comme constant et par conséquent le critère donné dans (4.5) devient quadratique. Sous la contrainte $\|\mathbf{h}\| = 1$, le canal \mathbf{h} est estimé comme le vecteur propre correspondant à la valeur propre minimale de la matrice $\mathcal{Y}^H \mathcal{R}^+ \mathcal{Y}$. L'algorithme IQML nécessite une bonne initialisation. Dans [Hua96], l'auteur initialise le canal par son estimée obtenue par la méthode SRM [BR94] et démontre que l'estimée du canal ainsi obtenue par la méthode IQML est consistante à des hauts SNRs. A des bas SNRs, la présence du bruit biaise l'estimée du canal obtenue par la méthode IQML et la performance de la méthode IQML reste médiocre même si l'initialisation est consistante. Nous proposons une méthode pour débruiter le critère DML : le critère débruité résolu selon la stratégie IQML aboutira à des estimées consistantes du canal.

Considérons la situation asymptotique, où le nombre de données M est infini. Par la loi des grands nombres, le critère DML est équivalent à sa moyenne statistique qui est $\text{tr}\{P_{\mathcal{T}^H(\mathbf{h}^\perp)}E(\mathbf{Y}\mathbf{Y}^H)\}$ et $E(\mathbf{Y}\mathbf{Y}^H) = \mathbf{X}\mathbf{X}^H + \sigma_v^2I$, où $\mathbf{X} = \mathcal{T}(\mathbf{h})A$ est la composante purement signal du signal reçu \mathbf{Y} . La stratégie de débruitage consiste à éliminer le terme asymptotique correspondant au bruit du critère DML qui devient :

$$\min_{\|\mathbf{h}\|=1} \text{tr}\{P_{\mathcal{T}^H(\mathbf{h}^\perp)}(\mathbf{Y}\mathbf{Y}^H - \sigma_v^2I)\}. \quad (4.9)$$

Notons que cette opération ne change pas la solution du critère DML puisque la quantité soustraite $\sigma_v^2\text{tr}\{P_{\mathcal{T}^H(\mathbf{h}^\perp)}\} = \sigma_v^2(M(m-1)-N+1)I$ est constante. Le problème (4.9) est résolu selon la stratégie IQML : en considérant \mathcal{R} comme constante, le problème de minimisation devient quadratique

$$\min_{\|\mathbf{h}\|=1} \mathbf{h}^H \{\mathcal{Y}^H\mathcal{R}^+\mathcal{Y} - \sigma_v^2D\} \mathbf{h} \quad (4.10)$$

où la matrice D vérifie $\mathbf{h}^H D \mathbf{h} = \text{tr}\{\mathcal{T}^H(\mathbf{h}^\perp)\mathcal{R}^+\mathcal{T}(\mathbf{h}^\perp)\}$.

Asymptotiquement en termes de nombre de données, l'algorithme DIQML est globalement convergent. En effet, asymptotiquement, il est équivalent au critère débruité suivant :

$$\min_{\|\mathbf{h}\|=1} \mathbf{h}^H \mathcal{X}^H \mathcal{R}^+ \mathcal{X} \mathbf{h} \quad (4.11)$$

où \mathcal{X} est la matrice remplie avec les échantillons de la composante purement signal du signal reçu \mathbf{Y} et qui vérifie $\mathcal{X}\mathbf{h} = \mathcal{T}(\mathbf{h}^\perp)\mathbf{X}$. La première itération de l'algorithme DIQML aboutit à une estimée consistante du canal quelle que soit la valeur de \mathcal{R} (et donc quelle que soit l'initialisation du canal), la solution de (4.11) est \mathbf{h}_o , le vrai canal (si $\text{Ker}(\mathcal{R}^+) \cap \text{Image}(\mathcal{X}) = 0$, ce qui peut être facilement garanti). La seconde itération aboutit au minimiseur global du critère ML. Evidemment, à des hauts SNRs, la convergence globale est assurée comme c'est le cas pour l'algorithme d'origine : IQML.

En pratique, avec un grand nombre (mais fini) de données M , la matrice centrale dans le critère (4.10) est indéfinie. Par conséquent, au lieu d'éliminer le terme asymptotique exact σ_v^2D , nous éliminons une quantité λD qui ne change pas le critère (comme nous l'avons déjà mentionné) et qui est suffisante pour rendre la matrice $\mathcal{Q}(\mathbf{h}) = \mathcal{Y}^H \mathcal{R}^+ \mathcal{Y} - \lambda D$ semi-définie positive avec une seule singularité. Le canal \mathbf{h} est solution de :

$$\min_{\|\mathbf{h}\|=1, \lambda} \mathbf{h}^H \{\mathcal{Y}^H \mathcal{R}^+ \mathcal{Y} - \lambda D\} \mathbf{h} \quad (4.12)$$

sous la contrainte “ $\mathcal{Q}(\mathbf{h})$ est semi-définie positive”. Le canal \mathbf{h} est le vecteur propre généralisé correspondant à la valeur propre généralisée minimale de $\mathcal{Y}^H \mathcal{R}^+ \mathcal{Y}$ et D , et λ est la valeur propre généralisée minimale. Asymptotiquement, la valeur $\lambda \rightarrow \sigma_v^2$, et le critère est équivalent à (4.11) : la convergence asymptotique est valable pour le canal \mathbf{h} et pour λ qui tend vers σ_v^2 .

4.3.2 Pseudo-Quadratic ML (PQML)

Le principe de PQML a été appliqué en premier lieu pour l'estimation des sinusôides dans le bruit [OS91] puis pour l'estimation aveugle de canal dans [HB96]. Le gradient de la fonction de coût DML peut être arrangé comme $\mathcal{P}(\mathbf{h})\mathbf{h}$, où la matrice $\mathcal{P}(\mathbf{h})$ est (idéalement) semi-définie positive. La solution du problème ML vérifie $\mathcal{P}(\mathbf{h})\mathbf{h} = 0$, qui est résolu sous la contrainte $\|\mathbf{h}\| = 1$, selon la stratégie PQML comme suit : $\mathcal{P}(\mathbf{h})\mathbf{h}$ est aussi le gradient de la fonction de coût quadratique $\mathbf{h}^H \mathcal{P}(\mathbf{h}) \mathbf{h}$ dans laquelle la matrice $\mathcal{P}(\mathbf{h})$ est considérée constante, et comme $\mathcal{P}(\mathbf{h})$ est semi-définie positive, l'estimée du canal \mathbf{h} est choisie dans [HB96] comme le vecteur propre correspondant à la valeur absolue de la valeur propre minimale de $\mathcal{P}(\mathbf{h})$. Cette solution est utilisée pour réévaluer $\mathcal{P}(\mathbf{h})$ et d'autres itérations peuvent ainsi être faites. La difficulté réside dans la manière de trouver la matrice adéquate $\mathcal{P}(\mathbf{h})$ et surtout avec la contrainte “ $\mathcal{P}(\mathbf{h})$ semi-définie positive”. Dans notre problème :

$$\mathcal{P}(\mathbf{h}) = \mathcal{Y}^H \mathcal{R}^+ \mathcal{Y} - \mathcal{B}^H \mathcal{B} \quad (4.13)$$

où \mathcal{B} est la matrice définie tel que $\mathcal{T}^H(\mathbf{h}^\perp) \mathcal{B} = \mathcal{B}^* \mathbf{h}^*$ avec $\mathcal{B} = [\mathcal{T}(\mathbf{h}^\perp) \mathcal{T}^H(\mathbf{h}^\perp)]^+ \mathcal{T}(\mathbf{h}^\perp) \mathbf{Y}$. Asymptotiquement, l'effet du second terme est d'éliminer la contribution du bruit présent dans le premier terme. Le critère est asymptotiquement globalement convergent : n'importe quelle initialisation de $\mathcal{P}(\mathbf{h})$ aboutit à une estimée consistante du canal et la seconde itération aboutit au minimiseur global.

La matrice $\mathcal{P}(\mathbf{h})$ est indéfinie pour un nombre de données fini M , et l'application directe de la stratégie PQML n'aboutira pas aux résultats attendus comme il est mentionné dans [HB96], sauf pour des hauts SNRs. Remarquons que la méthode PQML est en étroite relation avec la méthode DIQML puisque les premiers termes de (4.10) et (4.13) sont identiques et la relation entre les seconds termes respectifs s'exprime comme

$$E(\mathcal{B}^H \mathcal{B}) = \sigma_v^2 D. \quad (4.14)$$

En d'autres termes, le second terme du gradient dans la méthode PQML est *remplacé* par sa moyenne statistique dans l'approche DIQML. Par analogie avec la stratégie DIQML, nous introduisons un facteur arbitraire λ de façon à ce que le critère PQML devienne le problème de minimisation suivant :

$$\min_{\|\mathbf{h}\|=1, \lambda} \mathbf{h}^H \{ \mathcal{Y}^H \mathcal{R}^+ \mathcal{Y} - \lambda \mathcal{B}^H \mathcal{B} \} \mathbf{h} \quad (4.15)$$

avec la contrainte “semi-définie positive” imposée à la matrice centrale. L'estimée du canal \mathbf{h} est le vecteur propre généralisé correspondant à la valeur propre généralisée minimale de $\mathcal{Y}^H \mathcal{R}^+ \mathcal{Y}$ et $\mathcal{B}^H \mathcal{B}$, et λ est la valeur propre généralisée minimale. Asymptotiquement, il y a une convergence globale pour \mathbf{h} , et pour λ qui tend vers un. Les points stationnaires de PQML sont les mêmes que ceux de DML, c'est pourquoi l'algorithme PQML possède la même performance que DML. Asymptotiquement, l'algorithme PQML aboutit au minimiseur global du problème ML. Vu l'approximation introduite dans le second

terme de la méthode DIQML par rapport au second terme du gradient de l'approche PQML, l'algorithme DIQML n'aboutit pas au minimiseur global du problème ML et sa performance reste donc inférieure à celles de PQML et DML.

4.4 Extensions au Cas Semi-aveugle

Contrairement aux approches purement aveugles, les techniques d'estimation du canal semi-aveugle exploitent la connaissance de certains symboles dans le paquet de données. Ces méthodes apparaissent plus robustes que les approches purement aveugles et aboutissent à des performances supérieures [dCS97b]. Ces approches permettent l'estimation correcte du canal lorsque les méthodes d'estimation purement aveugles et celles basées sur les séquences d'apprentissage échouent séparément. Nous proposons deux techniques semi-aveugles basées sur le principe ML et qui combinent l'approche PQML que nous venons de décrire avec un critère basé sur une séquence d'apprentissage. Nous ne considérons plus l'approche DIQML puisque la méthode PQML aboutit à des performances meilleures. Nous supposons, pour des raisons de simplicité, que les symboles formant la séquence d'apprentissage sont regroupés et que cette même séquence est située au début du paquet de données. Le vecteur des symboles est $A = [A_k^H \ A_u^H]^H$, où la portion A_k regroupe M_k symboles connus, et la portion A_u est constituée des M_u symboles inconnus du paquet de données.

4.4.1 PQML Least-Squares (PQML-LS)

La première approche, que nous appelons PQML-LS, provient d'un point de vue purement déterministe. Nous appliquons l'approche DML à : $\mathbf{Y} = [\mathbf{Y}_{TS}^H \ \mathbf{Y}_B^H]^H$. La portion $\mathbf{Y}_{TS} = \mathcal{T}_{TS}(\mathbf{h})A_k + \mathbf{V}_{TS}$ regroupe les observations contenant uniquement les symboles connus. La portion \mathbf{Y}_B regroupe les observations contenant les symboles inconnus : ses premiers $N - 1$ composants incluent un mélange de symboles connus et inconnus. Nous n'exploitons pas la connaissance de ces symboles dans ce contexte et donc nous traiterons tous ces symboles comme inconnus. Notons que cette perte d'information de ces symboles peut être critique surtout lorsque la séquence d'apprentissage est trop courte, de longueur inférieure à N symboles !

Comme les vecteurs \mathbf{Y}_{TS} et \mathbf{Y}_B sont découplés en termes du bruit, le critère DML pour \mathbf{Y} est la somme des critères DML pour \mathbf{Y}_{TS} et \mathbf{Y}_B :

$$\min_{\mathbf{h}} \mathbf{Y}_B^H P_{\mathcal{T}^H(\mathbf{h}^\perp)} \mathbf{Y}_B + \|\mathbf{Y}_{TS} - \mathcal{T}_{TS}(\mathbf{h})A_k\|^2. \quad (4.16)$$

Ce critère semi-aveugle est résolu selon la méthode PQML. La stratégie PQML semi-aveugle s'opère comme suit : le gradient de la fonction de coût peut être écrit comme $\mathcal{Q}(\mathbf{h})\mathbf{h} + \mathcal{S}(\mathbf{h})$ où $\mathcal{Q}(\mathbf{h})$ est (idéalement) semi-définie positive. A chaque itération, nous supposons que les matrices $\mathcal{Q}(\mathbf{h})$ et $\mathcal{S}(\mathbf{h})$ sont constantes, et donc l'estimée du canal \mathbf{h}

est la solution d'un système linéaire qui sera utilisée pour réévaluer $\mathcal{Q}(\mathbf{h})$ et $\mathcal{S}(\mathbf{h})$ pour faire d'autres itérations. Les quantités qui nous intéressent sont les suivantes :

$$\begin{aligned}\mathcal{Q}(\mathbf{h}) &= \mathcal{Y}_B^H \mathcal{R}_B^+ \mathcal{Y}_B - \mathcal{B}_B^H \mathcal{B}_B + \mathcal{A}_{TS}^H \mathcal{A}_{TS}, \\ \mathcal{S}(\mathbf{h}) &= \mathcal{A}_{TS}^H \mathbf{Y}_{TS},\end{aligned}\tag{4.17}$$

où la matrice \mathcal{A}_{TS} est définie telle que $\mathcal{T}_{TS}(\mathbf{h})A_k = \mathcal{A}_{TS}\mathbf{h}$. La partie aveugle du critère est résolue selon la méthode PQML et celle correspondant à la séquence d'apprentissage est résolue comme étant un problème des moindres carrés (LS). Nous introduisons la même stratégie de valeur propre généralisée décrite dans la section précédente qui permet à la matrice $\mathcal{Q}(\mathbf{h})$ d'être semi-définie positive. A une itération donnée, l'expression de l'estimée du canal \mathbf{h} est :

$$\mathbf{h} = (\mathcal{Y}_B^H \mathcal{R}_B^+ \mathcal{Y}_B - \lambda \mathcal{B}_B^H \mathcal{B}_B + \mathcal{A}_{TS}^H \mathcal{A}_{TS})^{-1} \mathcal{A}_{TS}^H \mathbf{Y}_{TS}.\tag{4.18}$$

Les différentes quantités figurant dans l'expression (4.18) sont calculées grâce à l'itération précédente. Ce critère nécessite au moins N symboles connus pour que l'algorithme fonctionne.

4.4.2 PQML Weighted-Least-Squares (PQML-WLS)

L'approche PQML-WLS combine un point de vue déterministe et un point de vue stochastique. Dans le modèle gaussien [AS97], [dCS97b], les symboles d'entrée sont considérés comme étant des variables aléatoires gaussiennes. Cette hypothèse permet de renforcer le problème par rapport au modèle déterministe : un canal irréductible peut être estimé à un facteur de phase près et pour uniquement un seul symbole connu qui n'est pas situé sur les bords du paquet d'où n'importe quel canal (sans contrainte d'irréductibilité) peut être identifiable.

Décomposons le vecteur d'observation \mathbf{Y} comme $\mathbf{Y} = [\mathbf{Y}_{SB}^H \mathbf{Y}_B^H]^H$. La portion \mathbf{Y}_B regroupe toutes les observations où les symboles inconnus apparaissent. La portion \mathbf{Y}_{SB} regroupe toutes les observations contenant les symboles connus et $N - 1$ observations où un mélange de symboles connus et inconnus apparaît. Les symboles dans \mathbf{Y}_B sont traités comme déterministes et donc nous appliquons DML pour le vecteur \mathbf{Y}_B . Dans \mathbf{Y}_{SB} , les symboles connus sont traités comme déterministes et les symboles inconnus comme variables aléatoires gaussiennes i.i.d. de moyenne nulle et de variance σ_v^2 . L'approche GML appliquée à \mathbf{Y}_{SB} permet de prendre en compte les symboles connus, alors que la méthode PQML-LS ne le permettait pas. Comme les deux vecteurs \mathbf{Y}_{TS} et \mathbf{Y}_B sont découplés en termes du bruit, le critère ML combiné sera la somme du critère DML appliqué à \mathbf{Y}_B , que nous résolvons selon la stratégie PQML, et du critère GML appliqué à \mathbf{Y}_{SB} .

Le vecteur \mathbf{Y}_{SB} s'écrit comme $\mathbf{Y}_{SB} = \mathcal{T}_{SB}(\mathbf{h})A_{SB} + \mathbf{V}_{SB}$, où $A_{SB} = [A_k^H A_u']$ et A_u' sont les symboles connus dans \mathbf{Y}_{SB} . Le vecteur \mathbf{Y}_{SB} est gaussien : $\mathbf{Y}_{SB} \sim$

$\mathcal{N}(\mathcal{T}_{SB}(\mathbf{h})A_{SB}^{\circ}, C_{Y_{SB}Y_{SB}})$, A_{SB}° est la moyenne des symboles, C_{AA} leur matrice de covariance et $C_{Y_{SB}Y_{SB}} = \mathcal{T}_{SB}(\mathbf{h})C_{AA}\mathcal{T}_{SB}^H(\mathbf{h}) + \sigma_v^2 I$. L'approche GML considère l'estimation conjointe de \mathbf{h} et de σ_v^2 à partir du critère d'optimisation :

$$\begin{aligned}
 & \max_{\mathbf{h}, \sigma_v^2} f(\mathbf{Y} | \mathbf{h}, \sigma_v^2) \\
 & \Leftrightarrow \min_{\mathbf{h}, \sigma_v^2} \left\{ \ln \det C_{Y_{SB}Y_{SB}} + (\mathbf{Y}_{SB} - \mathcal{T}_{SB}(\mathbf{h})A_{SB}^{\circ})^H C_{Y_{SB}Y_{SB}}^{-1} (\mathbf{Y}_{SB} - \mathcal{T}_{SB}(\mathbf{h})A_{SB}^{\circ}) \right\}.
 \end{aligned} \tag{4.19}$$

Nous utilisons la stratégie PQML pour résoudre (4.19). Dans le gradient de la fonction de coût les deux termes provenant de la dérivation de $C_{Y_{SB}Y_{SB}}$ s'annulent entre eux asymptotiquement, et nous les négligeons. Les quantités qui nous intéressent sont :

$$\begin{aligned}
 \mathcal{P}(\mathbf{h}) &= \mathcal{A}_{SB}^H C_{Y_{SB}Y_{SB}}^{-1} \mathcal{A}_{SB}, \\
 \mathcal{S}(\mathbf{h}) &= C_{Y_{SB}Y_{SB}}^{-1} \mathbf{Y}_{SB}.
 \end{aligned} \tag{4.20}$$

Le critère PQML (approximatif) est équivalent au problème des moindres carrés optimalement pondérés (WLS) suivant :

$$\min_{\mathbf{h}, \sigma_v^2} \|\mathbf{Y}_{SB} - \mathcal{T}_{SB}(\mathbf{h})A_{SB}^{\circ}\|_{C_{Y_{SB}Y_{SB}}^{-1}}. \tag{4.21}$$

Ce critère aboutit à des performances supérieures à celles du critère LS décrit dans la section précédente : il contient toutes les équations du critère LS et permet d'incorporer l'information provenant des symboles connus. Le critère combiné a pour expression :

$$\min_{\mathbf{h}, \sigma_v^2} \mathbf{Y}_B^H P_{\mathcal{T}^H(\mathbf{h}^+)} \mathbf{Y}_B + \sigma_v^2 \|\mathbf{Y}_{SB} - \mathcal{T}_{SB}(\mathbf{h})A_{SB}^{\circ}\|_{C_{Y_{SB}Y_{SB}}^{-1}}^2. \tag{4.22}$$

Notons que l'unique information perdue provient des symboles inconnus présents dans la séquence combinée (décrite précédemment), ce qui est négligeable puisque le vecteur \mathbf{Y}_B est généralement choisi de grande taille. A chaque itération, la solution pour le vecteur \mathbf{h} est donnée par :

$$\mathbf{h} = (\mathcal{Y}_B^H \mathcal{R}_B^+ \mathcal{Y}_B - \lambda \mathcal{B}_B^H \mathcal{B}_B + \sigma_v^2 \mathcal{A}_{SB}^H C_{Y_{SB}Y_{SB}}^{-1} \mathcal{A}_{SB})^{-1} \mathcal{A}_{SB}^H \mathbf{Y}_{SB}. \tag{4.23}$$

Ce critère nécessite uniquement un symbole pour fonctionner.

PQML-WLS aboutira à une performance meilleure que celle donnée par PQML-LS parce que la séquence d'apprentissage du critère aboutit à une performance meilleure. L'étude de la convergence asymptotique peut se faire selon deux points de vue. Si nous considérons M_u comme asymptotique et M_k comme fini, les critères (4.16) et (4.22) sont équivalents au critère aveugle. Le facteur d'échelle n'étant pas identifiable de façon aveugle, il est estimé par l'intermédiaire de la séquence d'apprentissage : c'est pourquoi l'estimée obtenue n'est pas consistante. Si nous considérons, maintenant, M_u et M_k comme étant infinis (avec l'hypothèse $\frac{\sqrt{M_u}}{M_k} \rightarrow 0$), une première itération de l'algorithme donne une estimée consistante et la seconde aboutit au minimiseur global (ce résultat peut être obtenu en suivant le même raisonnement asymptotique décrit précédemment pour le cas purement aveugle).

4.5 Simulations

Dans le but d'illustrer les performances des algorithmes décrits dans ce chapitre, nous considérons un canal irréductible \mathbf{H}_1 et un canal mal-conditionné \mathbf{H}_2 , dont les sous-canaux possèdent pratiquement un zéro en commun. Les deux canaux considérés sont de longueur $N = 4$ avec $m = 2$ sous-canaux. La longueur du paquet de données est $M = 200$ et la séquence d'entrée est une séquence i.i.d. QPSK. Le rapport signal à bruit est défini comme $\text{SNR} = (\|\mathbf{h}\|^2 \sigma_a^2) / (m \sigma_v^2)$.

4.5.1 Algorithmes Aveugles

Les méthodes d'estimation aveugles aboutissent à une estimée du canal $\hat{\mathbf{h}}$ avec $\|\hat{\mathbf{h}}\| = 1$, nous ajustons le facteur scalaire α tel que $\mathbf{h}_o^H(\alpha \hat{\mathbf{h}}) = \mathbf{h}_o^H \mathbf{h}_o$ (voir [dCS97b]) : l'estimée finale du canal est $\hat{\mathbf{h}} = \alpha \hat{\mathbf{h}}$. Nous traçons les courbes correspondant à l'erreur quadratique moyenne normalisée (NMSE) (en prenant la moyenne sur 100 réalisations Monte-Carlo) définie comme :

$$\text{NMSE} = \frac{\|\mathbf{h} - \hat{\mathbf{h}}\|^2}{\|\mathbf{h}\|^2}. \quad (4.24)$$

Nous considérons uniquement le canal \mathbf{H}_1 . L'initialisation des algorithmes DIQML et PQML est faite par la méthode SRM. Dans la figure 4.1, nous illustrons la performance des algorithmes DIQML et PQML pour des valeurs de SNR égales à 10dB et 20dB et nous la comparons à la borne de Cramer-Rao aveugle calculée avec la contrainte $\mathbf{h}_o^H \hat{\mathbf{h}} = \mathbf{h}_o^H \mathbf{h}_o$ [dCS97b], qui correspond à la façon avec laquelle nous avons précédemment ajusté le facteur scalaire. Une amélioration par rapport à l'initialisation SRM peut être constatée pour les deux algorithmes DIQML et PQML. La performance de l'algorithme PQML est supérieure à celle de la méthode DIQML, et reste étroitement proche de la borne de Cramer-Rao déterministe aveugle. Dans la figure 4.2, nous avons tracé l'évolution de la fonction de coût DML. Nous constatons que pour l'algorithme DIQML, cette fonction de coût n'est pas monotone : de petites fluctuations après la première itération peuvent être observées. En ce qui concerne la méthode PQML, la fonction de coût est monotone (décroissante). Après une ou deux itérations, les algorithmes DIQML et PQML atteignent leur régime permanent. Dans la figure 4.3, nous avons tracé l'évolution de la valeur propre généralisée minimale moyenne λ_{gen} en fonction du nombre des itérations effectuées, dans le cas de l'algorithme DIQML, pour des valeurs de SNR égales à 10dB et 20dB : nous constatons que λ_{gen} tend vers la variance du bruit σ_v^2 (l'algorithme DIQML sous-estime, en moyenne, la variance du bruit) et que l'évolution de λ_{gen} en fonction du nombre des itérations effectuées n'est pas monotone. Dans la figure 4.4, nous avons réalisé la même simulation dans le cas de l'algorithme PQML, il est clair que l'évolution de la valeur propre généralisée minimale moyenne λ_{gen} vers 1 est monotone (croissante) : cette évolution est d'autant plus rapide (nécessite un nombre réduit d'itérations effectuées pour atteindre 1) lorsque la puissance du bruit décroît.

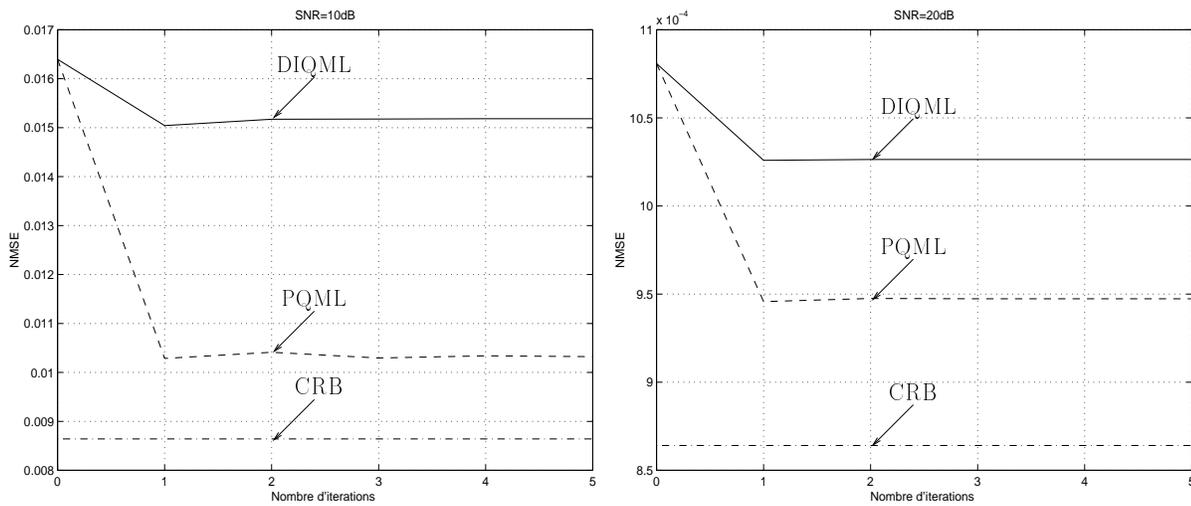


FIG. 4.1 – Performance des algorithmes DIQML et PQML : NMSE.

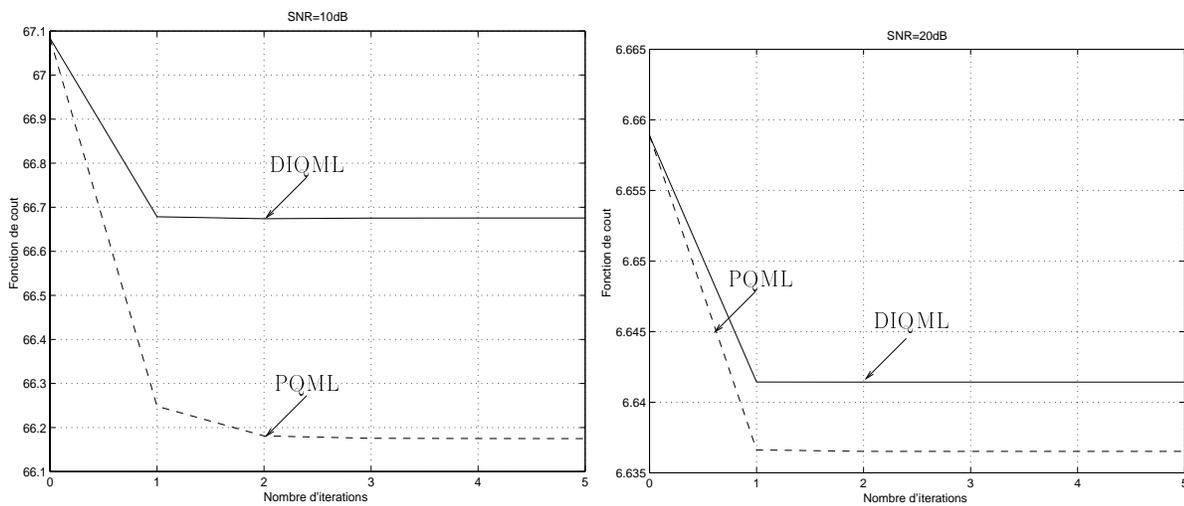


FIG. 4.2 – Performance des algorithmes DIQML et PQML : fonction de coût DML.

4.5.2 Algorithmes Semi-aveugles

Dans la figure 4.5, nous traçons l'erreur NMSE correspondant aux algorithmes PQML-LS et PQML-WLS en utilisant 10 symboles connus et une initialisation par la méthode SRM. Cette même initialisation est utilisée pour l'algorithme PQML pour lequel le facteur scalaire est ajusté par une estimation basée sur la séquence d'apprentissage. Le SNR étant égal à 10dB et le canal est \mathbf{H}_1 . La variance du bruit est supposée connue. Les deux algorithmes semi-aveugles aboutissent à des performances supérieures à celle de l'algorithme aveugle PQML. Nous constatons, aussi, que la performance de l'algorithme PQML-WLS est supérieure à celle de la méthode PQML-LS. Les différences entre les performances sont petites parce que les trois algorithmes sont étroitement proches de la borne de Cramer-Rao déterministe semi-aveugle [dCS97b].

Dans la figure 4.6, nous traçons les mêmes courbes dans le cas du canal mal condi-

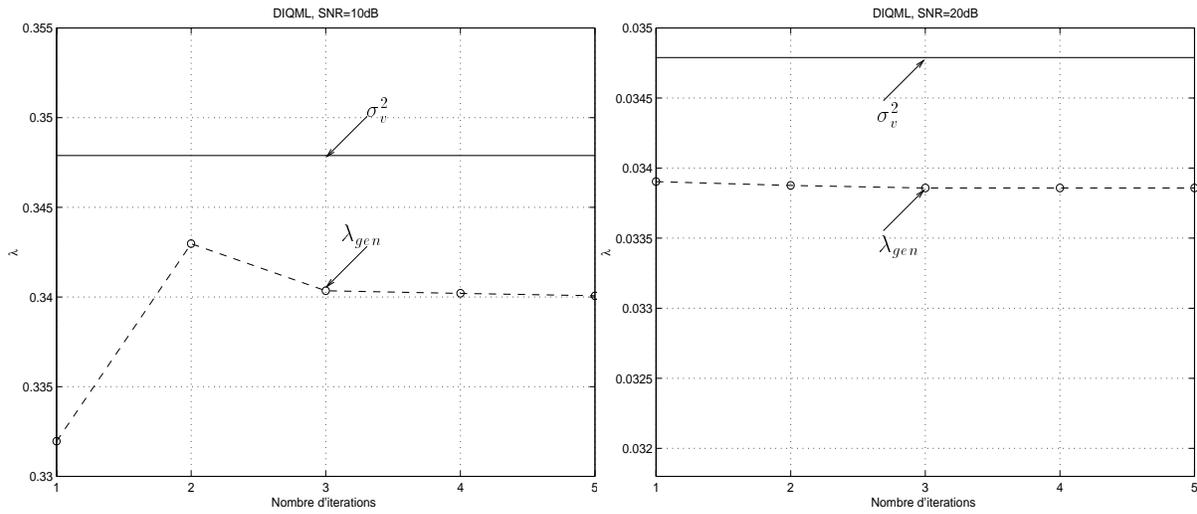


FIG. 4.3 – DIQML : Evolution de λ en fonction du nombre d'itérations effectuées.

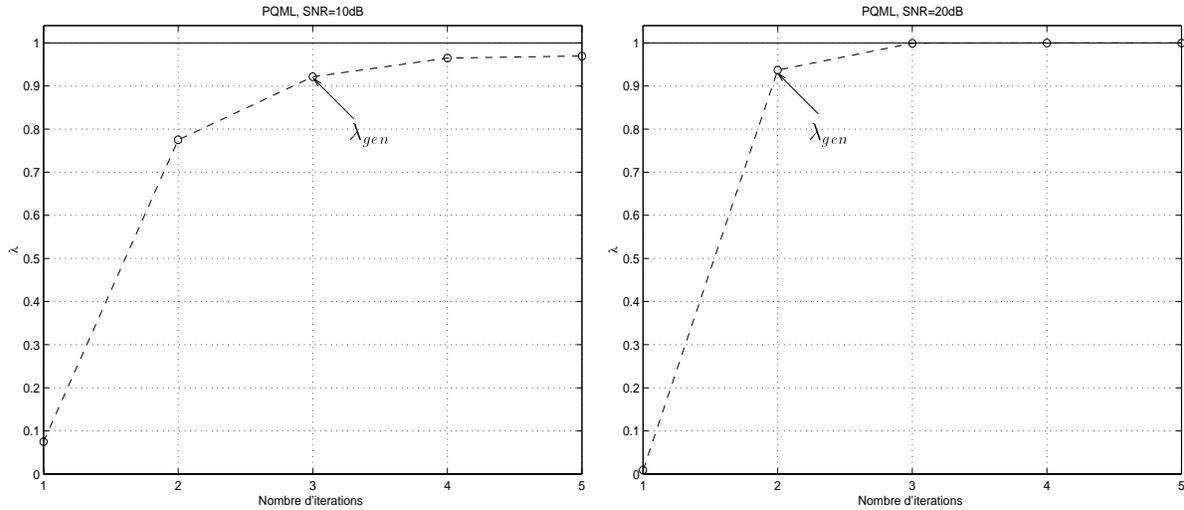


FIG. 4.4 – PQML : Evolution de λ en fonction du nombre d'itérations effectuées.

tionné \mathbf{H}_2 pour lequel les méthodes d'estimation aveugle échouent. Les algorithmes sont initialisés par une séquence d'apprentissage de longueur 10. De nouveau, les performances des deux algorithmes sont proches de la borne de Cramer-Rao et la performance de l'algorithme PQML-WLS est supérieure à celle de la méthode PQML-LS. Dans le cas extrême du canal \mathbf{H}_2 et pour une séquence d'apprentissage de longueur 3 pour laquelle l'algorithme PQML-LS ne fonctionne pas à un SNR = 20dB, l'algorithme PQML-WLS converge avec plus de difficultés que dans le cas d'un canal bien conditionné, mais sa performance reste proche de la borne de Cramer-Rao déterministe semi-aveugle. En ce qui concerne les fonctions de coût, elles atteignent pratiquement leur régime permanent après une itération, mais elles peuvent présenter des petites fluctuations.

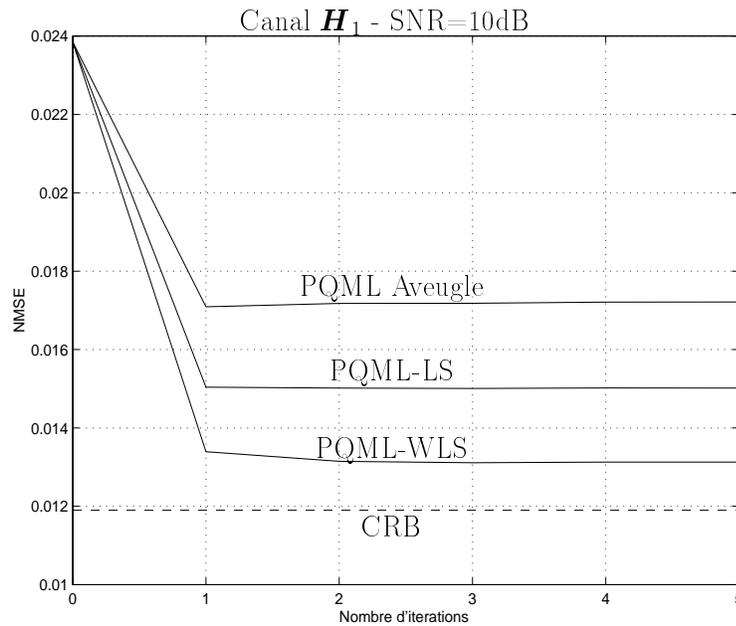


FIG. 4.5 – Performance des algorithmes PQML, PQML-LS et PQML-WLS : NMSE.

4.6 Discussion

Les approches de maximum de vraisemblance se présentent comme une alternative sérieuse pour l'estimation (aveugle) multi-canal. L'atout principal de ces méthodes réside dans leur optimalité en termes de performance : asymptotiquement l'estimateur ML atteint sa borne de Cramer-Rao et donc il est efficace. Ces approches se divisent en deux classes selon les hypothèses faites sur les symboles transmis : dans les cas où les symboles sont supposés être déterministes, les méthodes ML obtenues sont déterministes ; dans le cas où les symboles sont supposés être des variables aléatoires gaussiennes, les algorithmes ML dérivés sont qualifiés de gaussiens. Cette hypothèse stochastique qui porte sur les symboles permet de renforcer le problème d'estimation multi-canal par rapport au modèle déterministe. Cette robustesse se manifeste au niveau des bornes de performances correspondant aux méthodes d'estimation obtenues. En effet, le contexte gaussien permet d'aboutir à des bornes de Cramer-Rao plus basses que celles obtenues dans un contexte déterministe. Par conséquent, les méthodes gaussiennes, qui éventuellement peuvent être dérivées, sont susceptibles d'avoir des performances supérieures aux approches déterministes et donc d'aboutir à des estimées du canal plus précises.

4.6.1 Méthodes pour les Symboles Déterministes

Le chapitre précédent nous a permis de discuter une classe de méthodes d'estimation multi-canal basées sur les sous-espaces. Ces méthodes sont essentiellement : la méthode SRM, l'ajustement du SS et l'ajustement du SB. Asymptotiquement, les performances de ces méthodes restent inférieures à celles de la classe des méthodes ML déterministes.

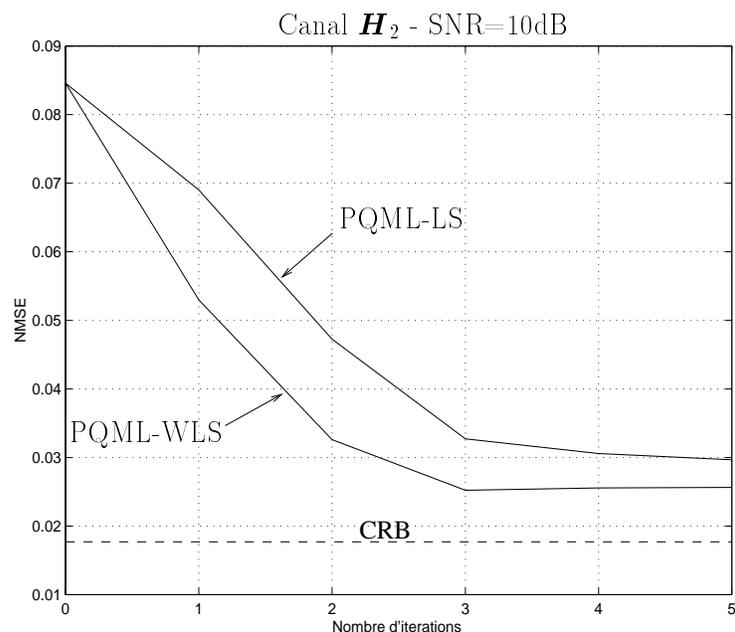


FIG. 4.6 – Performance des algorithmes PQML-LS et PQML-WLS : NMSE.

Comme le critère DML est un critère non linéaire en termes des paramètres à estimer, nous avons discuté des procédures itératives dont les méthodes SRM, SSF et NSF peuvent servir d'initialisation. Un facteur essentiel à considérer dans l'initialisation, à part la consistance de l'estimée, concerne la complexité de la méthode : de ce point de vue la méthode SRM, vu sa faible complexité, peut être utilisée pour l'initialisation de ces stratégies itératives. La première procédure itérative que nous avons présentée pour résoudre DML est IQML. Cette procédure est biaisée (à cause de la présence du bruit) à des bas SNRs, c'est pourquoi nous avons proposé de la débruiter. La méthode débruitée ainsi obtenue, appelée DIQML, permet d'obtenir des estimées consistantes du canal. Néanmoins, sa performance reste inférieure à celle de la méthode PQML dont la performance est asymptotiquement optimale. Donc, d'un point de vue performance de la méthode, l'approche PQML représente la meilleure méthode pouvant être envisagée dans le contexte déterministe. Le défi qui reste, dans ce contexte déterministe, est de dériver des méthodes déterministes ayant la même performance que l'algorithme PQML et dont la complexité est plus faible. Le chapitre 8 présentera une méthode d'estimation multi-canal basée sur une prédiction bidirectionnelle d'une portion du signal reçu, ayant la même performance que l'algorithme PQML, et de complexité plus faible que ce dernier.

4.6.2 Méthodes pour les Symboles Gaussiens

Contrairement au modèle des symboles déterministes où le canal peut être déterminé de façon aveugle à un facteur scalaire complexe près, dans le cas du modèle des symboles gaussiens, la norme du canal peut être estimée. Dans le cas gaussien, l'approche essentielle

est l'approche de maximum de vraisemblance gaussienne (GML). Dans ce cas $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, R_{YY})$ avec $R_{YY} = \sigma_a^2 \mathcal{T}(\mathbf{h}) \mathcal{T}^H(\mathbf{h}) + \sigma_v^2 I$ (et donc $R_{AA} = \sigma_a^2 I$). La fonction de coût à minimiser est

$$\mathcal{L}(\mathbf{h}) = c^t + \ln \det R_{YY} + \mathbf{Y}^H R_{YY}^{-1} \mathbf{Y} . \quad (4.25)$$

Les techniques d'optimisation standard telles que la méthode de Gauss-Newton et la méthode de scoring peuvent être utilisées. Une autre approche pour le modèle des symboles gaussiens est la méthode "covariance matching" [GH97].

Le critère donné par (4.25) est non linéaire et donc nécessite une initialisation consistante pour atteindre le minimum global. L'inconvénient des méthodes standard (Gauss-Newton ou scoring) est que ces méthodes sont des procédures itératives qui sont lourdes en termes de coût de calcul. Puisque les méthodes gaussiennes sont susceptibles d'aboutir à des performances supérieures à celles des méthodes déterministes, le défi consiste à dériver des méthodes gaussiennes non itératives (ou nécessitant un nombre faible d'itérations pour converger).

4.7 Exploitation de l'Alphabet Fini

Dans cette section, nous nous intéressons à l'exploitation de l'alphabet fini des symboles transmis dans le contexte de l'identification aveugle multi-canal (voir [Tal96]). L'incorporation de la connaissance de l'alphabet fini permet aux symboles transmis d'agir comme une séquence d'apprentissage pour l'estimation du canal. Par conséquent, une approche aveugle qui exploite l'alphabet fini des symboles transmis est susceptible d'aboutir à une performance supérieure à celle de sa version purement aveugle. Dans [AS98a], nous avons proposé d'incorporer la connaissance a priori de l'alphabet fini en combinant un critère d'estimation aveugle du canal avec un critère d'égalisation linéaire MMSE à décision dirigée (MMSE-DD). Les résultats de simulation illustrent l'amélioration significative de la performance de la méthode exploitant l'alphabet fini.

4.7.1 Critère d'Egalisation

Nous considérons un égaliseur qui combine l'information aveugle provenant du signal d'erreur d'une méthode aveugle d'estimation du canal (méthode SRM) et l'information inhérente à l'alphabet fini des symboles transmis. L'aspect "alphabet fini" est illustré par l'intermédiaire de la détection des symboles conformément à un critère à décision dirigée. Les symboles détectés sont donnés par (les indices sont omis pour la simplicité de la notation) :

$$\hat{\hat{A}} = \text{dec}(\hat{A}) \quad (4.26)$$

où $\text{dec}(\cdot)$ est l'opérateur de décision qui choisit l'élément de l'alphabet fini des symboles le plus proche de son argument et \hat{A} est la sortie de l'égaliseur MMSE linéaire :

$$\hat{A} = R_{AY} R_{YY}^{-1} \mathbf{Y} = \sigma_a^2 \mathcal{T}^H(\mathbf{h}) R_{YY}^{-1} \mathbf{Y} . \quad (4.27)$$

Si nous considérons $\widehat{A} = A$, alors l'erreur sur les estimées des symboles \hat{A} est

$$\begin{aligned}\tilde{A} &= A - \hat{A} \\ &= (I - \sigma_a^2 \mathcal{T}^H(\mathbf{h}) R_{YY}^{-1} \mathcal{T}(\mathbf{h})) A - \sigma_a^2 \mathcal{T}^H(\mathbf{h}) R_{YY}^{-1} \mathbf{V}.\end{aligned}\quad (4.28)$$

Par conséquent, nous construisons un critère en combinant le terme aveugle et le terme d'erreur de l'égaliseur DD, regroupés dans le vecteur E suivant :

$$E = \begin{bmatrix} \mathcal{T}(\mathbf{h}^\perp) \mathbf{Y} \\ A - \sigma_a^2 \mathcal{T}^H(\mathbf{h}) R_{YY}^{-1} \mathbf{Y} \end{bmatrix}.\quad (4.29)$$

L'aspect aléatoire de E peut être découpé en la contribution des symboles transmis et celle du bruit :

$$\begin{bmatrix} 0 \\ I - \sigma_a^2 \mathcal{T}^H(\mathbf{h}) R_{YY}^{-1} \mathcal{T}(\mathbf{h}) \end{bmatrix} A + \begin{bmatrix} \mathcal{T}(\mathbf{h}^\perp) \\ -\sigma_a^2 \mathcal{T}^H(\mathbf{h}) R_{YY}^{-1} \end{bmatrix} \mathbf{V}.\quad (4.30)$$

La minimisation de $E^H R_{EE}^{-1} E$ par rapport à \mathbf{h} aboutit à la minimisation du critère découpé suivant :

$$\min_{\mathbf{h}} \mathbf{Y}^H P_{\mathcal{T}^H(\mathbf{h}^\perp)} \mathbf{Y} + \|A - \sigma_a^2 \mathcal{T}^H(\mathbf{h}) R_{YY}^{-1} \mathbf{Y}\|_{\left(\mathcal{T}^H(\mathbf{h}) \mathcal{T}(\mathbf{h}) + \frac{\sigma_v^2}{\sigma_a^2} I\right)}^2.\quad (4.31)$$

Le premier terme dans (4.31) est le critère DML et le second terme correspond à un critère d'égalisation DD des moindres carrés optimalement pondérés (WLS).

4.7.2 PQML-DD-WLS

Nous utilisons la stratégie PQML pour résoudre le premier terme du critère (4.31) et nous résolvons le second terme au sens WLS où la matrice de pondération est

$$W = \left(\mathcal{T}^H(\mathbf{h}) \mathcal{T}(\mathbf{h}) + \frac{\sigma_v^2}{\sigma_a^2} I \right)\quad (4.32)$$

et la matrice R_{YY} est considérée constante. Puisque PQML résout le critère DML de façon optimale et la partie WLS est optimalement pondérée, la stratégie décrite précédemment pour résoudre le critère (4.31) est sensée aboutir à des bonnes performances. Le problème est donc résolu itérativement : à chaque itération, la solution pour \mathbf{h} est donnée par :

$$\mathbf{h} = \sigma_a^2 (\mathbf{y}^H \mathcal{R}^+ \mathbf{y} - \lambda \mathcal{B}^H \mathcal{B} + \sigma_a^4 \mathcal{X}^T W^T \mathcal{X}^*)^{-1} \mathcal{X}^T W^T A^*,\quad (4.33)$$

où la matrice \mathcal{X} est définie tel que $\mathcal{X} \mathbf{h}^* = \mathcal{T}^H(\mathbf{h}) X$ et $X = \sigma_a^2 R_{YY}^{-1} \mathbf{Y}$. L'initialisation de l'algorithme PQML-DD-WLS peut être faite par une méthode d'estimation de canal purement aveugle. Nous pouvons aussi initialiser PQML-DD-WLS par l'estimée du canal obtenue par l'algorithme PQML (qui est initialisé à son tour par l'intermédiaire de la méthode SRM).

La performance de l'algorithme peut être améliorée en utilisant des décisions douces sur les symboles à la place des décisions dures. Dans ce cas, nous obtenons un estimateur MMSE qui utilise des probabilités a posteriori sur les \tilde{a}_k :

$$\hat{\tilde{a}}_k = \frac{\sum_{i=1}^K e^{-\frac{|\tilde{a}_k - a_i|^2}{\sigma^2}} a_i}{\sum_{i=1}^K e^{-\frac{|\tilde{a}_k - a_i|^2}{\sigma^2}}}, \quad (4.34)$$

où $\{a_i, i = 1, \dots, K\}$ constitue l'alphabet des symboles et σ^2 est l'élément diagonal de la matrice $\sigma_a^2 (I - \sigma_a^2 \mathcal{T}^H(\mathbf{h}) R_Y^{-1} \mathcal{T}(\mathbf{h}))$. L'amélioration de performance se justifie par le fait que cet estimateur minimise l'erreur MSE tout en étant non linéaire (il n'y a pas de contrainte de linéarité imposée).

4.7.3 Decision-Directed Least-Squares (DD-LS)

Lorsque le vecteur des symboles A est connu, le critère $\min_{\mathbf{h}} \|\mathbf{Y} - \mathcal{T}(\mathbf{h})A\|^2$ est équivalent au problème de minimisation suivant :

$$\min_{\mathbf{h}} \|\mathbf{Y} - \mathcal{A}\mathbf{h}\|^2, \quad (4.35)$$

où la matrice \mathcal{A} est définie tel que $\mathcal{A}\mathbf{h} = \mathcal{T}(\mathbf{h})A$. La résolution de ce problème LS par rapport à \mathbf{h} permet de trouver le canal comme :

$$\mathbf{h} = (\mathcal{A}^H \mathcal{A})^{-1} \mathcal{A}^H \mathbf{Y}. \quad (4.36)$$

4.7.4 Simulations

Nous considérons un paquet de données de longueur $M = 200$, un canal \mathbf{H} complexe irréductible généré de façon aléatoire de longueur $N = 3$ avec $m = 2$ sous-canaux. Les symboles transmis sont i.i.d. QPSK. Le SNR considéré est le SNR moyen défini comme $\text{SNR} = (\|\mathbf{h}\|^2 \sigma_a^2) / (m \sigma_v^2)$ et la mesure de performance est faite par l'intermédiaire du NMSE (dont la définition est donnée dans (4.24)). Dans la figure 4.7, nous considérons un SNR=10dB, et nous effectuons une itération de l'algorithme PQML initialisé par la méthode SRM. L'estimée du canal obtenue est utilisée pour initialiser l'algorithme PQML-DD-WLS et la partie WLS du critère (4.31). Nous effectuons quatre itérations pour ces deux algorithmes. Nous constatons que la performance de l'algorithme PQML-DD-WLS est supérieure à celle de la méthode WLS particulièrement au niveau de la première itération. Ceci prouve que le critère combiné donné par (4.31) aboutit à une amélioration de performance par rapport à celles des deux méthodes à combiner. Nous avons aussi tracé, sur la même figure, les courbes correspondant à la méthode DD-LS dans les deux cas suivants : dans le premier cas les symboles sont détectés et dans le

second cas les symboles sont considérés connus (dans ce cas la performance obtenue peut être interprétée comme une borne de performance pour la méthode PQML-DD-WLS). Nous observons que la méthode DD-LS dans laquelle les symboles sont détectés aboutit à une bonne performance (l'erreur NMSE obtenue est proche de la borne donnée par la méthode DD-LS dans le cas où les symboles sont connus). Dans le but d'illustrer l'importance de l'initialisation dans la stratégie DD-LS, nous avons tracé dans la figure 4.8 les courbes correspondant à deux initialisations différentes : celle faite par la méthode SRM et celle effectuée par l'algorithme PQML. Il est clair que le bon fonctionnement de la stratégie DD-LS (où les symboles sont détectés) est dû à la bonne initialisation obtenue par l'intermédiaire de la méthode PQML.

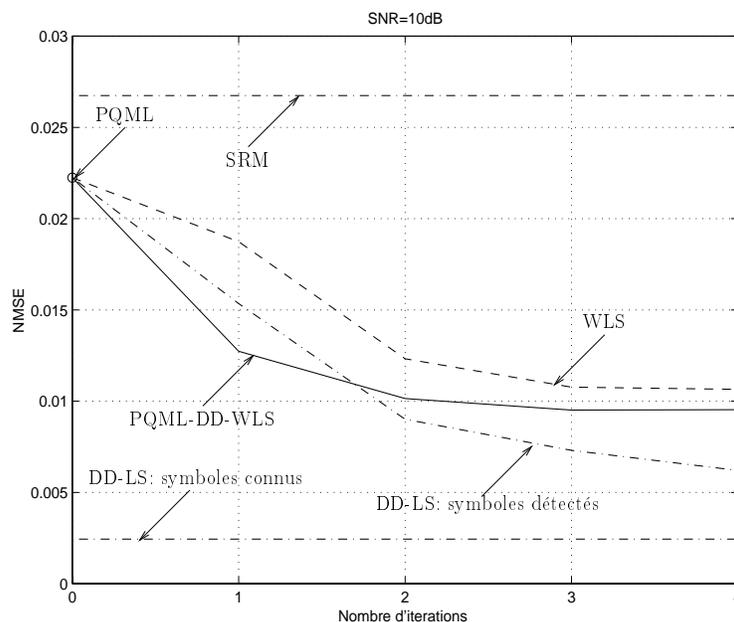


FIG. 4.7 – Performance des algorithmes PQML-DD-WLS, WLS, DD-LS (symboles détectés) et DD-LS (symboles connus) : NMSE.

4.8 Conclusions

Les méthodes d'estimation du canal basées sur le critère de maximum de vraisemblance présentent une classe de méthodes d'estimation aveugle multi-canal intéressante, aussi bien dans le contexte déterministe que dans le contexte gaussien. Un tel intérêt se justifie par leur optimalité en termes de performance asymptotique et donc une supériorité, de ce point de vue, par rapport à la classe des méthodes basées sur les sous-espaces. Une sous-classe des méthodes ML est celle qui regroupe les approches basées sur le critère ML gaussien. Cette dernière trouve encore son intérêt dans la robustesse des méthodes pouvant être dérivées (par rapport aux approches déterministes). En effet, ces méthodes s'avèrent robustes en termes de conditions d'identifiabilité (qui deviennent moins res-

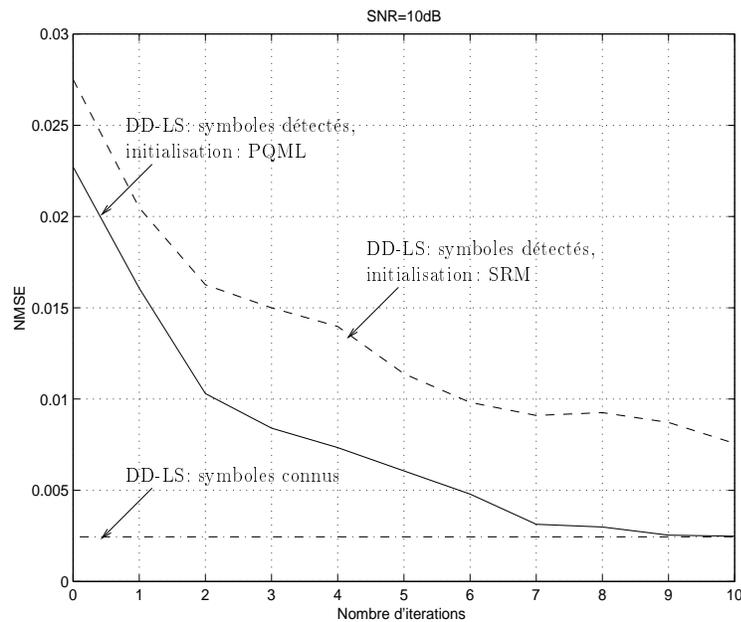


FIG. 4.8 – Importance de l'initialisation pour la méthode DD-LS (symboles détectés).

trictives : la classe des canaux pouvant être identifiés est élargie), de performance de la méthode (qui devient meilleure que celle correspondant à une méthode déterministe donnée) et de bornes de performance (qui deviennent plus basses dans le contexte gaussien). Ce chapitre a récapitulé les différentes méthodes basées sur le critère DML que nous avons proposées et analysées. La méthode la plus optimale de cette classe, qui est la méthode PQML a été étendue pour être combinée avec un critère basé sur une séquence d'apprentissage dans le but d'obtenir des méthodes semi-aveugles (plus robustes que les procédures purement aveugles) : PQML-LS et PQML-WLS. Les approches semi-aveugles obtenues aboutissent à des performances supérieures à celles de leurs versions purement aveugles.

La robustesse de l'estimation aveugle du canal peut être améliorée, de manière significative, en exploitant le fait que les symboles transmis ne sont pas des nombres complexes arbitraires, mais appartiennent à un alphabet fini. La comparaison des stratégies exploitant un tel alphabet : PQML-DD-WLS, WLS et DD-LS (où les symboles sont détectés) nous permet de constater que la stratégie DD-LS, bénéficiant d'une bonne initialisation, aboutit à la meilleure performance.

Chapitre 5

Exploitation de la Connaissance des Filtres de Transmission

Ce chapitre présente le concept de l'identification aveugle des canaux multiples à réponse impulsionnelle finie avec connaissance des filtres de transmission. Cette information *a priori* peut en réalité non seulement inclure le filtre de transmission (TX) (filtre de mise en forme) mais aussi le filtre de réception (RX). Ces filtres sont des entités typiques présentes dans les systèmes de communications numériques. L'exploitation de cette information interne permet à l'opération d'estimation de se concentrer sur la réponse impulsionnelle de la partie réelle du canal à estimer (canal de propagation). Par conséquent, l'opération d'estimation peut être faite de façon plus précise. En effet, l'utilisation de cette information permet souvent de réduire *de manière significative* le nombre de paramètres à estimer (la longueur du filtre à identifier est plus courte que dans le cas purement aveugle), ce qui nécessite une énergie moindre du signal ou une quantité moindre de données pour effectuer l'opération d'identification du canal.

Les chapitres précédents nous ont permis d'analyser une classe de méthodes d'estimation aveugle du canal qui sont paramétrisées directement par le canal incluant la méthode SRM, les méthodes d'ajustement de sous-espaces (SS et SB), et les méthodes de maximum de vraisemblance (DML, IQML, DIQML et PQML). Comme l'information *a priori* est exprimée en termes de la réponse impulsionnelle du canal, nous considérons l'extension des méthodes paramétrisées directement par le canal, pour incorporer l'information *a priori*. Toutes ces méthodes sont formulées pour un mode de transmission par paquets. Nous discutons aussi les limites de performance sous forme de bornes de Cramer-Rao (CRB). Les méthodes et les bornes de Cramer-Rao sont discutées dans deux contextes des symboles inconnus transmis : déterministe et gaussien. Les résultats des simulations indiquent que l'exploitation de l'information *a priori* peut conduire à des améliorations significatives de performance, à une capacité de la méthode étendue (celle qui exploite la connaissance *a priori* des filtres de TX/RX) d'identifier des canaux mal-conditionnés, et au fait que les méthodes de maximum de vraisemblance peuvent encore améliorer la performance.

5.1 Introduction

Le but de l'identification aveugle est d'identifier le canal inconnu en disposant uniquement du signal reçu (et de la description statistique de l'entrée). La majorité des travaux élaborés dans le contexte de l'identification aveugle considère le canal total (ou global) qui inclut le filtre de transmission, le canal réel (de propagation) et le filtre de réception (voir figure 5.1). Cependant, souvent l'entité inconnue dans cette cascade est le multi-trajets qui représente le canal de propagation. L'identification du canal aveugle exploitant la connaissance a priori des filtres de TX/RX a été introduite dans [SSG94] puis elle a été traitée dans [DM95],[Din97]. Une approche similaire mais utilisant la connaissance des filtres de transmission dans le contexte de l'estimation multi-canal par séquence d'apprentissage est présentée dans [NCP97],[CNP97],[NCP]. Dans [NGP98], les auteurs montrent que, grâce uniquement à l'exploitation de la connaissance des filtres de transmission, on peut spécifier de manière unique un égaliseur MMSE. Dans [TS98b], la connaissance a priori des filtres de TX/RX est exploitée pour égaliser partiellement le filtre de transmission afin de raccourcir la longueur des canaux utilisés, et des techniques d'annulation des interférences ont été testées. Dans [TS99], l'exploitation de la connaissance des filtres de TX/RX est utilisée pour estimer conjointement les canaux de tous les utilisateurs.

Si le suréchantillonnage est utilisé, le filtre de réception est habituellement un filtre passe-bande qui introduit peu de distorsions. Le filtre de transmission est cependant celui qui limite la largeur de bande. Par conséquent, dans le cas du suréchantillonnage, ces deux aspects rendent le problème d'estimation aveugle multi-canal mal-conditionné. Le canal peut être paramétrisé par sa réponse impulsionnelle \mathbf{h} ou par le filtre de prédiction $\mathbf{P}(z)$ et $\mathbf{h}(0)$ qui satisfont $\mathbf{P}(z)\mathbf{H}(z) = \mathbf{h}(0)$ [SP94]. Cependant, il n'est pas encore évident de savoir comment exprimer l'information a priori des filtres de TX/RX en terme du filtre de prédiction. Pour cela, nous ne considérons ici que l'extension des méthodes paramétrisées par le canal \mathbf{h} . Ce chapitre est organisé comme suit : dans la première section, nous introduisons le formalisme permettant d'exploiter la connaissance des filtres de TX/RX dans la formulation du problème d'optimisation, puis nous présentons dans la deuxième section les différentes méthodes étendues. La section suivante concerne les aspects d'identifiabilité que l'exploitation de cette connaissance a priori induit. Ensuite, nous dérivons dans la cinquième section les bornes de Cramer-Rao (déterministe et gaussienne) et nous présentons dans la sixième section les résultats des simulations pour illustrer les performances des différentes méthodes discutées.

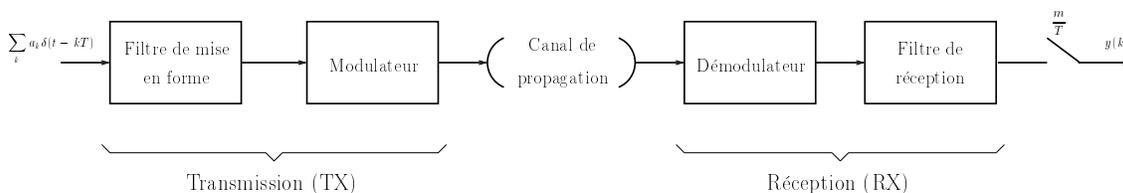


FIG. 5.1 – *Système typique de communications numériques.*

5.2 Connaissance des Filtres de Transmission/Réception

L'exploitation de la connaissance des filtres de TX/RX nécessite en principe que la fréquence d'échantillonnage satisfasse le critère de Nyquist et donc que le suréchantillonnage par rapport au rythme des symboles soit utilisé. Supposons, tout d'abord, que nous sommes dans le cas d'une seule antenne à la réception. Considérons la réponse impulsionnelle totale du canal $h(t) = c(t) * g(t)$ qui représente la convolution de deux systèmes : $g(t)$ est le filtre de TX (filtre de mise en forme) ou sa convolution avec le filtre de RX, et $c(t)$ est le canal de propagation convolué avec le filtre de RX ou respectivement le canal de propagation uniquement. Dans le domaine fréquentiel, nous obtenons

$$H(f) = C(f)G(f). \quad (5.1)$$

Considérons maintenant $g(t)$ à bande limitée. Donc $G(f) = G(f)F(f)$ où $F(f)$ est un filtre passe-bas idéal (filtre rectangle de longueur unité) avec une largeur de bande supérieure ou égale à celle de $G(f)$. L'expression (5.1) devient

$$H(f) = G(f)(F(f)C(f)) \quad (5.2)$$

où les deux facteurs $G(f)$ et $F(f)C(f)$ sont, maintenant, tous les deux à bande limitée. Avec un taux d'échantillonnage suffisant (qui satisfait essentiellement le critère de Nyquist pour $G(f)$), la version discrète de $h(t)$ peut être considérée comme la convolution de la version discrète de $g(t)$ avec une certaine représentation à temps discret pour $c(t)$ et qui correspond à une version discrète de $F(f)C(f)$. Notons que si le taux d'échantillonnage dépasse la largeur de bande $G(f)$, alors $F(f)$ n'est pas unique puisque sa largeur de bande peut être choisie de façon arbitraire entre la largeur de bande de $G(f)$ et la fréquence d'échantillonnage. Par conséquent, la représentation à temps discret pour $c(t)$ n'est pas unique.

5.2.1 Cas du Suréchantillonnage

Considérons, dans cette section, le cas d'une seule antenne à la réception. Considérons, maintenant, un facteur de suréchantillonnage m et soit $H(z) = C(z)G(z)$ la fonction de transfert suréchantillonnée du canal total qui est la cascade de deux fonctions de transfert suréchantillonnées : $C(z)$ et $G(z)$. Chacune de ces fonctions de transfert peut être décomposée en ses composantes polyphases, par exemple $H(z) = \sum_{i=0}^{m-1} z^{-i} H_i(z^m)$. Ces composantes peuvent aussi être représentées sous forme d'une expression matricielle, ainsi nous obtenons

$$\begin{aligned} \mathbf{G}(z) &= [G_1(z) \cdots G_m(z)]^T = \sum_{\substack{k=0 \\ L-1}}^{K-1} \mathbf{g}(k) z^{-k} \\ \mathbf{C}(z) &= [C_1(z) \cdots C_m(z)]^T = \sum_{k=0}^{L-1} \mathbf{c}(k) z^{-k} \end{aligned} \quad (5.3)$$

avec $K+L-1 = N$. Les relations entre les composantes polyphases peuvent être obtenues à partir de

$$\sum_{i=0}^{m-1} z^{-i} \mathbf{H}_i(z^m) = \left(\sum_{k=0}^{m-1} z^{-k} \mathbf{G}_k(z^m) \right) \left(\sum_{l=0}^{m-1} z^{-l} \mathbf{C}_l(z^m) \right) \quad (5.4)$$

En particulier, pour $m = 2$ nous obtenons

$$\begin{bmatrix} \mathbf{H}_0(z) \\ \mathbf{H}_1(z) \end{bmatrix} = \begin{bmatrix} \mathbf{G}_0(z) & z^{-1} \mathbf{G}_1(z) \\ \mathbf{G}_1(z) & \mathbf{G}_0(z) \end{bmatrix} \begin{bmatrix} \mathbf{C}_0(z) \\ \mathbf{C}_1(z) \end{bmatrix} = \begin{bmatrix} \mathbf{C}_0(z) & z^{-1} \mathbf{C}_1(z) \\ \mathbf{C}_1(z) & \mathbf{C}_0(z) \end{bmatrix} \begin{bmatrix} \mathbf{G}_0(z) \\ \mathbf{G}_1(z) \end{bmatrix} \quad (5.5)$$

ou sous forme compacte

$$\mathbf{H}(z) = \underline{\mathbf{G}}(z) \mathbf{C}(z) = \underline{\mathbf{C}}(z) \mathbf{G}(z). \quad (5.6)$$

Dans le domaine temporel, cette expression se traduit comme la convolution suivante

$$\mathcal{T}_M(\mathbf{H}) = \mathcal{T}_M(\underline{\mathbf{G}}) \mathcal{T}_{M+K-1}(\mathbf{C}), \quad (5.7)$$

où \mathbf{C} est similaire à \mathbf{H} et

$$\underline{\mathbf{G}} = [\underline{\mathbf{g}}(K-1) \cdots \underline{\mathbf{g}}(0)], \quad \underline{\mathbf{g}}(k) = \begin{bmatrix} g_0(k) & g_1(k-1) \\ g_1(k) & g_0(k) \end{bmatrix} \quad (5.8)$$

et nous supposons que $g_1(K-1) = 0$. Soit \mathbf{c} le vecteur contenant les coefficients du canal de propagation (défini de façon similaire à \mathbf{h}), $\mathbf{c} = [\mathbf{c}^T(L-1) \cdots \mathbf{c}^T(0)]^T$, la relation entre \mathbf{h} et \mathbf{c} s'écrit comme

$$\mathbf{h} = \mathcal{T}_L^T(\underline{\mathbf{G}}^t) \mathbf{c}, \quad (5.9)$$

où t désigne l'opération de transposition des blocs : $\underline{\mathbf{G}}^t = [\underline{\mathbf{g}}^T(K-1) \cdots \underline{\mathbf{g}}^T(0)]$.

5.2.2 Cas d'un Réseau d'Antennes

Dans les applications CDMA (Code Division Multiple Access), l'excès en largeur de bande est grand et donc de grands facteurs de suréchantillonnage peuvent être considérés. Dans les applications TDMA (Time Division Multiple Access), l'excès en largeur de bande est petit et par conséquent le facteur de suréchantillonnage sera typiquement limité à $m = 2$. Cependant, des sous-canaux supplémentaires peuvent être obtenus en exploitant des signaux provenant d'un réseau d'antenne.

Dans le cas d'un réseau de n antennes, $\mathbf{H}_i(z) = \mathbf{G}(z) \mathbf{C}_i(z)$ pour le signal au niveau de la $i^{\text{ème}}$ antenne, $i = 1 \dots n$, d'où

$$\mathbf{H}(z) = [\mathbf{H}_1^T(z) \cdots \mathbf{H}_n^T(z)]^T = \text{blockdiag}\{\mathbf{G}(z) \cdots \mathbf{G}(z)\} \mathbf{C}(z) \quad (5.10)$$

où maintenant $\mathbf{H}(z)$ et $\mathbf{C}(z)$ regroupent mn canaux et peuvent être exprimés comme suit

$$\mathbf{H}(z) = (I_n \otimes \mathbf{G}(z)) \mathbf{C}(z), \quad (5.11)$$

La convolution temporelle exprimée dans (5.7) devient dans ce cas

$$\mathcal{T}_M(\mathbf{H}) = \mathcal{T}_M([I_n \otimes \underline{\mathbf{g}}(K-1) \cdots I_n \otimes \underline{\mathbf{g}}(0)]) \mathcal{T}_{M+K-1}(\mathbf{C}). \quad (5.12)$$

La connaissance a priori des filtres de transmission est exploitée en exprimant $\mathbf{h} = \mathbf{G}\mathbf{c}$ et en estimant \mathbf{c} , où

$$\mathbf{G} = \mathcal{T}_L^T([I_n \otimes \underline{\mathbf{g}}^T(K-1) \cdots I_n \otimes \underline{\mathbf{g}}^T(0)]). \quad (5.13)$$

5.3 Méthodes Aveugles avec Connaissance des Filtrés de Transmission

L'exploitation de la connaissance des filtres de transmission se fait en exprimant le canal global en fonction du canal de propagation, $\mathbf{h} = \mathcal{T}_L^T(\underline{\mathbf{G}}^t)\mathbf{c}$ (nous considérons ici le cas du suréchantillonnage mais le raisonnement reste le même pour le cas d'un réseau d'antennes, la seule entité à changer est la matrice $\mathcal{T}_L^T(\underline{\mathbf{G}}^t)$ qui sera remplacée par la matrice \mathbf{G}), puis en résolvant le problème d'optimisation pour \mathbf{c} . Les méthodes d'identification aveugle multi-canal discutées dans les chapitres précédents ont, toutes, le facteur commun suivant : le problème de minimisation peut toujours se ramener à la formulation quadratique $\min_{\|\mathbf{h}\|=1} \mathbf{h}^H \mathbf{A} \mathbf{h}$.

5.3.1 Méthodes Aveugles

Dans le cas des méthodes d'estimation multi-canal basées sur la notion de sous-espaces : SRM, ajustement du SS, ajustement du SB et la méthode dérivée de la prédiction linéaire (MDPL), la matrice \mathbf{A} est une matrice constante qui ne dépend que du signal reçu \mathbf{Y} (il n'y a pas de dépendance du canal \mathbf{h}). En ce qui concerne les méthodes de maximum de vraisemblance (dont la résolution nécessite des procédures itératives) : IQML, DIQML et PQML, en plus de la dépendance du signal reçu \mathbf{Y} , à une itération k donnée la matrice \mathbf{A}_k du critère à optimiser est construite à partir de l'estimée du canal à l'itération précédente. Ces méthodes sont récapitulées dans le tableau 5.1.

5.3.2 Méthodes Etendues

Comme les méthodes récapitulées dans le tableau 5.1 sont de la forme $\min_{\|\mathbf{h}\|=1} \mathbf{h}^H \mathbf{A} \mathbf{h}$, en exprimant $\mathbf{h} = \mathcal{T}_L^T(\underline{\mathbf{G}}^t)\mathbf{c}$, nous obtenons

$$\min_{\mathbf{c}} \mathbf{c}^H \mathcal{T}^*(\underline{\mathbf{G}}^t) \mathbf{A} \mathcal{T}^T(\underline{\mathbf{G}}^t) \mathbf{c}. \quad (5.14)$$

L'exploitation des filtres de transmission permet donc d'obtenir, à partir de l'estimée du canal de propagation, l'estimée du canal total comme

$$\hat{\mathbf{h}} = \mathcal{T}_L^T(\underline{\mathbf{G}}^t) \hat{\mathbf{c}}. \quad (5.15)$$

Méthode	Matrice du critère à optimiser : A
SRM	$A = \sum_{i=1}^m \mathcal{S}_i \widehat{R}_{Y^*Y}^* \mathcal{S}_i^H, \mathcal{S}_i = I_N \otimes \mathcal{P}_i.$
Ajustement du SS	A est tel que $\text{tr} \{ \mathcal{T}^H(\mathbf{h}) P_{V_S}^\perp \mathcal{T}(\mathbf{h}) \} = \mathbf{h}^H A \mathbf{h}.$
Ajustement du SB	$A = \sum_{i=1}^m \mathcal{S}_i B^* \mathcal{S}_i^H,$ B est tel que $\text{tr} \{ \mathcal{T}(\mathbf{h}^\perp) P_{V_N}^\perp \mathcal{T}^H(\mathbf{h}^\perp) \} = \text{tr} \{ \mathbf{h}^\perp B \mathbf{h}^{\perp H} \}.$
MDPL	$A = \frac{1}{2\pi j} \oint \mathbf{Q}^\dagger(z) \overline{\mathbf{P}}^\dagger(z) \overline{\mathbf{P}}(z) \mathbf{Q}(z) \frac{dz}{z}, \mathbf{H}(z) = \mathbf{Q}(z) \mathbf{h}.$
IQML	$A_k = \begin{cases} \mathcal{Y}^H \mathcal{Y} & \text{pour } k = 1, \\ \mathcal{Y}^H \mathcal{R}^+(\widehat{\mathbf{h}}_{k-1}) \mathcal{Y} & \text{pour } k > 1. \end{cases}$
DIQML	$A_k = \begin{cases} \mathcal{Y}^H \mathcal{Y} & \text{pour } k = 1, \\ \mathcal{Y}^H \mathcal{R}^+(\widehat{\mathbf{h}}_{k-1}) \mathcal{Y} - \lambda D(\widehat{\mathbf{h}}_{k-1}) & \text{pour } k > 1. \end{cases}$
PQML	$A_k = \begin{cases} \mathcal{Y}^H \mathcal{Y} & \text{pour } k = 1, \\ \mathcal{Y}^H \mathcal{R}^+(\widehat{\mathbf{h}}_{k-1}) \mathcal{Y} - \lambda \mathcal{B}^H(\widehat{\mathbf{h}}_{k-1}) \mathcal{B}(\widehat{\mathbf{h}}_{k-1}) & \text{pour } k > 1. \end{cases}$

TAB. 5.1 – Matrice A du critère à minimiser des méthodes aveugles .

Dans toutes les méthodes, à l'exception de la méthode SRM, nous pouvons utiliser $\|\mathbf{c}\| = 1$ comme contrainte de non trivialité. Dans le cas de la méthode SRM, la contribution du bruit doit être proprement prise en compte pour éviter le biais. Une solution proposée dans [Smi96] consiste à traduire la contrainte sur le canal global $\|\mathbf{h}\|^2 = 1$ en la contrainte suivante sur le canal de propagation $\mathbf{c}^H \mathcal{T}^*(\underline{\mathbf{G}}^t) \mathcal{T}^T(\underline{\mathbf{G}}^t) \mathbf{c} = 1$. Cette translation ramène le problème d'optimisation à un problème de vecteur propre généralisé qui peut alternativement être transformé en un problème régulier de vecteur propre minimal. Cette solution consiste à contraindre le filtre de façon à ce qu'il n'ait pas d'influence sur la composante bruit. Dans [AS97], nous avons proposé une solution qui consiste à éliminer (asymptotiquement) complètement la contribution du bruit. En effet, pour une paramétrisation du SB équilibrée \mathbf{H}^\perp , la contribution du bruit dans EA est un multiple de l'identité alors que la contribution du signal est singulière. Par conséquent, la contribution du bruit peut être éliminée en considérant

$$\min_{\|\mathbf{c}\|=1} \mathbf{c}^H \mathcal{T}^*(\underline{\mathbf{G}}^t) (A - \lambda_{\min}(A) I) \mathcal{T}^T(\underline{\mathbf{G}}^t) \mathbf{c}. \quad (5.16)$$

Dans [PMD99], un algorithme du type sous-espace, se basant sur l'équation (5.14) et exploitant la dualité de cette équation en fonction des coefficients du canal de propagation et ceux du filtre de transmission, est proposé. En fait, cette dualité est déduite de l'équation (5.6), et l'algorithme fonctionne donc en deux temps : dans une première étape,

disposant d'une initialisation en termes des coefficients du filtre de transmission, les coefficients du canal de propagation sont estimés, et dans une seconde étape, ces mêmes coefficients serviront à initialiser la procédure permettant d'estimer les coefficients du filtre de transmission et le processus pourra ainsi être itéré.

Pour exploiter la connaissance des filtres de transmission dans le cas GML, il faut introduire (5.7) dans (4.25).

5.4 Aspects d'Identifiabilité

En se basant sur [BKAK78], les conditions d'identifiabilité lorsqu'on considère le canal global ont été établies : les sous-canaux ne doivent pas partager de zéros en commun [TXK91],[TXK94]. Une interprétation de cette condition d'identifiabilité est la suivante : si les sous-canaux possèdent des zéros en commun, ces termes peuvent être factorisés et la représentation du canal total introduira un mono-canal (correspondant aux termes factorisés) suivi d'une représentation multi-canal. Dans le cas du modèle des symboles déterministes, le mono-canal ne peut pas être séparé à partir de la séquence d'entrée. Lorsque le modèle des symboles est gaussien, l'aspect mono-canal rend les statistiques de second ordre insuffisantes pour identifier le canal (les termes factorisés ne sont pas à phase minimale en général).

Nous discutons ici les aspects d'identifiabilité lorsque la méthode d'estimation multi-canal exploite la connaissance des filtres de TX/RX. Il est clair que lorsque la partie connue ou inconnue du canal n'est pas identifiable ou que le produit de deux fonctions de transfert correspondantes aboutit à des zéros en commun pour le canal total, ce dernier n'est pas identifiable par les méthodes purement aveugles.

En exploitant la connaissance des filtres de TX/RX, la méthode d'estimation multi-canal peut identifier le canal même dans le cas où les filtres de TX/RX n'est pas identifiable et le canal de propagation est identifiable. Pour expliquer cet aspect, considérons le cas de deux sous-canaux ($m = 2$). Si les deux sous-canaux $H_1(z)$ et $H_2(z)$ satisfont les conditions d'identifiabilité aveugles alors il existe deux polynômes $F_1(z)$ et $F_2(z)$ tel que

$$F_1(z)H_1(z) + F_2(z)H_2(z) = 1 \quad (5.17)$$

En remplaçant dans (5.17) $H_1(z)$ et $H_2(z)$ par leur expressions données par (5.5), nous obtenons

$$F_1(z) (G_1(z)C_1(z) + z^{-1}G_2(z)C_2(z)) + F_2(z) (G_2(z)C_1(z) + G_1(z)C_2(z)) = 1, \quad (5.18)$$

ce qui peut s'écrire sous la forme

$$(F_1(z)G_1(z) + F_2(z)G_2(z)) C_1(z) + (z^{-1}F_1(z)G_2(z) + F_2(z)G_1(z)) C_2(z) = 1. \quad (5.19)$$

L'égalité (5.19) implique qu'il existe deux polynômes $Q_1(z) = F_1(z)G_1(z) + F_2(z)G_2(z)$ et $Q_2(z) = z^{-1}F_1(z)G_2(z) + F_2(z)G_1(z)$ tel que $Q_1(z)C_1(z) + Q_2(z)C_2(z) = 1$: cette

condition est la condition d'identifiabilité classique sur le canal de propagation ($C_1(z)$ et $C_2(z)$ sont premiers entre eux). Notons que dans (5.19) les deux polynômes $G_1(z)$ et $G_2(z)$ ne sont pas soumis à une condition particulière, ce qui signifie que même dans le cas où $G(z)$ n'est pas identifiable, nous pouvons identifier le canal.

5.5 Bornes de Cramer-Rao avec Connaissance des Filtres de TX/RX

Dans la section §3.3, nous avons calculé les bornes de Cramer-Rao correspondant au problème de l'estimation aveugle dans le cas du modèle des symboles déterministes et du modèle gaussien. La matrice d'information de Fisher sans contrainte et la borne de Cramer-Rao peuvent être facilement transformées en borne de Cramer-Rao avec connaissance des filtres de transmission avec contrainte. Puisque $\mathbf{h} = \mathcal{T}_L^T(\underline{\mathbf{G}}^t)\mathbf{c}$, le canal \mathbf{h} avec connaissance des filtres de transmission satisfait la contrainte

$$P_{\mathcal{T}_L^T(\underline{\mathbf{G}}^t)}^\perp \mathbf{h} = 0. \quad (5.20)$$

Ceci aboutit à la borne de Cramer-Rao pour l'erreur d'estimation sans biais $\tilde{\mathbf{h}} = \mathbf{h} - \hat{\mathbf{h}}$ avec connaissance des filtres de transmission ([dCS],[SN98],[AS98b],[AS]):

$$C_{\tilde{\mathbf{h}}} = E\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H \geq \left[P_{\mathcal{T}_L^T(\underline{\mathbf{G}}^t)} J(\mathbf{h}) P_{\mathcal{T}_L^T(\underline{\mathbf{G}}^t)} \right]^+. \quad (5.21)$$

Le fait que la pseudo-inverse de Moore-Penrose est utilisée veut dire que les composantes de \mathbf{h} qui sont dans $\ker \left\{ P_{\mathcal{T}_L^T(\underline{\mathbf{G}}^t)} J(\mathbf{h}) P_{\mathcal{T}_L^T(\underline{\mathbf{G}}^t)} \right\}$ sont connues. Ceci veut dire, tout d'abord, que les composantes de \mathbf{h} qui ne sont pas dans l'espace colonne de $\mathcal{T}_L^T(\underline{\mathbf{G}}^t)$ sont connues et que leur valeur vaut zéro (vu la contrainte). Notons, en plus, que la matrice $\mathcal{T}_L^T(\underline{\mathbf{G}}^t)$ peut ne pas être de rang colonne plein (si le taux d'échantillonnage est supérieur au taux de Nyquist). D'un autre côté, une autre singularité apparaît qui est inhérente à l'estimation aveugle de canal. Pour les méthodes déterministes, le canal peut uniquement être identifié à un facteur scalaire près. Nous ajustons le facteur scalaire en considérant comme estimée finale du canal $\hat{\mathbf{h}} = \alpha \hat{\mathbf{h}}$ où le scalaire α est obtenu à partir de $\min_\alpha \|\mathbf{h} - \alpha \hat{\mathbf{h}}\|$.

Pour le modèle des symboles gaussiens, nous avons besoin d'utiliser les paramètres réels associés, qui peuvent être représentés sous la forme vectorielle suivante (à partir de $\mathbf{h} = \mathcal{T}_L^T(\underline{\mathbf{G}}^t)\mathbf{c}$):

$$\mathbf{h}_R = \begin{bmatrix} \text{Re}(\mathbf{h}) \\ \text{Im}(\mathbf{h}) \end{bmatrix} = \begin{bmatrix} \text{Re}(\mathcal{T}_L^T(\underline{\mathbf{G}}^t)) & -\text{Im}(\mathcal{T}_L^T(\underline{\mathbf{G}}^t)) \\ \text{Im}(\mathcal{T}_L^T(\underline{\mathbf{G}}^t)) & \text{Re}(\mathcal{T}_L^T(\underline{\mathbf{G}}^t)) \end{bmatrix} \mathbf{c}_R = G_R \mathbf{c}_R. \quad (5.22)$$

La borne de Cramer-Rao gaussienne avec connaissance des filtres de transmission est donc identique à celle exprimée dans (5.21) mais avec les entités \mathbf{h} et $\mathcal{T}_L^T(\underline{\mathbf{G}}^t)$ remplacées

par \mathbf{h}_R et G_R . L'ambiguïté d'estimation dans le modèle gaussien, correspond à un facteur de phase, $\widehat{\mathbf{h}} = e^{j\phi} \mathbf{h}$, qui peut être ajusté en imposant que le scalaire $\mathbf{h}^H \widehat{\mathbf{h}}$ soit réel et positif.

A l'exception de l'interprétation appropriée de la pseudo-inverse, ces résultats pour le cas gaussien ont été, indépendamment, dérivés dans [Smi96], où certains exemples montrent que l'hypothèse gaussienne améliore considérablement la qualité de l'estimation dans certains cas.

5.6 Simulations

Nous considérons une séquence de symboles i.i.d. BPSK, un paquet de données de longueur $M = 162$ (ceci représente la longueur d'un paquet de données dans le standard TDMA IS-54) et un facteur de suréchantillonnage $m = 2$. Le canal global, illustré dans la figure 5.2, est la convolution d'une impulsion "raised cosine" limitée à $13T$ avec un facteur de roll-off $\alpha = 0.9$ (voir par exemple [Pro95]), et d'un canal à deux trajets $c(t) = \delta(t) - 0.82\delta(t - T)$. Sans exploitation des filtres de transmission le canal global n'est pas identifiable dans ce cas [TXK94]. La mesure de performance est faite par l'intermédiaire de l'erreur quadratique moyenne normalisée (NMSE), en prenant la moyenne sur 100 réalisations Monte-Carlo :

$$\text{NMSE} = \frac{1}{100} \sum_{i=1}^{100} \mathbf{h}^H P_{\widehat{\mathbf{h}}^{(i)}}^\perp \mathbf{h} / \|\mathbf{h}\|^2, \text{ where } \mathbf{h}^H P_{\widehat{\mathbf{h}}}^\perp \mathbf{h} = \min_{\alpha} \|\alpha \widehat{\mathbf{h}} - \mathbf{h}\|^2. \quad (5.23)$$

Le rapport signal à bruit est défini comme :

$$\text{SNR} = \frac{\sigma_a^2 \|\mathbf{h}\|^2}{\sigma_v^2}. \quad (5.24)$$

Les limites de performance considérées sont les bornes de Cramer-Rao déterministes, normalisées et calculées comme : $\text{tr}\{CRB_{\mathbf{h}}\} / \|\mathbf{h}\|^2$.

5.6.1 Techniques Basées sur les Sous-espaces

Dans la figure 5.3, nous comparons la performance (correspondant à l'estimation du canal \mathbf{h}) de la méthode SRM et celle de la méthode SRM avec exploitation de la connaissance du filtre de TX (SRM WTXFK) par rapport à leurs bornes de Cramer-Rao respectives. Les résultats de simulation montrent qu'en termes de CRB, l'approche WTXFK aboutit à des performances supérieures à l'approche purement aveugle. La différence entre la méthode SRM et SRM WTXFK est encore plus spectaculaire : l'algorithme SRM appliqué sur le canal global souffre des zéros qui sont (pratiquement) en commun, alors que l'algorithme SRM WTXFK fonctionne bien et aboutit à de bonnes performances. Dans le but d'illustrer que l'exploitation de la connaissance des filtres de transmission ne se réduit pas simplement à un problème sous-dimensionné (où le

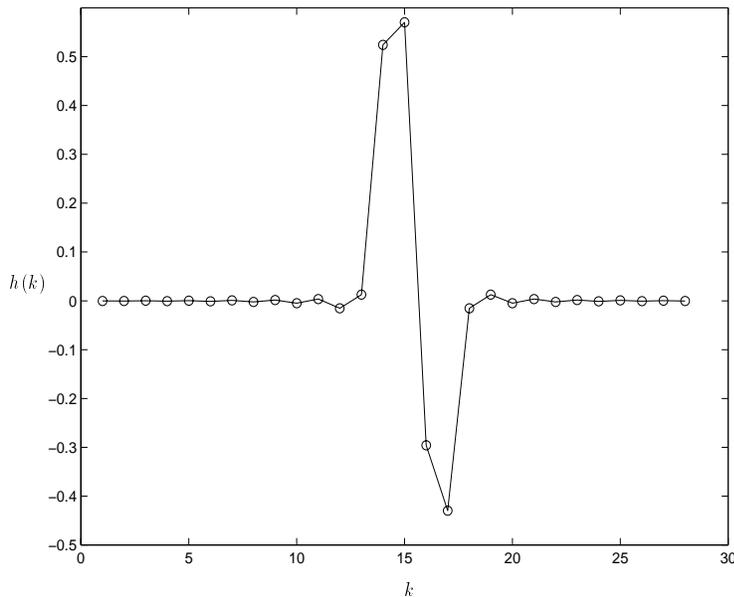


FIG. 5.2 – Canal global.

nombre de paramètres à estimer se trouve réduit), nous considérons la performance de la méthode SRM pondérée par le rapport du nombre de paramètres à estimer : $\frac{L}{N}$. La performance ainsi obtenue (courbe qui correspond à LD SRM : Lower Dimensional SRM) n'atteint pas la performance de SRM WTXFK. En effet, à part la réduction du nombre de paramètres à estimer, l'exploitation de la connaissance des filtres de transmission améliore considérablement le conditionnement de la matrice du critère d'optimisation. Ces deux éléments agissent simultanément pour améliorer la qualité des estimées du canal. Nous avons effectué les mêmes simulations dans le cas de la technique d'ajustement du SS (SSF), les résultats concernant les variantes de la technique d'ajustement du SS (SSF, SSF WTXFK et LD SSF), qui sont illustrés dans la figure 5.4, aboutissent aux mêmes constatations précédentes concernant le cas SRM, SRM WTXFK et LD SRM.

Dans la figure 5.5, avec les mêmes données, nous comparons les deux formes sans biais de la méthode SRM WTXFK : la première version est celle qui utilise $\|\mathbf{h}\| = 1$ comme contrainte de non trivialité alors que la seconde utilise comme contrainte $\|\mathbf{c}\| = 1$ mais en considérant la matrice $A - \lambda_{\min}(A)I$ au lieu de la matrice A . Il est clair que la seconde approche aboutit à des performances supérieures à celles de la première version (avec un facteur supérieur à 5 à un SNR=13dB).

5.6.2 Techniques de Maximum de Vraisemblance

Dans la simulation illustrée dans la figure 5.6, notre idée consiste à étudier le comportement des méthodes SRM WTXFK et IQML WTXFK (une itération de l'algorithme initialisée par l'estimée du canal $\hat{\mathbf{h}}$ obtenue par l'intermédiaire de la méthode SRM WTXFK) en fonction du conditionnement du canal de propagation \mathbf{c} . Nous utilisons

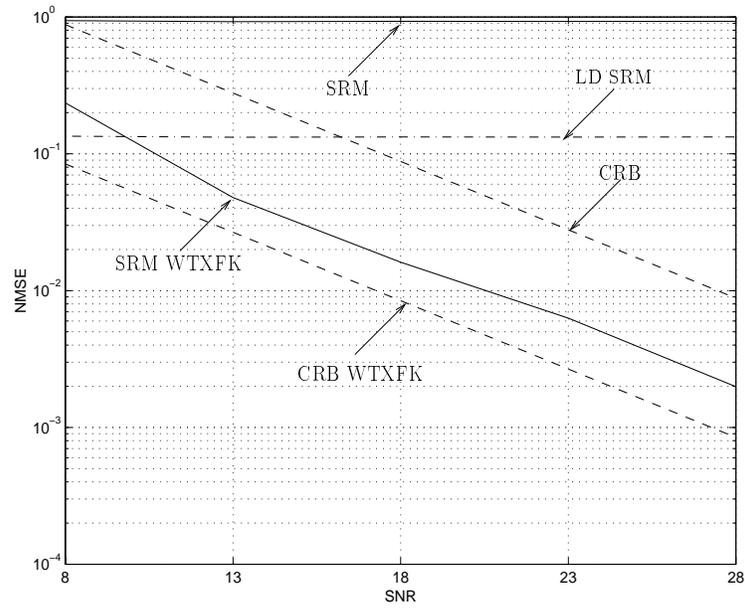


FIG. 5.3 – Performance des méthodes : SRM et SRM WTXFK.

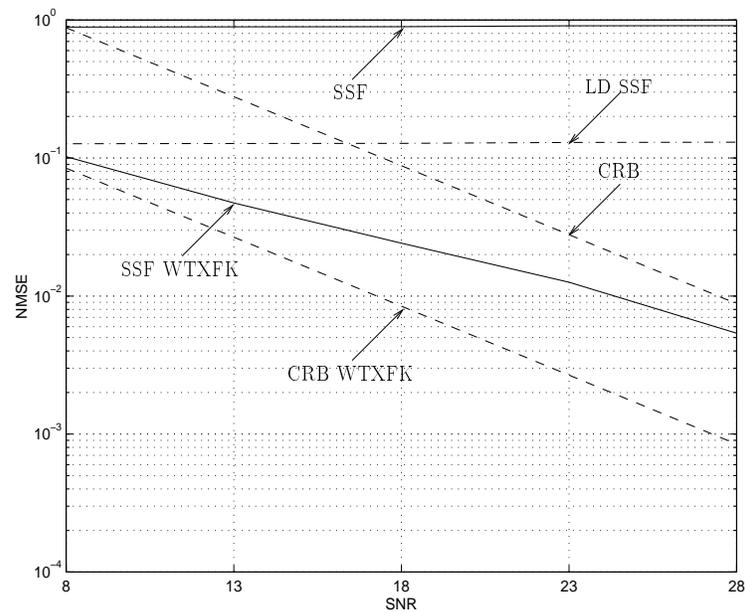


FIG. 5.4 – Performance des méthodes : SSF et SSF WTXFK.

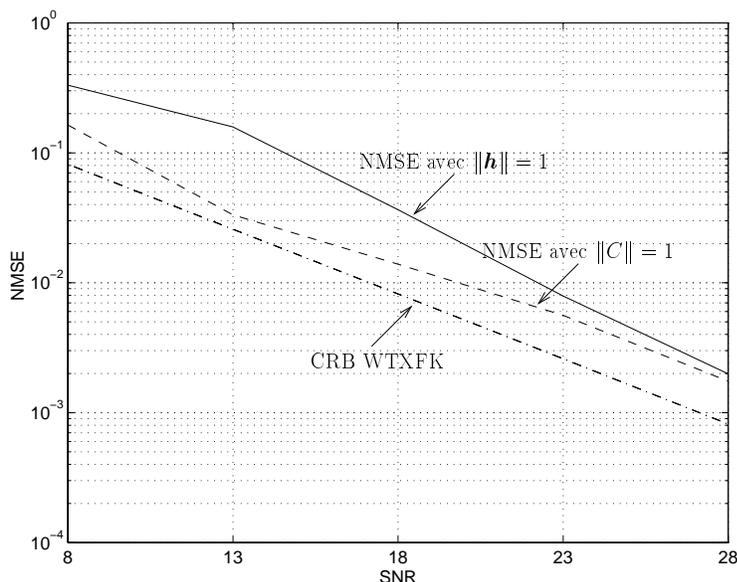


FIG. 5.5 – Comparaison des méthodes SRM WTXFK avec $\|\mathbf{c}\| = 1$ et $\|\mathbf{h}\| = 1$.

le même filtre de mise en forme (décrit précédemment) et nous considérons le canal de propagation défini comme suit :

$$\mathbf{C} = \begin{bmatrix} 1 & 1 \\ 1 & a \end{bmatrix}. \quad (5.25)$$

La variation du paramètre a caractérise donc le conditionnement du canal de propagation \mathbf{c} . Quand $a = 1$, les deux sous-canaux sont parallèles (zéro en commun), et lorsque $a = -1$ les deux sous-canaux sont orthogonaux (le canal est bien conditionné). Nous considérons un SNR=33dB, et un intervalle de variation de -1 à 0.8 pour le paramètre a . Nous constatons que la méthode IQML WTXFK, en moyenne, n'aboutit pas à des performances qui sont meilleures que celles de la méthode SRM WTXFK (la meilleure amélioration est de l'ordre de deux fois obtenue pour $a = 0.4$); mais nous avons noté que pour certaines réalisations les performances de la méthode IQML WTXFK sont nettement meilleures que celles de la méthode SRM WTXFK.

5.7 Conclusions

Les méthodes d'estimation aveugles multi-canal de second ordre peuvent aboutir à des performances meilleures en exploitant la connaissance a priori des filtres de TX et/ou de RX. Une certaine classe de ces méthodes qui regroupe celles qui sont directement paramétrisées par le canal, peuvent être facilement étendues en exploitant cette connaissance. Les méthodes étendues sont supérieures, d'un point de vue de performance, à leurs versions purement aveugles sur deux niveaux (bien que les simulations ont été élaborées dans le cas des méthodes SRM et SSF, les constatations établies restent valables pour toutes les méthodes d'estimation aveugles du canal décrites précédemment). Le premier

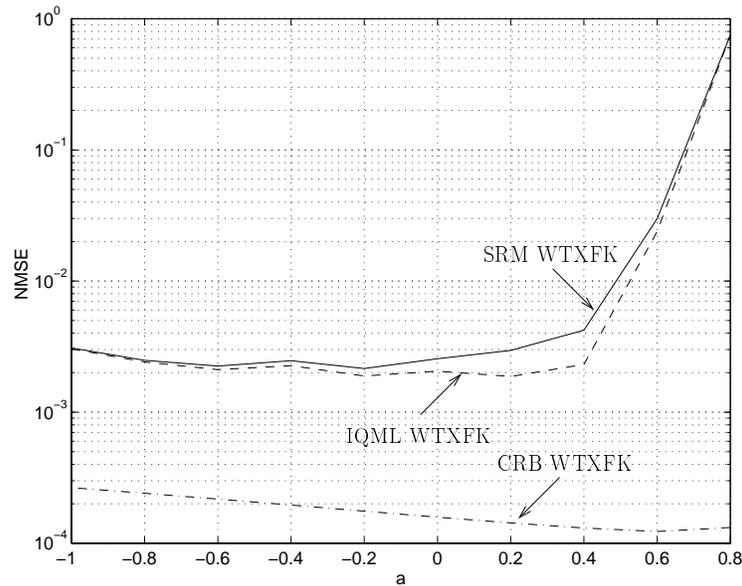


FIG. 5.6 – *Comparaison des méthodes : SRM WTXFK et IQML WTXFK.*

niveau concerne les bornes de performance : les CRBs correspondant aux méthodes qui exploitent cette connaissance a priori sont plus basses que celles correspondant aux versions purement aveugles et donc la borne pouvant être atteinte par un estimateur efficace est plus intéressante dans le cas de l'exploitation des connaissances des filtres de TX/RX que dans le cas purement aveugle. Le second niveau concerne la méthode elle-même : une technique exploitant la connaissance des filtres de TX/RX aboutit à des NMSEs étroitement proches de leur CRBs correspondantes et qui sont nettement inférieures aux NMSEs d'une méthode purement aveugle.

Les méthodes d'identification aveugle souffrent des limites d'identifiabilité (canaux qui possèdent des zéros proches). En pratique, pour être identifiée, cette classe de canaux nécessitera une quantité de données plus importante ou une puissance du signal plus grande (des hauts SNRs). En exploitant la connaissance a priori des filtres de TX/RX, le nombre de paramètres à estimer se trouve réduit et le problème d'optimisation se trouve mieux conditionné. La classe des méthodes d'estimation multi-canal exploitant la connaissance des filtres de TX/RX aboutit donc à des performances meilleures et remédie aux limites d'identifiabilité inhérentes à la classe des méthodes purement aveugles. Notons que deux éléments agissent simultanément pour aboutir à des performances améliorées pour les méthodes exploitant la connaissance des filtres de transmission : le premier élément concerne l'amélioration du conditionnement de la matrice du critère d'optimisation et le second élément agit au niveau de la réduction du nombre de paramètres à estimer et en conséquence, souvent, réduit la complexité de calcul de la méthode considérée.

Chapitre 6

Robustesse à la Surdétermination de l'Ordre du Canal

Les méthodes d'estimation multi-canal aveugle paramétrisées directement par le canal regroupant les méthodes basées sur les sous-espaces (SRM, ajustement du SS, ajustement du SB et MDPL) et les approches de maximum de vraisemblance (DML, IQML, DIQML et PQML) présentent, toutes, l'inconvénient du manque de robustesse à la surdétermination de l'ordre du canal. En d'autres termes, dès que l'ordre du canal est surestimé, ces méthodes n'aboutissent pas à une estimée consistante du canal. L'approche de prédiction linéaire est robuste à la surestimation de l'ordre du canal et produit les solutions de tous les ordres, et en conséquence une détection de l'ordre peut être aisément incorporée. Nous nous inspirons de cette approche pour dériver des mécanismes adéquats permettant de rendre les méthodes d'identification aveugle paramétrisées directement par le canal robustes à la surdétermination de l'ordre. Dans le but de bien définir le critère d'optimisation dans ces approches, une contrainte portant sur les coefficients du canal doit être ajoutée. Typiquement, la contrainte d'un canal de norme unité est utilisée. L'utilisation de cette contrainte provoque l'absence de robustesse à la surestimation de l'ordre du canal. Pour résoudre ce problème, nous proposons d'*ancrer* la réponse impulsionnelle du canal à son premier coefficient en contraignant le premier coefficient du canal à être de norme unité et d'effectuer un traitement approprié sur la matrice du critère à optimiser. Ceci aboutit à une reparamétrisation du canal dans laquelle un polynôme similaire au prédicteur (de la prédiction linéaire) apparaît. Ce polynôme peut à nouveau être construit de manière récursive fournissant les polynômes de tous les ordres.

6.1 Introduction

Dans les chapitres précédents, les méthodes discutées ainsi que les simulations effectuées supposent que l'ordre correct du modèle (ou la longueur du canal) est connu avant l'application de toute procédure d'estimation du canal. En pratique l'ordre du canal n'est pas connu et des procédures permettant sa détection/détermination sont donc né-

cessaires. La qualité de l'estimée du canal dépend de l'efficacité de la procédure utilisée : une surestimation de l'ordre du canal ne permet pas d'aboutir à une estimée consistante du canal. En fait ce problème n'est pas inhérent à l'identification de systèmes mais il peut être aussi rencontré dans plusieurs applications en traitement du signal tel que le traitement d'antennes. Plusieurs approches ont été proposées pour résoudre ce problème : certaines utilisent des critères issus de la théorie d'information tels que "Minimum Description Length" (MDL) et "Akaike's Information theoretic Criterion" (AIC), d'autres utilisent des tests introduisant des seuils de décision (souvent subjectifs) [KY88]. Une procédure de détection d'ordre du canal tel que MDL peut être adoptée dans le cas des méthodes d'estimation aveugle du canal (récapitulées dans le tableau 5.1) ainsi que leurs versions exploitant la connaissance des filtres de TX/RX. Néanmoins, nous nous concentrons, dans ce chapitre, sur des procédures inspirées de la prédiction linéaire proposée dans [Slo94c]. L'idée clé consiste à ancrer la réponse impulsionnelle du canal à son premier coefficient [ADS97b], [ADS97a].

6.2 Ancrage de la Réponse Impulsionnelle

Dans ce qui suit, nous nous intéressons à la formulation du problème de l'estimation du canal de manière à avoir une situation duale à la prédiction linéaire. Nous proposons d'aborder le problème d'estimation du canal de façon à avoir une formulation analogue à celle obtenue dans le cas de la prédiction linéaire où le problème de minimisation se ramène à $\min_P PR_{YY}P^H$ où $P = \begin{bmatrix} p(L-1) & \cdots & P(1) & I_m \end{bmatrix}$ est le prédicteur d'ordre $L-1$.

6.2.1 Prédiction Linéaire

Dans le cas scalaire, par robustesse de la prédiction linéaire, nous entendons : quand on essaye de prédire un processus autorégressif (AR) avec un ordre supérieur à l'ordre correct du processus AR en question, on aura comme coefficients du prédicteur des coefficients du modèle AR jusqu'à l'ordre correct suivis automatiquement par des zéros. Considérons, maintenant, le cas multivariable sans bruit (donc singulier), nous pouvons déterminer $\mathbf{P}(z)$ de façon robuste à la surdétermination de son ordre [Slo94c]. Pour cela, dans les équations normales, au lieu de considérer la pseudo-inverse de R_{YY} , il faut prendre un inverse généralisé lié à la factorisation triangulaire LDU de R_{YY} :

$$R_{YY} = LDL^H. \quad (6.1)$$

Si la matrice R_{YY} est singulière, la matrice D sera aussi singulière. Si la matrice R_{YY} est Toeplitz, les éléments de D seront décroissants et donc tous les zéros arriveront à la fin. Par conséquent, en prenant comme inverse de R_{YY} la matrice

$$R_{YY}^\# = L^{-H} D^+ L^{-1}, \quad (6.2)$$

ou D^+ est la pseudo-inverse habituelle, nous obtenons la solution du filtre qui n'est pas minimale en norme mais minimale en longueur (nombre des coefficients non nuls) : c'est bien cette propriété qui engendre la robustesse à la surdétermination de l'ordre.

Si R_{YY} est block Toeplitz, comme R_{YY} est une matrice singulière, il faut considérer la représentation

$$LR_{YY}L^H = D. \quad (6.3)$$

Les lignes du facteur triangulaire L sont des prédicteurs et les éléments de la matrice D sont des variances d'erreur de prédiction. A partir du bon ordre, la prédiction devient singulière et à partir de ce moment là, les matrices D et L deviennent block Toeplitz, et la ligne de la matrice L qui devient répétée correspond au prédicteur cherché (de longueur minimale).

6.2.2 Approche de Base

Pour aboutir à une situation duale à la prédiction linéaire que nous venons d'expliquer, l'approche que nous considérons consiste à écrire $\mathbf{H}(z)$ comme

$$\mathbf{H}(z) = \mathbf{Q}(z)\mathbf{h}(0) \quad (6.4)$$

ou

$$\mathbf{h} = [\mathbf{h}^H(N-1) \cdots \mathbf{h}^H(1) \mathbf{h}^H(0)]^H = \mathbf{Q}\mathbf{h}(0) = [\mathbf{q}^H(N-1) \cdots \mathbf{q}^H(1) I_m]^H \mathbf{h}(0) \quad (6.5)$$

où la matrice carrée (dont le premier coefficient est I_m) $\mathbf{Q}(z)$ est, en quelque sorte, analogue au polynôme de la prédiction linéaire $\mathbf{P}(z)$. L'idée clé consiste à *ancrer* la réponse impulsionnelle du canal à son premier coefficient. Ceci représente une des propriétés qui conduisent à la robustesse à la surestimation de l'ordre. L'approche que nous proposons peut être utilisée dans les techniques d'estimation aveugle décrites précédemment (récapitulées dans le tableau 5.1). En effet, le facteur commun de ces méthodes est la possibilité de formuler le problème de minimisation sous la forme $\mathbf{h}^H B \mathbf{h}$, sous la contrainte $\|\mathbf{h}\|_2 = 1$. Dans le cas sans bruit, B est singulière avec un noyau de dimension 1 dans les méthodes décrites précédemment. En présence du bruit, nous pouvons rendre B singulière en remplaçant B par $A = B - \lambda_{\min}(B)I$. La minimisation de $\mathbf{h}^H B \mathbf{h}$ ou $\mathbf{h}^H A \mathbf{h}$ sous la contrainte $\|\mathbf{h}\|_2 = 1$ conduit à $\mathbf{h} = V_{\min}(B) = V_{\min}(A)$. La valeur correspondante du critère $\mathbf{h}^H A \mathbf{h}$ est zéro. D'autre part, lorsqu'on minimise $\mathbf{h}^H A \mathbf{h}$ sous la contrainte $\|\mathbf{h}(0)\|_2 = 1$, il est clair que $\mathbf{h} = \alpha V_{\min}(A)$ avec α choisi de façon à ce que $\|\mathbf{h}(0)\|_2 = 1$ annule $\mathbf{h}^H A \mathbf{h}$ et donc $\mathbf{h} = \alpha V_{\min}(A)$ minimise $\mathbf{h}^H A \mathbf{h}$ sous la contrainte $\|\mathbf{h}(0)\|_2 = 1$. En d'autres termes, en remplaçant B par A , les problèmes $\min_{\|\mathbf{h}\|_2=1} \mathbf{h}^H B \mathbf{h}$ et $\min_{\|\mathbf{h}(0)\|_2=1} \mathbf{h}^H A \mathbf{h}$ conduisent à des solutions proportionnelles et donc équivalentes pour \mathbf{h} . Le problème $\min_{\|\mathbf{h}(0)\|_2=1} \mathbf{h}^H A \mathbf{h}$ peut être formulé comme suit

$$\min_{\|\mathbf{h}(0)\|_2=1} \mathbf{h}^H(0) \left(\min_{\mathbf{q}(0)=I_m} \mathbf{Q}^H A \mathbf{Q} \right) \mathbf{h}(0). \quad (6.6)$$

6.3 Robustesse

L'opération d'ancrage de la réponse impulsionnelle du canal à son premier coefficient, elle-même, ne permet pas d'obtenir la propriété de robustesse à la surdétermination de l'ordre du canal. Pour obtenir cette robustesse, nous proposons de faire suivre l'étape d'ancrage de la réponse impulsionnelle du canal par une étape dans laquelle nous exploitons, de façon appropriée, la structure de la matrice A du critère à optimiser. L'exploitation de cette structure peut se faire selon deux approches : la première approche consiste en une détection implicite de l'ordre du canal cependant la seconde approche constitue un mécanisme de détermination explicite de cet ordre.

6.3.1 Première Approche : Détermination Implicite de l'Ordre

Dans [ADS97b], une première approche utilisant la formulation du problème donnée par (6.6) a été introduite. L'obtention de la robustesse à la surestimation d'ordre ($N' > N$) se fait en deux étapes. Dans une première étape, nous réécrivons A et \mathbf{h} sous la forme matricielle suivante :

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} \tilde{\mathbf{h}} \\ \mathbf{h}(0) \end{bmatrix} \quad (6.7)$$

où A_{11} est $m(N' - 1) \times m(N' - 1)$ et A_{22} est $m \times m$. La solution de (6.6) est :

$$\begin{aligned} \mathbf{h}(0) &= V_{min} (A_{22} - A_{21}A_{11}^{-1}A_{12}), \\ \tilde{\mathbf{h}} &= \left[\mathbf{h}^H(N'-1) \dots \mathbf{h}^H(1) \right]^H = -A_{11}^{-1}A_{12}\mathbf{h}(0). \end{aligned} \quad (6.8)$$

L'ancrage de la réponse impulsionnelle, sans traitement de la matrice du critère d'optimisation, ne permet pas d'obtenir la robustesse à la surestimation d'ordre : A_{11} est singulière quand $N' > N$ et l'utilisation de la pseudo-inverse de Moore-Penrose A_{11}^{\dagger} aboutirait à des problèmes de surestimation de l'ordre du canal. La deuxième étape consiste donc d'une détection implicite de l'ordre N : nous considérons la factorisation triangulaire UDL pour $A_{11} = UDU^H$, nous forçons les éléments bloc diagonaux singuliers qui se répètent dans D à zéro, nous mettons les éléments des colonnes correspondantes dans U , qui se trouvent en haut des éléments diagonaux, à zéro, et nous utilisons $A_{11}^{\#} = U^{-H}D^+U^{-1}$ comme inverse généralisé de A_{11} .

Pour expliquer comment la démarche précédente permet d'obtenir la robustesse à la surestimation de l'ordre du canal, considérons le cas où l'ordre du canal est surdéterminé et la méthode SRM (les conclusions qui seront établies sont aussi valables pour toutes les méthodes aveugles paramétrisées par \mathbf{h} et qui sont récapitulées dans le tableau 5.1). Soit $\mathbf{H}^{\perp\dagger}(z)$ la paramétrisation $\mathbf{H}_{bal}^{\perp\dagger}(z)$ définie dans (2.36). Dans le domaine fréquentiel, le critère SRM est

$$\frac{1}{M} \frac{1}{2\pi j} \oint \|\hat{\mathbf{H}}^{\perp\dagger}(z) \mathbf{y}(z)\|^2 \frac{dz}{z} \xrightarrow{M \rightarrow \infty} \text{tr} \left\{ \frac{1}{2\pi j} \oint \hat{\mathbf{H}}^{\perp\dagger}(z) S_{\mathbf{y}\mathbf{y}}(z) \hat{\mathbf{H}}^{\perp}(z) \frac{dz}{z} \right\} = \hat{\mathbf{h}}^T B^* \hat{\mathbf{h}}^* \quad (6.9)$$

où B est la matrice définie dans (3.27). En considérant la $i^{\text{ème}}$ ligne : $\widehat{\mathbf{H}}_i^\perp(z)$ de $\widehat{\mathbf{H}}_{bal}^{\perp\dagger}(z)$, l'expression donnée par (6.9) devient

$$\begin{aligned}
& \text{tr} \left\{ \frac{1}{2\pi j} \oint \widehat{\mathbf{H}}^{\perp\dagger}(z) S_{\mathbf{y}\mathbf{y}}(z) \widehat{\mathbf{H}}^\perp(z) \frac{dz}{z} \right\} \\
&= \sum_{i=1}^m \frac{1}{2\pi j} \oint \widehat{\mathbf{H}}_i^{\perp\dagger}(z) S_{\mathbf{y}\mathbf{y}}(z) \widehat{\mathbf{H}}_i^\perp(z) \frac{dz}{z} \\
&= \frac{1}{2\pi j} \oint \widehat{\mathbf{H}}^T(z) \left(\sum_{i=1}^m \mathcal{P}_i S_{\mathbf{y}\mathbf{y}}(z) \mathcal{P}_i^T \right) \widehat{\mathbf{H}}^{T\dagger}(z) \frac{dz}{z} \\
&= \frac{1}{2\pi j} \oint \widehat{\mathbf{H}}^T(z) \left(2\sigma_v^2 I_m + \sum_{i=1}^m \mathcal{P}_i \mathbf{H}(z) \mathbf{H}^\dagger(z) \mathcal{P}_i^T \right) \widehat{\mathbf{H}}^{T\dagger}(z) \frac{dz}{z}
\end{aligned} \tag{6.10}$$

où la matrice $B = A + 2\sigma_v^2 I$ est une matrice bloc Toeplitz. La matrice $\sum_{i=1}^m \mathcal{P}_i \mathbf{H}(z) \mathbf{H}^\dagger(z) \mathcal{P}_i^T$

est de rang $m-1$ (pour n'importe quel choix de $\mathbf{H}^\perp(z)$). Si la longueur de l'estimée du canal $\widehat{\mathbf{h}}$ est N' (la longueur du canal \mathbf{h} est N), alors les matrices A et A_{11} sont singulières avec des degrés de singularités respectifs égaux à $N'-N+1$ et $N'-N$. Le choix de la solution de norme minimale pour $\widehat{\mathbf{h}}$ (correspondant à l'utilisation de la pseudo-inverse de Moore-Penrose A_{11}^\dagger pour A_{11}) engendrerait des problèmes de surdétermination de l'ordre du canal. Considérons, maintenant, la factorisation triangulaire UDL de A_{11} (de dimension bloc $N'-1$), $A_{11} = UDU^H$. La matrice A_{11} étant hermitienne et bloc Toeplitz, Les matrices U et D deviennent bloc Toeplitz après N blocs. Par conséquent, le dernier élément bloc diagonal dans D apparaît $N'-N$ fois et sa diagonale est composée de $m-1$ éléments non nuls suivis par un zéro. Pour résoudre le problème de la surestimation de l'ordre du canal, ces éléments bloc diagonaux singuliers (qui se repètent) de la matrice D doivent être remplacés par des zéros. En plus, les éléments des colonnes correspondantes dans la matrice U , qui se trouvent en haut des éléments diagonaux, doivent être forcés à zéro. L'étape suivante consiste à utiliser la matrice $A_{11}^\# = U^{-H} D^+ U^{-1}$ comme inverse généralisé de la matrice A_{11} . Par conséquent, La solution de (6.6), robuste à la surdétermination de l'ordre du canal, est :

$$\begin{aligned}
\mathbf{h}(0) &= V_{\min} \left(A_{22} - A_{21} A_{11}^\# A_{12} \right), \\
\widehat{\mathbf{h}} &= \left[\mathbf{h}^H(N'-1) \cdots \mathbf{h}^H(1) \right]^H = -A_{11}^\# A_{12} \mathbf{h}(0).
\end{aligned} \tag{6.11}$$

Cette approche introduisant le polynôme $\mathbf{Q}(z)$ constitue le dual de l'approche introduite dans [Slo94c] pour le prédicteur $\mathbf{P}(z)$, pour rendre l'approche de prédiction linéaire robuste à la surdétermination de l'ordre.

6.3.2 Deuxième Approche : Détermination Explicite de l'Ordre

Le noyau de la matrice A est, dans le cas de statistiques exactes, de dimension $N' - N + 1$, où N' est la longueur (sur)estimée. En plus, la matrice A est Toeplitz par

blocs, ce qui permet d'en effectuer une factorisation triangulaire rapide par l'algorithme de Levinson multi-canal adapté au cas singulier : $LAL^H = D$ où L est une matrice triangulaire inférieure par blocs, et D est une matrice bloc diagonale. Les blocs de D représentent l'équivalent des matrices de covariance des erreurs de prédiction (dans le cas de la prédiction linéaire). Pour l'ordre $n \geq N - 1$, l'élément de la matrice bloc diagonale est singulier, et le calcul des inverses dans l'algorithme de Levinson est remplacé par un calcul de pseudo-inverses. Ce faisant, de manière similaire à [Slo94c], les éléments $\mathbf{q}(n), n \geq N$ deviennent nuls. D'autre part, l'élément $\mathbf{h}(0)$ est déterminé comme étant le vecteur propre minimal du dernier élément bloc de la matrice bloc diagonale et $\mathbf{h} = \mathbf{Q}\mathbf{h}(0)$. La robustesse à la surdétermination de l'ordre s'en déduit immédiatement. Dans le cas de statistiques estimées, le seuil intervenant dans le calcul de la pseudo-inverse sera fonction du SNR (qui peut être déduit de $\lambda_{\min}(A)$). On notera que l'apparition du premier élément bloc diagonal singulier correspond à l'ordre exact et constitue un mécanisme de détermination explicite de l'ordre du canal.

6.4 Simulations

Pour illustrer la robustesse à la surdétermination de l'ordre du canal dans le cas des deux approches décrites précédemment, nous mesurons la performance des méthodes d'estimation par l'intermédiaire de l'erreur quadratique moyenne normalisée (NMSE) (en prenant la moyenne sur 300 réalisations Monte Carlo) donnée par

$$\text{NMSE} = \frac{1}{300} \sum_{i=1}^{300} \mathbf{h}^H P_{\hat{\mathbf{h}}^{(i)}}^\perp \mathbf{h} / \|\mathbf{h}\|^2 \text{ où } \mathbf{h}^H P_{\hat{\mathbf{h}}}^\perp \mathbf{h} = \min_{\alpha} \|\alpha \hat{\mathbf{h}} - \mathbf{h}\|^2. \quad (6.12)$$

Nous utilisons un canal complexe, généré de façon aléatoire \mathbf{H} , avec $N = 3$ et $m = 3$, donné come suit

$$\mathbf{H} = \begin{bmatrix} 0.0591 - 0.3600j & 0.3516 + 1.2460j & 1.1650 + 0.8717j \\ 1.7971 - 0.1356j & -0.6965 - 0.6390j & 0.6268 - 1.4462j \\ 0.2641 - 1.3493j & 1.6961 + 0.5774j & 0.0751 - 0.7012j \end{bmatrix}. \quad (6.13)$$

Les symboles sont i.i.d. BPSK, et la longueur de la trame des données est $M = 100$. Le rapport signal à bruit est défini comme : $\text{SNR} = (\|\mathbf{h}\|^2 \sigma_a^2) / (m \sigma_v^2)$. Le canal est supposé avoir un ordre surestimé égal à N' .

6.4.1 Cas de la Première Approche

Dans la figure 6.1, en utilisant l'ordre correct du canal, nous comparons la performance des estimées du canal obtenues par l'intermédiaire de la méthode SRM utilisant les contraintes de normalisation $\|\mathbf{h}\| = 1$ et $\|\mathbf{h}(0)\| = 1$. Nous comparons aussi l'erreur normalisée obtenue à la borne de Cramer-Rao normalisée [SP94]. Les deux contraintes de normalisation aboutissent à des performances comparables pour l'estimation du canal

surtout à des hauts SNRs. Notons que lorsque la contrainte de normalisation $\|\mathbf{h}\| = 1$ est utilisée, la soustraction de la contribution du bruit n'a pas d'influence sur la performance de l'opération d'estimation du canal.

Dans la figure 6.2, le problème de la surdétermination de l'ordre du canal est illustré : nous appliquons la méthode SRM en supposant que l'ordre du canal est $N' = 4 > 3 = N$. Nous pouvons observer que, lorsque le plus petit élément de la matrice D est mis à zéro, l'utilisation de la contrainte de normalisation $\|\mathbf{h}(0)\| = 1$ aboutit à un léger accroissement du NMSE par rapport à l'utilisation de la contrainte de normalisation $\|\mathbf{h}\| = 1$ (et en considérant l'ordre correct du canal). Lorsque m éléments de la matrice D sont forcés à zéro, nous observons une diminution du NMSE obtenu avec la contrainte de normalisation $\|\mathbf{h}(0)\| = 1$ par rapport à celui obtenu avec la contrainte $\|\mathbf{h}\| = 1$ (et en considérant l'ordre correct du canal). L'échec de l'utilisation de la contrainte de normalisation $\|\mathbf{h}\| = 1$ lorsque l'ordre du canal est surestimé est aussi illustré dans la figure 6.2.

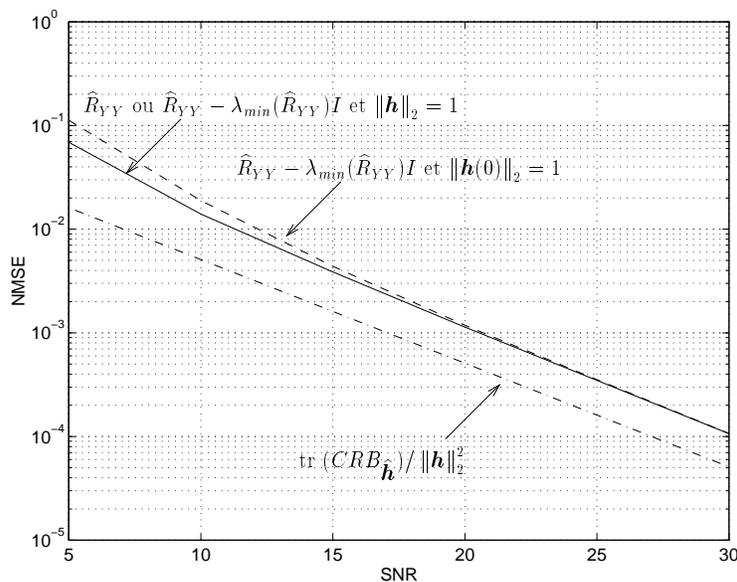


FIG. 6.1 – *SRM - première approche : comparaison des erreurs (NMSEs) obtenues avec les contraintes $\|\mathbf{h}\|_2 = 1$ et $\|\mathbf{h}(0)\|_2 = 1$.*

Nous considérons dans ce qui suit la méthode d'ajustement du SS (SSF). Lorsque la longueur du canal est supposée être égale à $N' > N$, le sous-espace signal correspondant est de dimension $M+N'-1$ (qui est supérieure à la dimension du SS correspondant à la longueur correcte du canal). Dans la méthode d'ajustement du SS, l'estimée du canal est obtenue en exprimant l'orthogonalité entre ce SS et un SB de dimension réduite, égale à $(m-1)M-N'+1$. Ceci ne pose aucun problème fondamental puisque l'expression de l'orthogonalité du SS à $m-1$ vecteurs du SB aboutit à l'identifiabilité [SP94] (notons, toutefois, que pour la méthode d'ajustement du SB, si une longueur du canal surestimée est utilisée, la dimension du SS ne doit pas être surestimée). Dans la figure 6.3,

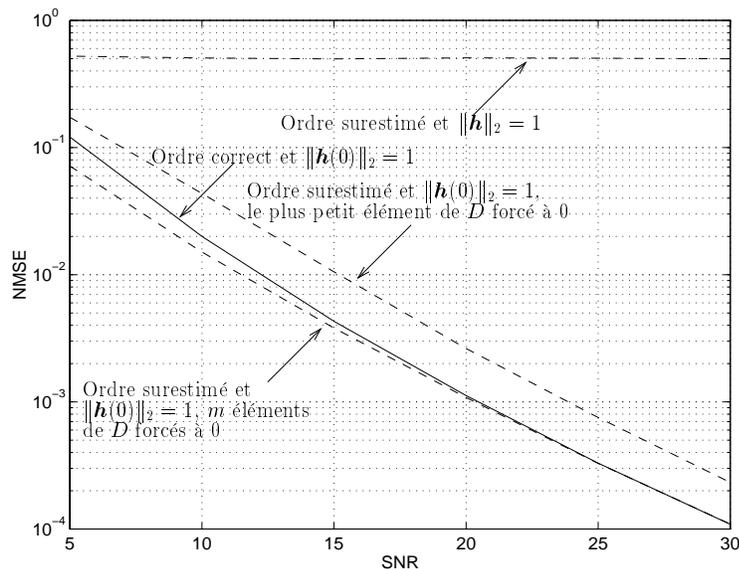


FIG. 6.2 – SRM - première approche : robustesse à la surestimation de l'ordre du canal avec les contraintes $\|\mathbf{h}\|_2 = 1$ et $\|\mathbf{h}(0)\|_2 = 1$.

nous appliquons la méthode d'ajustement du SS en supposant que l'ordre du canal est $N' = 5 > 3 = N$ et nous utilisons la procédure décrite précédemment dans laquelle nous ne forçons pas tous les m éléments dans le bloc singulier de la matrice D à zéro, nous mettons uniquement le dernier élément de chaque bloc excédent à zéro. Nous observons que l'utilisation de la contrainte de normalisation $\|\mathbf{h}(0)\| = 1$ aboutit à un léger accroissement du NMSE par rapport à l'utilisation de la contrainte de normalisation $\|\mathbf{h}\| = 1$ (et en considérant l'ordre correct du canal). L'échec de l'utilisation de la contrainte de normalisation $\|\mathbf{h}\| = 1$ lorsque l'ordre du canal est surestimé est clair.

Dans la figure 6.4, nous supposons que nous avons le même ordre surestimé du canal et nous estimons \mathbf{h} conformément à la procédure décrite précédemment. Nous forçons, dans ce cas, tous les m éléments dans chaque bloc singulier de la matrice D à zéro. Les courbes correspondant à la contrainte de normalisation $\|\mathbf{h}(0)\| = 1$ dans le cas où l'ordre du canal est surdéterminé et la contrainte $\|\mathbf{h}\| = 1$ dans le cas où l'ordre du canal est correct deviennent superposées pour un $\text{SNR} \geq 10\text{dB}$.

6.4.2 Cas de la Deuxième Approche

Les coefficients $\mathbf{q}(N'-1) \cdots \mathbf{q}(1)$ sont calculés par l'intermédiaire de l'algorithme de Levinson multicanal. Dans la figure 6.5, nous illustrons la robustesse de notre approche dans le cas de la méthode SRM. Nous avons utilisé la technique SRM pour estimer le canal \mathbf{H} en supposant que sa longueur est $N' = 4 > 3 = N$, nous constatons que l'utilisation de la contrainte $\|\mathbf{h}(0)\| = 1$ conduit à une légère augmentation du NMSE, alors que l'échec de l'utilisation de la contrainte $\|\mathbf{h}\| = 1$ pour préserver la robustesse de la méthode à la surdétermination de la longueur du canal est clair. Dans la figure 6.6,

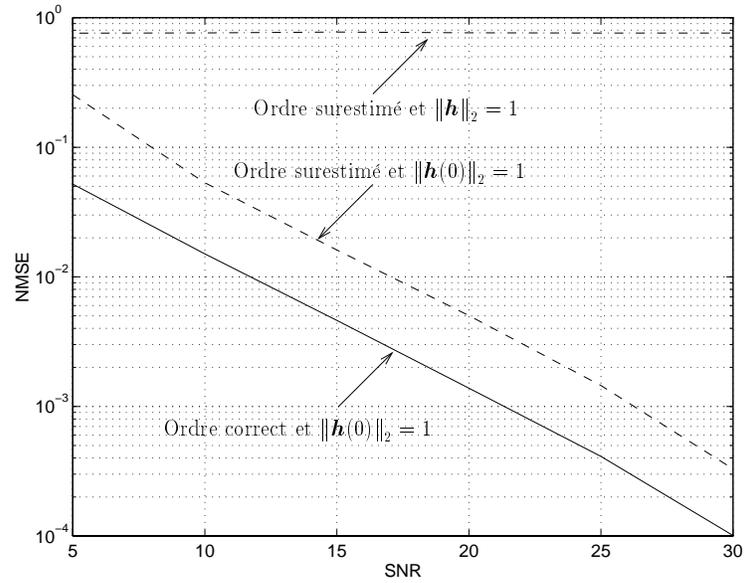


FIG. 6.3 – *SSF - première approche : robustesse à la surestimation de l'ordre du canal avec les contraintes $\|\mathbf{h}\|_2 = 1$ et $\|\mathbf{h}(0)\|_2 = 1$ en utilisant une solution intermédiaire.*

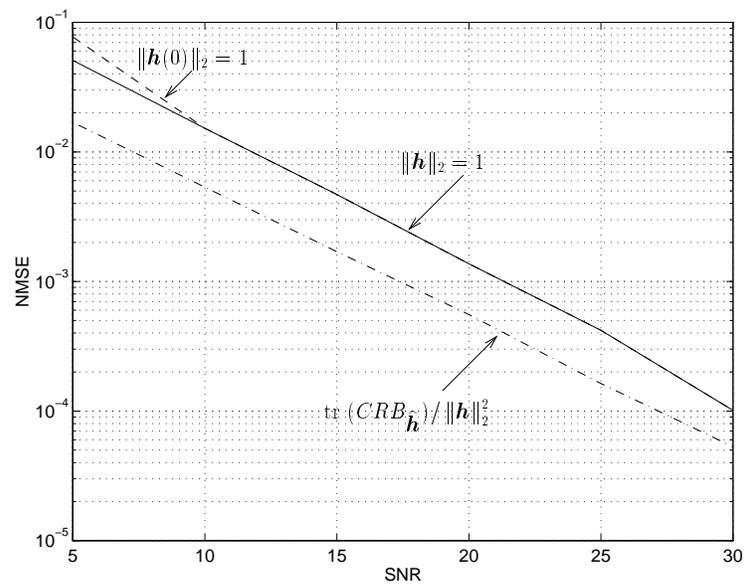


FIG. 6.4 – *SSF - première approche : robustesse à la surestimation de l'ordre du canal avec les contraintes $\|\mathbf{h}\|_2 = 1$ et $\|\mathbf{h}(0)\|_2 = 1$ en utilisant la procédure proposée.*

la robustesse de notre approche est illustrée dans le cas de la méthode d'ajustement du SS (SSF). Nous avons utilisé un ensemble incomplet de vecteurs de l'espace bruit pour estimer \mathbf{H} en supposant que sa longueur est $N' = 5 > 3 = N$ et nous avons implémenté la stratégie décrite précédemment. Cette dernière donne l'ordre correct du canal jusqu'à un SNR de 5 dB. Nous constatons que l'utilisation de la contrainte $\|\mathbf{h}(0)\| = 1$ conduit à une légère augmentation du NMSE par rapport à la méthode classique (dans laquelle, pour le cas de l'ordre correct, nous avons utilisé l'ensemble complet des vecteurs du sous-espace bruit), alors que l'utilisation de la contrainte $\|\mathbf{h}\| = 1$ échoue à préserver la robustesse de la méthode à la surdétermination de la longueur du canal.

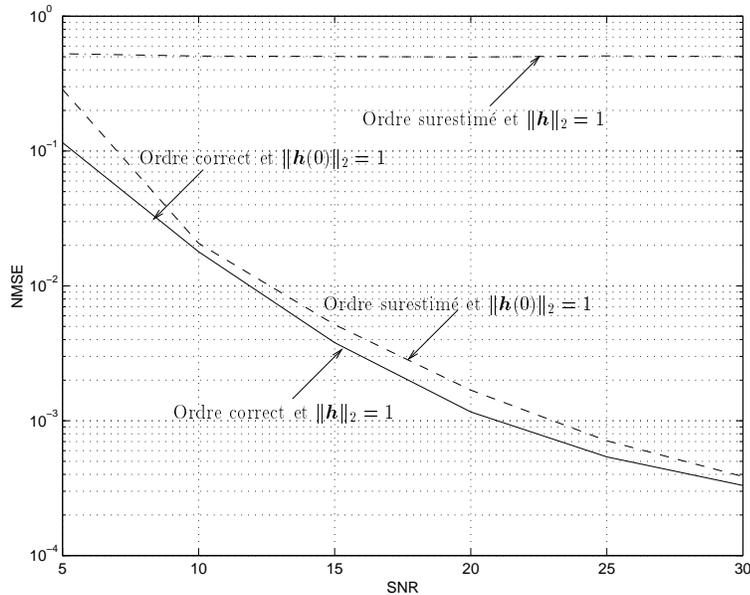


FIG. 6.5 – SRM - deuxième approche : robustesse à la surestimation de l'ordre du canal avec les contraintes $\|\mathbf{h}\|_2 = 1$ et $\|\mathbf{h}(0)\|_2 = 1$.

6.5 Conclusions

La classe des méthodes d'estimation aveugle multi-canal souffre du manque de robustesse à la surdétermination de l'ordre du canal. En s'inspirant de la prédiction linéaire (qui présente bien cet aspect de robustesse), nous avons pu établir un formalisme permettant d'aboutir à la robustesse des méthodes d'estimation aveugle multi-canal. En effet, dans le cas de l'estimation aveugle multi-canal, nous avons reformulé le problème d'optimisation de façon à obtenir une situation duale à celle obtenue dans le cas de la prédiction linéaire. La formulation que nous avons établie se base essentiellement sur l'ancrage de la réponse impulsionnelle du canal en son premier coefficient. Pour obtenir la robustesse à la surdétermination de l'ordre, cette étape d'ancrage doit être suivie par un traitement approprié sur la matrice du critère à optimiser. Ce traitement peut s'opérer

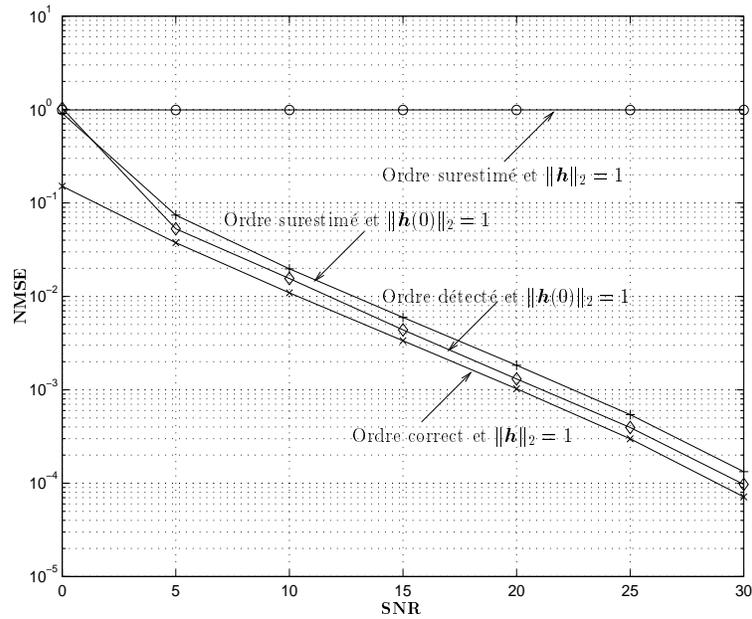


FIG. 6.6 – SSF - deuxième approche : robustesse à la surestimation de l'ordre du canal avec les contraintes $\|\mathbf{h}\|_2 = 1$ et $\|\mathbf{h}(0)\|_2 = 1$.

selon deux approches : dans la première approche, l'ordre du canal est déterminé de façon implicite, et dans la seconde approche, la détermination de cet ordre est faite de façon explicite. Les résultats des simulations obtenus ont confirmé l'efficacité des approches proposées pour préserver la robustesse des méthodes d'estimation aveugles multi-canal.

Chapitre 7

Estimation Aveugle en Présence d'un Bruit Stationnaire de Couleur Inconnue

Les méthodes d'estimation aveugles multi-canal décrites dans les chapitres précédents utilisent, toutes, le même modèle de données où le bruit additif est supposé blanc ou de couleur connue. Dans ce chapitre, nous nous intéressons au cas où le bruit additif n'est plus considéré comme blanc mais uniquement stationnaire. Nous proposons, dans ce contexte, une technique d'estimation de canal aveugle qui se base sur l'idée suivante : dans le cas du suréchantillonnage, le signal d'intérêt est *cyclostationnaire* cependant le bruit est *stationnaire*. A ce niveau, nous pouvons différencier les deux composantes signal et estimer leurs paramètres conjointement. De façon alternative, nous établissons une approche dans laquelle la contribution du bruit se trouve éliminée pour laisser uniquement l'information sur le signal d'intérêt. Cette approche est liée aux statistiques cycliques dont l'avantage est l'élimination de la source stationnaire (bruit) ou les autres sources cyclostationnaires ayant des cycles différents [Gar91b],[Gar93].

7.1 Introduction

Nous considérons le cas où les canaux multiples FIR sont obtenus à partir d'un réseau de n antennes et d'un suréchantillonnage du signal reçu avec un facteur m . Dans ce cas, la matrice de covariance du bruit stationnaire est bloc Toeplitz dont les blocs sont de dimension $n \times n$, cependant les matrices de covariance de la partie signal et du signal reçu total sont bloc Toeplitz dont les blocs sont de dimension $mn \times mn$. Dans ce chapitre, nous nous concentrons essentiellement sur une méthode d'ajustement de sous-espace basée sur une structure de déplacement appropriée de la matrice de covariance dans laquelle la contribution du bruit disparaît. La technique établie aboutit à des performances acceptables (en termes d'estimation du canal) en comparaison avec sa borne de Cramer-Rao. Comme la méthode proposée est basée sur une paramétrisation

linéaire en termes de la réponse impulsionnelle du canal, la connaissance a priori des filtres de transmission peut être facilement incorporée [AS97]. La méthode établie peut donc exploiter cette information et sa performance peut, en conséquence, être améliorée.

Le cas de l'estimation des directions d'arrivée (DOA) dans un bruit coloré a été analysé dans [Gör97a], où certaines méthodes et bornes de performances sont présentées pour estimer les directions d'arrivée. Lorsque le bruit n'est pas modélisé (la couleur du bruit n'est pas prise en compte), la qualité de l'estimation se dégrade parce que les estimées obtenues sont fortement biaisées [Vib93]. Dans le contexte de l'estimation aveugle multi-canal, une méthode basée sur une modification de l'approche sous-espace classique a été proposée dans le cas où la matrice de covariance du bruit spatialement coloré est inconnue [AMHLM97]. Dans [AS98b], nous exploitons une structure de déplacement de la matrice de covariance du signal reçu pour dériver une méthode d'ajustement de sous-espace identifiant le canal lorsque le bruit additif est supposé uniquement stationnaire.

7.2 Méthode d'Identification

Nous considérons mn canaux obtenus à partir d'un suréchantillonnage avec un facteur m , et n antennes réceptrices. La relation entrée/sortie qui régit le système est

$$\mathbf{y}(k) = \sum_{i=0}^{N-1} \mathbf{h}(i)a(k-i) + \mathbf{v}(k) = \mathbf{H}A_N(k) + \mathbf{v}(k). \quad (7.1)$$

Cette relation n'est autre que celle donnée par (1.6) mais où les entités vectorielles de taille m qui y sont associées deviennent de taille mn . En effet,

$$\mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ \vdots \\ y_{mn}(k) \end{bmatrix}, \mathbf{h}(i) = \begin{bmatrix} h_1(i) \\ \vdots \\ h_{mn}(i) \end{bmatrix}, \mathbf{v}(k) = \begin{bmatrix} v_1(k) \\ \vdots \\ v_{mn}(k) \end{bmatrix}. \quad (7.2)$$

Soit $\mathbf{H}(z) = \sum_{i=0}^{N-1} \mathbf{h}(i)z^{-i} = [\mathbf{H}_1(z) \cdots \mathbf{H}_{mn}(z)]^T$ la fonction de transfert du canal mono-entrée/multi-sortie. Considérons les symboles i.i.d. et le bruit $\mathbf{v}(k)$ additif indépendant (gaussien) avec $r_{\mathbf{v}\mathbf{v}}(k-i) = E \mathbf{v}(k)\mathbf{v}(i)^H$. Notons que l'hypothèse de stationnarité du bruit fait que les coefficients de covariance $r_{\mathbf{v}\mathbf{v}}(k)$ sont bloc Toeplitz dont les blocs sont de dimension $n \times n$. Si M échantillons sont reçus, la relation vectorielle qui régit le système est la suivante

$$\mathbf{Y} = \mathcal{T}(\mathbf{h})A + \mathbf{V}, \quad (7.3)$$

où les quantités \mathbf{Y} , $\mathcal{T}(\mathbf{h})$, A et \mathbf{V} sont définies de la même façon que celles de l'expression (1.11). Nous supposons que $mnM > M+N-1$, auquel cas la matrice de convolution $\mathcal{T}(\mathbf{h})$ possède plus de lignes que de colonnes. Si $\mathbf{H}_i(z)$, $i = 1, \dots, mn$ n'ont pas de zéros communs (le canal est irréductible), alors la matrice $\mathcal{T}(\mathbf{h})$ est de rang colonne plein (hypothèse que nous adoptons dans ce qui suit).

7.2.1 Structure de Déplacement

L'idée de base consiste à remarquer que la matrice de covariance du bruit coloré R_{VV} est bloc Toeplitz dont les blocs sont de dimension $n \times n$. D'un autre côté, $\mathcal{T}(\mathbf{h})R_{AA}\mathcal{T}^H(\mathbf{h})$ est bloc Toeplitz dont les blocs sont de dimension $mn \times mn$ (nous supposons que les symboles transmis a_k sont stationnaires et donc la matrice R_{AA} est Toeplitz). Par conséquent, la matrice de covariance du signal reçu

$$R_{YY} = \mathcal{T}(\mathbf{h})R_{AA}\mathcal{T}^H(\mathbf{h}) + R_{VV} \quad (7.4)$$

est bloc Toeplitz dont les blocs sont de dimension $mn \times mn$. Nous considérons le déplacement de la matrice R_{YY} au niveau des blocs de dimension $n \times n$. Cette structure de déplacement est réalisée par l'extraction de deux sous-matrices \underline{R}_{YY} et \overline{R}_{YY} , de dimension $(mnM-n) \times (mnM-n)$, définies comme suit : \underline{R}_{YY} (respectivement \overline{R}_{YY}) est une sous-matrice de R_{YY} de laquelle les n dernières (respectivement les n premières) lignes et les n dernières (respectivement les n premières) colonnes ont été omises (voir figure 7.1). La matrice \underline{R}_{YY} peut être écrite comme

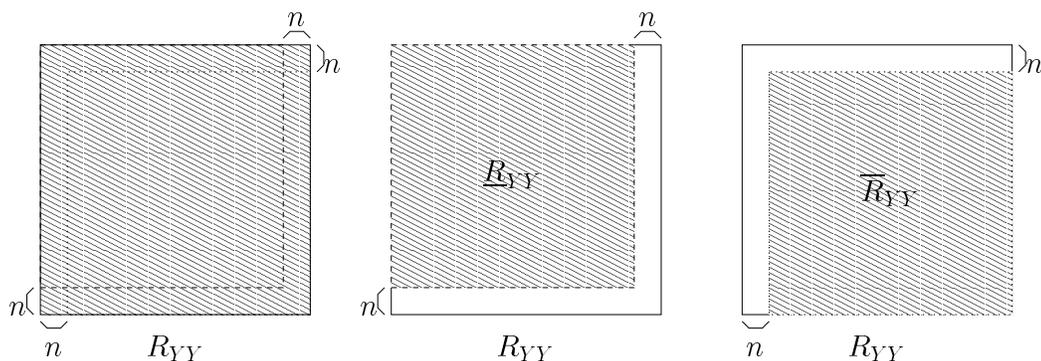


FIG. 7.1 – Les matrices R_{YY} , \underline{R}_{YY} et \overline{R}_{YY} .

$$\underline{R}_{YY} = \underline{\mathcal{T}}(\mathbf{h}) R_{AA} \underline{\mathcal{T}}^H(\mathbf{h}) + \underline{R}_{VV}, \quad (7.5)$$

où $\underline{\mathcal{T}}(\mathbf{h})$ correspond à la matrice $\mathcal{T}(\mathbf{h})$ de laquelle nous avons omis les n dernières lignes. De façon similaire, en considérant $\overline{\mathcal{T}}(\mathbf{h})$ comme la matrice $\mathcal{T}(\mathbf{h})$ de laquelle nous avons omis les n premières lignes, la matrice \overline{R}_{YY} peut être écrite comme

$$\overline{R}_{YY} = \overline{\mathcal{T}}(\mathbf{h}) R_{AA} \overline{\mathcal{T}}^H(\mathbf{h}) + \overline{R}_{VV}. \quad (7.6)$$

Par conséquent, la structure de déplacement $\underline{R}_{YY} - \overline{R}_{YY}$ est donnée par

$$\underline{R}_{YY} - \overline{R}_{YY} = \underline{\mathcal{T}}(\mathbf{h}) R_{AA} \underline{\mathcal{T}}^H(\mathbf{h}) - \overline{\mathcal{T}}(\mathbf{h}) R_{AA} \overline{\mathcal{T}}^H(\mathbf{h}), \quad (7.7)$$

qui est bien paramétrisée linéairement par le canal \mathbf{h} .

7.2.2 Approche d'Ajustement de Sous-espace

Considérons la décomposition propre de la matrice $\underline{R}_{YY} - \overline{R}_{YY}$ dans laquelle les valeurs propres sont ordonnées dans un ordre décroissant :

$$\underline{R}_{YY} - \overline{R}_{YY} = \sum_{i=1}^{mnM-n} \lambda_i V_i V_i^H = \mathcal{V}_+ \Lambda_+ \mathcal{V}_+^H + \mathcal{V}_- \Lambda_- \mathcal{V}_-^H, \quad (7.8)$$

où Λ_+ (respectivement Λ_-) désigne l'ensemble des valeurs propres positives (respectivement négatives) de la matrice $\underline{R}_{YY} - \overline{R}_{YY}$, et \mathcal{V}_+ (respectivement \mathcal{V}_-) leurs vecteurs propres correspondants. Dans ce qui suit, nous supposons que R_{AA} est de rang plein et que le nombre de canaux est au moins égal à 3 : $mn \geq 3$. Dans ce cas, la matrice $\underline{R}_{YY} - \overline{R}_{YY}$ possède un sous-espace bruit pour M suffisamment grand, la dimension du sous-espace signal positif (nombre de colonnes de \mathcal{V}_+) est égale au rang colonne plein de $\underline{\mathcal{I}}(\mathbf{h})$. Cependant, la dimension du sous-espace signal négatif (nombre de colonnes de \mathcal{V}_-) est égale au rang colonne plein de $\overline{\mathcal{T}}(\mathbf{h})$. Les deux sous-espaces signal (positif et négatif) possèdent la même dimension. L'espace colonne de \mathcal{V}_+ peut être interprété comme les composantes de l'espace colonne de $\underline{\mathcal{I}}(\mathbf{h})$ qui sont orthogonales à \mathcal{V}_- (l'interprétation est similaire dans le cas de \mathcal{V}_- et $\overline{\mathcal{T}}(\mathbf{h})$). Par conséquent, nous avons les équivalences suivantes :

$$\begin{aligned} \text{Image} \{ \mathcal{V}_+ \} &= \text{Image} \{ P_{\mathcal{V}_-}^\perp \underline{\mathcal{I}}(\mathbf{h}) \}, \\ \text{Image} \{ \mathcal{V}_- \} &= \text{Image} \{ P_{\mathcal{V}_+}^\perp \overline{\mathcal{T}}(\mathbf{h}) \}. \end{aligned} \quad (7.9)$$

Un critère naturel d'ajustement de sous-espace signal peut être formulé comme suit :

$$\min_{\mathbf{h}, T_1, T_2} \| P_{\mathcal{V}_-}^\perp \underline{\mathcal{I}}(\mathbf{h}) - \mathcal{V}_+ T_1 \|_F^2 + \| P_{\mathcal{V}_+}^\perp \overline{\mathcal{T}}(\mathbf{h}) - \mathcal{V}_- T_2 \|_F^2. \quad (7.10)$$

Après optimisation par rapport à T_1, T_2 , le problème d'optimisation (7.10) devient

$$\min_{\|\mathbf{h}\|=1} \text{tr} \{ \underline{\mathcal{I}}^H(\mathbf{h}) P_{\mathcal{V}_-}^\perp P_{\mathcal{V}_+}^\perp P_{\mathcal{V}_-}^\perp \underline{\mathcal{I}}(\mathbf{h}) \} + \text{tr} \{ \overline{\mathcal{T}}^H(\mathbf{h}) P_{\mathcal{V}_+}^\perp P_{\mathcal{V}_-}^\perp P_{\mathcal{V}_+}^\perp \overline{\mathcal{T}}(\mathbf{h}) \}, \quad (7.11)$$

laquelle expression peut être simplifiée à

$$\min_{\|\mathbf{h}\|=1} \text{tr} \{ \underline{\mathcal{I}}^H(\mathbf{h}) P_{\mathcal{V}_N} \underline{\mathcal{I}}(\mathbf{h}) \} + \text{tr} \{ \overline{\mathcal{T}}^H(\mathbf{h}) P_{\mathcal{V}_N} \overline{\mathcal{T}}(\mathbf{h}) \} = \min_{\|\mathbf{h}\|=1} \mathbf{h}^H (A_1 + A_2) \mathbf{h}, \quad (7.12)$$

où \mathcal{V}_N représente les vecteurs propres correspondant au sous-espace bruit ($P_{\mathcal{V}_N} = I - P_{\mathcal{V}_+} - P_{\mathcal{V}_-}$), et les matrices A_1, A_2 peuvent être déterminées à partir de $P_{\mathcal{V}_N}, \underline{\mathcal{I}}(\mathbf{h})$ et $\overline{\mathcal{T}}(\mathbf{h})$. Sous la contrainte $\|\mathbf{h}\| = 1$, la solution \mathbf{h} est donc le vecteur propre correspondant à la valeur propre minimale de la matrice $A_1 + A_2$.

7.3 Bornes de Cramer-Rao

Dans ce qui suit, nous discutons les bornes de Cramer-Rao (gaussienne et déterministe) correspondant à ce cas spécifique où le bruit additif est supposé être uniquement stationnaire. La définition des bornes de Cramer-Rao est donnée dans la section

§3.3. Nous rappelons ici l'expression de la matrice d'information de Fisher réelle ainsi que son expression complexe. Soit θ le vecteur des paramètres complexes à estimer et $\theta_R = \left[\operatorname{Re}(\theta)^H \quad \operatorname{Im}(\theta)^H \right]^H$ ses paramètres réels associés. La matrice d'information de Fisher associée à θ_R est :

$$J_{\theta_R \theta_R} = E_{\mathbf{Y}|\theta_R} \left(\frac{\partial \ln f(\mathbf{Y}|\theta_R)}{\partial \theta_R} \right) \left(\frac{\partial \ln f(\mathbf{Y}|\theta_R)}{\partial \theta_R} \right)^T. \quad (7.13)$$

Pour les paramètres complexes φ, ψ (qui sont des parties de θ), la matrice d'information de Fisher $J_{\varphi\psi}$ est définie comme

$$J_{\varphi\psi} = E_{\mathbf{Y}|\theta} \left(\frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \varphi^*} \right) \left(\frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \psi^*} \right)^H, \quad (7.14)$$

où l'opération de dérivation complexe par rapport à $\theta = \alpha + j\beta$, est définie comme $\frac{\partial}{\partial \theta} = \frac{1}{2} \left(\frac{\partial}{\partial \alpha} - j \frac{\partial}{\partial \beta} \right)$. Le vecteur des paramètres complexes étendu est défini comme $\theta_C = [\theta^T \quad \theta^H]^T$. La matrice $J_{\theta_C \theta_C}$ contient *la même information* que la matrice $J_{\theta_R \theta_R}$. La quantité qui nous intéresse, généralement, est l'erreur quadratique moyenne (MSE) définie comme

$$\text{MSE} = E\|\tilde{\theta}\|^2 = E\|\tilde{\theta}_R\|^2 = \frac{1}{2} E\|\tilde{\theta}_C\|^2 \geq \frac{1}{2} \operatorname{tr}\{J_{\theta_C \theta_C}^{-1}\} \geq \operatorname{tr}\{J_{\theta\theta}^{-1}\} \quad (7.15)$$

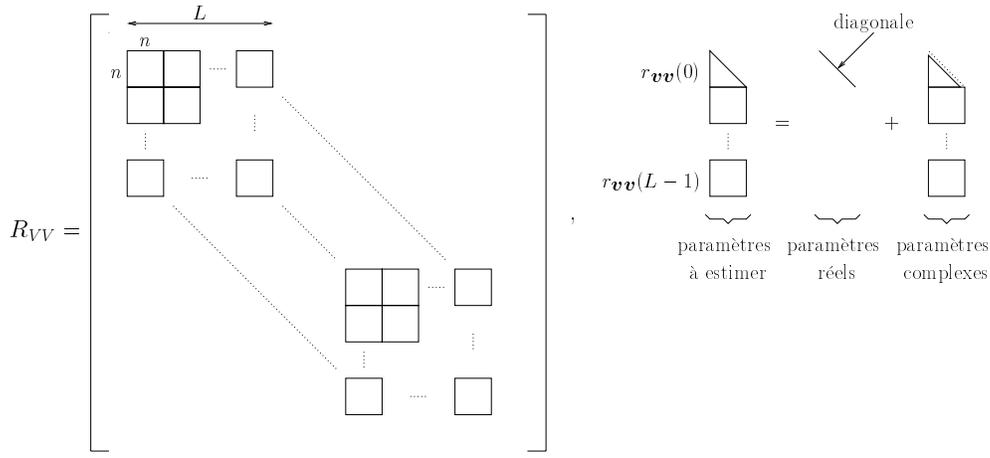
où $\tilde{\theta} = \theta - \hat{\theta}$.

7.3.1 Borne de Cramer-Rao Gaussienne

Dans le cas où les symboles sont gaussiens et circulaires, nous avons $\mathbf{Y} \sim \mathcal{N}(0, R_{YY})$ avec $R_{YY} = \sigma_a^2 \mathcal{T}(\mathbf{h}) \mathcal{T}^H(\mathbf{h}) + R_{VV}$ (donc $R_{AA} = \sigma_a^2 I$). La fonction de coût de maximum de vraisemblance à minimiser est

$$\mathcal{L}(\mathbf{h}, R_{VV}) = c^t + \ln \det R_{YY} + \mathbf{Y}^H R_{YY}^{-1} \mathbf{Y}. \quad (7.16)$$

Nous avons à estimer, conjointement, les coefficients du canal et les coefficients de covariance du bruit. Comme la matrice de covariance du bruit coloré, R_{VV} , est bloc Toeplitz dont les blocs sont de dimension $n \times n$. Les paramètres de la matrice R_{VV} à estimer sont : les éléments de la partie triangulaire supérieure ou inférieure de $r_{\mathbf{v}\mathbf{v}}(0)$ et les éléments des matrices $r_{\mathbf{v}\mathbf{v}}(1) \cdots r_{\mathbf{v}\mathbf{v}}(L-1)$ (chaque matrice est de dimension $n \times n$), où L désigne la longueur du filtre FIR utilisé pour générer le bruit coloré MA. On peut remarquer que les éléments diagonaux de la matrice $r_{\mathbf{v}\mathbf{v}}(0)$ sont réels et par conséquent on est en présence d'un mélange de paramètres réels et complexes. Nous considérons le vecteur des paramètres suivant : $\theta = [\theta_1^H \quad \theta_2^H]^H$, dans lequel θ_1 désigne le vecteur obtenu en empilant les éléments diagonaux de $r_{\mathbf{v}\mathbf{v}}(0)$ et θ_2 désigne le vecteur obtenu en concaténant les coefficients du canal \mathbf{h} et le vecteur formé en empilant les colonnes de la matrice $\left[\underline{r}_{\mathbf{v}\mathbf{v}}^H(0) \quad r_{\mathbf{v}\mathbf{v}}^H(1) \quad \cdots \quad r_{\mathbf{v}\mathbf{v}}^H(L-1) \right]^H$, où $\underline{r}_{\mathbf{v}\mathbf{v}}(0)$ désigne la partie triangulaire


 FIG. 7.2 – Paramètres de R_{VV} à estimer.

inférieure stricte de $r_{\mathbf{v}\mathbf{v}}(0)$ (ces paramètres de R_{VV} à estimer sont illustrés dans la figure 7.2).

Soit $\theta_C = [\theta_1^T \ \theta_2^T \ \theta_2^H]^T$ le vecteur des paramètres complexes étendu. Nous obtenons comme matrice d'information de Fisher associée à θ_C

$$J_{\theta_C \theta_C} = \begin{bmatrix} J_{\theta_1 \theta_1} & J_{\theta_1 \theta_2} & J_{\theta_1 \theta_2^*} \\ J_{\theta_2 \theta_1} & J_{\theta_2 \theta_2} & J_{\theta_2 \theta_2^*} \\ J_{\theta_2^* \theta_1} & J_{\theta_2^* \theta_2} & J_{\theta_2^* \theta_2^*} \end{bmatrix} \quad (7.17)$$

où la matrice $J_{\varphi\psi}$ est donnée par (7.14) et

$$J_{\varphi^* \psi^*} = J_{\varphi\psi}^*. \quad (7.18)$$

En exploitant cette propriété et la symétrie hermitienne de $J_{\theta_C \theta_C}$, nous avons à calculer uniquement les matrices $J_{\theta_1 \theta_1}$, $J_{\theta_1 \theta_2}$, $J_{\theta_2 \theta_2}$ et $J_{\theta_2 \theta_2^*}$. La quantité qui nous intéresse est l'erreur quadratique moyenne (MSE) sur les estimées du canal $\hat{\mathbf{h}}$. Soit \mathcal{P} une matrice de permutation tel que $\mathcal{P}\theta_C = [\mathbf{h}_C^T \ \eta_C^T]^T$ où η_C représente les paramètres qui ne nous intéressent pas dans l'opération d'estimation. Nous avons

$$\mathcal{P} J_{\theta_C \theta_C} \mathcal{P}^H = \begin{bmatrix} J_{\mathbf{h}_C \mathbf{h}_C} & J_{\mathbf{h}_C \eta_C} \\ J_{\eta_C \mathbf{h}_C} & J_{\eta_C \eta_C} \end{bmatrix}. \quad (7.19)$$

Par conséquent, nous obtenons la borne de Cramer-Rao suivante

$$E\|\tilde{\mathbf{h}}\|^2 = \frac{1}{2} E\|\tilde{\mathbf{h}}_C\|^2 \geq \frac{1}{2} \text{tr} \left\{ \left(J_{\mathbf{h}_C \mathbf{h}_C} - J_{\mathbf{h}_C \eta_C} J_{\eta_C \eta_C}^{-1} J_{\eta_C \mathbf{h}_C} \right)^+ \right\}. \quad (7.20)$$

7.3.2 Borne de Cramer-Rao Déterministe

Dans le modèle déterministe, le canal \mathbf{h} et les symboles A sont considérés comme des quantités déterministes. Le vecteur des paramètres complexe θ est $\theta = \begin{bmatrix} \mathbf{r}^H & \mathbf{h}^H & A^H \end{bmatrix}^H$.

Le vecteur A contient les symboles d'entrée, \mathbf{h} est le canal et \mathbf{r} désigne le vecteur contenant les coefficients de covariance du bruit coloré à estimer. La fonction de densité de probabilité est donnée par :

$$f(\mathbf{Y}|\theta) = \frac{1}{\pi^{mM} \det R_{VV}} e^{-[\mathbf{Y} - \mathbf{Y}^{(S)}]^H R_{VV}^{-1} [\mathbf{Y} - \mathbf{Y}^{(S)}]} \quad (7.21)$$

où $\mathbf{Y}^{(S)} = \mathcal{T}(\mathbf{h})A$ désigne la partie signal de \mathbf{Y} . La fonction de coût de maximum de vraisemblance à minimiser est

$$\mathcal{L}(\theta) = c^t + \ln \det R_{VV} + (\mathbf{Y} - \mathcal{T}(\mathbf{h})A)^H R_{VV}^{-1} (\mathbf{Y} - \mathcal{T}(\mathbf{h})A). \quad (7.22)$$

En suivant le même raisonnement que dans le cas gaussien, nous pouvons écrire le vecteur θ comme $\theta = \begin{bmatrix} \theta_1^H & \theta_2^H \end{bmatrix}^H$, dans lequel θ_1 désigne le vecteur obtenu en empilant les éléments diagonaux de la matrice $r_{\mathbf{v}\mathbf{v}}(0)$ et θ_2 désigne le vecteur formé en concaténant les symboles A , les coefficients du canal \mathbf{h} et les corrélations complexes du bruit coloré. Nous considérons, de nouveau, la matrice d'information de Fisher associée au vecteur des paramètres complexes étendu θ_C . La matrice $J_{\theta_C \theta_C}$ est de nouveau donnée par l'expression (7.17), où les différentes matrices $J_{\varphi\psi}$ sont calculées en dérivant la fonction de coût $\mathcal{L}(\theta)$ définie dans (7.22). De nouveau, la quantité qui nous intéresse est l'erreur MSE sur les estimées du canal.

7.4 Connaissance du Filtre de Transmission (TX)

Comme l'élimination du bruit stationnaire dans l'approche d'ajustement de sous-espace, que nous avons décrite précédemment, est basée sur le suréchantillonnage et par conséquent sur l'excès en largeur de bande, l'utilisation de la connaissance a priori sur le filtre de transmission (qui limite l'excès en largeur de bande) devrait être utile. Nous avons décrit le formalisme de l'exploitation de la connaissance des filtres de TX/RX, dans le cas du suréchantillonnage et/ou d'un réseau d'antennes dans la section §5.2. Dans ce qui suit, nous utilisons ce formalisme pour exploiter la connaissance à priori sur le filtre de transmission dans la méthode d'ajustement de sous-espace que nous avons introduite.

7.4.1 Méthode avec Connaissance du Filtre de TX (WTXFK)

Nous considérons un facteur de suréchantillonnage m et un réseau de n antennes : les signaux reçus aux n différentes antennes sont donc suréchantillonnés avec un facteur m . Nous supposons que le nombre de canaux obtenus est au moins égal à trois : $mn \geq 3$, dans ce cas la matrice $\underline{R}_{YY} - \overline{R}_{YY}$ possède un sous-espace bruit pour M suffisamment grand. Puisque la méthode d'ajustement de sous-espace se ramène au problème de minimisation exprimé dans (7.12) : $\min_{\|\mathbf{h}\|=1} \mathbf{h}^H (A_1 + A_2) \mathbf{h}$, l'exploitation de la connaissance du filtre de transmission aboutit au problème d'optimisation suivant

$$\min_{\mathbf{c}} \mathbf{c}^H \mathbf{G}^H (A_1 + A_2) \mathbf{G} \mathbf{c}, \quad (7.23)$$

où $\mathbf{G} = \mathcal{T}_L^T([I_n \otimes \underline{\mathbf{g}}^T(K-1) \cdots I_n \otimes \underline{\mathbf{g}}^T(0)])$ et les matrices $\underline{\mathbf{g}}(i), i = 0 \cdots K-1$ sont définies dans (5.8). Sous la contrainte $\|\mathbf{c}\| = 1$, l'estimée du canal de propagation, $\hat{\mathbf{c}}$, est trouvée comme le vecteur propre associé à la valeur propre minimale de la matrice $\mathbf{G}^H(A_1 + A_2)\mathbf{G}$ et l'estimée du canal global est $\hat{\mathbf{h}} = \mathbf{G}\hat{\mathbf{c}}$. La méthode obtenue est donc une méthode d'identification du canal avec connaissance du filtre de TX (WTFXK).

7.4.2 Bornes de Cramer-Rao WTFXK

Avec l'exploitation de la connaissance du filtre de TX, nous obtenons $\hat{\mathbf{h}} = \mathbf{G}\hat{\mathbf{c}}$ à partir de $\hat{\mathbf{c}}$. La matrice d'information de Fisher pour \mathbf{c} peut être obtenue à partir de la matrice d'information de Fisher pour \mathbf{h} non structuré (que nous avons trouvée précédemment). A partir de $\mathbf{h} = \mathbf{G}\mathbf{c}$, nous obtenons $\mathbf{h}_C = \mathbf{G}_C\mathbf{c}_C$ où $\mathbf{G}_C = \text{blockdiag}\{\mathbf{G}, \mathbf{G}^*\}$. Donc

$$\frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \mathbf{c}_C^*} = \frac{\partial \mathbf{h}_C^H}{\partial \mathbf{c}_C^*} \frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \mathbf{h}_C^*} = \mathbf{G}_C^H \frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \mathbf{h}_C^*}, \quad (7.24)$$

ce qui implique que

$$\mathbf{J}_{\mathbf{c}_C\varphi} = \mathbf{G}_C^H \mathbf{J}_{\mathbf{h}_C\varphi}. \quad (7.25)$$

Par conséquent, la matrice d'information de Fisher $\mathbf{J}_{\mathbf{c}_C\mathbf{c}_C}$ pour \mathbf{c}_C , en éliminant les paramètres η_C , devient

$$\mathbf{J}_{\mathbf{c}_C\mathbf{c}_C} = \mathbf{G}_C^H \left(\mathbf{J}_{\mathbf{h}_C\mathbf{h}_C} - \mathbf{J}_{\mathbf{h}_C\eta_C} \mathbf{J}_{\eta_C\eta_C}^{-1} \mathbf{J}_{\eta_C\mathbf{h}_C} \right) \mathbf{G}_C \quad (7.26)$$

où la matrice du milieu est celle qui apparaît dans (7.20). Maintenant, \mathbf{c} est en fait une quantité intermédiaire dans l'estimation de \mathbf{h} . Pour trouver la matrice d'information de Fisher pour \mathbf{c}_C , nous obtenons à partir de $\mathbf{h}_C = \mathbf{G}_C\mathbf{c}_C$ l'expression $\mathbf{c}_C = \mathbf{G}_C^+ \mathbf{h}_C$. Donc, nous avons

$$\mathbf{J}_{\mathbf{h}_C\varphi} = \mathbf{G}_C^{+H} \mathbf{J}_{\mathbf{c}_C\varphi}. \quad (7.27)$$

Puisque $\mathbf{G}_C \mathbf{G}_C^+ = P_{\mathbf{G}_C}$, nous obtenons l'expression de la borne de Cramer-Rao suivante

$$E\|\tilde{\mathbf{h}}\|^2 \geq \frac{1}{2} \text{tr} \left\{ \left[P_{\mathbf{G}_C} \left(\mathbf{J}_{\mathbf{h}_C\mathbf{h}_C} - \mathbf{J}_{\mathbf{h}_C\eta_C} \mathbf{J}_{\eta_C\eta_C}^{-1} \mathbf{J}_{\eta_C\mathbf{h}_C} \right) P_{\mathbf{G}_C} \right]^+ \right\}. \quad (7.28)$$

Pour le cas de l'exploitation de la connaissance du filtre de TX. Cette borne doit être comparée à celle donnée dans (7.20) et qui correspond au cas où \mathbf{h} est non structuré. La pseudo-inverse dans l'expression (7.28) indique que le facteur d'ambiguïté est fixé au niveau de $\hat{\mathbf{h}}$.

7.5 Simulations

Dans cette section, nous illustrons la performance de la méthode d'ajustement de sous-espace. Cette performance est évaluée par l'intermédiaire du comportement de l'erreur d'estimation du canal ainsi que sa comparaison à la borne de Cramer-Rao correspondante. Puis, nous nous intéressons à l'apport, en termes de performance, de la méthode exploitant la connaissance du filtre de TX par rapport à sa version purement aveugle.

7.5.1 Méthode d'Ajustement de Sous-espace

Nous considérons un paquet de données de longueur 100 périodes symboles, un canal \mathbf{H} généré de façon aléatoire, de longueur $N = 4$. Le nombre d'antennes est $n = 2$ et le facteur de suréchantillonnage est $m = 2$. Les symboles d'entrées sont i.i.d. BPSK. Le bruit coloré est un processus MA généré par filtrage d'un bruit vectoriel de dimension $n \times 1$, complexe, blanc et gaussien. Le filtre utilisé pour la génération de ce bruit est un filtre \mathbf{f} FIR de dimension $n \times n$ et dont la longueur est égale à 3. Le rapport signal à bruit est défini comme suit:

$$\text{SNR} = \frac{(\|\mathbf{h}\|^2/mn)\sigma_a^2}{(\|\mathbf{f}\|^2/n)\sigma_v^2} = \frac{\|\mathbf{h}\|^2\sigma_a^2}{m\|\mathbf{f}\|^2\sigma_v^2}. \quad (7.29)$$

Nous utilisons une matrice de covariance estimée \hat{R}_{YY} de taille $M = 20$. Les méthodes d'estimation aveugle aboutissent à des estimées du canal $\hat{\mathbf{h}}$ avec $\|\hat{\mathbf{h}}\| = 1$, nous ajustons le facteur scalaire α tel que $\mathbf{h}_o^H(\alpha\hat{\mathbf{h}}) = \mathbf{h}_o^H\mathbf{h}_o$ où \mathbf{h}_o est le vrai canal (voir [dCS97b]): l'estimée finale est $\hat{\mathbf{h}} = \alpha\hat{\mathbf{h}}$. Nous considérons, comme mesure de performance, l'erreur quadratique moyenne normalisée (NMSE) (en prenant la moyenne sur 100 réalisations Monte-Carlo) définie comme

$$\text{NMSE} = \text{E}\|\mathbf{h} - \hat{\mathbf{h}}\|^2/\|\mathbf{h}\|^2. \quad (7.30)$$

Dans la figure 7.3, nous avons tracé l'erreur d'estimation sur le canal (obtenue par la méthode d'ajustement de sous-espace) en fonction du SNR: la décroissance de la courbe au fur et à mesure que le SNR augmente montre bien que la méthode proposée (SSF) fonctionne. Sur la même figure nous avons tracé la courbe correspondant à la borne de Cramer-Rao (CRB) gaussienne normalisée, calculée comme: $\text{tr}\{CRB_{\mathbf{h}}\}/\|\mathbf{h}\|^2$. Bien que la méthode soit déterministe, et puisque la borne de Cramer-Rao Gaussienne est plus basse que la borne de Cramer-Rao déterministe, nous pouvons constater que la courbe du NMSE n'est pas proche de la borne de Cramer-Rao gaussienne (pour un SNR ≥ 20 dB, la courbe devient étroitement proche de la borne de Cramer-Rao). Ceci implique que, potentiellement, un gain de performance considérable peut être obtenu en exploitant l'information de second ordre des symboles transmis.

7.5.2 Méthode d'Ajustement de Sous-espace WTXFK

Dans la figure 7.5, nous illustrons l'exploitation de la connaissance a priori du filtre de TX. Nous mesurons l'erreur NMSE obtenue avec la méthode d'ajustement de sous-espace WTXFK en fonction du SNR. Nous considérons un paquet de données de longueur 200 périodes symboles. Le signal reçu sur une antenne ($n = 1$) est suréchantillonné avec un facteur $m = 3$. Le canal de propagation est de longueur 3 et le filtre de TX est un filtre GMSK linéarisé tronqué sur 4 périodes symboles (voir figure 7.4). Les résultats de simulation montrent que la méthode d'ajustement de sous-espace purement aveugle ne permet pas d'identifier ce canal: vu l'excès en largeur de bande limité, ce canal possède

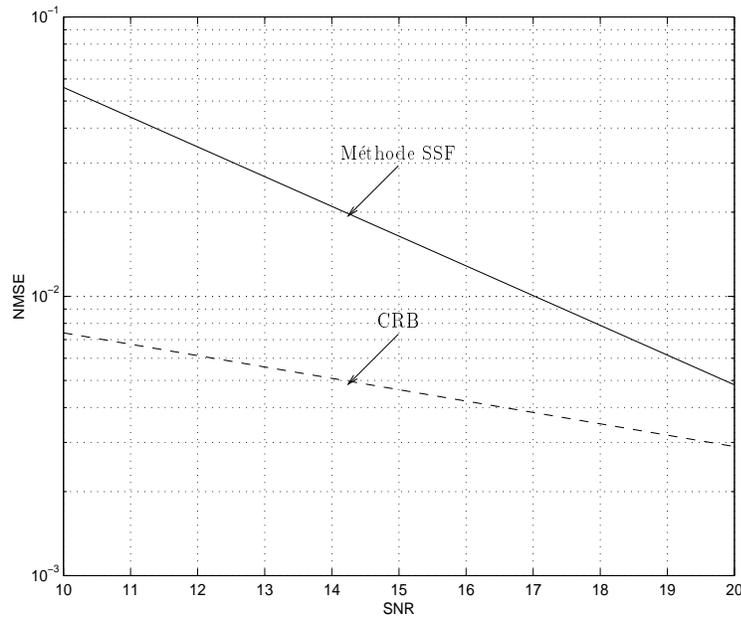
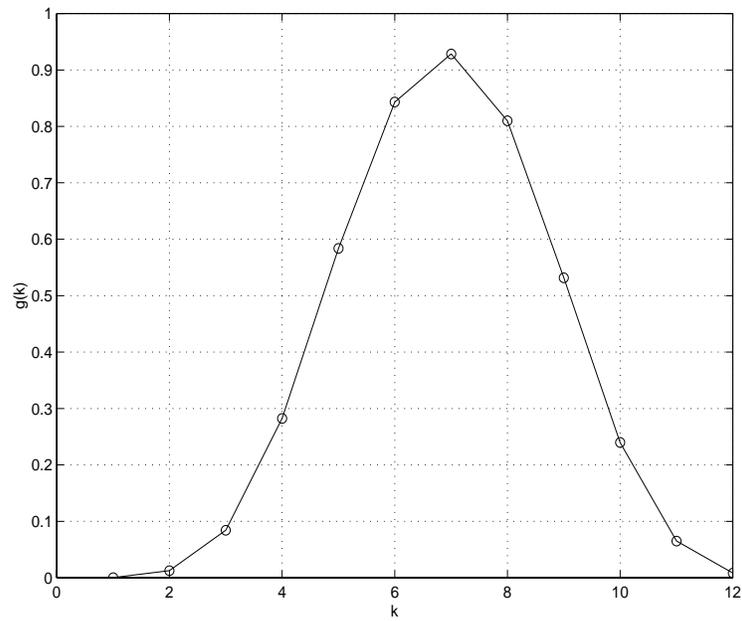
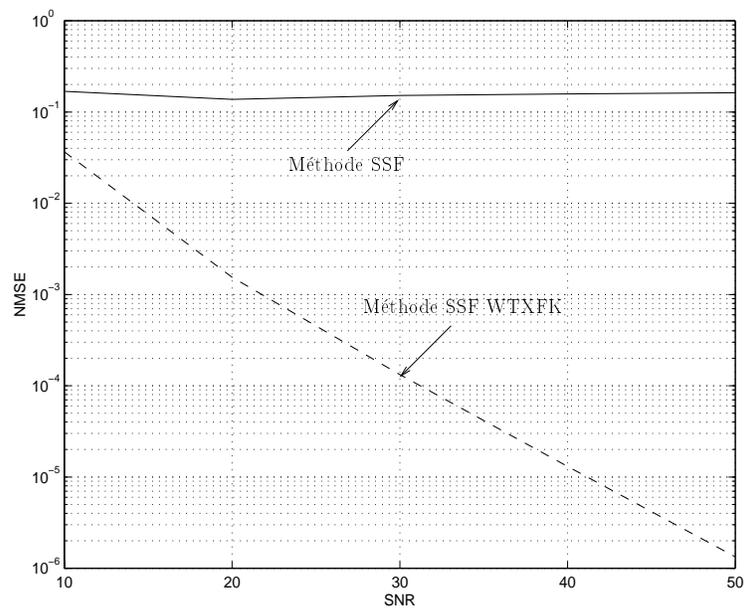


FIG. 7.3 – Performance de la méthode d'ajustement de sous-espace.

des zéros qui sont pratiquement en commun. Cependant, la méthode d'ajustement de sous-espace WTXFK (SSF WTXFK) permet d'identifier le canal et aboutit à des erreurs d'estimation du canal qui décroissent au fur et à mesure que le SNR augmente. Ceci confirme les conclusions, déjà établies dans la section §5.4, concernant la capacité des méthodes exploitant la connaissance a priori des filtres de TX/RX d'identifier des canaux qui ne peuvent pas être identifiés en utilisant des méthodes d'estimation purement aveugles.

7.6 Conclusions

Dans ce chapitre, nous avons traité l'estimation aveugle multi-canal dans le cas où le bruit additif est supposé être uniquement stationnaire. Nous avons considéré la situation dans laquelle l'aspect multi-canal provient à la fois du suréchantillonnage et d'un réseau d'antennes. En se basant sur une structure de déplacement appropriée de la matrice de covariance du signal reçu, nous avons dérivé une technique d'ajustement de sous-espace permettant d'identifier le canal. L'utilisation de la connaissance a priori sur le filtre de TX permet de rendre la méthode proposée plus robuste (en termes d'identifiabilité et de performance). Nous avons aussi calculé les bornes de Cramer-Rao (gaussienne et déterministe) correspondant à ce cas spécifique du bruit coloré. Les résultats des simulations illustrent le bon fonctionnement de la méthode d'ajustement de sous-espace proposée et montrent que l'exploitation de la connaissance a priori du filtre de TX est d'un intérêt capital dans ce cas.

FIG. 7.4 – *Filtre GMSK linéarisé.*FIG. 7.5 – *Performance de la méthode d'ajustement de sous-espace WTXFK.*

Chapitre 8

Estimation Multi-canal Aveugle par Egalisation MMSE ZF

Dans ce chapitre nous décrivons une nouvelle méthode d'estimation multi-canal que nous avons présentée dans [AS99c]. Notre approche est basée sur l'égalisation aveugle MMSE ZF (Minimum Mean Square Error Zero Forcing) et elle est liée à la méthode proposée par Tsatsanis *et al.* [TX97] qui correspond à une égalisation MMSE non biaisée. Nous interprétons cette approche en termes de prédiction linéaire bidirectionnelle, appelée "smoothing" par Tong [TZ98]. Nous établissons les liens existants entre les approches MMSE, MOE (Minimum Output Energy) et MMSE ZF, et nous prouvons l'équivalence entre elles sous la contrainte d'absence du biais dans le cas sans bruit. Notre analyse indique comment appliquer de façon correcte le principe de Capon [Cap69] pour la formation de voies LCMV (Linearly Constrained Minimum Variance) à l'égalisation multi-canal. Nous démontrons aussi que l'application du principe de Capon dans la méthode de Tsatsanis et l'estimation de canal dans l'approche de Tong deviennent non biaisées uniquement à des hauts SNRs. Alors que le but est de faire MMSE ZF, il est plus facile mais équivalent d'aborder le problème par UMMSE (Unbiased MMSE) appliqué sur la composante signal du signal reçu. La matrice de covariance du signal reçu doit donc être "débruitée" avant son utilisation dans la méthode d'estimation aveugle. Nous proposons une approche sans décomposition propre de la matrice de covariance qui présente des performances excellentes. En plus, nous présentons une manière simple et efficace pour incorporer la détection de l'ordre du canal dans la méthode proposée [AS99b],[AS99a]. Les résultats de simulation sont présentés pour démontrer les performances des approches discutées.

8.1 Introduction

En absence de séquence d'apprentissage, l'égalisation de canal peut être réalisée en deux étapes : dans la première étape, on estime le canal (en utilisant une des méthodes aveugles que nous avons décrites dans les chapitres précédents) et dans une seconde

étape, on construit l'égaliseur en se basant sur l'estimée du canal obtenue. D'autres approches essaient de concevoir directement l'égaliseur. Dans ce contexte, une méthode basée sur la prédiction linéaire multi-canal a été proposée dans [AMDG⁺95]. Tsatsanis *et al.* (voir [TX97] et [TX99]) ont proposé de s'inspirer du principe de Capon [Cap69] pour la formation de voies LCMV dans le but de l'appliquer au problème de l'estimation aveugle pour dériver directement des égaliseurs linéaires multi-canal. Bien qu'elle soit proposée pour dériver des égaliseurs multi-canal, la méthode de Tsatsanis *et al.* peut être vue comme étant une technique d'estimation du canal. Un travail parallèle mais équivalent est celui effectué par Tong *et al.* dans [TZ98] pour le cas mono-utilisateur et dans [ZT98] pour le cas multi-utilisateur. Dans [TZ98], Tong *et al.* dérivent une méthode d'estimation du canal appelée "smoothing" où ils se basent sur un point de vue déterministe. La technique d'égalisation proposée par Zhu *et al.* [ZDC99], lorsqu'elle est considérée dans le cas mono-utilisateur (contexte mono-entrée/multi-sortie), devient aussi équivalente aux deux approches précédentes.

Ce chapitre est organisé comme suit : dans la première section nous expliquons le principe de Capon dans le cas de l'estimation des directions d'arrivées puis nous déduisons les similitudes qui existent entre ce contexte et celui de l'estimation multi-canal. Dans la deuxième section, nous établissons les relations entre les critères MSE et MOE et nous déduisons, en conséquence, les équivalences entre les différents égaliseurs dérivés. La troisième section traite l'égalisation aveugle MMSE ZF, nous nous intéressons tout d'abord au problème de minimisation en soi, puis nous interprétons l'algorithme proposé en termes de prédiction linéaire bidirectionnelle, ensuite nous analysons la méthode dans le cas sans bruit et nous mettons en lumière la relation de notre approche avec d'autres approches proposés et dans la dernière partie de cette section nous expliquons le mécanisme de détection de l'ordre du canal que nous proposons. Dans la section suivante, nous présentons les résultats de simulations pour illustrer les performances des approches discutées.

Nous conservons les différentes entités définies dans (1.6) et (1.9) sauf pour le vecteur des coefficients du canal, qui sera défini comme

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}(0) \\ \vdots \\ \mathbf{h}(N-1) \end{bmatrix}. \quad (8.1)$$

8.2 Méthode de Capon

Un principe bien connu en traitement d'antennes lorsque la direction d'arrivée et la réponse du réseau d'antennes sont connues, est le formateur de voies MVDR (Minimum Variance Distortionless Response). Ce dernier minimise les signaux des interféreurs sans apporter de distortion au signal de l'utilisateur d'intérêt. La méthode de Capon [Cap69] part du principe MVDR pour dériver des solutions aveugles (sans avoir à connaître la direction).

8.2.1 Estimation des Directions d'Arrivées

Considérons tout d'abord le cas de l'estimation des directions d'arrivées. Dans le cas d'une seule source dans la direction de θ_0 plus un bruit blanc spatialement, le signal reçu peut être écrit comme

$$\mathbf{Y}(k) = S(\theta_0)a(k) + \mathbf{V}(k). \quad (8.2)$$

La sortie du formateur de voies est donnée par $\hat{a}(k) = F^H \mathbf{Y}(k)$. La contrainte de distorsion nulle du signal étant $F^H S(\theta_0) = 1$, le formateur de voies LCMV satisfait le critère de minimisation suivant :

$$\begin{aligned} \min_{F: F^H S(\theta_0)=1} E \|\hat{a}(k)\|^2 &\Leftrightarrow \min_{F: F^H S(\theta_0)=1} F^H (E \mathbf{Y}(k) \mathbf{Y}(k)^H) F \Leftrightarrow \min_{F: F^H S(\theta_0)=1} F^H R_{YY} F \\ &\Rightarrow F = \frac{1}{S^H(\theta_0) R_{YY}^{-1} S(\theta_0)} R_{YY}^{-1} S(\theta_0) \\ &\Rightarrow MV = (S^H(\theta_0) R_{YY}^{-1} S(\theta_0))^{-1} \end{aligned} \quad (8.3)$$

où MV désigne la variance minimale obtenue. La méthode de Capon stipule, si l'entité θ_0 est, maintenant, inconnue, de prendre alors $\hat{\theta} = \arg \max_{\theta} MV$, ce qui revient à sélectionner la direction d'arrivée ayant la variance de sortie maximale. La justification de ce choix provient du fait qu'on doit maximiser la composante signal à la sortie (après que les interférences ont été supprimées), donc nous avons

$$\begin{aligned} \hat{\theta} = \arg \max_{\theta} (S^H(\theta) R_{YY}^{-1} S(\theta))^{-1} &= \arg \min_{\theta} S^H(\theta) R_{YY}^{-1} S(\theta) \\ &= S^{-1}(V_{max}(R_{YY})) = \theta_0. \end{aligned} \quad (8.4)$$

8.2.2 Estimation Multi-canal

Dans le cas de l'estimation multi-canal, le formateur de voies devient un égaliseur qui restitue le signal transmis avec un délai éventuel de d échantillons : $\hat{a}(k-d) = F^H \mathbf{Y}(k)$ et le paramètre inconnu devient le canal. Pour appliquer le principe de Capon, la question essentielle est maintenant : quelle est la bonne contrainte de distorsion nulle ? Pour Tsatsanis [TX97], c'est l'absence du biais, ce qui est exprimé par

$$F^H \mathcal{T} e_d = F^H \begin{bmatrix} 0 \\ \mathbf{h} \\ 0 \end{bmatrix} \triangleq F^H \tilde{\mathbf{h}} = 1 \text{ et } e_d = \underbrace{[0 \cdots 0]_d}_{d} 10 \cdots 0]^T \quad (8.5)$$

où \mathcal{T} est la matrice $\mathcal{T}(\mathbf{h})$ utilisée dans l'expression (1.11). L'absence du biais correspond à une distorsion nulle pour le symbole que l'égaliseur estime. En d'autre terme, le symbole en question doit apparaître dans la sortie de l'égaliseur avec un coefficient égal à un. Cependant lorsqu'on applique le principe de Capon sous la contrainte d'absence du biais, les interférences entre symboles (IES) faussent l'estimation du canal en présence du bruit. La contrainte correcte de distorsion nulle à considérer, est en fait : *le forçage à zéro des IES*, donnée par

$$F^H \mathcal{T} = e_d^T. \quad (8.6)$$

Le forçage à zéro (ZF) implique, non seulement que le symbole à estimer apparaît à la sortie de l'égaliseur sans être distordu, mais aussi que ceci reste vrai pour toute la séquence de symboles transmise. Par conséquent, le formateur de voies LCMV à qui le principe de Capon peut être appliqué devient, dans ce cas, l'égaliseur MMSE ZF.

8.3 Relations entre Critères & Egaliseurs

Dans cette section, nous montrons d'abord l'équivalence entre les critères MMSE et MOE dans le cas de l'absence du biais. Puis, nous établissons l'égalité entre l'égaliseur MMSE et MOE sous la contrainte ZF. Nous décrivons, ensuite, la dérivation de l'expression de l'égaliseur MMSE non biaisé ainsi que celle de l'égaliseur MMSE ZF et nous prouvons l'égalité entre ces deux égaliseurs dans le cas sans bruit (ou dans le cas des statistiques débruitées).

8.3.1 Relations entre Critères

Le critère MMSE non biaisé est équivalent au critère MOE non biaisé. Ceci est démontré comme suit :

$$\begin{aligned}
 \text{MSE} &= \sigma_a^2 = E|a_{k-d} - \hat{a}|^2 = E|a_{k-d} - F^H \mathbf{Y}|^2 \\
 &= \sigma_a^2 - F^H R_{Y_a} - R_{a_Y} F + F^H R_{YY} F \\
 &= \sigma_a^2 - \sigma_a^2 F^H \tilde{\mathbf{h}} - \sigma_a^2 \tilde{\mathbf{h}}^H F + \underbrace{F^H R_{YY} F}_{\text{OE}} \\
 &= \text{OE} - \sigma_a^2 \text{ si non biaisé} \\
 \Rightarrow \arg \min_{F: F^H \tilde{\mathbf{h}}=1} \text{MSE} &= \arg \min_{F: F^H \tilde{\mathbf{h}}=1} \text{OE} \\
 \Rightarrow F_{\text{UMMSE}} &= F_{\text{UMOE}}.
 \end{aligned} \tag{8.7}$$

Puisque la contrainte ZF implique la contrainte d'absence de biais, la relation décrite par (8.7) implique l'égalité suivante :

$$F_{\text{MMSEZF}} = F_{\text{MOEZF}}. \tag{8.8}$$

Pour un récepteur linéaire F , le rapport signal à bruit plus interférences (SINR) à la sortie est défini comme

$$\text{SINR} = \frac{F^H R_{Y_S Y_S} F}{F^H R_{Y_{IN} Y_{IN}} F}, \tag{8.9}$$

où $\mathbf{Y}_S = \tilde{\mathbf{h}} a_{k-d}$ représente la partie signal de \mathbf{Y} et $\mathbf{Y}_{IN} = \mathbf{Y} - \mathbf{Y}_S$ représente la partie bruit plus interférence de \mathbf{Y} . Dans le cas où les symboles sont non corrélés, nous avons

$$\text{SINR} = \frac{\sigma_a^2 F^H \tilde{\mathbf{h}} \tilde{\mathbf{h}}^H F}{F^H (R_{YY} - \sigma_a^2 \tilde{\mathbf{h}} \tilde{\mathbf{h}}^H) F} \Rightarrow \text{SINR}^{-1} = \frac{\sigma_a^2 F^H \tilde{\mathbf{h}} \tilde{\mathbf{h}}^H F}{\sigma_a^2 |F^H \tilde{\mathbf{h}}|^2} - 1, \tag{8.10}$$

donc

$$\max_F \text{SINR} \Leftrightarrow \min_F \text{SINR}^{-1} \Leftrightarrow \min_F \frac{\sigma_a^2 F^H R_{YY} F}{\sigma_d^2 |F^H \tilde{\mathbf{h}}|^2} \Leftrightarrow \min_{F: F^H \tilde{\mathbf{h}}=1} F^H R_{YY} F. \quad (8.11)$$

D'où

$$F_{MSINR} = F_{UMOE} \text{ (Max SINR)}. \quad (8.12)$$

8.3.2 Egaliseurs

Considérons, dans ce qui suit, la dérivation de l'égaliseur F_{UMMSE} . Cet égaliseur est obtenu en résolvant le problème de minimisation, $\min_{F: F^H \mathcal{T} e_d=1} F^H R_{YY} F$, qui peut être écrit comme :

$$\min_{F: F^H R_{YY}^{\frac{1}{2}} R_{YY}^{-\frac{1}{2}} \mathcal{T} e_d=1} \left(F^H R_{YY}^{\frac{1}{2}} \right) \left(R_{YY}^{\frac{1}{2}} F \right). \quad (8.13)$$

Introduisons la matrice $D \triangleq R_{YY}^{\frac{1}{2}} F$. Le problème d'optimisation (8.13) devient

$$\min_{D: D^H R_{YY}^{-\frac{1}{2}} \mathcal{T} e_d=1} D^H D. \quad (8.14)$$

d'où

$$D = R_{YY}^{-\frac{1}{2}} \mathcal{T} e_d \left(e_d^H \mathcal{T}^H R_{YY}^{-1} \mathcal{T} e_d \right)^{-1} \Rightarrow F_{UMMSE} = R_{YY}^{-\frac{1}{2}} D = R_{YY}^{-1} \mathcal{T} e_d \left(e_d^H \mathcal{T}^H R_{YY}^{-1} \mathcal{T} e_d \right)^{-1}. \quad (8.15)$$

En appliquant le lemme d'inversion matricielle sur la matrice $R_{YY}^{-1} = (\sigma_a^2 \mathcal{T} \mathcal{T}^H + \sigma_v^2 I)^{-1}$, nous obtenons $R_{YY}^{-1} \mathcal{T} = \mathcal{T} (\sigma_a^2 \mathcal{T}^H \mathcal{T} + \sigma_v^2 I)^{-1}$. En utilisant cette expression de $R_{YY}^{-1} \mathcal{T}$ dans (8.15), nous obtenons

$$\begin{aligned} F_{UMMSE} &= \left(e_d^H \mathcal{T}^H \mathcal{T} (\sigma_a^2 \mathcal{T}^H \mathcal{T} + \sigma_v^2 I)^{-1} e_d \right)^{-1} \mathcal{T} (\sigma_a^2 \mathcal{T}^H \mathcal{T} + \sigma_v^2 I)^{-1} e_d \\ &= \left(e_d^H \left(\sigma_a^2 I + \sigma_v^2 (\mathcal{T}^H \mathcal{T})^{-1} \right)^{-1} e_d \right)^{-1} \mathcal{T} (\sigma_a^2 \mathcal{T}^H \mathcal{T} + \sigma_v^2 I)^{-1} e_d \\ &= \left(e_d^H \left(I + \frac{\sigma_v^2}{\sigma_a^2} (\mathcal{T}^H \mathcal{T})^{-1} \right)^{-1} e_d \right)^{-1} \mathcal{T} \left(\mathcal{T}^H \mathcal{T} + \frac{\sigma_v^2}{\sigma_a^2} I \right)^{-1} e_d. \end{aligned} \quad (8.16)$$

Considérons, maintenant, l'égaliseur MMSE ZF qui est dérivé à partir de la résolution du problème de minimisation suivant :

$$\min_{F: F^H \mathcal{T} = e_d^T} F^H R_{YY} F = \min_{F: F^H \mathcal{T} = e_d^T} \{ \sigma_a^2 + \sigma_v^2 F^H F \} \Leftrightarrow \min_{F: F^H \mathcal{T} = e_d^T} F^H F. \quad (8.17)$$

Une approche traitant ce problème consiste à introduire \mathcal{T}^\perp : le complément orthogonal de la matrice \mathcal{T} qui vérifie $\mathcal{T}^H \mathcal{T}^\perp = 0$ et $P_{\mathcal{T}^\perp} = P_{\mathcal{T}}^\perp$. L'égaliseur F peut être écrit comme

$$F = \begin{bmatrix} \mathcal{T} & \mathcal{T}^\perp \end{bmatrix} \begin{bmatrix} F_1^T & F_2^T \end{bmatrix}^T = \mathcal{T} F_1 + \mathcal{T}^\perp F_2. \quad (8.18)$$

La contrainte ZF, exprimée par $F^H \mathcal{T} = e_d^T = F_1^H \mathcal{T}^H \mathcal{T}$, implique que $F_1 = (\mathcal{T}^H \mathcal{T})^{-1} e_d$, et donc $F = \mathcal{T} (\mathcal{T}^H \mathcal{T})^{-1} e_d + \mathcal{T}^\perp F_2$, où F_2 n'est pas soumis à une contrainte particulière. Appliquons maintenant le critère MMSE ZF

$$\begin{aligned} \min_{F: F^H \mathcal{T} = e_d^T} F^H F &= e_d^H (\mathcal{T}^H \mathcal{T})^{-1} e_d + \min_{F_2} F_2^H \mathcal{T}^\perp \mathcal{T}^\perp F_2 \\ &\Rightarrow F_2 = 0 \Rightarrow F_{MMSEZF} = \mathcal{T} (\mathcal{T}^H \mathcal{T})^{-1} e_d \end{aligned} \quad (8.19)$$

Notons que F_{MMSEZF} est insensible à la variance du bruit. Dans le cas sans bruit $\sigma_v^2 = 0$ donc la matrice de covariance débruitée est $R_{YY}^d = \sigma_a^2 \mathcal{T} \mathcal{T}^H$. Nous pouvons, immédiatement, constater à partir de (8.16) et (8.19) que dans le cas sans bruit (ou dans le cas des statistiques débruitées) nous avons

$$F_{UMMSE,d} = F_{MMSEZF}. \quad (8.20)$$

A partir des égalités précédentes, nous avons aussi

$$F_{UMOEd} = F_{MOEZF} = F_{UMMSE,d} = F_{MMSEZF}. \quad (8.21)$$

Une démonstration alternative de ces égalités pour les égaliseurs de longueur infinie est présentée dans l'annexe A. Puisqu'il est plus facile d'appliquer le principe de Capon avec la contrainte d'absence de biais au lieu de la contrainte ZF (qui amènerait au problème d'optimisation suivant : $\max_{\mathbf{h}: \|\mathbf{h}\|=1} e_d^H (\mathcal{T}^H R_{YY}^{-1} \mathcal{T})^{-1} e_d$), nous appliquons la contrainte d'absence de biais en utilisant des covariances débruitées.

8.4 Egalisation Aveugle MMSE ZF

Le problème de minimisation, $\min_F F^H R_{YY} F$, (dans lequel la matrice R_{YY} est débruitée mais nous omettons le d pour la simplicité de notation) est résolu en deux étapes selon le principe de Capon :

- étape 1 : La résolution du critère d'optimisation $\min_{F: F^H \tilde{\mathbf{h}}=1} F^H R_{YY} F$ donne comme

$$\text{solution } F = \frac{1}{\tilde{\mathbf{h}}^H R_{YY}^{-1} \tilde{\mathbf{h}}} R_{YY}^{-1} \tilde{\mathbf{h}} : \text{égaliseur MMSE non biaisé.}$$

- étape 2 :

$$\max_{\mathbf{h}: \|\mathbf{h}\|=1} \min_{F: F^H \tilde{\mathbf{h}}=1} F^H R_{YY} F = \max_{\mathbf{h}: \|\mathbf{h}\|=1} \left(\tilde{\mathbf{h}}^H R_{YY}^{-1} \tilde{\mathbf{h}} \right)^{-1} \text{ ou } \min_{\mathbf{h}: \|\mathbf{h}\|=1} \tilde{\mathbf{h}}^H R_{YY}^{-1} \tilde{\mathbf{h}}. \quad (8.22)$$

Comme $\tilde{\mathbf{h}} = \begin{bmatrix} 0 \\ \mathbf{h} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} \mathbf{h} = T_1 \mathbf{h}$, le problème d'optimisation exprimé dans (8.22) devient

$$\min_{\mathbf{h}: \|\mathbf{h}\|=1} \mathbf{h}^H (T_1^H R_{YY}^{-1} T_1) \mathbf{h} \Rightarrow \mathbf{h} = V_{min} (T_1^H R_{YY}^{-1} T_1). \quad (8.23)$$

Le canal est évidemment estimé à un facteur scalaire près.

8.4.1 Interprétation en Termes de Prédiction Linéaire Bidirectionnelle

Alors que dans la prédiction linéaire progressive (avant) on prédit linéairement un échantillon donné à partir de ses échantillons passés, et dans la prédiction linéaire rétrograde (arrière) l'échantillon considéré est prédit à partir des échantillons qui viennent immédiatement après ; l'approche de prédiction linéaire bidirectionnelle consiste à prédire un vecteur d'échantillons à partir, simultanément, de ses échantillons passés et futurs.

Considérons la matrice T_2 définie comme $T_2 = \begin{bmatrix} I & 0 \\ 0 & 0 \\ 0 & I \end{bmatrix}$. La matrice T_2 peut être

interprétée comme la matrice engendrant le complément orthogonal de T_1 : $T_2 = T_1^\perp$, puisque $T_2^H T_1 = 0$ et $P_{T_2} + P_{T_1} = I$ ($P_{T_2} = P_{T_1}^\perp$), où $P_X^\perp = I - P_X = I - X(X^H X)^{-1} X^H$. A partir du lemme d'inversion matricielle, la matrice $T_1^H R_{YY}^{-1} T_1$ qui apparaît dans (8.23) peut être écrite comme

$$\left(T_1^H R_{YY} T_1 - T_1^H R_{YY} T_2 (T_2^H R_{YY} T_2)^{-1} T_2^H R_{YY} T_1 \right)^{-1}. \quad (8.24)$$

Considérons, maintenant, $\tilde{\mathbf{Y}} = T_1^H \mathbf{Y} - Q T_2^H \mathbf{Y}$ (voir figure 8.1), où la matrice considérée $Q = (T_1^H R_{YY} T_2) (T_2^H R_{YY} T_2)^{-1}$ est optimisée pour minimiser $R_{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}}$. $\tilde{\mathbf{Y}}$ est l'erreur de prédiction de la partie $T_1^H \mathbf{Y}$ de \mathbf{Y} à partir des parties voisinantes $T_2^H \mathbf{Y}$. Alors la matrice qui apparaît dans (8.24) est $R_{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}}^{-1}$. Ceci démontre que $\mathbf{h} = V_{max}(R_{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}})$.

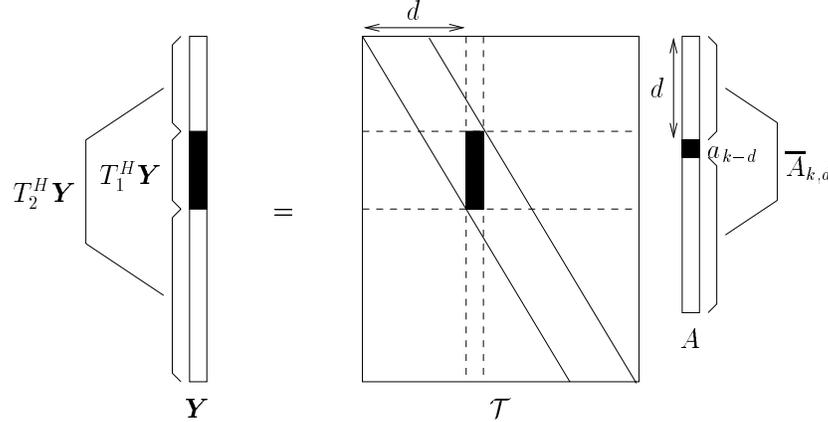


FIG. 8.1 – Prédiction linéaire bidirectionnelle.

8.4.2 Analyse dans le Cas sans Bruit

Dans le cas sans bruit $\mathbf{Y} = \mathcal{T} \mathbf{A}$, les vecteurs $T_1^H \mathbf{Y}$ et $T_2^H \mathbf{Y}$ sont illustrés dans la figure 8.1. Considérons la décomposition

$$\mathbf{Y} = \mathcal{T} \mathbf{A} = \tilde{\mathbf{h}} a_{k-d} + \overline{\mathcal{T}} \overline{\mathbf{A}}_{k,d}, \quad (8.25)$$

où les entités $\overline{\mathcal{T}}$ and $\overline{A}_{k,d}$ sont illustrées dans la figure 8.1. Nous avons, donc, $T_1^H \mathbf{Y} = \mathbf{h}a_{k-d} + T_1^H \overline{\mathcal{T}} \overline{A}_{k,d}$ et $T_2^H \mathbf{Y} = T_2^H \overline{\mathcal{T}} \overline{A}_{k,d}$. Si chacune de deux matrices $\begin{bmatrix} I & 0 & 0 \end{bmatrix} \mathcal{T}$ et $\begin{bmatrix} 0 & 0 & I \end{bmatrix} \mathcal{T}$ est de rang colonne plein, l'estimation de $T_1^H \mathbf{Y}$ à partir de $T_2^H \mathbf{Y} = T_2^H \overline{\mathcal{T}} \overline{A}_{k,d}$ est équivalente à l'estimation à partir de $\overline{A}_{k,d}$. Notons que la partie $T_1^H \overline{\mathcal{T}} \overline{A}_{k,d}$ peut être parfaitement estimée à partir de $\overline{A}_{k,d}$. En conséquence, l'erreur d'estimation pour $T_1^H \mathbf{Y}$ est

$$\tilde{\mathbf{Y}} = \mathbf{h}\tilde{a}_{k-d} |_{\overline{A}_{k,d}} \Rightarrow R_{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}} = \sigma_a^2 \mathbf{h}\mathbf{h}^H, \quad (8.26)$$

donc, en effet $\mathbf{h} = V_{max}(R_{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}})$. Dans le cas de symboles non corrélés (dans ce cas $\tilde{a}_{k-d} |_{\overline{A}_{k,d}} = a_{k-d}$), la variance σ_a^2 est remplacée par σ_a^2 . Notons que les corrélations considérées ici sont des moyennes statistiques. Cependant, le même raisonnement s'applique avec des moyenne temporelles, même en prenant la moyenne sur un nombre fini de données.

8.4.3 Relation avec d'Autres Approches

Dans cette section, nous établissons les liens reliant notre méthode d'égalisation aveugle MMSE ZF avec d'autres approches récemment proposées. Notre formulation constitue l'application correcte du principe de Capon pour la formation de voies LCMV, à l'égalisation multi-canal, alors que la formulation de Tsatsanis devient correcte uniquement pour des SNRs élevés. Notre prédiction linéaire bidirectionnelle correspond à une solution en une étape de l'approche multi-étape de Gesbert et Duhamel [GD97]. La méthode "Least-Squares Smoothing" (LSS) introduite par Tong *et al.* [TZ98] aboutit à des estimées biaisées du canal: la méthode garde uniquement le sous-espace signal (le sous-espace bruit est éliminé par l'opération SVD) mais la contribution du bruit aux valeurs propres correspondant à l'espace signal n'est pas éliminée. Notre prédiction linéaire bidirectionnelle correspond à une version de la méthode de Tong, où le biais a été éliminé, et qui évite le calcul de la SVD de la matrice de covariance du signal reçu. En fait, la présence de biais dans la méthode de Tong montre que le point de vue déterministe sur lequel sa dérivation est basée est inapproprié.

8.4.4 Détection de l'Ordre du Canal

Dans [TZ], Tong *et al.* ont modifié leur méthode (Least-Squares Smoothing) pour incorporer la détection de l'ordre du canal. L'algorithme dérivé estime conjointement le canal et son ordre et se base sur le fait que le système d'équations à résoudre pour identifier le canal, aboutit à une solution non triviale (qui est la canal à un facteur scalaire près) au bon ordre et à des solutions triviales si l'ordre est sous-estimé (ou éventuellement surestimé). Remarquons que cette procédure nécessite que l'opération d'estimation du canal soit effectuée à tous les ordres possibles.

Nous proposons une procédure simple et efficace pour incorporer la détection de l'ordre du canal dans notre méthode. En effet, la méthode proposée permet de trouver le canal

comme étant le vecteur propre correspondant à la valeur propre maximale d'une matrice de covariance d'erreurs de prédiction $R_{\hat{Y}\hat{Y}}$. Avec les statistiques exactes, $R_{\hat{Y}\hat{Y}} = \sigma_a^2 \mathbf{h} \mathbf{h}^H$ ce qui implique que le rang de $R_{\hat{Y}\hat{Y}}$ est 1. Quand la longueur de \mathbf{h} est surestimée de p , le rang de $R_{\hat{Y}\hat{Y}}$ devient $p + 1$. Par contre, quand cette longueur est sous-estimée, $R_{\hat{Y}\hat{Y}}$ est idéalement zéro (et donc n'a pas, en pratique, de structure particulière). Par conséquent, une méthode efficace pour détecter l'ordre du canal consiste à trouver l'ordre pour lequel $R_{\hat{Y}\hat{Y}}$ s'approche le plus d'une matrice de rang un. Dans le cas des statistiques exactes, en considérant les valeurs propres de $R_{\hat{Y}\hat{Y}}$ ordonnées dans un ordre décroissant : $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{mN}$ et en calculant le rapport $r = \lambda_1/\lambda_2$, comme $R_{\hat{Y}\hat{Y}}$ est de rang 1, $r \rightarrow \infty$. Lorsque le bruit est présent et lorsque l'ordre du canal est sous-estimé (ou surestimé), la valeur de r sera plus petite que sa valeur à l'ordre correct du canal. Donc la variation de r en fonction des différentes valeurs de l'ordre du canal est une fonction qui présente un maximum à l'ordre correct du canal. En effet, la matrice $R_{\hat{Y}\hat{Y}}$ peut être écrite comme la somme de sa meilleure approximation de rang un ($\hat{R}_{\hat{Y}\hat{Y}}$) et de l'erreur d'approximation ($\tilde{R}_{\hat{Y}\hat{Y}}$) :

$$R_{\hat{Y}\hat{Y}} = \lambda_1 V_1 V_1^H + \sum_{i>1} \lambda_i V_i V_i^H = \hat{R}_{\hat{Y}\hat{Y}} + \tilde{R}_{\hat{Y}\hat{Y}} \quad (8.27)$$

et

$$r = \frac{\lambda_1}{\lambda_2} = \frac{\|\lambda_1 V_1 V_1^H\|_2}{\|\sum_{i>1} \lambda_i V_i V_i^H\|_2} = \frac{\|\hat{R}_{\hat{Y}\hat{Y}}\|_2}{\|\tilde{R}_{\hat{Y}\hat{Y}}\|_2}. \quad (8.28)$$

8.4.5 Débruitage de la Matrice de Covariance

En présence de bruit, la matrice de covariance du signal reçu doit être débruitée et donc les égaliseurs MMSE ZF sont dérivés via le critère MMSE utilisant la matrice de covariance débruitée $R_{\hat{Y}\hat{Y}}^d$ (d désigne la matrice obtenue après l'opération de débruitage). Nous considérons deux opérations alternatives de débruitage de la matrice de covariance. Dans la première opération, nous éliminons la contribution du bruit estimée comme étant la valeur propre minimale de $\hat{R}_{\hat{Y}\hat{Y}}$. Nous constatons que la valeur propre minimale de $\hat{R}_{\hat{Y}\hat{Y}}$ sous-estime la puissance du bruit et par conséquent, la partie du bruit que nous éliminons de $\hat{R}_{\hat{Y}\hat{Y}}$ n'est pas suffisante. Nous proposons donc, dans une opération alternative de débruitage, d'estimer la puissance du bruit par la méthode SRM (voir section §3.4.2) [AS97] : $\hat{\sigma}_v^2 = \lambda_{\min}(SRM)$. Avec un nombre fini de données cette méthode va aussi sous-estimer la puissance du bruit mais l'estimée qu'elle donne est proche de la vraie valeur. Le débruitage avec une estimée $\hat{\sigma}_v^2 > \lambda_{\min}(\hat{R}_{\hat{Y}\hat{Y}})$ rend la matrice débruitée $\hat{R}_{\hat{Y}\hat{Y}} - \hat{\sigma}_v^2 I$ indéfinie. Puisqu'une matrice de covariance est nécessairement semi-définie positive, on ne garde que la partie semi-définie positive de la matrice débruitée : $[\hat{R}_{\hat{Y}\hat{Y}} - \hat{\sigma}_v^2 I]_+$ (cette opération peut être effectuée par une décomposition LDU). En fait, pour bien débruiter $\hat{R}_{\hat{Y}\hat{Y}}$, nous introduisons un facteur d'ajustement α : $\hat{\sigma}_v^2 = \alpha \lambda_{\min}(SRM)$. Le facteur α est idéalement choisi tel que

$$\sigma_v^2 \lesssim \hat{\sigma}_v^2 < \lambda_{\min}^{SS} \quad (8.29)$$

où λ_{min}^{SS} désigne la plus petite valeur propre de l'ensemble des valeurs propres de R_{YY} correspondant au sous-espace signal. On peut remarquer que si $\hat{\sigma}_v^2$ satisfait ces contraintes, le sous-espace signal n'est pas affecté.

8.5 Simulations

Nous considérons une longueur de trame des données $M = 200$, un canal complexe \mathbf{H} généré de façon aléatoire de longueur $N = 3$ avec $m = 2$ sous-canaux. Les symboles sont i.i.d. QPSK. Le SNR est défini comme $(\|\mathbf{h}\|^2\sigma_a^2)/(m\sigma_v^2)$. Les méthodes d'estimation aveugles donnent une estimée du canal $\hat{\mathbf{h}}$ avec $\|\hat{\mathbf{h}}\| = 1$, nous ajustons le facteur d'échelle α tel que $\mathbf{h}_o^H(\alpha\hat{\mathbf{h}}) = \mathbf{h}_o^H\mathbf{h}_o$ où \mathbf{h}_o est le vrai canal (voir [dCS97b]) : l'estimée finale est $\hat{\mathbf{h}} = \alpha\hat{\mathbf{h}}$. La mesure de performance est faite par l'intermédiaire de l'erreur quadratique moyenne normalisée (NMSE) (en prenant la moyenne sur 100 réalisations Monte-Carlo) définie comme $NMSE = \|\mathbf{h} - \hat{\mathbf{h}}\|^2/\|\mathbf{h}\|^2$.

8.5.1 Performance de la Méthode

Dans la figure 8.2, nous comparons notre approche de prédiction linéaire bidirectionnelle avec d'autres méthodes d'estimation du canal et avec sa borne de Cramer-Rao (CRB). Nous utilisons une matrice empirique \hat{R}_{YY} de longueur $L = 3N$. La méthode de Tsatsanis [TX97] donne la performance la plus mauvaise et une amélioration dans cette même performance peut être constatée au fur et à mesure que le SNR augmente (cette constatation est conforme avec le fait que la formulation de Tsatsanis devient correcte uniquement pour un SNR élevé). La méthode LSS de Tong [TZ98] donne des performances comparables à celle de la méthode de Tsatsanis à cause de la présence du biais dans l'estimée du canal. Nous avons tracé les courbes qui correspondent aux différentes stratégies proposées pour débruiter la matrice \hat{R}_{YY} et qui sont expliquées dans la section précédente. Nous estimons la variance du bruit comme la valeur propre minimale de \hat{R}_{YY} puis par la méthode SRM [AS97] et nous utilisons deux facteurs d'ajustement $\alpha = 1.5$ et $\alpha = 2$. Il est clair qu'avec $\alpha = 2$ nous atteignons la même performance que celle de la méthode PQML [AdCS98] ce qui implique que \hat{R}_{YY} est parfaitement débruitée et en conséquence la performance de la prédiction linéaire bidirectionnelle est optimale. Notons que l'élimination de la contribution du bruit des valeurs propres dans la méthode LSS aboutit aussi à la même performance optimale.

8.5.2 Détection de l'Ordre du Canal

Pour le même canal, nous avons considéré le problème de la surestimation ou sous-estimation de l'ordre du canal (l'ordre correct du canal étant égal à 3). Dans la figure 8.3, nous avons tracé, en considérant des valeurs différentes du SNR, le rapport $r = \lambda_1/\lambda_2$ en fonction des différentes valeurs supposées pour l'ordre du canal. Les courbes sont

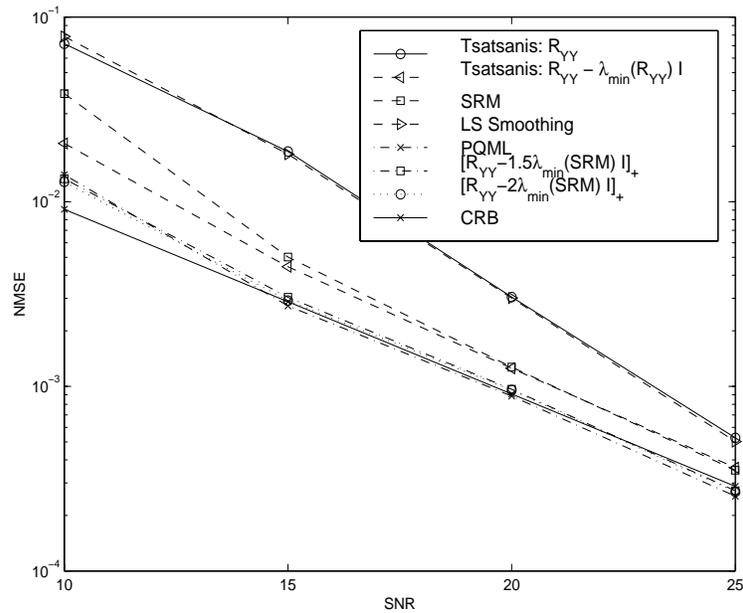


FIG. 8.2 – Performance des différents algorithmes.

obtenues en prenant la moyenne sur 100 réalisations Monte-Carlo. Il est clair que la fonction tracée donne un maximum net à l'ordre correct du canal.

8.6 Conclusions

L'équivalence entre les critères d'optimisation MMSE non biaisé et MOE non biaisé nous a permis d'établir que des égaliseurs ($F_{UMOE,d}$, F_{MOEZF} , $F_{UMMSE,d}$ et F_{MMSEZF}) qui sont dérivés au sens de critères d'optimisation différents, sont identiques dans le cas sans bruit. Puisqu'il est plus facile d'appliquer le principe de Capon avec la contrainte d'absence de biais au lieu de la contrainte ZF, la contrainte d'absence de biais est appliquée en utilisant des covariances débruitées. La méthode d'estimation multi-canal aveugle par égalisation MMSE ZF, que nous avons décrite dans ce chapitre, constitue la procédure correcte de l'application du principe de Capon pour la formation de voies LCMV à l'identification multi-canal. L'intérêt de cette méthode se manifeste sur deux plans. D'une part, la méthode présente des performances optimales (équivalentes à celle de l'algorithme PQML). D'autre part, le coût de calcul faible de la méthode qui permet d'estimer le canal sans exiger la décomposition propre de la matrice de covariance du signal reçu (laquelle décomposition est indispensable dans les approches d'ajustement de sous-espaces ou la technique "smoothing" de Tong *et al.*). La méthode proposée peut être augmentée pour incorporer une procédure simple de détection de l'ordre du canal qui se base sur la détermination du rang d'une matrice de covariance d'erreurs de prédiction. Les résultats de simulations confirment la supériorité, en termes de performance, de l'approche proposée par rapport aux autres techniques qui y sont reliées ainsi que

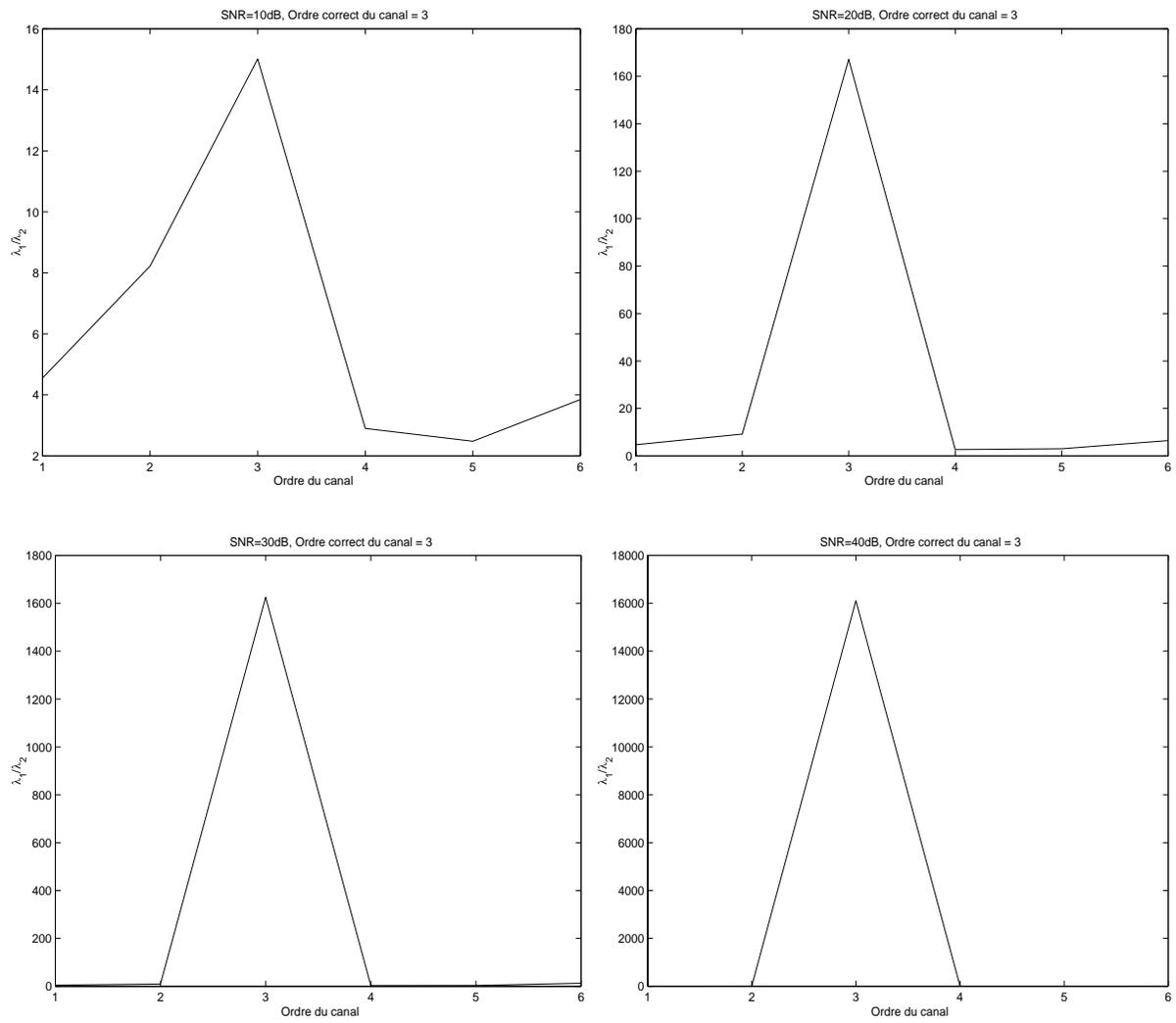


FIG. 8.3 – Détection de l'ordre du canal.

l'efficacité de la stratégie de détection de l'ordre du canal.

Conclusion Générale

Les méthodes d'estimation aveugles proposées dans cette thèse adoptent un modèle des symboles déterministe dans lequel aucune information a priori sur les symboles transmis n'est exploitée. L'avantage de ces méthodes déterministes est qu'elles permettent d'aboutir à des solutions analytiques (ou nécessitent des procédures itératives à convergence rapide) pour le problème d'estimation de canal. Leur inconvénient est parfois leur coût de calcul élevé et surtout leur manque de robustesse à la surdétermination de l'ordre du canal. Le modèle gaussien exploite l'information sur les symboles transmis. Les algorithmes d'estimation qui se basent sur ce modèle auront des performances meilleures que ceux qui utilisent le modèle déterministe. Toutefois, cette amélioration de performance s'accompagne d'un accroissement du degré de non convexité de la fonction de coût considérée et donc l'apparition du problème des minima locaux. Les méthodes de résolution qui peuvent être envisagées, dans ce cas de modèle, sont itératives et "lourdes" en termes de complexité numérique. Trouver des techniques analytiques pour résoudre le problème d'estimation du canal lorsque le modèle des symboles est gaussien reste un défi ouvert ! En fait, la relation entre l'exploitation de la connaissance a priori sur le système à identifier et l'accroissement du degré de non convexité de la fonction de coût considérée est une relation de cause à effet : plus on exploite cette information, plus le degré de non convexité augmente.

Dans le cadre des méthodes d'estimation déterministes, nous avons proposé des méthodes alternatives pour les techniques d'ajustement de sous-espaces. Ces méthodes alternatives, évitent la décomposition propre de la matrices de covariance (et en conséquence réduisent le coût de calcul des versions classiques) tout en ayant, asymptotiquement, les mêmes performances. Une autre classe de méthodes déterministes est celle qui est basée sur le critère de maximum de vraisemblance. La stratégie IQML pour résoudre le critère ML est biaisée et une procédure pour enlever ce biais est l'algorithme DIQML. La performance de cet algorithme est supérieure à celles des autres algorithmes déterministes, mais reste inférieure à celle de l'algorithme PQML. Asymptotiquement, les deux algorithmes, DIQML et PQML, sont globalement convergents et leur convergence nécessite quelques itérations. Nous avons essayé de bénéficier de la robustesse aveugle de l'agortihme PQML pour dériver deux algorithmes semi-aveugles : PQML-LS et PQML-WLS dont la performance du second reste fortement encourageante. Nous avons aussi utilisé l'algorithme PQML dans un contexte d'exploitation de l'alphabet fini effectuée au niveau de l'opération d'égalisation. Nous constatons qu'une simple stratégie d'égalisation

sation à décisions dirigées, DD-LS, lorsqu'elle est bien initialisée aboutit à la meilleure performance.

Les algorithmes d'identification purement aveugles et les approches basées exclusivement sur les séquences d'apprentissage correspondent à deux cas extrêmes de méthodes possibles. Combiner les idées appartenant à ces deux classes de méthodes aboutira à des approches semi-aveugles dont la séquence d'apprentissage se trouve réduite et qui bénéficient de la robustesse des techniques aveugles à un coût de calcul faible.

Sur un autre niveau, l'exploitation de la connaissance a priori des filtres de TX/RX constitue une piste prometteuse. En effet, les bornes de performances pouvant être atteintes dans ce cas sont beaucoup plus basses que celles correspondant au cas purement aveugle. Les critères qui exploitent cette connaissance se trouvent mieux conditionnés et ainsi des canaux qui ne sont pas identifiables par des méthodes purement aveugles peuvent être identifiés grâce à l'exploitation de cette connaissance. Les algorithmes d'estimation de canal qui incorporent la connaissance a priori des filtres de TX/RX aboutissent à des performances supérieures à celles des algorithmes aveugles. Une piste intéressante consiste à exploiter cette connaissance a priori dans des méthodes semi-aveugles.

En pratique, l'ordre du canal n'est pas connu et des procédures de sa détection s'imposent. Pour préserver la robustesse des méthodes d'estimation aveugles déterministes, nous avons proposé deux procédures alternatives. La première étape de ces deux procédures est une opération d'ancrage de la réponse impulsionnelle en son premier coefficient vectoriel. Dans la deuxième étape, la première approche exploite la structure de la matrice du critère à optimiser ; cependant la deuxième approche se base sur l'algorithme de Levinson multi-canal adapté au cas singulier. En fait, la première approche que nous proposons peut être vue comme un mécanisme de détection implicite de l'ordre du canal, par contre la seconde approche détecte cet ordre de façon explicite.

Dans un contexte de bruit additif qui n'est plus supposé blanc mais uniquement stationnaire de couleur inconnue, nous avons proposé une méthode d'ajustement de sous-espace pour identifier le canal. Nous avons considéré le cas où les différents canaux utilisés proviennent à partir d'une combinaison du suréchantillonnage et d'un réseau d'antennes. La méthode proposée exploite une structure de déplacement de la matrice de covariance du signal reçu dans laquelle la contribution du bruit disparaît. Comme l'élimination du bruit stationnaire dans l'approche d'ajustement de sous-espace est basée sur le suréchantillonnage, et par conséquent sur l'excès en largeur de bande, nous avons étendu la méthode proposée pour exploiter la connaissance a priori des filtres de TX/RX afin d'améliorer sa performance.

Nous avons aussi proposé une méthode d'estimation de canal basée sur l'égalisation aveugle MMSE ZF. Nous avons prouvé, dans ce contexte, l'équivalence entre plusieurs critères d'égalisation dans le cas sans bruit et sous la contrainte d'absence du biais. Par conséquent, nous avons déduit qu'une étape de débruitage de la matrice de covariance du signal reçu est nécessaire avant son utilisation dans la méthode d'estimation aveugle. L'avantage de la technique proposée est qu'elle évite la décomposition propre

de la matrice de covariance tout en aboutissant à des performances identiques à celles de l'algorithme PQML. Nous avons, ensuite, proposé une procédure simple et efficace pour incorporer la détection de l'ordre du canal dans notre approche.

Une de nos perspectives concerne l'exploitation de la connaissance des filtres de TX/RX. Dans le cadre de cette thèse, nous avons considéré que le filtre de TX/RX et le signal reçu sont suréchantillonnés avec le même facteur de suréchantillonnage. Nous proposons d'étudier le cas où le filtre de TX/RX est suréchantillonné avec un facteur de suréchantillonnage différent de celui du signal reçu (tout en satisfaisant le critère de Nyquist), d'établir les conditions d'identifiabilité et d'évaluer les performances des méthodes exploitant cette connaissance. Un suréchantillonnage fractionnel du signal reçu est aussi à envisager dans cette étude. Ces aspects font l'objet de nos travaux de recherches actuels.

D'autres perspectives de ce travail de thèse sont envisageables. Les approches que nous avons étudiées traitent le cas mono-utilisateur, leurs extensions éventuelles au cas multi-utilisateur sont à considérer. Il sera aussi intéressant de tester les algorithmes établis dans un environnement réaliste et de trouver des versions de complexité réduite pour l'algorithme qui présente le meilleur comportement.

Annexe A

Démonstrations pour les Egaliseurs de Longueur Infinie dans le Domaine Fréquentiel

Soit $\mathbf{h}(z) = [H_1(z) \cdots H_m(z)]^T$ la fonction de transfert mono-entrée/multi-sortie du canal et soit $\mathbf{f}(z)$ le filtre de dimension $1 \times m$ tel que $\hat{a}_k = \mathbf{f}(q)\mathbf{y}(k)$, l'énergie à la sortie (OE) est donnée par

$$\text{OE} = \frac{1}{2\pi j} \oint \mathbf{f}(z) S_{yy}(z) \mathbf{f}^\dagger(z) \frac{dz}{z} = E|\hat{a}_k|^2, \quad (\text{A.1})$$

où $\mathbf{f}^\dagger(z) = \mathbf{f}^H(1/z^*)$.

A.1 Contrainte d'Absence du Biais

Dans le domaine fréquentiel, la contrainte d'absence du biais (qui correspond à l'approche de Tsatsanis) est

$$\oint \mathbf{f}(z) \mathbf{h}(z) \frac{dz}{z} = 1. \quad (\text{A.2})$$

La densité spectrale de puissance peut être écrite comme $S_{yy}(z) = S_{yy}^{\frac{1}{2}}(z) S_{yy}^{\frac{1}{2}}(z)$. Soit $g^\dagger(z) = \mathbf{f}(z) S_{yy}^{\frac{1}{2}}(z)$ et considérons le produit scalaire défini comme

$$\langle u(z), v(z) \rangle = \frac{1}{2\pi j} \oint u^\dagger(z) v(z) \frac{dz}{z}. \quad (\text{A.3})$$

Nous avons

$$\begin{aligned} \min_{g(z): \langle g(z), S_{yy}^{-\frac{1}{2}}(z) \mathbf{h}(z) \rangle = 1} \|g(z)\|^2 &\Rightarrow g(z) = \alpha S_{yy}^{-\frac{1}{2}}(z) \mathbf{h}(z), \\ \alpha \text{ est tel que } \langle g(z), S_{yy}^{-\frac{1}{2}}(z) \mathbf{h}(z) \rangle = 1 &= \alpha^* \frac{1}{2\pi j} \oint \mathbf{h}^\dagger(z) S_{yy}^{-1} \mathbf{h}(z) \frac{dz}{z} \\ \Rightarrow \mathbf{f}(z) = g^\dagger(z) S_{yy}^{-\frac{1}{2}}(z) &= \frac{1}{\frac{1}{2\pi j} \oint \mathbf{h}^\dagger(z) S_{yy}^{-1} \mathbf{h}(z) \frac{dz}{z}} \mathbf{h}^\dagger(z) S_{yy}^{-1}(z). \end{aligned} \quad (\text{A.4})$$

et donc

$$\text{MOE} = \left(\frac{1}{2\pi j} \oint \frac{dz}{z} \mathbf{h}^\dagger(z) S_{yy}^{-1}(z) \mathbf{h}(z) \right)^{-1}. \quad (\text{A.5})$$

Maintenant, le principe de Capon consiste au problème d'optimisation suivant :

$$\max_{\hat{\mathbf{h}}(z): \|\hat{\mathbf{h}}(z)\|=1} \text{MOE}, \quad (\text{A.6})$$

en utilisant (A.5) nous obtenons

$$\begin{aligned} & \max_{\hat{\mathbf{h}}(z): \|\hat{\mathbf{h}}(z)\|=1} \left(\min_{\mathbf{f}(z): \oint \mathbf{f}(z) \hat{\mathbf{h}}(z) \frac{dz}{z} = 1} \frac{1}{2\pi j} \oint \mathbf{f}(z) S_{yy}(z) \mathbf{f}^\dagger(z) \frac{dz}{z} \right) \\ &= \max_{\hat{\mathbf{h}}(z): \|\hat{\mathbf{h}}(z)\|=1} \frac{1}{\frac{1}{2\pi j} \oint \hat{\mathbf{h}}^\dagger(z) S_{yy}^{-1}(z) \hat{\mathbf{h}}(z) \frac{dz}{z}} \\ &= \min_{\hat{\mathbf{h}}(z): \|\hat{\mathbf{h}}(z)\|=1} \frac{1}{2\pi j} \oint \hat{\mathbf{h}}^\dagger(z) S_{yy}^{-1}(z) \hat{\mathbf{h}}(z) \frac{dz}{z}. \end{aligned} \quad (\text{A.7})$$

En utilisant le lemme d'inversion matricielle sur $S_{yy}^{-1}(z)$, nous obtenons

$$S_{yy}^{-1}(z) = (\sigma_v^2 I + \sigma_a^2 \mathbf{h}(z) \mathbf{h}(z)^\dagger)^{-1} = \sigma_v^{-2} I - \sigma_v^{-2} \mathbf{h}(z) \left(\mathbf{h}^\dagger(z) \mathbf{h}(z) + \frac{\sigma_v^2}{\sigma_a^2} \right)^{-1} \mathbf{h}^\dagger(z), \quad (\text{A.8})$$

donc $\sigma_v^2 S_{yy}^{-1}(z) \xrightarrow{\sigma_v^2 \rightarrow 0} P_{\mathbf{h}(z)}^\perp$. Par conséquent,

$$\begin{aligned} & \min_{\hat{\mathbf{h}}(z): \|\hat{\mathbf{h}}(z)\|=1} \frac{1}{2\pi j} \oint \hat{\mathbf{h}}^\dagger(z) S_{yy}^{-1}(z) \hat{\mathbf{h}}(z) \frac{dz}{z} \xrightarrow{\text{SNR} \rightarrow \infty} \min_{\hat{\mathbf{h}}(z): \|\hat{\mathbf{h}}(z)\|=1} \frac{1}{2\pi j \sigma_v^2} \oint \hat{\mathbf{h}}^\dagger(z) P_{\mathbf{h}(z)}^\perp \hat{\mathbf{h}}(z) \frac{dz}{z} \\ & \Rightarrow \hat{\mathbf{h}}(z) = \mathbf{h}(z) \beta(z), \text{ où } \beta(z) \text{ est un filtre scalaire.} \end{aligned} \quad (\text{A.9})$$

Si $\mathbf{h}(z)$ est FIR et $\hat{\mathbf{h}}(z)$ est FIR ayant la même longueur alors $\hat{\mathbf{h}}(z) = \mathbf{h}(z) \beta$, où β est un facteur scalaire: $\beta = \frac{1}{\sqrt{\frac{1}{2\pi j} \oint \mathbf{h}^\dagger(z) \mathbf{h}(z) \frac{dz}{z}}}$. Donc la méthode de Capon avec

la distorsion interprétée comme une absence du biais aboutit à l'estimée correcte du canal lorsque le SNR $\rightarrow \infty$. Par conséquent une interprétation appropriée de l'absence de distorsion est le ZF.

A.2 Contrainte ZF

Considérons maintenant la contrainte ZF, $\mathbf{f}(z) \mathbf{h}(z) = 1$, le problème MOE est formulé comme

$$\min_{\mathbf{f}(z): \mathbf{f}(z) \mathbf{h}(z) = 1} \frac{1}{2\pi j} \oint \mathbf{f}(z) S_{yy}(z) \mathbf{f}^\dagger(z) \frac{dz}{z}, \quad (\text{A.10})$$

où $\mathbf{f}(z)\mathbf{h}(z) = 1$ est une contrainte pour toute les fréquences et la quantité $\mathbf{f}(z)S_{yy}(z)\mathbf{f}^\dagger(z)$ somme les contributions de toutes les fréquences. Par conséquent, de façon équivalente, à n'importe quelle fréquence, nous avons

$$\min_{\mathbf{f}(z):\mathbf{f}(z)\mathbf{h}(z)=1} \mathbf{f}(z)S_{yy}(z)\mathbf{f}^\dagger(z). \quad (\text{A.11})$$

En utilisant $g^\dagger(z) = \mathbf{f}(z)S_{yy}^{-\frac{1}{2}}(z)$, le problème de minimisation (A.11) devient

$$\min_{g^\dagger(z)S_{yy}^{-\frac{1}{2}}(z)\mathbf{h}(z)=1} g^\dagger(z)g(z), \quad (\text{A.12})$$

donc

$$\begin{aligned} g(z) &= \alpha S_{yy}^{-\frac{1}{2}}(z)\mathbf{h}(z), \quad \alpha \text{ est tel que } \alpha^*\mathbf{h}^\dagger(z)S_{yy}^{-1}(z)\mathbf{h}(z) = 1 \\ \Rightarrow \mathbf{f}(z) &= g^\dagger(z)S_{yy}^{-\frac{1}{2}}(z) = \frac{1}{\mathbf{h}^\dagger(z)S_{yy}^{-1}(z)\mathbf{h}(z)}\mathbf{h}^\dagger(z)S_{yy}^{-1}(z). \end{aligned} \quad (\text{A.13})$$

Par conséquent,

$$\text{MOE}(\widehat{\mathbf{h}}(z)) = \frac{1}{2\pi j} \oint \frac{1}{\widehat{\mathbf{h}}^\dagger(z)S_{yy}^{-1}(z)\widehat{\mathbf{h}}(z)} \frac{dz}{z}. \quad (\text{A.14})$$

Le principe de Capon consiste, maintenant, à $\max_{\widehat{\mathbf{h}}(z):\|\widehat{\mathbf{h}}(z)\|=1} \text{MOE}(\widehat{\mathbf{h}}(z))$, en utilisant (A.14), nous obtenons

$$\begin{aligned} &\max_{\widehat{\mathbf{h}}(z):\|\widehat{\mathbf{h}}(z)\|=1} \text{MOE}(\widehat{\mathbf{h}}(z)) \\ &= \max_{\widehat{\mathbf{h}}(z):\|\widehat{\mathbf{h}}(z)\|=1} \left(\min_{\mathbf{f}(z):\mathbf{f}(z)\widehat{\mathbf{h}}(z)=1} \frac{1}{2\pi j} \oint \mathbf{f}(z)S_{yy}\mathbf{f}^\dagger(z) \frac{dz}{z} \right) \\ &= \max_{\widehat{\mathbf{h}}(z):\|\widehat{\mathbf{h}}(z)\|=1} \frac{1}{2\pi j} \oint \frac{1}{\widehat{\mathbf{h}}^\dagger(z)S_{yy}^{-1}(z)\widehat{\mathbf{h}}(z)} \frac{dz}{z}. \end{aligned} \quad (\text{A.15})$$

En utilisant le lemme d'inversion matricielle sur $(\sigma_a^2\mathbf{h}(z)\mathbf{h}^\dagger(z) + \sigma_v^2I)^{-1}$, l'expression (A.15) devient

$$\max_{\widehat{\mathbf{h}}(z):\|\widehat{\mathbf{h}}(z)\|=1} \frac{\sigma_v^2}{2\pi j} \oint \frac{\mathbf{h}^\dagger(z)\mathbf{h}(z) + \frac{\sigma_v^2}{\sigma_a^2}}{\widehat{\mathbf{h}}^\dagger(z)\widehat{\mathbf{h}}(z)\mathbf{h}^\dagger(z)\mathbf{h}(z) - \widehat{\mathbf{h}}^\dagger(z)\mathbf{h}(z)\mathbf{h}^\dagger(z)\widehat{\mathbf{h}}(z) + \frac{\sigma_v^2}{\sigma_a^2}\widehat{\mathbf{h}}^\dagger(z)\widehat{\mathbf{h}}(z)} \frac{dz}{z}. \quad (\text{A.16})$$

Soit $\widehat{\mathbf{h}}(z) = \widehat{\mathbf{h}}(z)\beta(z)$, où β est un scalaire tel que

$$\widehat{\mathbf{h}}^\dagger(z)\widehat{\mathbf{h}}(z) = 1 \Rightarrow \widehat{\mathbf{h}}^\dagger(z)\widehat{\mathbf{h}}(z) = \beta^\dagger(z)\beta(z). \quad (\text{A.17})$$

Donc le problème de maximisation (A.16) devient

$$\begin{aligned} &\max_{\widehat{\mathbf{h}}(z),\beta(z)} \frac{\sigma_v^2}{2\pi j} \oint \frac{\mathbf{h}^\dagger(z)\mathbf{h}(z) + \frac{\sigma_v^2}{\sigma_a^2}}{\beta^\dagger(z)\beta(z) \left[\left(1 - \widehat{\mathbf{h}}^\dagger(z)P_{\mathbf{h}(z)}\widehat{\mathbf{h}}(z)\right) \mathbf{h}^\dagger(z)\mathbf{h}(z) + \frac{\sigma_v^2}{\sigma_a^2} \right]} \frac{dz}{z} \\ &\Rightarrow \widehat{\mathbf{h}}(z) = \bar{\mathbf{h}}(z) = (\mathbf{h}^\dagger(z)\mathbf{h}(z))^{-\frac{1}{2}} \mathbf{h}(z). \end{aligned} \quad (\text{A.18})$$

Par conséquent, l'expression (A.18) devient

$$\beta(z): \frac{1}{2\pi j} \oint \beta^\dagger(z)\beta(z) \frac{dz}{z} = 1 \quad \frac{\sigma_v^2}{2\pi j} \oint \frac{\mathbf{h}^\dagger(z)\mathbf{h}(z) + \frac{\sigma_v^2}{\sigma_a^2}}{\beta^\dagger(z)\beta(z)} \frac{dz}{z} \quad (\text{A.19})$$

Si $\mathbf{h}(z)$ est FIR et $\hat{\mathbf{h}}(z)$ est FIR ayant la même longueur alors $\beta(z) = \gamma (\mathbf{h}^\dagger(z)\mathbf{h}(z))^{\frac{1}{2}}$, donc $\hat{\mathbf{h}}(z) = \gamma \mathbf{h}(z)$ et $\gamma = \frac{1}{\sqrt{\frac{1}{2\pi j} \oint \mathbf{h}^\dagger(z)\mathbf{h}(z) \frac{dz}{z}}}$ tel que $\|\hat{\mathbf{h}}(z)\| = 1$. En conclusion, la

méthode de capon avec l'absence de distorsion interprétée comme ZF fonctionne dans le cas du bruit blanc.

Bibliographie

- [AdCS98] J. Ayadi, E. de Carvalho, and D.T.M. Slock. “Blind and Semi-blind Maximum Likelihood Methods for FIR Multichannel Identification”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Seattle, WA, May 12-15 1998.
- [ADS97a] J. Ayadi, L. Deneire, and D.T.M. Slock. “Approches Robustes a la Sur-determination de l’Ordre du Canal pour l’Estimation Aveugle de Canaux Multiples RIF”. In *Proc. of the 16^{ème} colloque GRETSI sur le traitement du signal et des images*, Grenoble, France, September 1997.
- [ADS97b] J. Ayadi, L. Deneire, and D.T.M. Slock. “Channel-based Blind Multichannel Identification without Order Overestimation Problems”. In *Proc. of the 13th international conference on Digital Signal Processing: DSP 97*, Santorini, Greece, July 1997.
- [AMCG⁺97] K. Abed-Meraim, J.-F. Cardoso, A. Gorokhov, P. Loubaton, and E. Moulines. “Subspace Methods for Blind Identification of SIMO-FIR Systems”. *IEEE Trans. SP*, SP-45:42–55, Jan. 1997.
- [AMDG⁺95] K. Abed-Meraim, P. Duhamel, D. Gesbert, P. Loubaton, S. Mayrargue, E. Moulines, and D. Slock. “Prediction Error Methods for Time-Domain Blind Identification of Multichannel FIR Filters”. In *Proc. ICASSP 95 Conf.*, Detroit, Michigan, May 1995.
- [AMH97a] K. Abed-Meraim and Y. Hua. “Improved TSML Algorithm for Multichannel Blind Identification”. In *IEEE Signal Processing Advances in Wireless Communications Workshop*, Paris, Apr. 16-18 1997.
- [AMH97b] K. Abed-Meraim and Y. Hua. “New TSML Algorithm for Multichannel Blind Identification”. In *Proc. SYSID 97 Conf.*, Japan, July 1997.
- [AMHLM97] K. Abed-Meraim, Y. Hua, P. Loubaton, and E. Moulines. “Subspace Method for Blind Identification of Multichannel FIR Systems in Noise Field with Unknown Spatial Covariance”. *IEEE Signal Processing Letters*, 4(5):135–137, May 1997.

- [AMVW91] S. Andersson, M. Millnert, M. Viberg, and B. Whalberg. “An Adaptive Array for Mobile Communication Systems”. *IEEE Transactions on Vehicular Technology*, 40(1):230–236, 1991.
- [AS] J. Ayadi and D.T.M. Slock. “Blind Channel Estimation Exploiting Transmission Filter Knowledge”. To appear in Signal Processing (EURASIP) - Special Issue on Signal Processing Technologies for Short-Burst Wireless Communications.
- [AS97] J. Ayadi and D.T.M. Slock. “Cramer-Rao Bounds and Methods for Knowledge Based Estimation of Multiple FIR Channels”. In *IEEE Signal Processing Advances in Wireless Communications Workshop*, pages 353–356, Paris, Apr. 16-18 1997.
- [AS98a] J. Ayadi and D.T.M. Slock. “Blind Multichannel Estimation Exploiting the Finite Symbol Alphabet”. In *Proc. of the European Signal Processing Conference: EUSIPCO 98*, Island of Rhodes, Greece, Sept. 1998.
- [AS98b] J. Ayadi and D.T.M. Slock. “Blind Multichannel Identification in the Stationary Colored Noise Case”. In *Proc. of the 9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing: SSAP 98*, Portland, Oregon, USA, Sept. 1998.
- [AS99a] J. Ayadi and D.T.M. Slock. “Blind Multichannel Estimation and Joint Order Detection by MMSE ZF Equalization”. In *Proc. of the IEEE Vehicular Technology Conference: VTC 99*, Amsterdam, Netherlands, Sept. 1999.
- [AS99b] J. Ayadi and D.T.M. Slock. “Estimation Multi-canal Aveugle par Egalisation MMSE ZF avec Détection Conjointe de l’Étalement Temporel”. In *Proc. of the 17^{ème} colloque GRETSI sur le traitement du signal et des images*, Vannes, France, September 1999.
- [AS99c] J. Ayadi and D.T.M. Slock. “Multichannel Estimation by Blind MMSE ZF Equalization”. In *IEEE Signal Processing Advances in Wireless Communications Workshop*, Annapolis, Maryland, USA, May 9-12 1999.
- [BK83] G. Bienvenu and L. Kopp. “Optimality of High Resolution Array Processing Using the Eigensystem Approach”. *IEEE Trans. ASSP*, 31(5):1235–1248, Oct. 1983.
- [BKAK78] R.R. Bitmead, S.Y. Kung, B.D.O. Anderson, and T. Kailath. “Greatest Common Divisors via Generalized Sylvester and Bezout Matrices”. *IEEE Trans. Autom. Control*, 23(6):1943–1947, Dec. 1978.

- [BM86] Y. Bresler and A. Macovski. “Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise”. *IEEE Trans. Acoust., Speech and Sig. Proc.*, ASSP-34:1081–1089, Oct. 1986.
- [Böh84] J.F. Böhme. “Estimation of Source Parameters by Maximum Likelihood and Nonlinear Regression”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, volume 1, pages 731–734, San Diego, CA, March 1984.
- [BR94] L. A. Baccala and S. Roy. “A New Blind Time-domain Channel Identification Method Based on Cyclostationarity”. *IEEE Signal Processing Letters*, 1(6):89–91, June 1994.
- [BS92] P. Balaban and J. Salz. “Optimum Diversity Combining and Equalization in Digital Data Transmission with Applications to Cellular Mobile Radio - Part I: Theoretical Considerations”. *IEEE Transactions on Communications*, 40(5):885–894, May 1992.
- [Cap69] J. Capon. “High-resolution Frequency-wavenumber Spectrum Analysis”. *Proceedings of the IEEE*, 57(8):1408–1419, August 1969.
- [CNP97] M. Cedervall, B. Ng, and A. Paulraj. “Structured Methods for Blind Multichannel Identification”. In *Proc. of the 13th international conference on Digital Signal Processing: DSP 97*, Santorini, Greece, July 1997.
- [Com86] P. Comon. “Estimation Multivariable Complexe”. *Traitement du Signal*, 3(2):97–101, 1986.
- [DAS97a] L. Deneire, J. Ayadi, and D.T.M. Slock. “Subspace Fitting Without Eigendecomposition”. In *Proc. of the 13th international conference on Digital Signal Processing: DSP 97*, Santorini, Greece, July 1997.
- [DAS97b] L. Deneire, J. Ayadi, and D.T.M. Slock. “Weighted and Unweighted Subspace Fitting without Eigendecomposition”. In *In Proc. of the IEEE 5th Symposium on Communications and Vehicular Technology in the Benelux*, Twente, Netherlands, Oct. 1997.
- [dCS] E. de Carvalho and D.T.M. Slock. “Cramér–Rao Bounds for Blind Multichannel Estimation”. In preparation, for *IEEE Transactions on Information Theory*.
- [dCS96] E. de Carvalho and D.T.M. Slock. “Maximum-Likelihood Blind Equalization of Multiple FIR Channels”. In *Proc. ICASSP 96 Conf.*, pages 2451–2454, Atlanta, Georgia, May 1996.

- [dCS97a] E. de Carvalho and D.T.M. Slock. “Asymptotic Performance of ML Methods for Semi-blind Channel Estimation”. In *Proc. 31th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 3-5 1997.
- [dCS97b] E. de Carvalho and D.T.M. Slock. “Cramer-Rao Bounds for Semi-Blind, Blind and Training Sequence Based Channel Estimation”. In *IEEE Signal Processing Advances in Wireless Communications Workshop*, pages 129–132, Paris, Apr. 16-18 1997.
- [Din97] Z. Ding. “Multipath Channel Identification Based on Partial System Information”. *IEEE Trans. SP*, SP-45:235–240, Jan. 1997.
- [DM95] Z. Ding and Z. Mao. “Knowledge Based Identification of Fractionally Sampled Channels”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Detroit, Michigan, May 1995.
- [For75] G.D. Forney. “Minimal Bases of Rational Vector Spaces, with Applications to Multivariable Linear Systems”. *SIAM Journal of Control*, 13(3):493–520, May 1975.
- [Fuc96] J.J. Fuchs. “Rectangular Pisarenko Method Applied to Source Localization”. *IEEE Trans. SP*, SP-44(10):2377–2383, Oct. 1996.
- [Gar91a] W.A. Gardner. “A New Method of Channel Identification”. *IEEE Trans. Communications*, 39(6):813–817, June 1991.
- [Gar91b] W.A. Gardner. “Exploitation of Spectral Redundancy in Cyclostationary Signals”. *IEEE Signal Processing Magazine*, 8(2):14–37, April 1991.
- [Gar93] W.A. Gardner. *An Introduction to Cyclostationary signals: Cyclostationarity in Communications and Signal Processing*. IEEE Press, 1993.
- [GC98] O. Grellier and P. Comon. “Blind Separation and Equalization of a Channel with MSK Inputs”. In *SPIE Conference*, pages 26–34, San Diego, July 19–24 1998.
- [GD97] D. Gesbert and P. Duhamel. “Robust Blind Identification and Equalization Based on Multi-step Predictors”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Munich, Germany, April 1997.
- [GH97] G. Giannakis and S.D. Halford. “Asymptotically Optimal Blind Fractionally Spaced Channel Estimation and Performance Analysis”. *IEEE Transactions on Signal Processing*, 45(7):1815–1830, July 1997.
- [GN95] M.I. Gurelli and C.L. Nikias. “EVAM: An Eigenvector-based Algorithm for Multichannel Blind Deconvolution of Input Colored Signals”. *IEEE Transactions on Signal Processing*, 43(1):134–149, Jan. 1995.

- [God80] D.N. Godard. “Self-recovering Equalization and Carrier Tracking in Two Dimensional Data Communication Systems”. *IEEE Transactions on Communications*, 28(11):1867–1875, Nov. 1980.
- [Gör97a] B. Göransson. *On Parametric Methods for Source Localization*. PhD thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, July 1997.
- [Gor97b] A. Gorokhov. *Séparation Autodidacte de Mélanges Convolutifs : Méthodes du Second Ordre*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, Paris, France, May 1997.
- [GvdV96] J. Götze and Alle-Jan van der Veen. “On-line Subspace Estimation Using a Schur-type Method”. *IEEE Trans. SP*, SP-44(6):1585–1589, June 1996.
- [Hay94] S. Haykin, editor. *Blind Deconvolution*. Prentice-Hall, Englewood Cliffs, NJ, 1994.
- [HB96] G. Harikumar and Y. Bresler. “Analysis and Comparative Evaluation of Techniques for Multichannel Blind Deconvolution”. In *Proc. 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pages 332–335, Corfu, Greece, June 24-26 1996.
- [Hua96] Y. Hua. “Fast Maximum Likelihood for Blind Identification of Multiple FIR Channels”. *IEEE Trans. SP*, SP-44(3):661–672, Mar. 1996.
- [JSE⁺98] C.R. Johnson, P. Schniter, T.J. Enres, J.D. Behm, D.R. Brown, and R.A. Casas. “Blind Equalization Using the Constant Modulus Criterion: A Review”. *Proceedings of the IEEE*, 86(10):1927–1950, Oct. 1998.
- [Kai80] T. Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, NJ, 1980.
- [Kay93] S.M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [KJO] M. Kristensson, M. Jansson, and B. Ottersten. “Modified IQML and Weighted Subspace Fitting without Eigendecomposition”. Submitted to *Signal Processing*.
- [KJO98] M. Kristensson, M. Jansson, and B. Ottersten. “Modified IQML and a Statistically Efficient Method for Direction Estimation without Eigendecomposition”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Seattle, WA, May 12-15 1998.
- [KM95] D. Kundu and A. Mitra. “Estimating the Parameters of Exponentially Damped/Undamped Sinusoids in Noise: A Non-iterative Approach”. *Signal Processing (EURASIP)*, 46:363–368, 1995.

- [KOS96] M. Kristensson, B.E. Ottersten, and D.T.M. Slock. "Blind Subspace Identification of a BPSK Communication Channel". In *Proc. 30th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1996.
- [Kri98] M. Kristensson. *On Parameter Estimation in Wireless Communications, Sensor Array Processing and Spectral Analysis*. PhD thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, 1998.
- [KSS86] R. Kumaresan, L.L. Sharf, and A.K. Shaw. "An Algorithm for Pole-zero Modeling and Spectral Analysis". *IEEE Trans. Acoust., Speech and Sig. Proc.*, ASSP-34(3):637–640, June 1986.
- [KY88] K. Konstantinides and K. Yao. "Statistical Analysis of Effective Singular Values in Matrix Rank Determination". *IEEE Trans. ASSP*, pages 757–763, May 1988.
- [LX94a] H. Liu and G. Xu. "A deterministic Approach to Blind Symbol Estimation". In *Proc. of the 28th Asilomar Conference on Signals, Systems & Computers*, Pacific Grove, Oct. 31 - Nov. 2 1994.
- [LX94b] H. Liu and G. Xu. "A deterministic Approach to Blind Symbol Estimation". *IEEE Signal Processing Letters*, pages 205–207, Dec. 1994.
- [LXTK96] H. Liu, G. Xu, L. Tong, and T. Kailath. "Recent Developments in Blind Channel Equalization: From Cyclostationarity to Subspaces". *Signal Processing*, 50(12):83–99, April 1996.
- [MDCM95] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue. "Subspace Methods for the Blind Identification of Multichannel FIR filters". *IEEE Trans. SP*, SP-43(2):516–526, Feb. 1995.
- [NCP] B.C. Ng, M. Cedervall, and A. Paulraj. "A Structured Channel Estimator for Maximum Likelihood Sequence Detection in Multipath Fading Channels". submitted to *IEEE Transactions on Vehicular Technology*, Jan. 1997.
- [NCP97] B.C. Ng, M. Cedervall, and A.J. Paulraj. "A Structured Channel Estimation for Maximum-likelihood Sequence Detection". *IEEE Communications Letters*, 1(2), March 1997.
- [NGP98] B.C. Ng, D. Gesbert, and A. Paulraj. "A Semi-blind Approach to Structured Channel Equalization". In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Seattle, WA, May 12-15 1998.
- [OS91] M.R. Osborne and G.K. Smyth. "A Modified Prony Algorithm for Fitting Functions Defined by Difference Equations". *SIAM J. Sci. Stat. Comput.*, 12(2):362–382, 1991.

- [Ott96] B. Ottersten. “Array Processing for Wireless Communications”. In *Proc. 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pages 466–473, Corfu, Greece, June 24-26 1996.
- [OVSN93] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai. “Exact and Large Sample ML Techniques for Parameter Estimation and detection in Array Processing”. Technical Report TRITA-SB-9302 Report, Royal Institute of Technology, Stockholm, Sweden, Feb. 1993.
- [PMDCM99] L. Perros-Meilhac, P. Duhamel, P. Chevalier, and E. Moulines. “Blind Knowledge Based Algorithms Based on Second Order Statistics”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Phoenix, Arizona, USA, Mar. 15-19 1999.
- [Pro95] J.G. Proakis. *Digital Communications*. McGraw-Hill, 1995.
- [Sat75] Y. Sato. “A Method of Self-recovering Equalization for Multilevel Amplitude-modulation Systems”. *IEEE Transactions on Communications*, 23:679–682, June 1975.
- [SBEM90] S.C. Swales, M.A. Beach, D.J. Edwards, and J.P. McGeehan. “The Performance Enhancement of Multibeam Adaptive Base-station Antennas for Cellular Land Mobile Radio Systems”. *IEEE Transactions on Vehicular Technology*, 39(1):56–67, Feb. 1990.
- [Sch91] L.L. Scharf. *Statistical Signal Processing*. Addison-Wesley, Reading, MA, 1991.
- [Slo93] Dirk T.M. Slock. “Blind Fractionally-Spaced Equalization and Channel Identification from Second-Order Statistics or Data”. Technical Report RR N^o 93-002, Institut Eurécom, Sophia Antipolis, France, 17 Aug. 1993.
- [Slo94a] D.T.M. Slock. “Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Adelaide, Australia, April 1994.
- [Slo94b] D.T.M. Slock. “Blind Joint Equalization of Multiple Synchronous Mobile Users Using Oversampling and/or Multiple Antennas”. In *Proc. 28th Asilomar Conf. on Signals, Systems and Computers*, pages 1154–1158, Pacific Grove, CA, Oct. 31 - Nov. 2 1994.
- [Slo94c] D.T.M. Slock. “Subspace Techniques in Blind Mobile Radio Channel Identification and Equalization using Fractional Spacing and/or Multiple Antennas”. In *Proc. 3rd International Workshop on SVD and Signal Processing*, Leuven, Belgium, Aug. 22-25 1994.

- [Slo95] D.T.M. Slock. “Blind Joint Equalization of Multiple Synchronous Mobile Users for Spatial Division Multiple Access”. In E. Biglieri and M. Luise, editors, *Signal Processing in Telecommunications, Proc. of the 7th Tyrrhenian Int. Workshop on Dig. Communications*. Springer-Verlag, 1995.
- [Slo97] D.T.M. Slock. “From Sinusoids in Noise to Blind Deconvolution in Communications”. In A. Paulraj, V. Roychowdhury, and C.D. Schaper, editors, *Communications, Computation, Control and Signal Processing. A tribute to Thomas Kailath*. Kluwer Academic Publishers, Boston, 1997.
- [SLS97] P. Stoica, J. Li, and T. Söderström. “On the Inconsistency of IQML”. *Signal Processing (EURASIP)*, 56(2):185–190, Jan. 1997.
- [Smi96] D.L. Smith. *Blind Channel Identification and Equalization Using Second-order Cyclostationarity*. PhD thesis, Pennsylvania State University, Aug. 1996.
- [SN98] P. Stoica and B.C. Ng. “On the Cramer-Rao Bound Under Parametric Constraints”. *IEEE Signal Processing Letters*, 5(7):177–179, July 1998.
- [SP94] D.T.M. Slock and C.B. Papadias. “Blind Fractionally-Spaced Equalization Based on Cyclostationarity”. In *Proc. Vehicular Technology Conf.*, pages 1286–1290, Stockholm, Sweden, June 1994.
- [SP95] D.T.M. Slock and C.B. Papadias. “Further Results on Blind Identification and Equalization of Multiple FIR Channels”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Detroit, Michigan, May 1995.
- [SSG94] S.V. Schell, D.L. Smith, and W. Gardner. “Blind Channel Identification Using 2nd-Order Cyclostationary Statistics”. In *Proc. EUSIPCO 94, VII European Signal Processing Conference*, Edinburgh, Scotland, Sept. 13-16 1994.
- [TA83] J.R. Treichler and B.G. Agee. “A New Approach to Multipath Correction of Constant Modulus Signals”. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31:459–472, April 1983.
- [Tal96] S. Talwar. *Blind Space-time Algorithms for Wireless Communication Systems*. PhD thesis, Stanford University, Stanford, CA, January 1996.
- [TP98] L. Tong and S. Perreau. “Multichannel Blind Identification: From Subspace to Maximum Likelihood Methods”. *Proceedings of the IEEE*, 86(10):1951–1968, Oct. 1998.

- [TS98a] H. Trigui and D.T.M. SLock. “Cochannel Interference Cancellation within the Current GSM Standard”. In *Proc. of the IEEE International Conference on Universal Personal Communications (ICUPC)*, Florence, Italy, Oct. 5-9 1998.
- [TS98b] H. Trigui and D.T.M. SLock. “Training Sequence Aided Multichannel Identification in the Presence of Interference and Noise”. In *Proc. of the IEEE Global Telecommunications Conference (Globecom)*, Sydney, Australia, Nov. 8-12 1998.
- [TS99] H. Trigui and D.T.M. SLock. “Training Sequence Based Multiuser Channel Identification for Cochannel Interference Cancellation in GSM”. In *Proc. of the Signal Processing Advances in Wireless Communications (SPAWC)*, Annapolis, MD, USA, May 9-12 1999.
- [TX97] M.K. Tsatsanis and Z. Xu. “Constrained Optimization Methods for Blind Equalization of Multiple FIR Channels”. In *Proc. 31th Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 3-5 1997.
- [TX99] M.K. Tsatsanis and Z. Xu. “Constrained Optimization Methods for Direct Blind Equalization”. *IEEE Journal on Selected Areas in Communications*, 17(3):424–433, March 1999.
- [TXK91] L. Tong, G. Xu, and T. Kailath. “A New Approach to Blind Identification and Equalization of Multipath Channels”. In *Proc. of the 25th Asilomar Conference on Signals, Systems & Computers*, pages 856–860, Pacific Grove, CA, Nov. 1991.
- [TXK94] L. Tong, G. Xu, and T. Kailath. “Blind Identification and Equalization Based on Second Order Statistics: a Time-domain Approach”. *IEEE Trans. on Information Theory*, 40(2):340–349, March 1994.
- [TZ] L. Tong and Q. Zhao. “Joint Order Detection and Blind Channel Estimation by Least Squares Smoothing”. submitted to *IEEE Trans. SP* on Feb. 1998.
- [TZ98] L. Tong and Q. Zhao. “Blind Channel Estimation by Least Squares Smoothing”. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Seattle, WA, May 12-15 1998.
- [Vib93] M. Viberg. “Sensitivity of Parametric Direction Finding to Colored Noise Fields and Undermodeling”. *Signal Processing (EURASIP)*, 34(2), 1993.
- [XLTK95] G. Xu, H. Liu, L. Tong, and T. Kailath. “A Least-Squares Approach to Blind Channel Identification”. *IEEE Trans. SP*, SP-43(12):2982–2993, Dec. 1995.

- [ZDC99] J. Zhu, Z. Ding, and X. Cao. “Column-Anchored Zeroforcing Blind Equalization for Multiuser Wireless FIR Channels”. *IEEE Journal on Selected Areas in Communications*, 17(3):411–423, March 1999.
- [ZT97] H.H. Zeng and L. Tong. “Blind Channel Estimation Using the Second-order Statistics: Asymptotic Performance and Limitations”. *IEEE Transactions on Signal Processing*, 45(8):2060–2071, Aug. 1997.
- [ZT98] Q. Zhao and L. Tong. “Semi-blind Equalization by Least Squares Smoothing”. In *Proc. 32nd Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1998.